Lectures EEL 6545

1. Introduction. The Additive Law (1/9/2017)
2. Conditional Probability, Permutations/Combination, pdfs, delta Functions (1/18/2107)
3. pdfs, Cumulative distributions, Expected Values, Change of Variable (1/23/17)
4. pdf Formulation of Conditional Probability, Gaussian Distribution, Bivariate pdf Theory, Bivariate Gaussian (1/25/17)
5. Bivariate Gaussian, Sums of RandomVariables, Central Limit Theorem (1/30/17)
6. Statistical Description of Random Processes, Examples (2/1/17)
7. Prediction. Stationarity and Ergodicity of Data. (2/6/17)
8. Convergence in Probability. Estimating Autocorrelations from Data (2/8/17) (ok)
9. Time Averages versus Ensemble Averages, Impulse Response and Convolution, Fourier Theory (2/15/17)
10. Fourier Analysis in One Hour (2/20/17)
11. Wiener-Khitchine Theory, Bartlett's Method (2/22/17)
12. Power Spectral Density, Impulse Response/Transfer Function, White Noise (2/27/17)
13. ARMA, Yule-Walker (3/1/17)
14. Random Sine Wave, Bernoulli, Binomial, Poisson, Shot Noise (3/6/17)
15. Queueing Theory (3/8/17)
16. Frequency Domain, Random Arrivals (3/20/17)
17. Random Walk, Wiener Process, LMSE Prediction for a Random Process (3/22/17)
18. Review of Power and Noise in Electrical Circuits, Wiener Filter, Kalman Filter (3/29/17)
19. More Kalman Filter (4/3/17)
20. Transfer Function review (4/5/17 "first copy")
21. Still More Kalman Filter (4/5/17 "second copy")
22. Markov Processes (4/12/17)

Lecture 1
(1/9/2017)

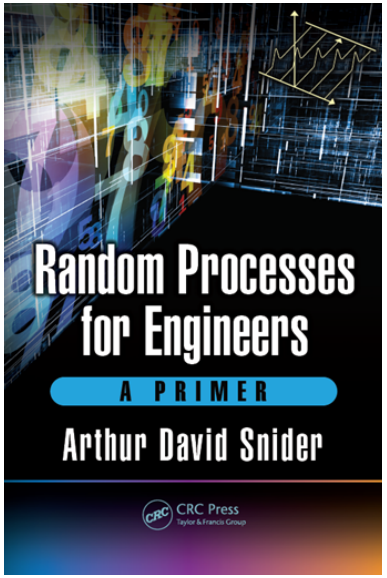

Random Processes for Engineers
A PRIMER
Arthur David Snider
CRC Press
Taylor & Francis Group

Figure 1.1 Billiard balls and sky

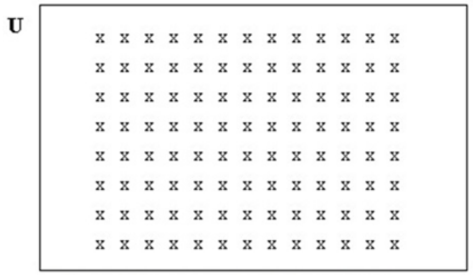


Figure 1.2 The Universe of Elemental Events

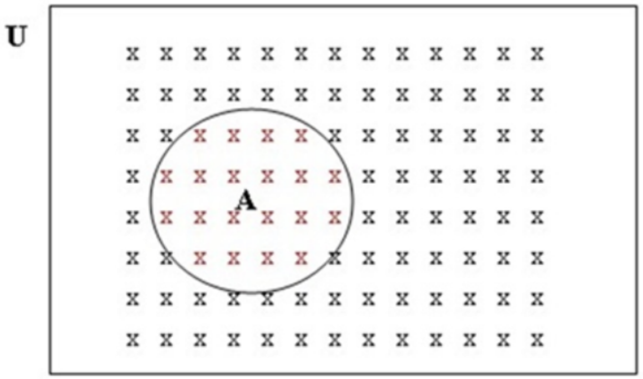


Figure 1.3 The truth set of statement A

Lecture 2
(1/18/2917)

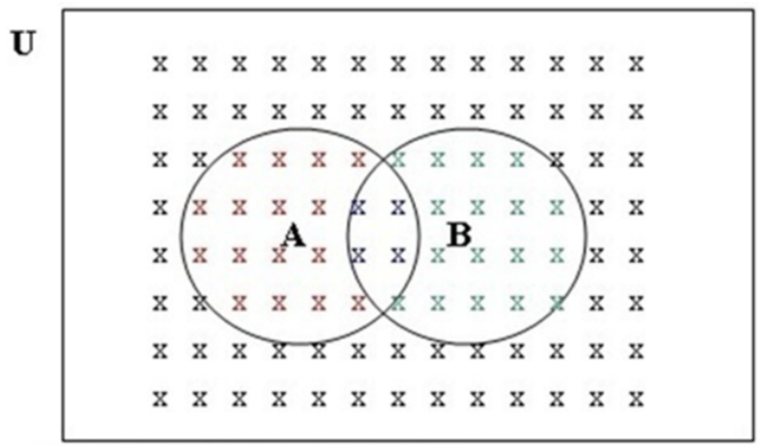


Figure 1.4 Additive Law of Probability

$$p(A \cup B) = p(A) + p(B) - p(A \cap B)$$

Prob(ball is prime) = 6/15

(2, 3, 5, 7, 11, 13)

Pr(ball is prime, given ball is striped) = 2/7

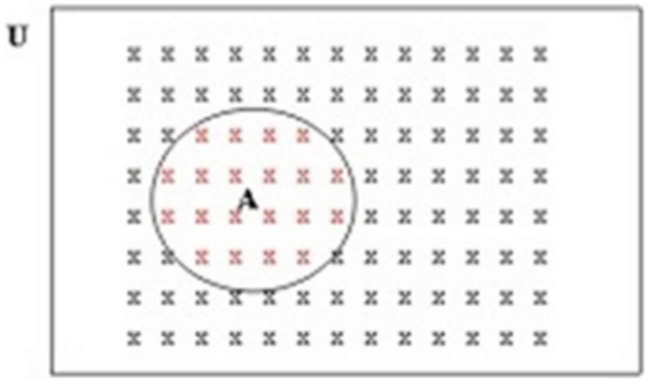


Figure 1.5 Universe

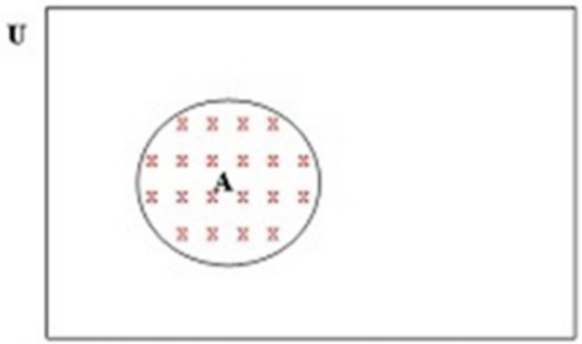


Figure 1.6 Conditional universe, given A

$$p(\mathbf{A}) = \sum_{x \in \mathbf{A}} p(x) \qquad p(\mathbf{A} \mid \mathbf{A}) = \sum_{x \in \mathbf{A}} p(x \mid \mathbf{A}) = 1$$

$$\sum_{x \in A} \frac{p(x)}{p(\mathbf{A})} = 1 \qquad p(x \mid \mathbf{A}) = \frac{p(x)}{p(A)}$$

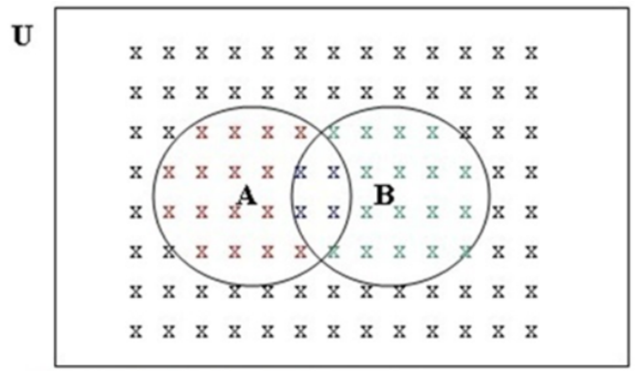


Figure 1.7 a priori Universe

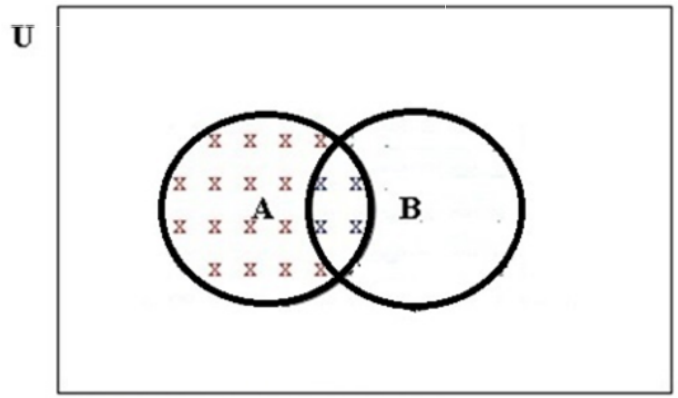


Figure 1.8 Conditional Universe

$$p(B|A) = \sum_{x \in A \cap B} p(x|A) = \sum_{x \in A \cap B} \frac{p(x)}{p(A)}$$

$$= \frac{1}{p(A)} \sum_{x \in A \cap B} p(x) = \frac{p(A \cap B)}{p(A)}$$

$$p(B \mid A) = \sum_{x \in A \cap B} p(x \mid A) = \sum_{x \in A \cap B} \frac{p(x)}{p(A)}$$

$$= \frac{1}{p(A)} \sum_{x \in A \cap B} p(x) = \frac{p(A \cap B)}{p(A)}$$

$$p(A \cap B) = p(A)p(B \mid A) = p(B)p(A \mid B)$$

$$p(B \mid A) = \frac{p(B)p(A \mid B)}{p(A)}$$

Bayes' Theorem

(Remember)

Prob(ball is prime) = 6/15

(2, 3, 5, 7, 11, 13)

Pr(ball is prime, given ball is striped) = 2/7

$$p(\textit{prime number} \mid \textit{striped ball})$$

$$= \frac{p(\textit{prime number} \,\&\, \textit{striped ball})}{p(\textit{striped ball})}$$

$$= \frac{p(\textit{balls } 11 \textit{ or } 13)}{p(\textit{balls } 9 \textit{ through } 15)}$$

$$= \frac{2/15}{7/15} = \frac{2}{7}$$

A and B are independent when any of the following equivalent conditions are true:

$$p(B|A) = p(B)$$

$$p(A|B) = p(A)$$

$$p(A \cap B) = p(A)p(B)$$

$$p(\mathbf{A} \cup \mathbf{B}) =$$
$$p(\mathbf{A}) + p(\mathbf{B}) - p(\mathbf{A} \cap \mathbf{B})$$

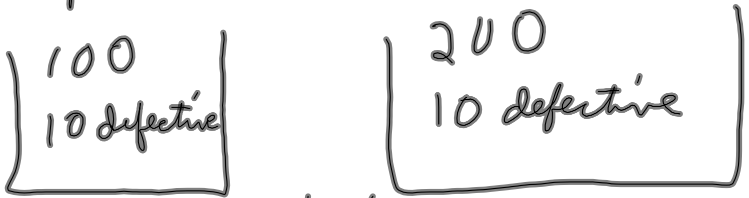
$$p(A \cap B) = p(A)p(B \mid A) = p(B)p(A \mid B)$$

$$p(A \mid B) = \frac{p(B \mid A)p(A)}{p(B)}$$

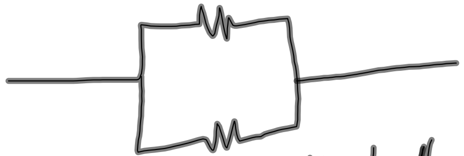
Independence

$$p(A \cap B) = p(A)p(B)$$

Example 3



Pick a box
pick 2 resistors



What is prob. that this conducts?

Yes if either resistor is good.

$P = P(1^{st} \text{ conducts or } 2^{nd} \text{ conducts})$

$= P(1^{st} \checkmark) + P(2^{nd} \checkmark) - P(\text{both conduct})$

$P(1^{st} \text{ conducts}) = P(1^{st} \text{ box \& conduct})$
$+ P(2^{nd} \text{ box \& " })$
$- P(1^{st} \text{ \& } 2^{nd} \text{ box \& conduct})$ → 0

$P(1^{st} \text{ box \& conducts}) = P(1^{st} \text{ box}) \times P(\text{conducts} \mid \text{given } 1^{st} \text{ box})$

$= \frac{1}{2} \frac{90}{100}$

$P(2^{nd} \text{ box \& cond}) = \frac{1}{2} \times \frac{190}{200}$

p(Box 1) = p(Box 2) = .5

If Box 1 is selected, p(1st resistor conducts) = 90/100; same for p(2nd resistor conducts)

But p(2nd resistor conducts | 1st resistor conducts) = 89/99

So p(both resistors conduct) = [90/100] [89/99]

$$p(A \cap B) = p(A)p(B \mid A) = p(B)p(A \mid B)$$

p(one or the other conducts) =

90/100 + 90/100 - [90/100][89/99]

p([Box 1] & [1st *or* 2nd resistor conducts])
+ p([Box 2] & [1st *or* 2nd resistor conducts])

$$\frac{1}{2}\left\{\left(\frac{90}{100}\right)+\left(\frac{90}{100}\right)-\left(\frac{90}{100}\right)\left(\frac{89}{99}\right)\right\}$$

$$+\frac{1}{2}\left\{\left(\frac{190}{200}\right)+\left(\frac{190}{200}\right)-\left(\frac{190}{200}\right)\left(\frac{189}{199}\right)\right\}$$

= Prob(parallel circuit conducts)

If the parallel branch does not conduct, what is the probability that the resistors come from box 1?

$$p(B \mid A) = \frac{p(B)p(A \mid B)}{p(A)}$$

p(Box 1 chosen | parallel doesn't conduct)

=

p(parallel doesn't conduct | Box 1)

x

p(Box 1)

÷

p(parallel doesn't conduct)

p(Box 1 chosen | parallel doesn't conduct)

=

p(parallel doesn't conduct | Box 1)

1 -

{90/100 + 90/100 - [90/100][89/99]}

x

p(Box 1) 1/2

÷

p(parallel doesn't conduct)

1 -

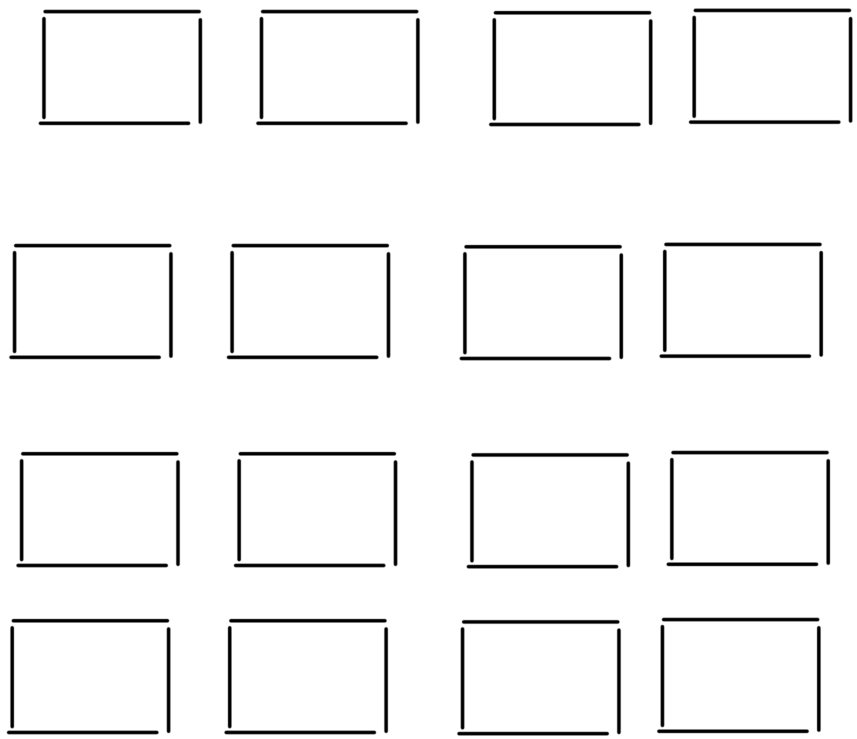
$$\frac{1}{2}\left\{\left(\frac{90}{100}\right)+\left(\frac{90}{100}\right)-\left(\frac{90}{100}\right)\left(\frac{89}{99}\right)\right\}$$

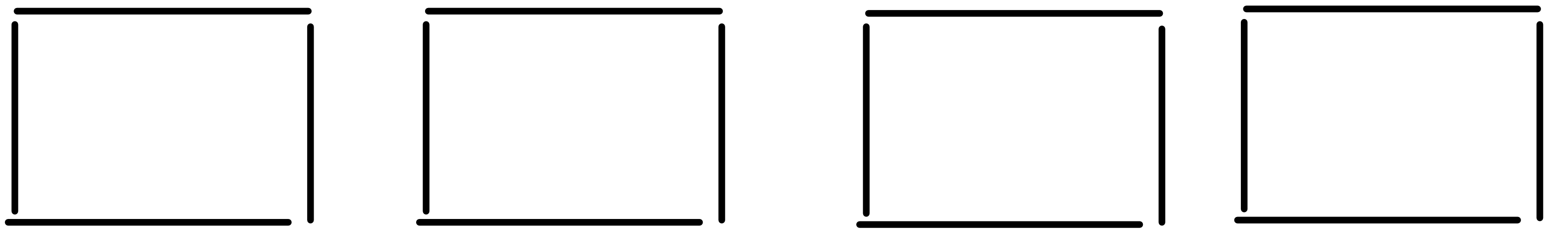
$$+\frac{1}{2}\left\{\left(\frac{190}{200}\right)+\left(\frac{190}{200}\right)-\left(\frac{190}{200}\right)\left(\frac{189}{199}\right)\right\}$$

$$\frac{\left[1-\left\{\left(\frac{90}{100}\right)+\left(\frac{90}{100}\right)-\left(\frac{90}{100}\right)\left(\frac{89}{99}\right)\right\}\right]\left[\frac{1}{2}\right]}{1-\left[\frac{1}{2}\left\{\left(\frac{90}{100}\right)+\left(\frac{90}{100}\right)-\left(\frac{90}{100}\right)\left(\frac{89}{99}\right)\right\}+\frac{1}{2}\left\{\left(\frac{190}{200}\right)+\left(\frac{190}{200}\right)-\left(\frac{190}{200}\right)\left(\frac{189}{199}\right)\right\}\right]}$$

Assignment (not graded):
Problem 3b and 3d.

How many ways
are there of ordering
the set of numbers
$\{1, 2, 3, 4\}$?





[1 x x x]   [2 x x x]   [3 x x x]   [4 x x x]

[1 x x x]   [2 x x x]   [3 x x x]   [4 x x x]

[1 2 × ×]   [2 1 × ×]   [3 1 × ×]   [4 1 × ×]

[1 3 × ×]   [2 3 × ×]   [3 2 × ×]   [4 2 × ×]

[1 4 × ×]   [2 4 × ×]   [3 4 × ×]   [4 3 × ×]

[1 x x x]	[2 x x x]	[3 x x x]	[4 x x x]
[1 2 × ×]	[2 1 × ×]	[3 1 × ×]	[4 1 × ×]
[1 3 × ×]	[2 3 × ×]	[3 2 × ×]	[4 2 × ×]
[1 4 × ×]	[2 4 × ×]	[3 4 × ×]	[4 3 × ×]
[1 2 3 4]	[2 1 3 4]	[3 1 2 4]	[4 1 2 3]
[1 2 4 3]	[2 1 4 3]	[3 1 4 2]	[4 1 3 2]
[1 3 2 4]	[2 3 1 4]	[3 2 1 4]	[4 2 1 3]
[1 3 4 2]	[2 3 4 1]	[3 2 4 1]	[4 2 3 1]
[1 4 2 3]	[2 4 1 3]	[3 4 1 2]	[4 3 1 2]
[1 4 3 2]	[2 4 3 1]	[3 4 2 1]	[4 3 2 1]

4 x 3 x 2 (x 1) = 4!

Permutations The number of permutations of n objects equals $n!$

There are about one trillion ways of lining up all 15 billiard balls:

$$15! \approx 1.3077 \times 10^{12}.$$

How many ways are there of selecting a *pair* (like 1 and 3) from [1 2 3 4]?

[1 2 : 3 4]	[2 1 : 3 4]	*[3 1 : 2 4]*	[4 1 : 2 3]
[1 2 : 4 3]	[2 1 : 4 3]	*[3 1 : 4 2]*	[4 1 : 3 2]
[1 3 : 2 4]	[2 3 : 1 4]	[3 2 : 1 4]	[4 2 : 1 3]
[1 3 : 4 2]	[2 3 : 4 1]	[3 2 : 4 1]	[4 2 : 3 1]
[1 4 : 2 3]	[2 4 : 1 3]	[3 4 : 1 2]	[4 3 : 1 2]
[1 4 : 3 2]	[2 4 : 3 1]	[3 4 : 2 1]	[4 3 : 2 1]

Number of permutations
divided by
number of times a pair gets counted.

Combinations The number of combinations of n objects, taking n_1 at a time, equals

$$\frac{n!}{n_1!(n-n_1)!}$$

Among a set of 15 unpainted billiard balls, there are 6,435 ways of selecting the 7 balls to be painted with stripes:

$15!/(7!8!) = 6,435.$

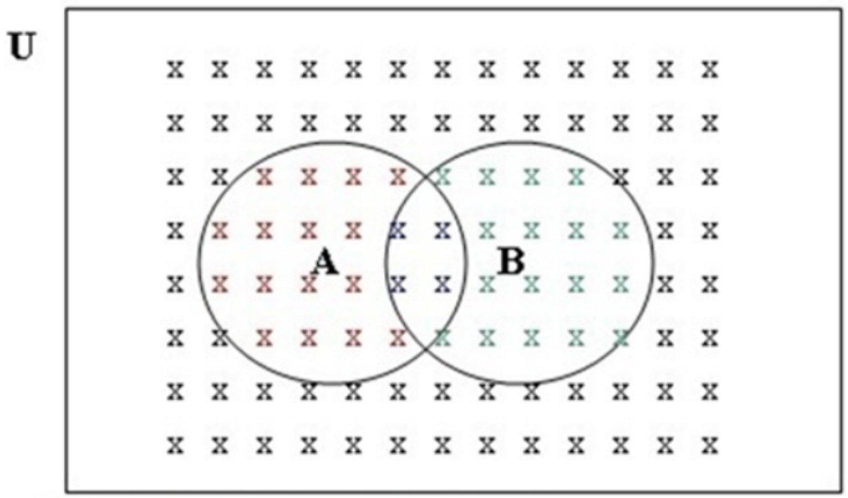
Partitions The number of ways of partitioning n objects into sets of size $n_1, n_2, \ldots, n_k$ (where $n = n_1 + n_2 + \ldots + n_k$) equals

$$\frac{n!}{n_1!n_2!\cdots n_k!}$$

{2 4 5 : 10 11 13 14 15 : 1 3 6 7 8 9 12}

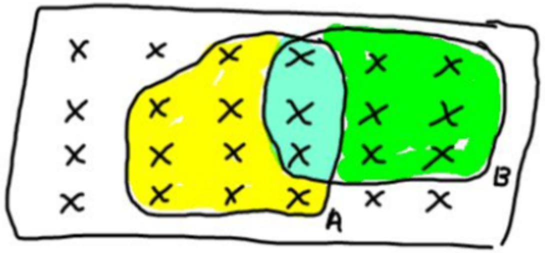
$$15!/(3!5!7!) = 360{,}360$$

Lecture 3
(1/23/2017)



Independence

$$p(A \cap B) = p(A)\,p(B)$$



$$P(A \cap B) = P(A) \times P(B)$$

$$\sum_{x \in A \cap B} p(x) = \left[\sum_{x \in A} p(x)\right] \times \left[\sum_{x \in B} p(x)\right]$$

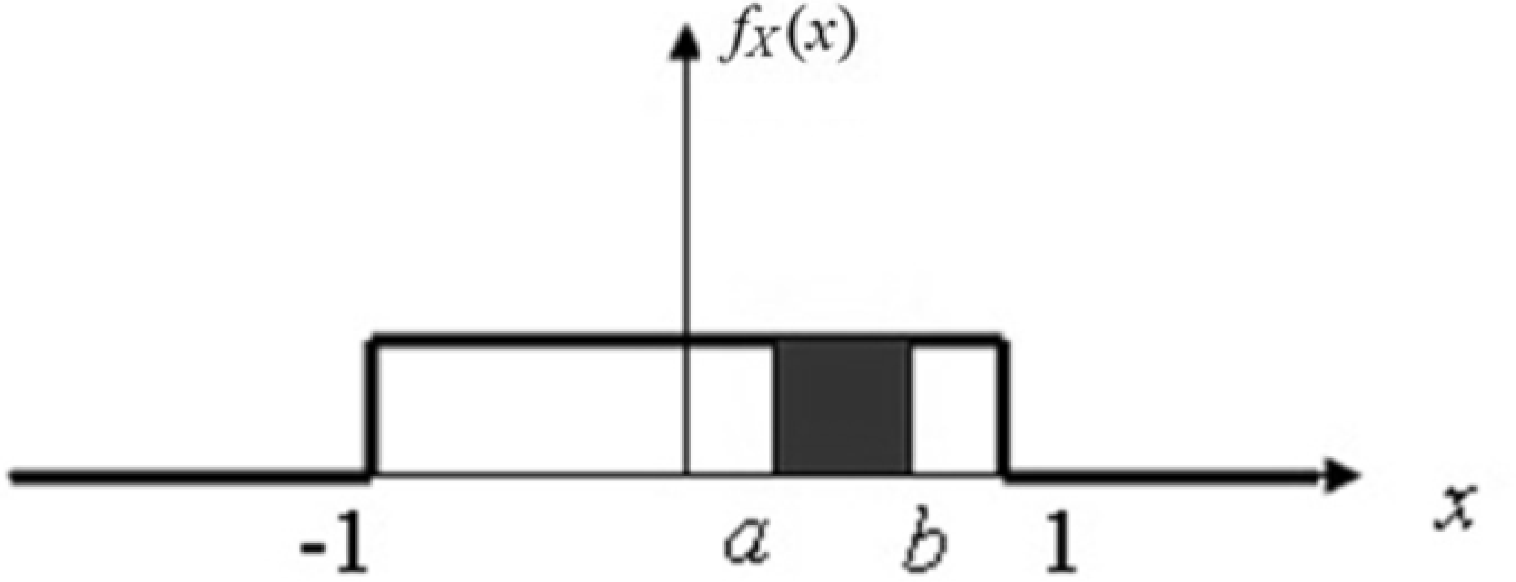


Figure 1.9 Probability density function

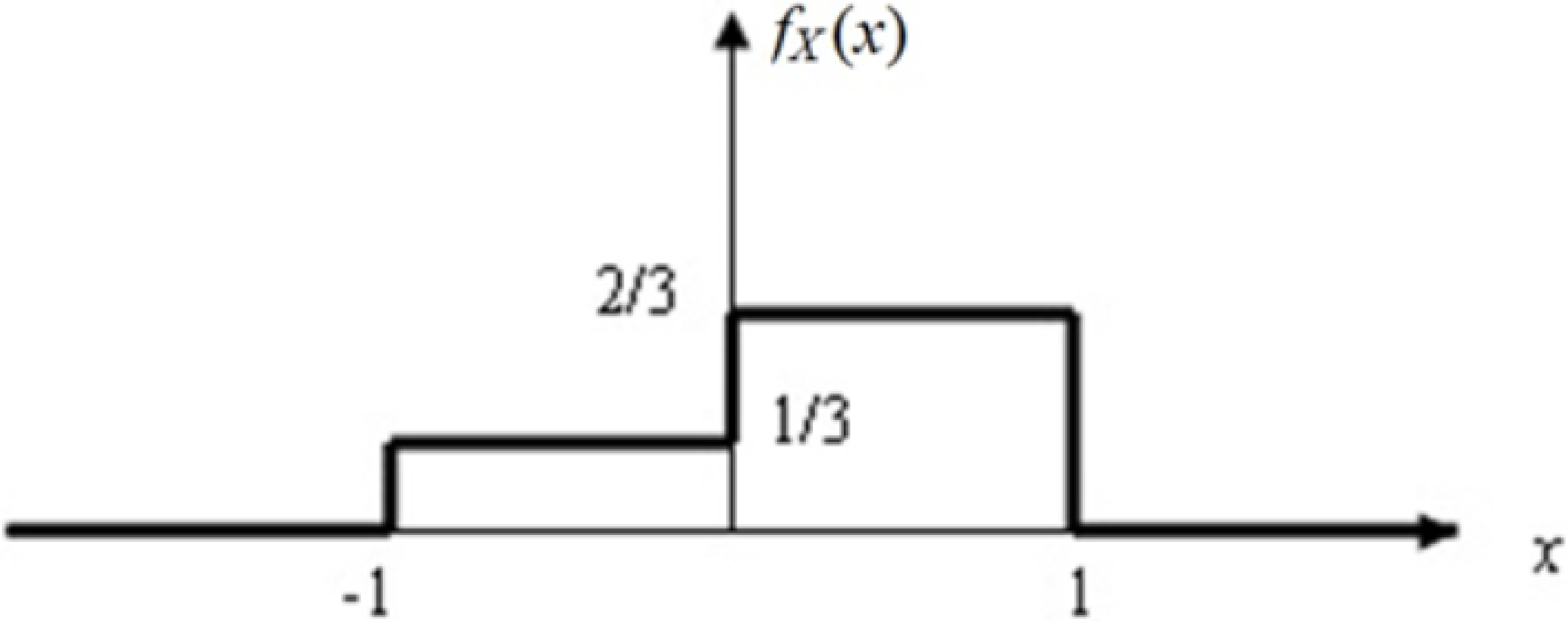


Figure 1.10 Skewed pdf

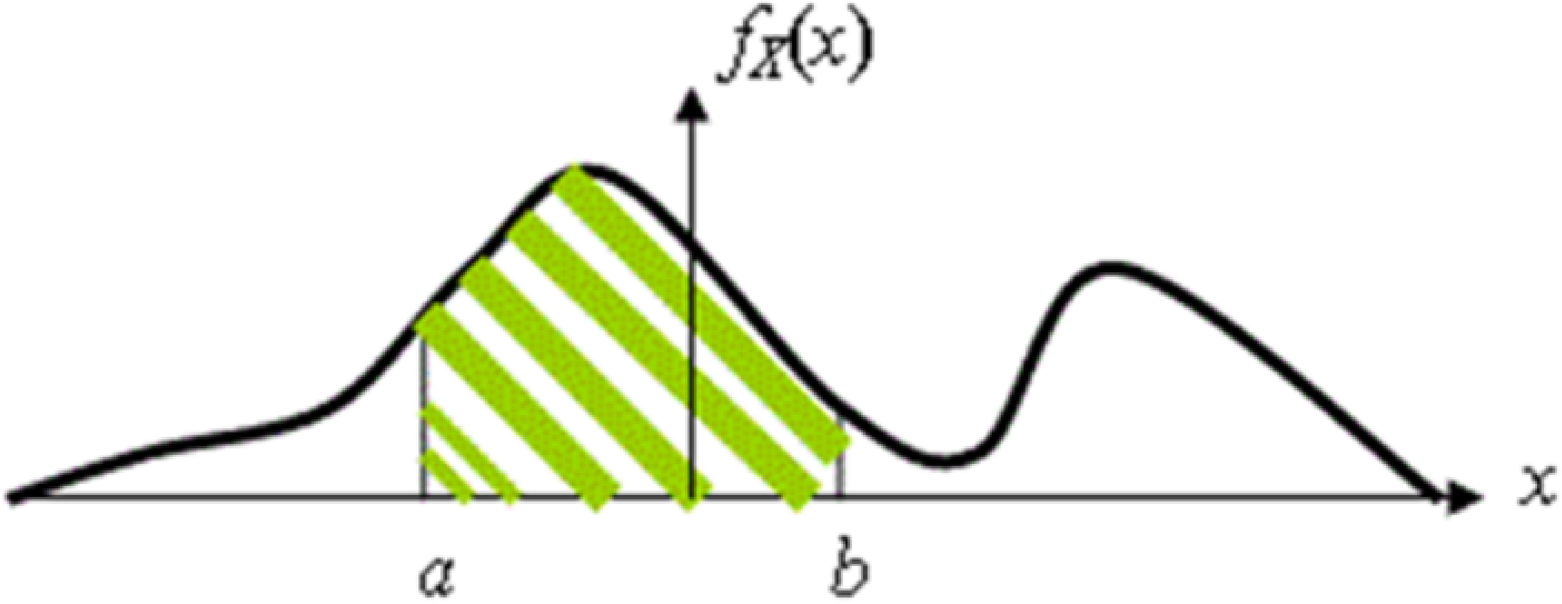
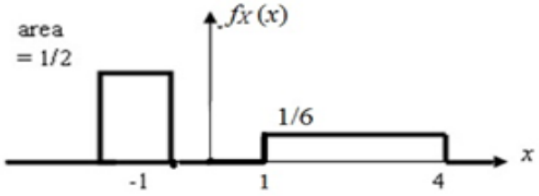
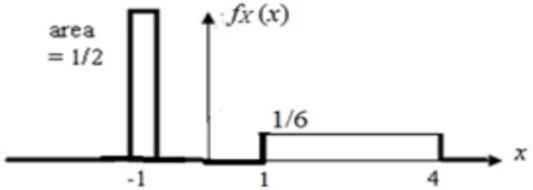
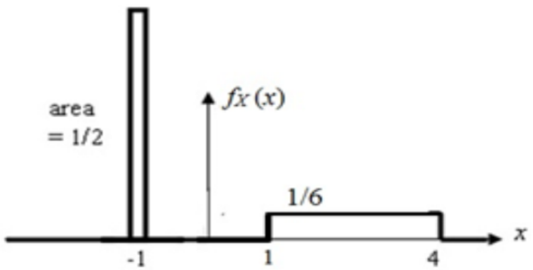


Figure 1.11. $f_X(x) \geq 0, \int_{-\infty}^{\infty} f_X(x)\,dx = 1$

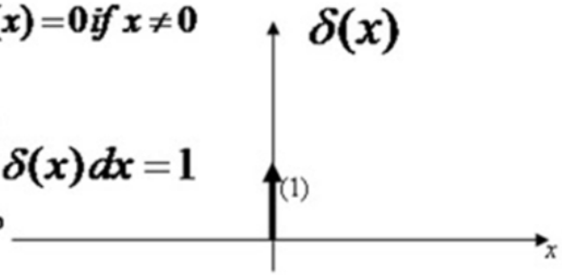






$$\delta(x) = 0 \; if \; x \neq 0$$

$$\int_{-\infty}^{\infty} \delta(x)\,dx = 1$$



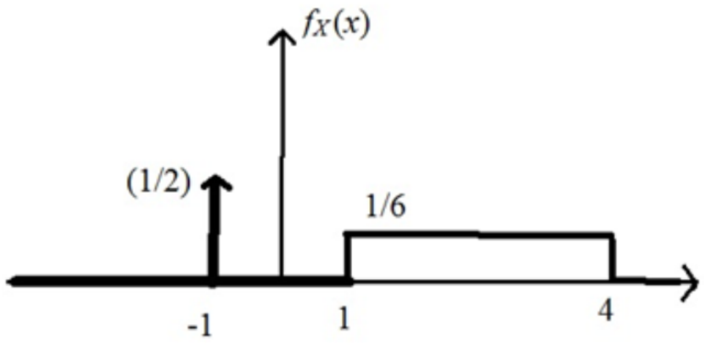
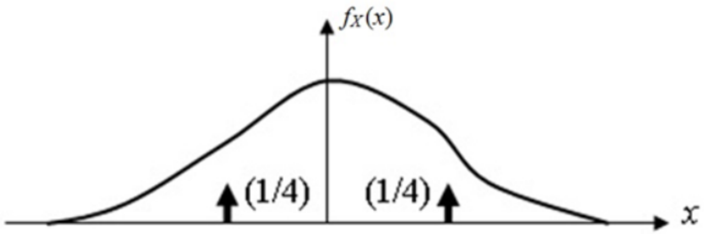


Figure 1.14 and 1.15
Mixed discrete and continuous pdf



$$f(x)\ \delta(x) \equiv f(0)\ \delta(x)$$

$$f(x)\ \delta(x-x_0) \equiv f(x_0)\ \delta(x-x_0)$$

$$\int_{-\infty}^{\infty} f(x)\delta(x-x_0)\,dx = \int_{-\infty}^{\infty} f(x_0)\delta(x-x_0)\,dx = f(x_0)$$

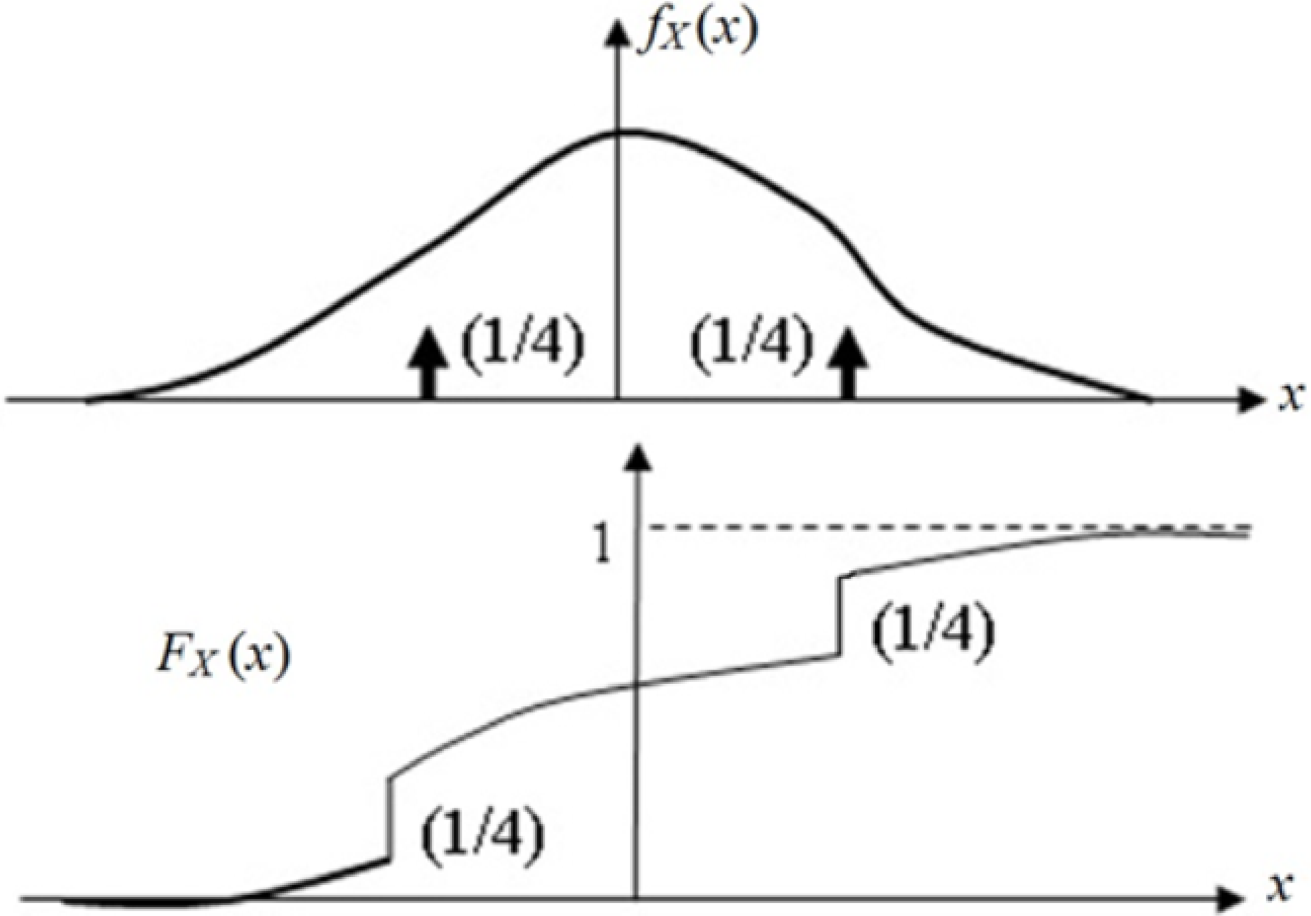


Figure 1.16 Cumulative distribution function

Summary: Important Facts about Continuous Random Variables

$f_X(x)$ is the probability density function ("pdf")

$f_X(x)\,dx$ equals the probability that X lies in the infinitesimal interval $x < X < x + dx$.

The delta-function property: $\int_{-\infty}^{\infty} f(x)\,\delta(x - x_0)\,dx = \int_{-\infty}^{\infty} f(x_0)\,\delta(x - x_0)\,dx = f(x_0)$

Cumulative Distribution Function: $F_X(x) = \int_{-\infty}^{x} f_X(\xi)\,d\xi$

Conditional pdf:

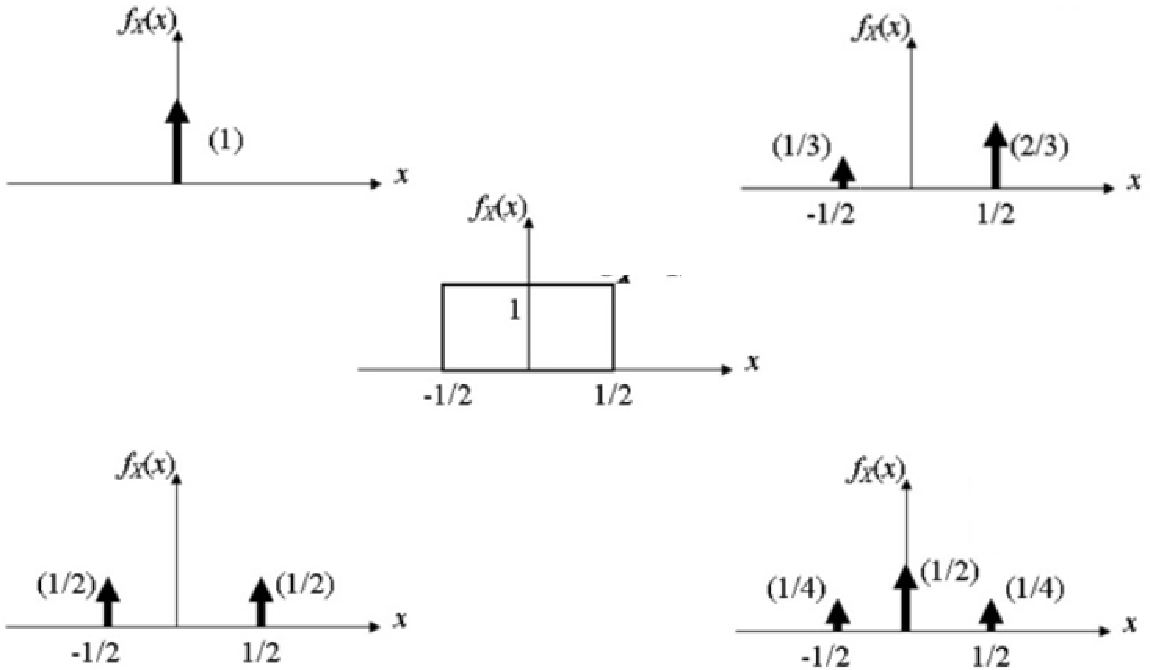
$$f_{X|A}(x) = \begin{cases} \frac{f_X(x)}{p(A)} & \text{if } x \text{ is in the truth set of } A \\ 0 & \text{otherwise} \end{cases}$$

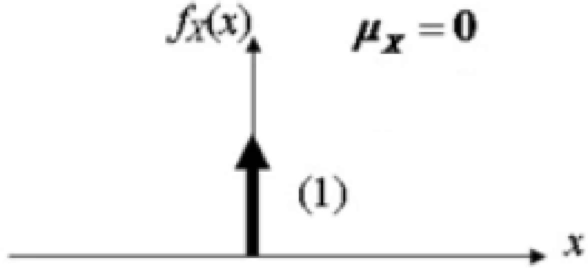
#1	.5000000000…
#2	.3333333333…
#3	.70710678118…
#4	.78539816339…
…	…
#n	.$xxxxxxxxxxx$…
…	…

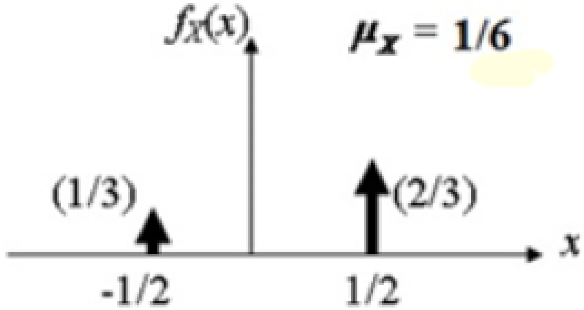
Figure 1.17 Proposed enumeration of (0,1)

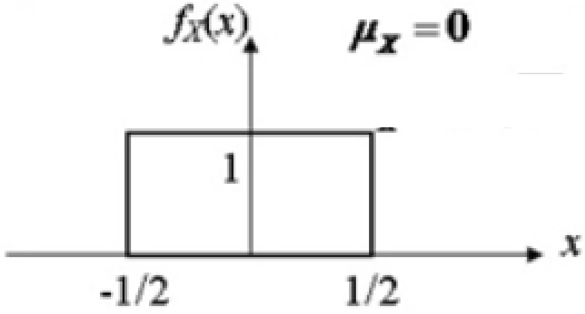
The first moment is the *mean* of X and is often denoted μ_X.

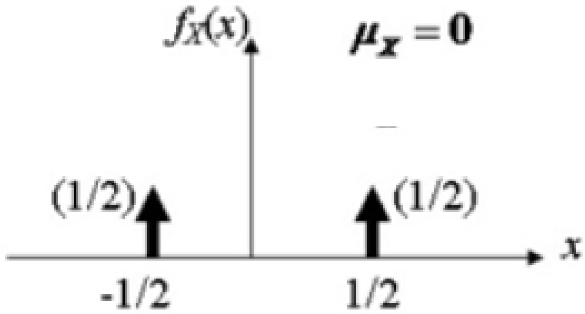
$$\int_{-\infty}^{\infty} x\, f_X(x)\, dx = \mu_X = \overline{X} = E\{X\}$$

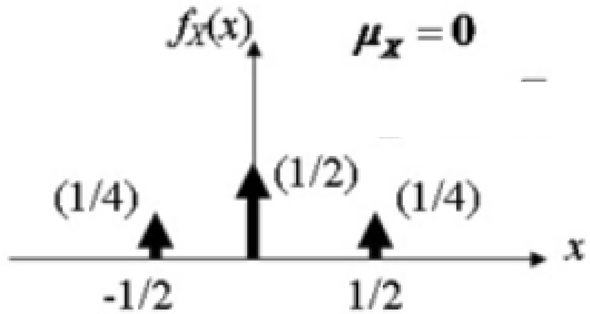



$f_X(x)$
$\mu_X = 0$
(1)
x


$f_X(x)$
$\mu_X = 1/6$
(1/3)
(2/3)
-1/2
1/2
x


$f_X(x)$
$\mu_X = 0$
1
-1/2
1/2
x


$f_X(x)$
$\mu_X = 0$
(1/2)
(1/2)
-1/2
1/2
x


$f_X(x)$
$\mu_X = 0$
(1/4)
(1/2)
(1/4)
-1/2
1/2
x

The second moment of X is its mean square,

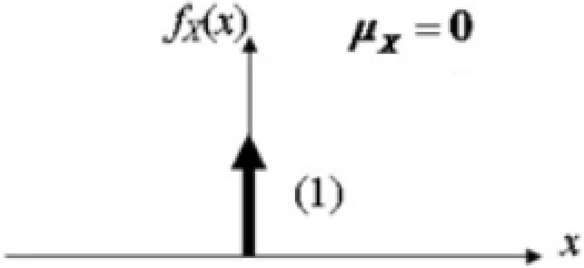
$$\int_{-\infty}^{\infty} x^2 f_X(x)\,dx = \overline{X^2} = E\{X^2\}.$$

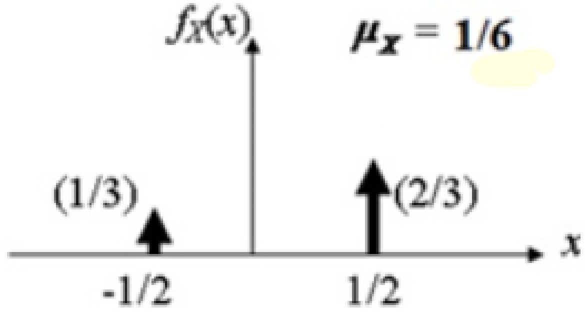
To get the "spread" of the curve,
we want the average distance
from the mean.
But that's zero.

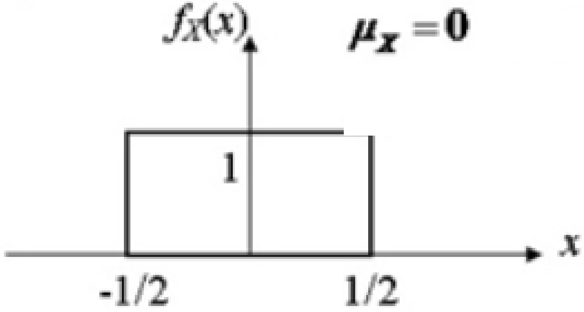
$$E\{X-\mu_X\} = \int_{-\infty}^{\infty} (x-\mu_X)\ f_X(x)\ dx = 0\ .$$

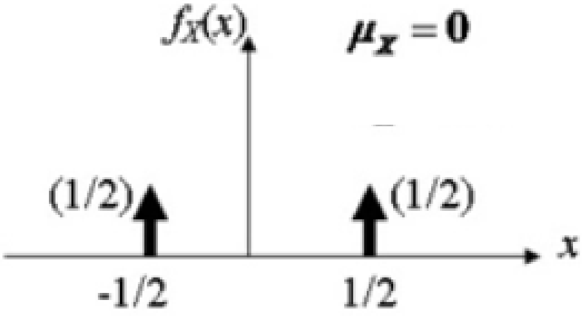
So we take the expected value of $(X-\mu_X)^2$
(which is always positive), and then
take the square root.
"root-mean-square" is called the standard deviation σ_X:

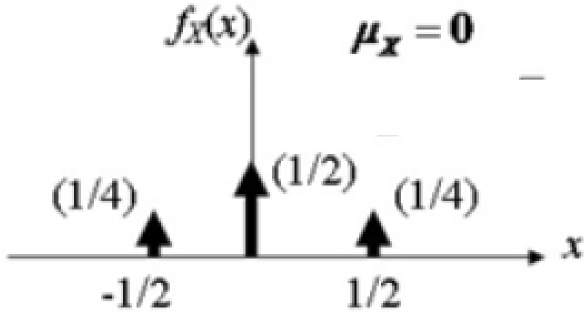
$$\sigma_X = \sqrt{\int_{-\infty}^{\infty} (x-\mu_X)^2\ f_X(x)\ dx} = \sqrt{E\{(X-\mu_X)^2\}}\ .$$

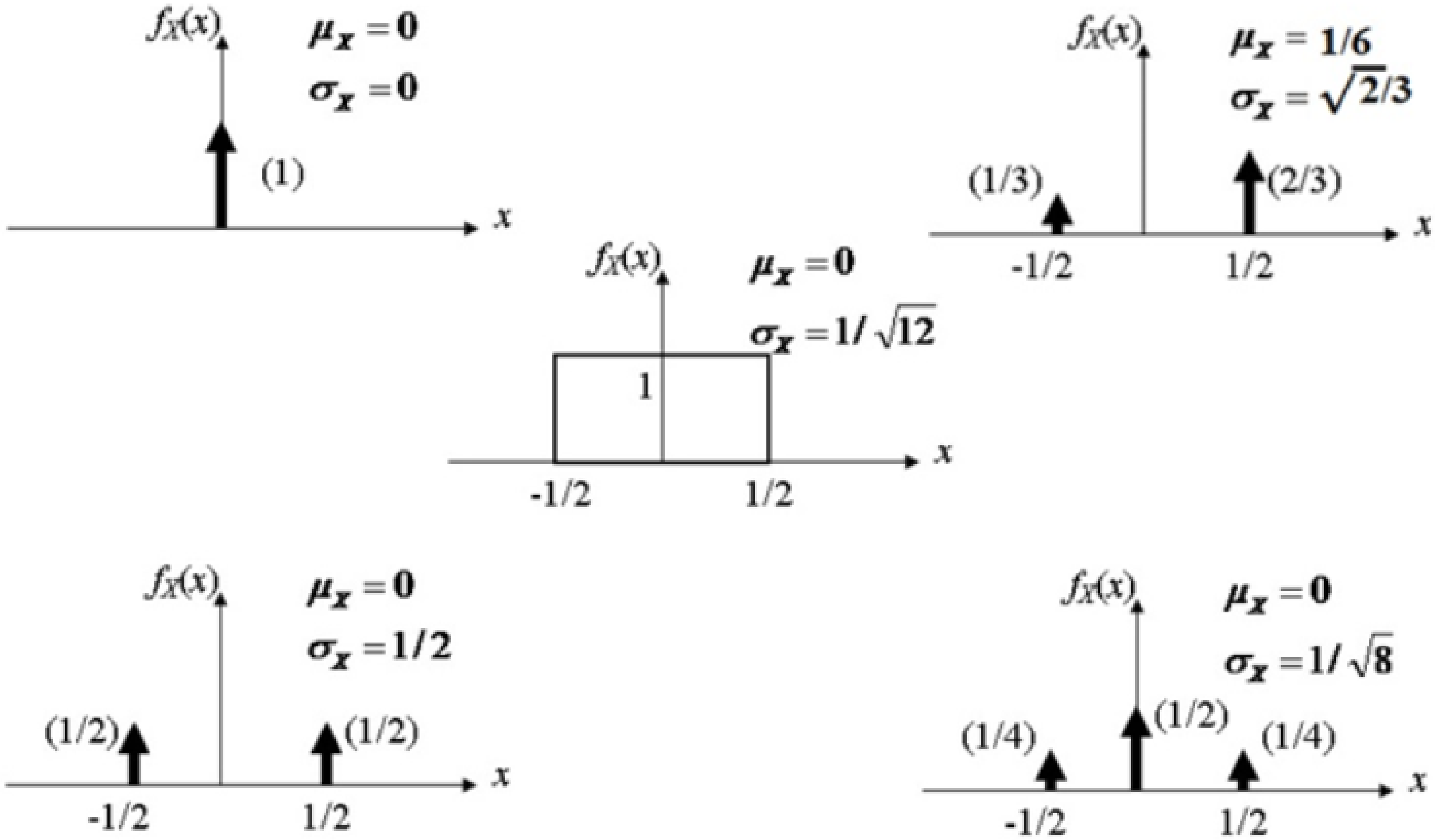

$f_X(x)$
$\mu_X = 0$
(1)
x


$f_X(x)$
$\mu_X = 1/6$
(1/3)
(2/3)
x
-1/2
1/2


$f_X(x)$
$\mu_X = 0$
1
x
-1/2
1/2


$f_X(x)$
$\mu_X = 0$
(1/2)
(1/2)
x
-1/2
1/2


$f_X(x)$
$\mu_X = 0$
(1/2)
(1/4)
(1/4)
x
-1/2
1/2



fX(x)
μX = 0
σX = 0
(1)
x
fX(x)
μX = 1/6
σX = √2/3
(1/3)
(2/3)
x
-1/2
1/2
fX(x)
μX = 0
σX = 1/ √12
1
x
-1/2
1/2
fX(x)
μX = 0
σX = 1/2
(1/2)
(1/2)
x
-1/2
1/2
fX(x)
μX = 0
σX = 1/ √8
(1/4)
(1/2)
(1/4)
x
-1/2
1/2

Second-moment identity

$$\overline{X^2} = \overline{X}^2 + \sigma_X^2 \equiv \mu_X^2 + \sigma_X^2 \ .$$

$$\sigma_X^2 = E\{(X-\mu_x)^2\}$$

$$= E\{X^2 - 2X\mu_X + \mu_X^2\}$$

$$= E\{X^2\} - 2E\{X\}\mu_X + \mu_X^2$$

$$= E\{X^2\} - 2\mu_X^2 + \mu_X^2$$

$$= \overline{X^2} - \mu_X^2 .$$

characteristic function of X, $\Phi_X(\omega)$:

$$\Phi_X(\omega) = \int_{-\infty}^{\infty} e^{i\omega x} f_X(x)\,dx.$$

$$\Phi_X(\omega)|_{\omega=0} = \int_{-\infty}^{\infty} e^{i\omega x} f_X(x)\,dx|_{\omega=0} = 1;$$

$$\Phi'_X(\omega)|_{\omega=0} = \int_{-\infty}^{\infty} ix e^{i\omega x} f_X(x)\,dx|_{\omega=0} = iE\{X\}$$

$$\Phi''_X(\omega)|_{\omega=0} = \int_{-\infty}^{\infty} (ix)^2 e^{i\omega x} f_X(x)\,dx|_{\omega=0} = i^2E\{X^2\}$$

$$\Phi'''_X(\omega)|_{\omega=0} = \int_{-\infty}^{\infty} (ix)^3 e^{i\omega x} f_X(x)\,dx|_{\omega=0} = i^3E\{X^3\}$$

Summary: Important Facts about Expected Value and Moments

Expected Value: $\int_{-\infty}^{\infty} g(x)\, f_X(x)\, dx = \overline{g(X)} = E\{g(X)\}$

Generating Function: $\Phi_X(\omega) = \int_{-\infty}^{\infty} e^{i\omega x} f_X(x)\, dx = E\{e^{i\omega x}\}$

Moments: $E\{X^n\} == \int_{-\infty}^{\infty} x^n f_X(x)\, dx = i^{-n} \Phi_X^{(n)}(\omega)\, |_{\omega=0}$

$$E\{X\} = \mu_X = \overline{X}\,,\ E\{X^2\} = \overline{X^2}$$

Standard Deviation: $\sigma_X = \sqrt{\int_{-\infty}^{\infty} (x - \mu_X)^2\ f_X(x)\ dx} = \sqrt{E\{(X - \mu_X)^2\}}$.

Second Moment Identity: $\overline{X^2} = \mu_X^2 + \sigma_X^2$.

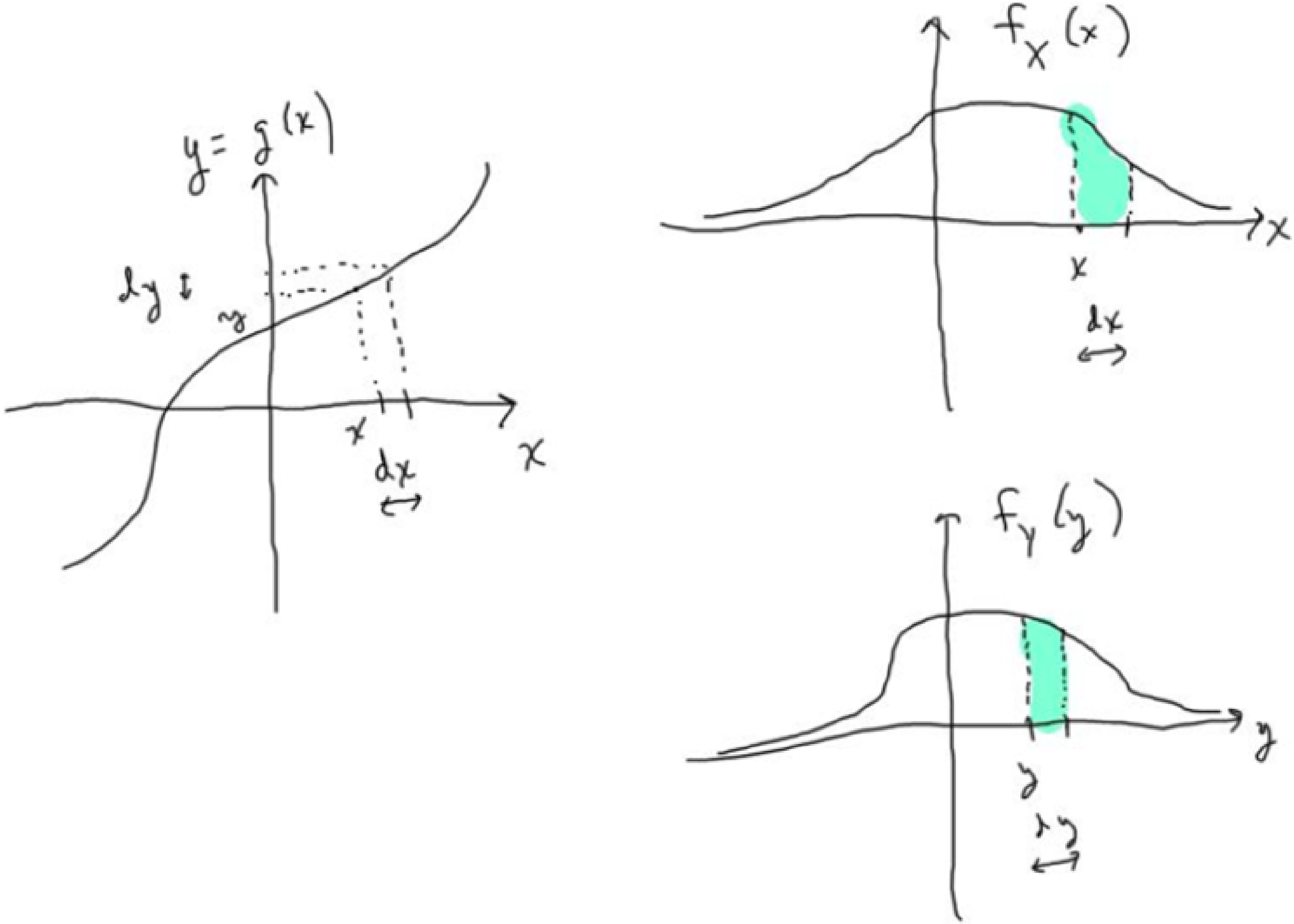
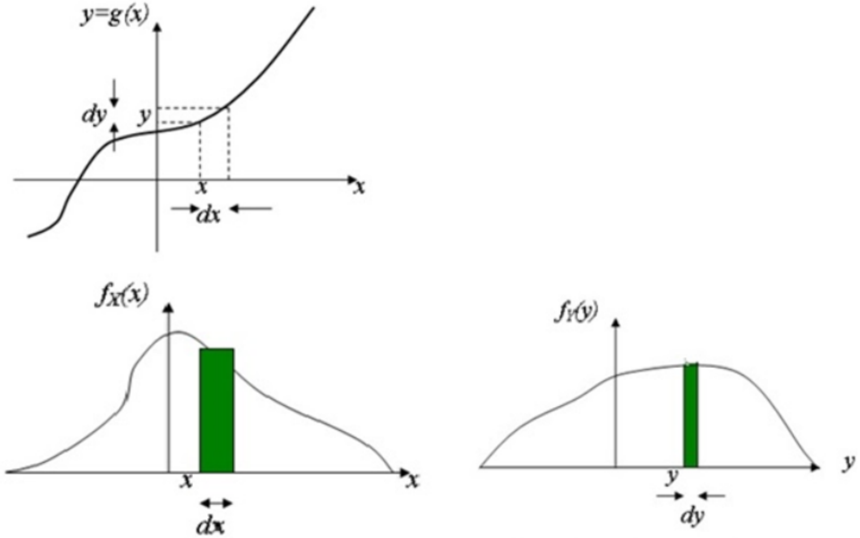


Figure 16. Change of (random) variable

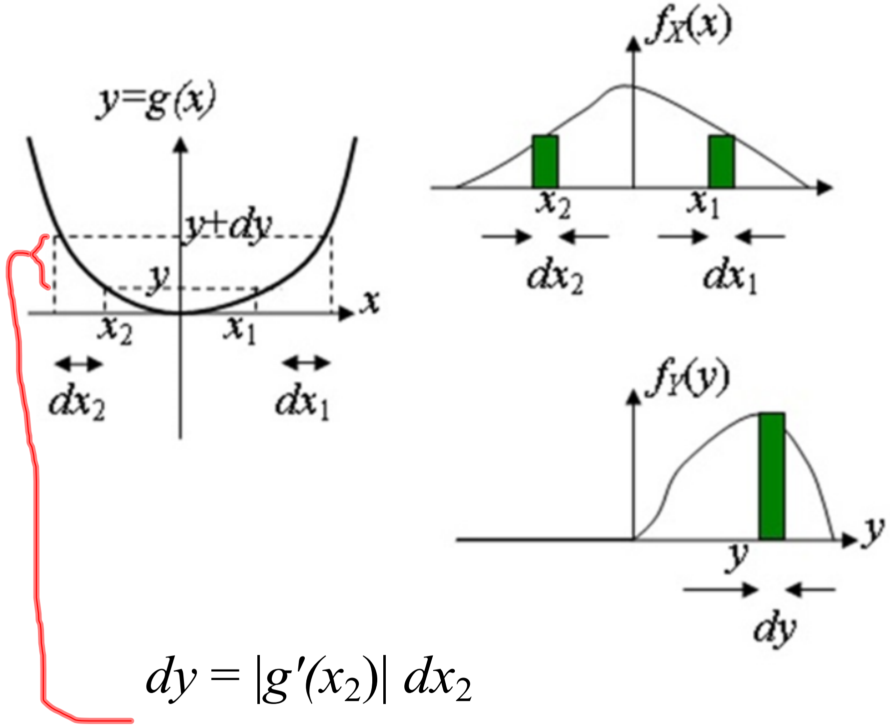


$$dy = g'(x)\, dx$$

$$f_Y(y)\, dy = f_X(x)\, dx$$

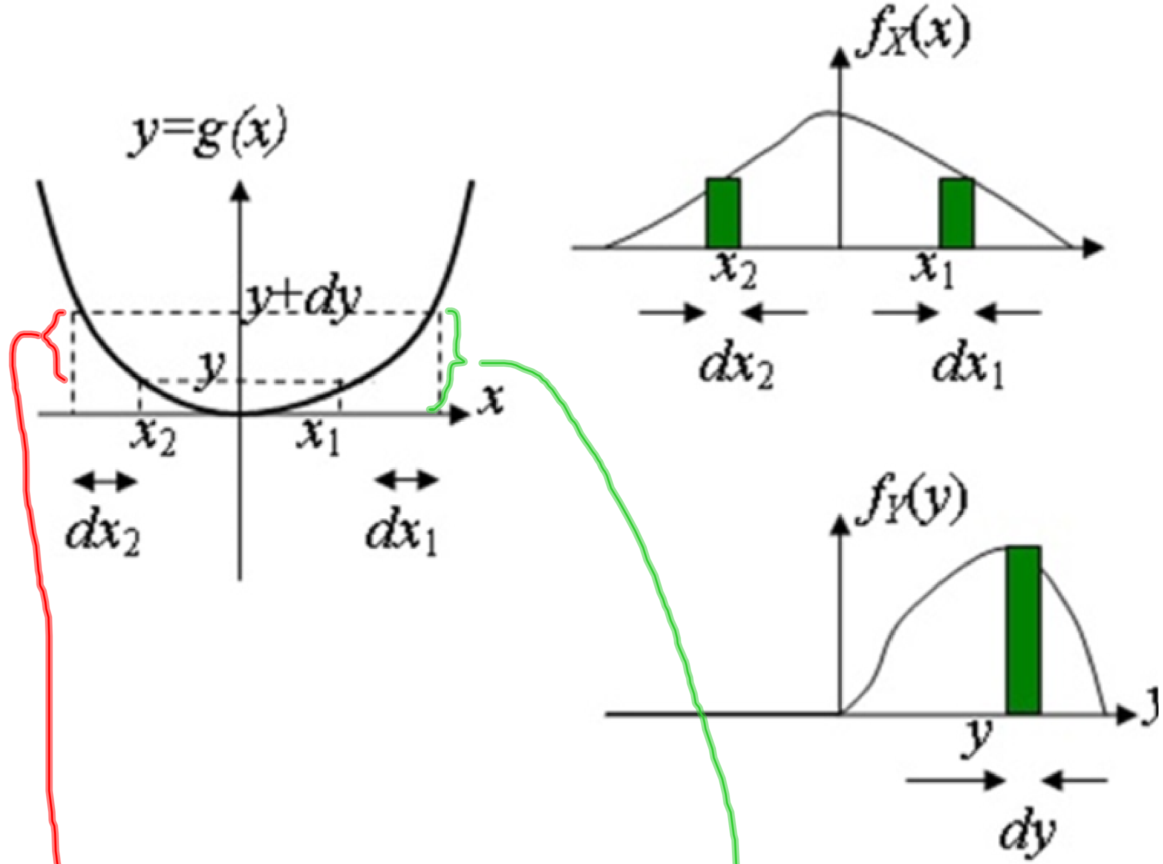
$$f_Y(y) = f_X(x)\ /\ g'(x)$$

Figure 1.19 Change of (random) variable



$$dy = |g'(x_2)|\, dx_2$$

Figure 1.20 Changing variables



$$dy = |g'(x_2)|\, dx_2$$

$$+\, g'(x_1)\, dx_1$$

ary: Important Facts about Change of Variable

pdf of $y = g(x)$: $$f_Y(y) = \sum_{\substack{preimages \\ of\ y}} \frac{f_X(x_i)}{|g'(x_i)|}$$

$$E\{aX+b\} = a\mu_X + b\ ,\ \ \sigma_{aX+b} = |a|\ \sigma_X\ .$$

ry: Important Facts about Change of Variable

pdf of $y = g(x)$: $f_Y(y) = \sum_{\substack{preimages \\ of\ y}} \frac{f_X(x_i)}{|g'(x_i)|}$

$E\{aX+b\} = a\mu_X + b$, $\sigma_{aX+b} = |a|\,\sigma_X$.

expressed in terms of y :

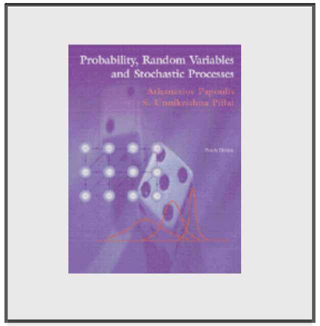
$x = g^{-1}(x)$

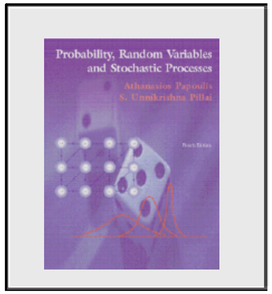
Lecture 4
(1/25/2017)

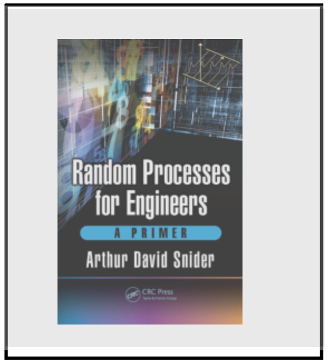
1. How many ways are there of placing a 2.5"x3.5"x0.0115" playing card in a 2.6"x3.6"x0.1115" box? (Backwards, uspide down, …?)

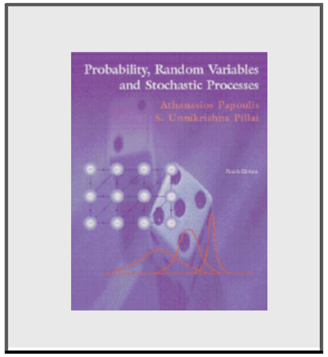
2. How many ways are there of placing a 1" cube in a 1.1" cubic box?

The game show problem.

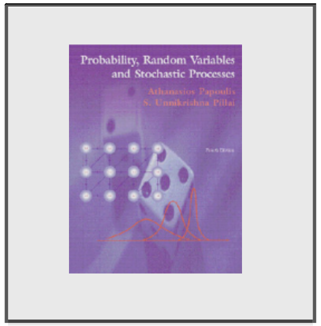


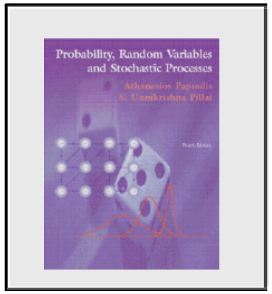


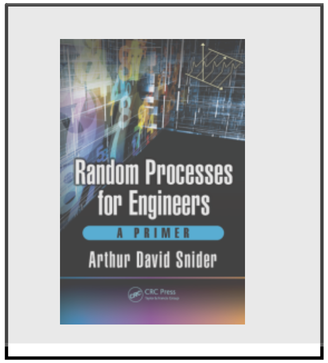


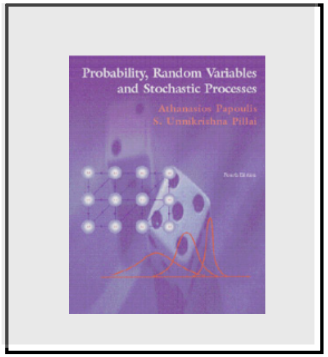


The game show problem.

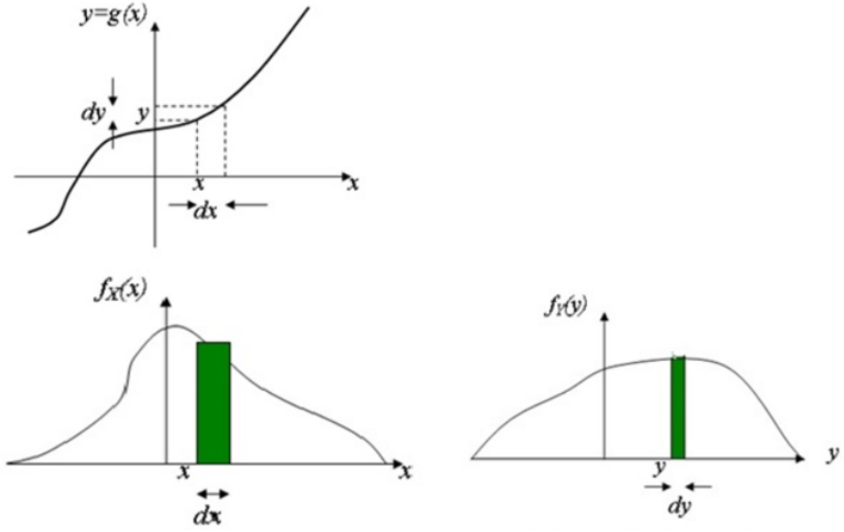








The game show problem.



$$dy = g'(x)\, dx$$

$$f_Y(y)\, dy = f_X(x)\, dx$$

$$f_Y(y) = f_X(x) \;/\; g'(x)$$

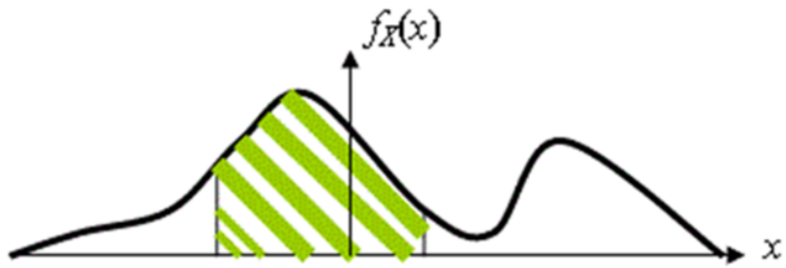
Figure 1.19 Change of (random) variable

ary: Important Facts about Change of Variable

pdf of $y = g(x)$: $f_Y(y) = \sum_{\substack{preimages \\ of\ y}} \frac{f_X(x_i)}{|g'(x_i)|}$

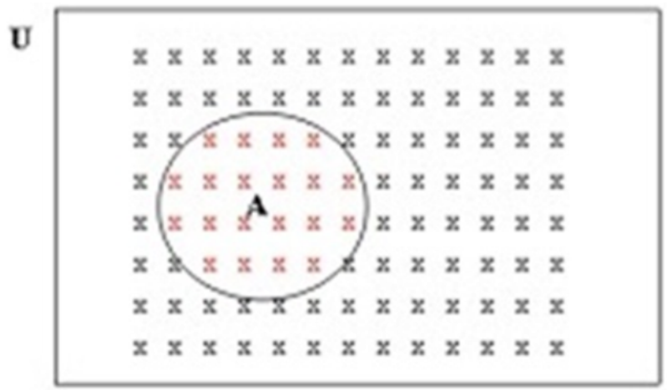
$E\{aX+b\} = a\mu_X + b$, $\sigma_{aX+b} = |a|\,\sigma_X$.

When $y = ax + b = g(x)$



$$p(x<X<x+dx) = f_X(x)\, dx$$

What is $f_{X|A}(x)$?



The pdf formulation of conditional probability

p(A and $x<X<x+dx$)

= p(A) x p($x<X<x+dx$ GIVEN A)

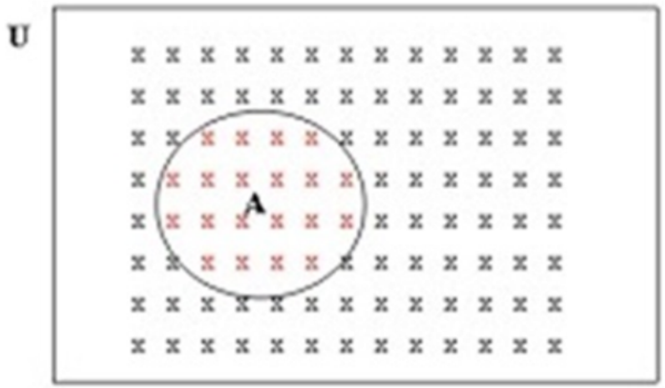
SO

p($x<X<x+dx$ GIVEN A) =

p(A and $x<X<x+dx$) / p(A)

p($x<X<x+dx$ GIVEN A) =

p(A and $x<X<x+dx$) / p(A)



If x is in the truth set of A then

p(A and $x<X<x+dx$)

is $f_X(x)\, dx$.

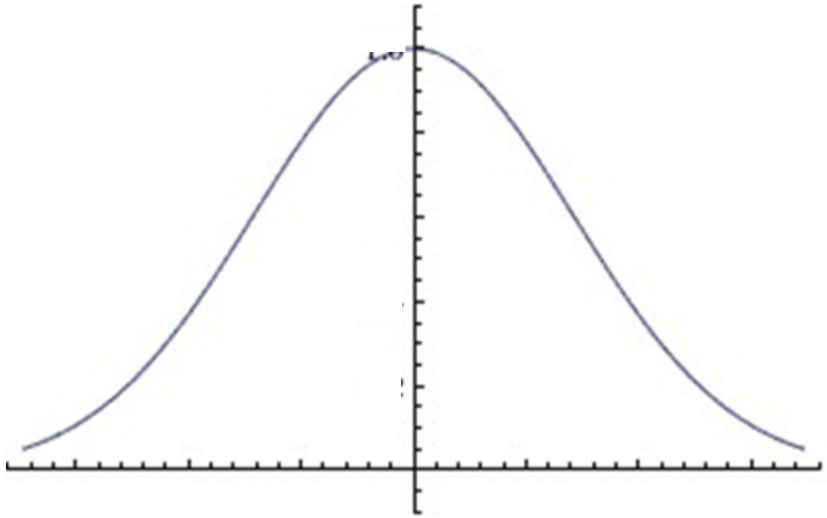
Otherwise it is zero. So

$$f_{X|A}(x)dx = \begin{cases} \frac{f_X(x)dx}{p(A)} & \text{if } x \text{ is in the truth set of } A \\ 0 & \text{otherwise} \end{cases}$$

$$f_{X|A}(x) = \begin{cases} \frac{f_X(x)}{p(A)} & \text{if } x \text{ is in the truth set of } A \\ 0 & \text{otherwise} \end{cases}$$

Suppose A is the statement "$x > 5$".
Then

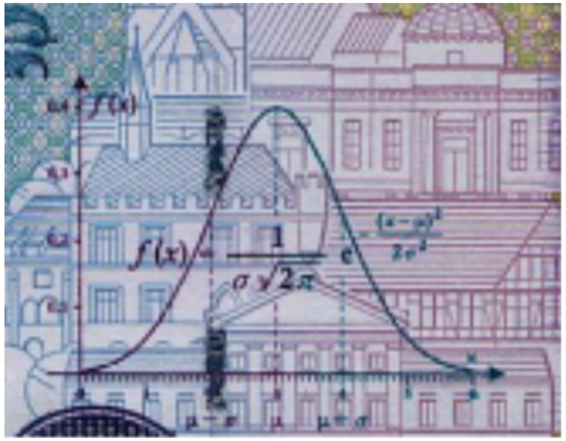
$$f_{X|A}(x) = \begin{cases} \frac{f_X(x)}{\int_5^\infty f_X(\xi)d\xi} & \text{if } x > 5 \\ 0 & \text{otherwise} \end{cases}$$

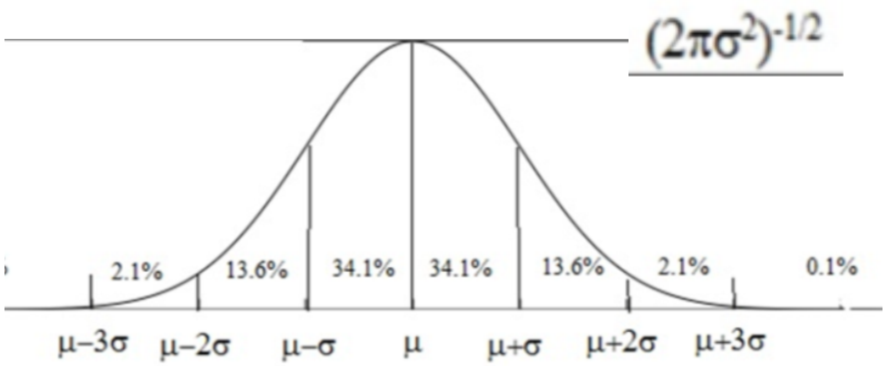


Bell curve

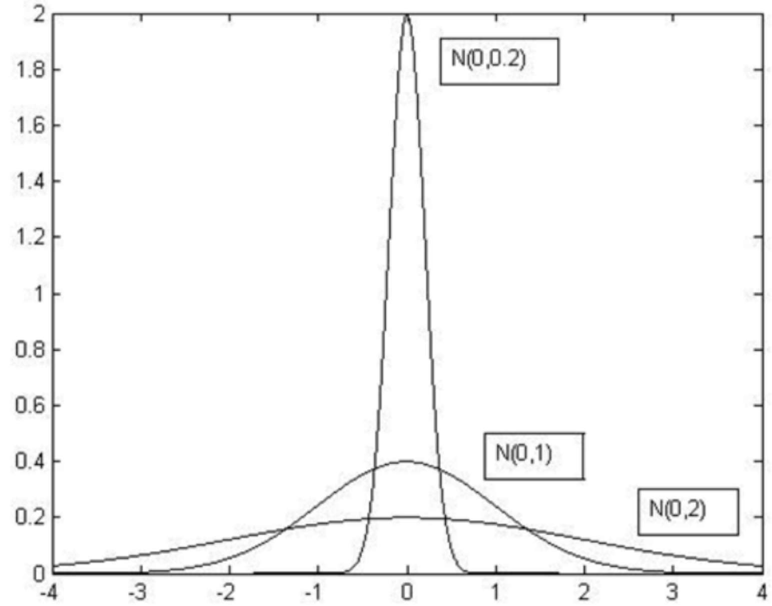
Normal or Gaussian pdf


GN4480100S8
DEUTSCHE BUNDESBANK
Banknote
10
ZEHN DEUTSCHE MARK
1777 - 1855 Carl Friedr. Gauß
Deutsche Bundesbank
Frankfurt am Main
1. September 1999
GN4480100S8


f(x)
f(x) = 1/(σ√2π) e



$$N(\mu,\sigma) = f_X(x) = e^{-\frac{(x-\mu)^2}{2\sigma^2}} \Big/ \sqrt{2\pi\sigma^2}$$


N(0,0.2)
N(0,1)
N(0,2)
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
-4
-3
-2
-1
0
1
2
3
4

$$N(\mu,\sigma) = f_X(x) = e^{-\frac{(x-\mu)^2}{2\sigma^2}} \Big/ \sqrt{2\pi\sigma^2}$$

(i) the denominator is simply the normalizing factor that makes the total area equal to 1;

(ii) the numerator is the exponential of a *quadratic* in x. In fact *every* function of the form e^{ax^2+bx+c}, with the appropriately scaled denominator, is a normal probability density function as long as $a<0$.

$$N(\mu,\sigma) = f_X(x) = e^{-\frac{(x-\mu)^2}{2\sigma^2}} \Big/ \sqrt{2\pi\sigma^2}$$

Every function of the form e^{ax^2+bx+c}

is a normal probability density function

(Complete the square to express ax^2+bx+c as

$a(x\text{-}p)^2+q$, identify μ as p,

identify σ as $1/\sqrt{-2a}$, and

incorporate the factor e^q into the normalization.

$$N(\mu,\sigma) = f_X(x) = e^{-\frac{(x-\mu)^2}{2\sigma^2}} \Big/ \sqrt{2\pi\sigma^2}$$

Characteristic function for the normal distribution:

$$\Phi_X(\omega) = \int_{-\infty}^{\infty} e^{i\omega x}\, e^{-\frac{(x-\mu)^2}{2\sigma^2}} / \sqrt{2\pi\sigma^2}\, dx$$

$$= e^{-(\sigma^2\omega^2/2) + i\mu\omega}$$

ıary: Important Facts about Change of Variable

$$\text{pdf of } y = g(x): \quad f_Y(y) = \sum_{\substack{preimages \\ of\ y}} \frac{f_X(x_i)}{|\,g'(x_i)\,|}$$

$$E\{\,aX+b\} = a\mu_X + b\,, \quad \sigma_{aX+b} = |a|\,\sigma_X\,.$$

So by shifting and rescaling you can express N(p,q) in terms of the "standard" N(0,1) and use tables of the latter to get your integrals.

"Univariate" pdfs

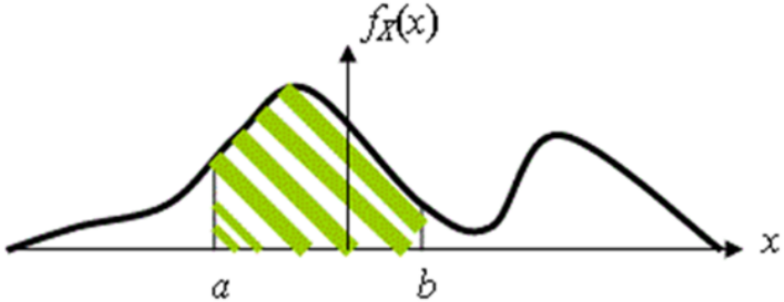
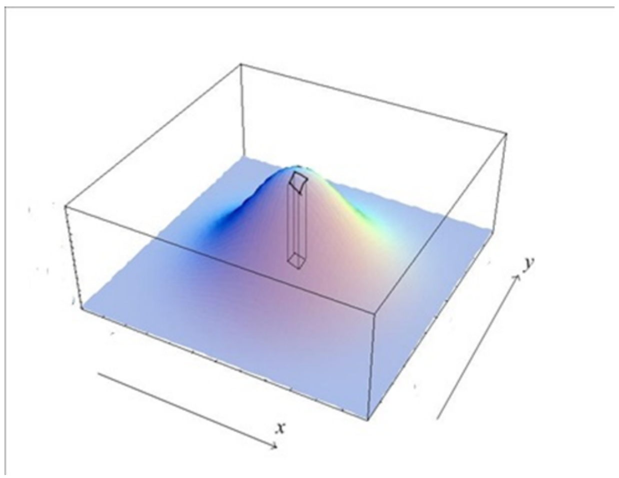


Figure 1.11. $f_X(x) \geq 0, \int_{-\infty}^{\infty} f_X(x)\,dx = 1$

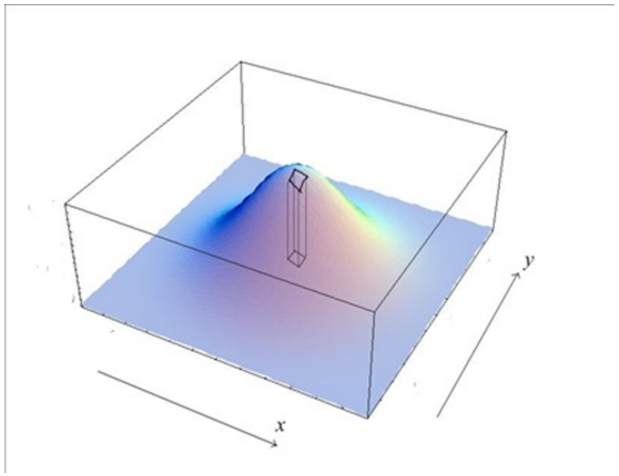
"Bivariate" pdfs



p(*X* lies in the interval [*a*,*b*] AND *Y* lies in [*c*,*d*])

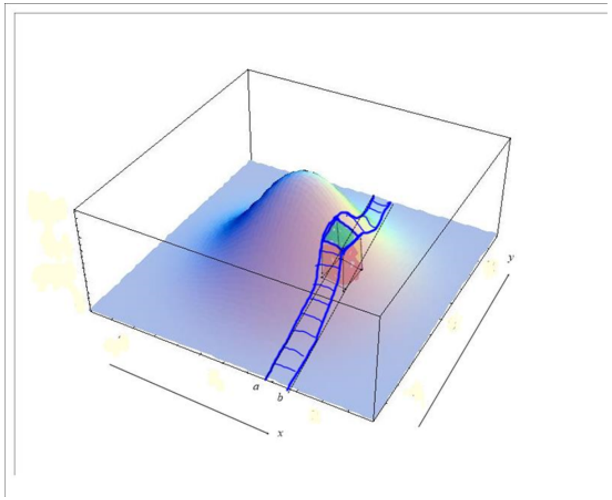
$$\int_a^b \int_c^d f_{XY}(x, y)\, dy\, dx$$

"bivariate probability density function"



p(*X* lies between *x* and *x*+*dx* *and* that *Y* lies between *y* and *y*+*dy*)

$f_{XY}(x,y)\ dx\ dy$



probability that X lies in the interval $[a,b]$

$$= \int_{x=a}^{b} \int_{y=-\infty}^{\infty} f_{XY}(x,y)\,dy\,dx$$

$$= \int_a^b f_X(x)\,dx \quad \text{if} \quad f_X(x) = \int_{y=-\infty}^{\infty} f_{XY}(x,y)\,dy$$

“marginal” probability density function.

Y \ X	x_1	x_2	x_3	x_4	$p_Y(Y)\downarrow$
y_1	4⁄32	2⁄32	1⁄32	1⁄32	**8⁄32**
y_2	2⁄32	4⁄32	1⁄32	1⁄32	**8⁄32**
y_3	2⁄32	2⁄32	2⁄32	2⁄32	**8⁄32**
y_4	8⁄32	0	0	0	**8⁄32**
$p_X(X) \rightarrow$	**16⁄32**	**8⁄32**	**4⁄32**	**4⁄32**	**32⁄32**

Joint and marginal distributions of a pair of discrete, random variables X,Y having nonzero mutual information I(X; Y). The values of the joint distribution are in the 4×4 square, and the values of the marginal distributions are along the right and bottom margins.

Expected Values for the bivariate case

$$E\{g(X,Y)\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)\, f_{XY}(x,y)\, dx\, dy$$

The bivariate mean of x:

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x\, f_{XY}(x,y)\, dx\, dy = \int_{-\infty}^{\infty} x\{\int_{-\infty}^{\infty} f_{XY}(x,y)\, dy\} dx$$

$$= \int_{-\infty}^{\infty} x\, f_X(x)\, dx = \mu_X$$

(The same good news holds for σ_x or any function of x alone!)

Remember the second moment identity?

Second-moment identity

$$\overline{X^2} = \overline{X}^2 + \sigma_X^2 \equiv \mu_X^2 + \sigma_X^2 \ .$$

$$\sigma_X^2 = E\{(X-\mu_x)^2\}$$

$$= E\{X^2 - 2X\mu_X + \mu_X^2\}$$

$$= E\{X^2\} - 2E\{X\}\mu_X + \mu_X^2$$

$$= E\{X^2\} - 2\mu_X^2 + \mu_X^2$$

$$= \overline{X^2} - \mu_X^2 .$$

$$\overline{X^2} = \overline{X}^2 + \sigma_X^2 \equiv \mu_X^2 + \sigma_X^2$$

The bivariate second moment identity

$$E\{XY\} = \mu_X \mu_Y + E\{(X-\mu_X)(Y-\mu_Y)\}$$

$$\overline{X^2} = \overline{X}^2 + \sigma_X^2 \equiv \mu_X^2 + \sigma_X^2$$

The bivariate second moment identity

correlation ↓ covariance ↓

$$E\{XY\} = \mu_X \mu_Y + E\{(X-\mu_X)(Y-\mu_Y)\}$$

INDEPENDENCE of bivariate variables

$$p(A|B) = p(A)$$

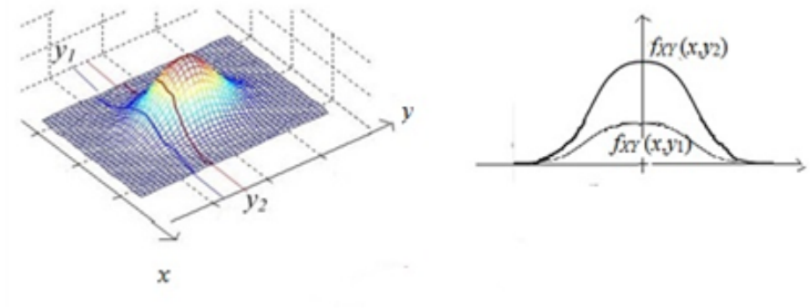
$$p(A \cap B) = p(A)\,p(B)$$

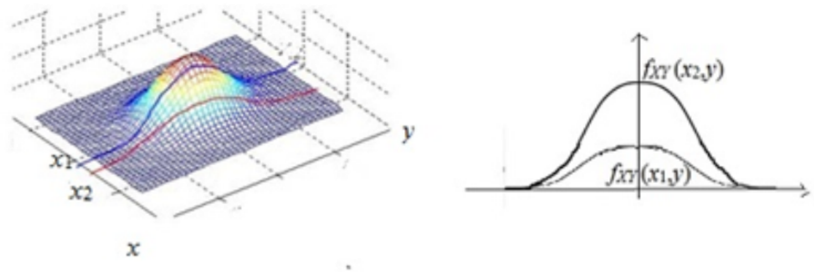
INDEPENDENCE of bivariate variables

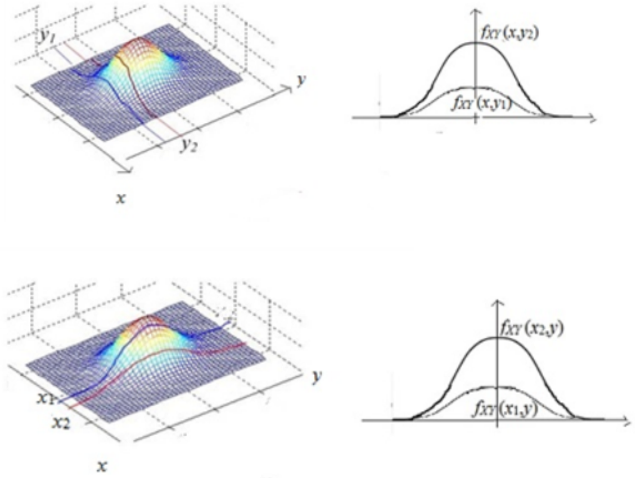
$f_{XY}(x,y)\, dx\, dy$ *equals* $\{f_X(x)\, dx\}$ *times* $\{f_Y(y)\, dy\}$

$f_{XY}(x,y) = f_X(x)\, f_Y(y)$ for independent variables.

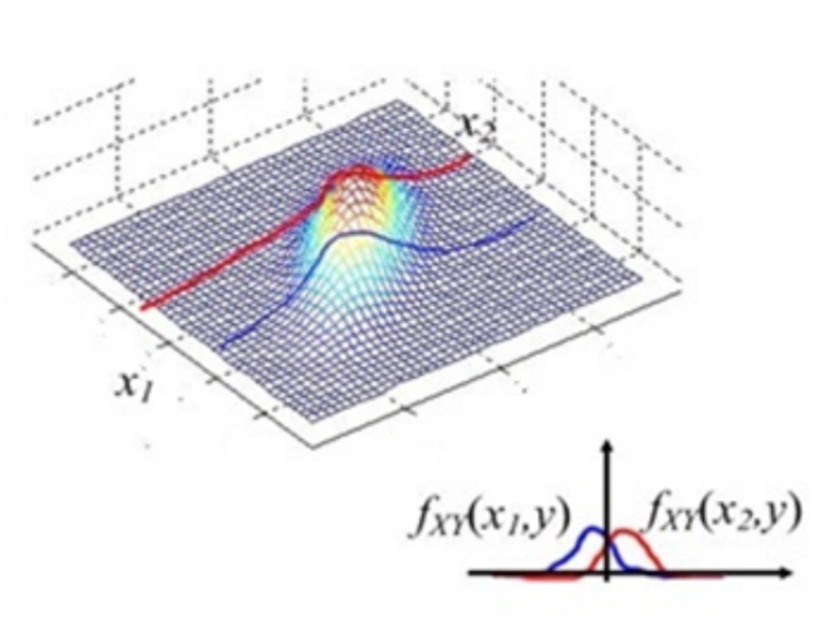
Independent






y1
y
y2
x
fXY(x,y2)
fXY(x,y1)
x1
x2
y
x
fXY(x2,y)
fXY(x1,y)

Dependent


x2
x1
fXY(x1,y)
fXY(x2,y)

Lecture 5
(1/30/2017)

Card trick probability:

$$52/52^2 = 1/52$$

$$f_{X|A}(x) = \begin{cases} \frac{f_X(x)}{p(A)} & \text{if } x \text{ is in the truth set of } A \\ 0 & \text{otherwise} \end{cases}$$

Suppose A is the statement "$x > 5$".
Then

$$f_{X|A}(x) = \begin{cases} \frac{f_X(x)}{\int_5^{\infty} f_X(\xi) d\xi} & \text{if } x > 5 \\ 0 & \text{otherwise} \end{cases}$$

Conditional pdf's

The conditional pdf for X, given that Y

takes the value y, is denoted $f_{X|Y}(x|y)$;

p($x < X < x+dx$, given that $Y=y$) is $f_{X|Y}(x|y)\,dx$

"$Y=y$" $\leftrightarrow$ "Y lies between y and $y+dy$"

$$f_{XY}(x,y)\,dx\,dy = f_Y(y)\,dy \times f_{X|Y}(x \mid y)\,dx$$

$$f_{X|Y}(x \mid y) = \frac{f_{XY}(x,y)}{f_Y(y)} = \frac{f_{XY}(x,y)}{\int_{-\infty}^{\infty} f_{XY}(x,y)\,dx} \,.$$

If X,Y are independent

$$f_{X|Y}(x \mid y) = \frac{f_X(x) f_Y(y)}{f_Y(y)} = f_X(x)$$

Summary of Important Equations for Bivariate Random Variables

Bivariate pdf: probability that X lies between x and $x+dx$ and that Y lies between y and $y+dy$ equals $f_{XY}(x,y)\,dx\,dy$.

Marginal pdf for X is $f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x,y)\,dy$.

Bivariate mean $\equiv$ marginal mean: $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x\, f_{XY}(x,y)\,dx\,dy = \int_{-\infty}^{\infty} x\, f_X(x)\,dx = \mu_X$.

Independence: $f_{XY}(x,y) = f_X(x)\, f_Y(y)$, which implies $E\{g_1(X)\, g_2(Y)\} = E\{g_1(X)\}\, E\{g_2(Y)\}$.

covariance = $\mathrm{cov}(X,Y) = E\{(X\text{-}\mu_X)(Y\text{-}\mu_Y)\}$.

correlation coefficient = $\rho_{XY} = \mathrm{cov}(X,Y)/(\sigma_X\, \sigma_Y)$, $-1 \le \rho_{XY} \le 1$.

correlation = $E\{XY\}$.

covariance = $\mathrm{cov}(X,Y) = E\{(X\text{-}\mu_X)(Y\text{-}\mu_Y)\}$

correlation coefficient

$$\rho_{XY} = \mathrm{cov}(X,Y)/(\sigma_X\, \sigma_Y)$$

$$-1 \le \rho_{XY} \le 1$$

correlation = $E\{XY\}$

Extended second-moment identity:

$$E\{XY\} = \mu_X \mu_Y + E\{(X-\mu_X)(Y-\mu_Y)\} = \mu_X \mu_Y + \text{cov}(X,Y) = \mu_X \mu_Y + \rho_{XY} \sigma_X \sigma_Y .$$

Conditional pdf: $f_{X|Y}(x \mid y) = \dfrac{f_{XY}(x,y)}{f_Y(y)}$.

Conditional mean: $\mu_{X|Y}(y) = \int_{-\infty}^{\infty} x\, f_{X|Y}(x \mid y)\, dx$.

Conditional standard deviation: $\sigma_{X|Y}(y) = \sqrt{\int_{-\infty}^{\infty} [x - \mu_{X|Y}(y)]^2 f_{X|Y}(x \mid y)\, dx}$.

Normal or Gaussian distribution

$$N(\mu,\sigma) = f_X(x) = e^{-\frac{(x-\mu)^2}{2\sigma^2}} \Big/ \sqrt{2\pi\sigma^2}$$

If you were to $\int_{-\infty}^{\infty} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \Big/ \sqrt{2\pi\sigma^2} \; dx$ 

If you were to $\int_{-\infty}^{\infty} x \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}} \Big/ \sqrt{2\pi\sigma^2} \; dx$ 

If you were to $\int_{-\infty}^{\infty} (x-\mu)^2 \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}} \Big/ \sqrt{2\pi\sigma^2} \; dx$ 

Normal or Gaussian distribution

$$N(\mu,\sigma) = f_X(x) = e^{-\frac{(x-\mu)^2}{2\sigma^2}} \Big/ \sqrt{2\pi\sigma^2}$$

If you were to $\int_{-\infty}^{\infty} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \Big/ \sqrt{2\pi\sigma^2} \, dx$  1

If you were to $\int_{-\infty}^{\infty} x \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}} \Big/ \sqrt{2\pi\sigma^2} \, dx$  μ

If you were to $\int_{-\infty}^{\infty} (x-\mu)^2 \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}} \Big/ \sqrt{2\pi\sigma^2} \, dx$  σ^2

Gaussian, or normal, bivariate distribution

If X and Y were independent and normal, then

$$f_{XY}(x,y) = e^{-\frac{(x-\mu_X)^2}{2\sigma_X^2}} e^{-\frac{(y-\mu_Y)^2}{2\sigma_Y^2}} \Big/ \sqrt{2\pi\sigma_X^2}\sqrt{2\pi\sigma_Y^2}$$

$$= e^{-\frac{1}{2}\left[\frac{(x-\mu_X)^2}{\sigma_X^2}+\frac{(y-\mu_Y)^2}{\sigma_Y^2}\right]} \Big/ 2\pi\sigma_X\sigma_Y$$

$$Ke^{Ax^2+By^2+Cx+Dy+E}$$

If they are NOT independent, then

$$e^{Ax^2+By^2+Cx+Dy+E+\mathbf{Fxy}}$$

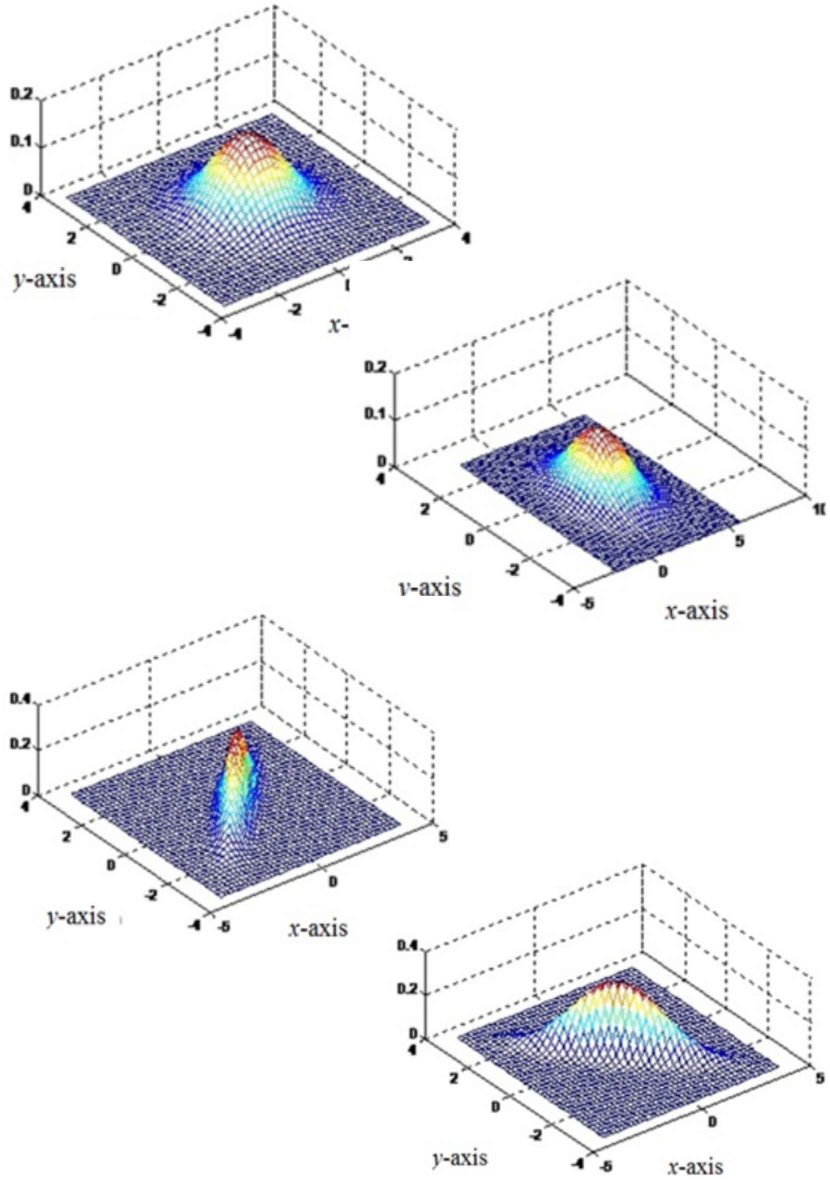
Independent Normal

$$f_{XY}(x,y) = e^{-\frac{(x-\mu_X)^2}{2\sigma_X^2}} e^{-\frac{(y-\mu_Y)^2}{2\sigma_Y^2}} \Big/ \sqrt{2\pi\sigma_X^2}\sqrt{2\pi\sigma_Y^2}$$

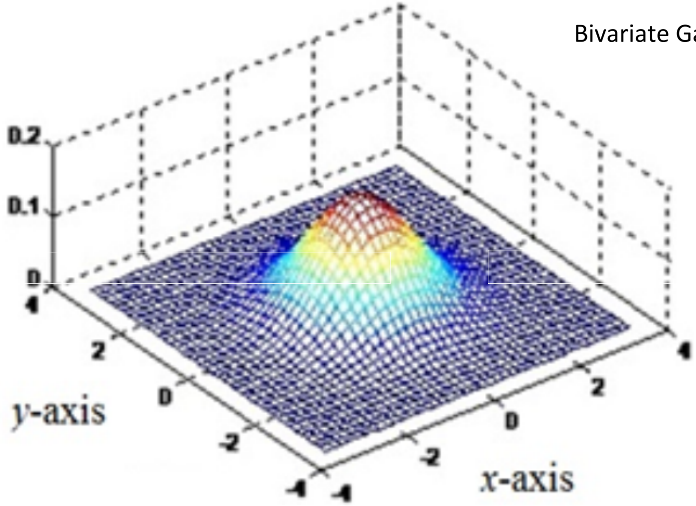
$$= e^{-\frac{1}{2}\left[\frac{(x-\mu_X)^2}{\sigma_X^2}+\frac{(y-\mu_Y)^2}{\sigma_Y^2}\right]} \Big/ 2\pi\sigma_X\sigma_Y$$

Not Independent Normal

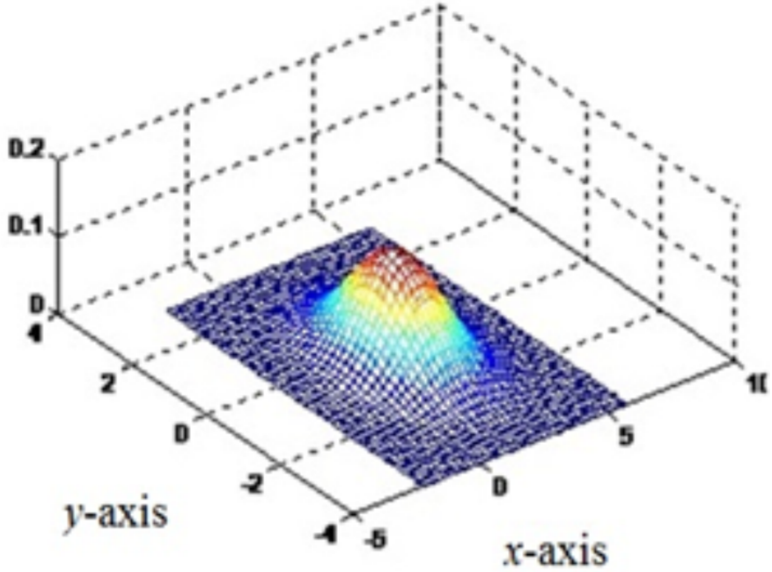
$$\frac{e^{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_X)^2}{\sigma_X^2}-2\rho\frac{(x-\mu_X)}{\sigma_X}\frac{(y-\mu_Y)}{\sigma_Y}+\frac{(y-\mu_Y)^2}{\sigma_Y^2}\right]}}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}}$$


0.2
0.1
0
y-axis
0.2
0.1
0
y-axis
x-axis
0.4
0.2
0
y-axis
x-axis
0.4
0.2
0
y-axis
x-axis

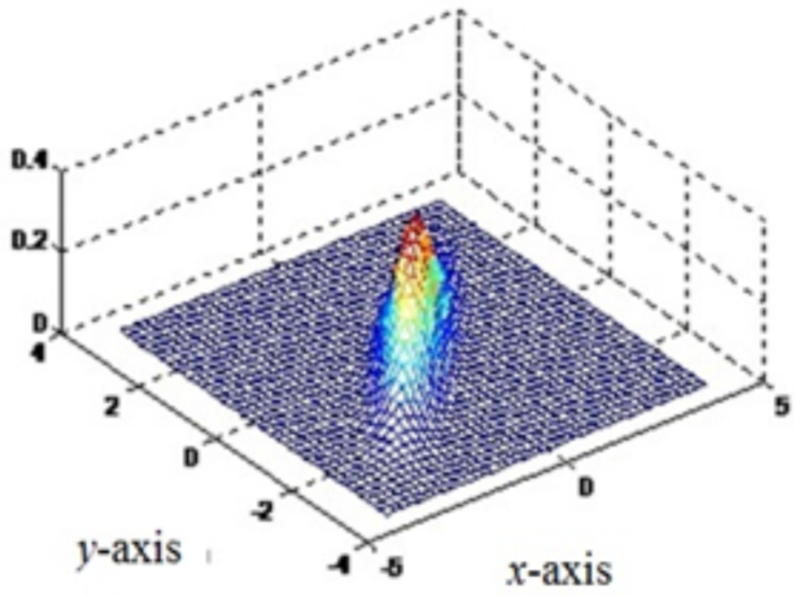
Bivariate Gaussian



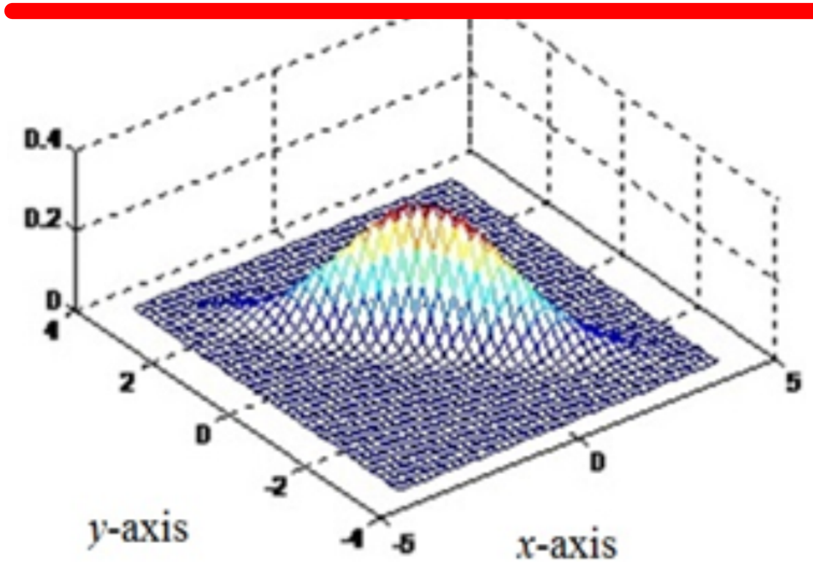
$\mu_X = \mu_Y = 0,\ \sigma_X = \sigma_Y = 1,\ \rho = 0$



$\mu_X = 1.8,\ \mu_Y = -0.5,\ \sigma_X = 1.1,\ \sigma_Y = 1,\ \rho = 0$



$\mu_X = \mu_Y = 0,\ \sigma_X = 1.1,\ \sigma_Y = 1,\ \rho = 0.9$



$\mu_X = \mu_Y = 0,\ \sigma_X = 1.1,\ \sigma_Y = 1,\ \rho = -0.9$

Gaussian, or normal, bivariate distribution

$$\frac{e^{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_X)^2}{\sigma_X^2}-2\rho\frac{(x-\mu_X)}{\sigma_X}\frac{(y-\mu_Y)}{\sigma_Y}+\frac{(y-\mu_Y)^2}{\sigma_Y^2}\right]}}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}}$$

If you were to

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_X)^2}{\sigma_X^2}-2\rho\frac{(x-\mu_X)}{\sigma_X}\frac{(y-\mu_Y)}{\sigma_Y}+\frac{(y-\mu_Y)^2}{\sigma_Y^2}\right]} \Big/ 2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}\ \ dx\,dy$$

Gaussian, or normal, bivariate distribution

$$f_{XY}(x,y) = e^{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_X)^2}{\sigma_X^2} - 2\rho\frac{(x-\mu_X)}{\sigma_X}\frac{(y-\mu_Y)}{\sigma_Y} + \frac{(y-\mu_Y)^2}{\sigma_Y^2}\right]} \Big/ 2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}$$

If you were to

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_X)^2}{\sigma_X^2} - 2\rho\frac{(x-\mu_X)}{\sigma_X}\frac{(y-\mu_Y)}{\sigma_Y} + \frac{(y-\mu_Y)^2}{\sigma_Y^2}\right]} \Big/ 2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2} \; dx\, dy$$

1

Gaussian, or normal, bivariate distribution

$$f_{XY}(x,y) = e^{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_X)^2}{\sigma_X^2} - 2\rho\frac{(x-\mu_X)}{\sigma_X}\frac{(y-\mu_Y)}{\sigma_Y} + \frac{(y-\mu_Y)^2}{\sigma_Y^2}\right]} \Big/ 2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}$$

marginal pdf for X is $N(\mu_x, \sigma_x)$

If you were to

$$\int_{-\infty}^{\infty} f_{XY}(x,y) = e^{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_X)^2}{\sigma_X^2} - 2\rho\frac{(x-\mu_X)}{\sigma_X}\frac{(y-\mu_Y)}{\sigma_Y} + \frac{(y-\mu_Y)^2}{\sigma_Y^2}\right]} \Big/ 2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2} \quad dy$$

If you were to

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{x} \cdot e^{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_X)^2}{\sigma_X^2} - 2\rho\frac{(x-\mu_X)}{\sigma_X}\frac{(y-\mu_Y)}{\sigma_Y} + \frac{(y-\mu_Y)^2}{\sigma_Y^2}\right]} \Big/ 2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2} \quad dx\, dy$$

Gaussian, or normal, bivariate distribution

$$f_{XY}(x,y) = e^{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_X)^2}{\sigma_X^2} - 2\rho\frac{(x-\mu_X)}{\sigma_X}\frac{(y-\mu_Y)}{\sigma_Y} + \frac{(y-\mu_Y)^2}{\sigma_Y^2}\right]} \Big/ 2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}$$

marginal pdf for X is $N(\mu_x, \sigma_x)$

If you were to

$$\int_{-\infty}^{\infty} f_{XY}(x,y) = e^{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_X)^2}{\sigma_X^2} - 2\rho\frac{(x-\mu_X)}{\sigma_X}\frac{(y-\mu_Y)}{\sigma_Y} + \frac{(y-\mu_Y)^2}{\sigma_Y^2}\right]} \Big/ 2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2} \quad dy$$

$$e^{-\frac{(x-\mu_X)^2}{2\sigma_X^2}} \Big/ \sqrt{2\pi\sigma_X^2}$$

If you were to

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{x} \cdot e^{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_X)^2}{\sigma_X^2} - 2\rho\frac{(x-\mu_X)}{\sigma_X}\frac{(y-\mu_Y)}{\sigma_Y} + \frac{(y-\mu_Y)^2}{\sigma_Y^2}\right]} \Big/ 2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2} \quad dx\, dy$$

μ_x

$$\frac{e^{-\frac{1}{2(1-\rho^2)}[\frac{(x-\mu_X)^2}{\sigma_X^2}-2\rho\frac{(x-\mu_X)}{\sigma_X}\frac{(y-\mu_Y)}{\sigma_Y}+\frac{(y-\mu_Y)^2}{\sigma_Y^2}]}}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}}$$

ρ is the correlation coefficient

$$\rho = E\{(X-\mu_X)(Y-\mu_Y)\}/(\sigma_X \sigma_Y)$$

$$e^{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_X)^2}{\sigma_X^2}-2\rho\frac{(x-\mu_X)}{\sigma_X}\frac{(y-\mu_Y)}{\sigma_Y}+\frac{(y-\mu_Y)^2}{\sigma_Y^2}\right]} \Big/ 2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}$$

Uncorrelated Gaussian variables are independent.

All independent variables are uncorrelated.

$\mathrm{cov}(X,Y) = E[(X-\mu_X)(Y-\mu_Y)] = 0$

$$
e^{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_X)^2}{\sigma_X^2}-2\rho\frac{(x-\mu_X)}{\sigma_X}\frac{(y-\mu_Y)}{\sigma_Y}+\frac{(y-\mu_Y)^2}{\sigma_Y^2}\right]} \Big/ 2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}
$$

The conditional pdf of X given Y,

$$
f_{X|Y}(x \mid y) = \frac{f_{XY}(x, y)}{f_Y(y)}
$$

, is $N(\mu_{X|Y}(y), \sigma_{X|Y}(y))$

$$
\mu_{X|Y}(y) = \mu_X + \rho\frac{\sigma_X}{\sigma_Y}(y-\mu_Y), \quad \sigma_{X|Y}(y) = \sigma_X\sqrt{1-\rho^2}
$$

$$e^{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_X)^2}{\sigma_X^2}-2\rho\frac{(x-\mu_X)}{\sigma_X}\frac{(y-\mu_Y)}{\sigma_Y}+\frac{(y-\mu_Y)^2}{\sigma_Y^2}\right]} \Big/ 2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}$$

The conditional pdf of X given Y,

$$f_{X|Y}(x\,|\,y)=\frac{f_{XY}(x,y)}{f_Y(y)}$$, is $N(\mu_{X|Y}(y),\ \sigma_{X|Y}(y))$

$$\mu_{X|Y}(y) = \mu_X + \rho\frac{\sigma_X}{\sigma_Y}(y-\mu_Y),\quad \sigma_{X|Y}(y) = \sigma_X\sqrt{1-\rho^2}$$

If you divided

$$e^{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_X)^2}{\sigma_X^2}-2\rho\frac{(x-\mu_X)}{\sigma_X}\frac{(y-\mu_Y)}{\sigma_Y}+\frac{(y-\mu_Y)^2}{\sigma_Y^2}\right]} \Big/ 2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}$$

by $\dfrac{e^{-(y-\mu_y)^2/2\sigma_y^2}}{\sqrt{2\pi\sigma_y^2}}$,

$$e^{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_X)^2}{\sigma_X^2}-2\rho\frac{(x-\mu_X)}{\sigma_X}\frac{(y-\mu_Y)}{\sigma_Y}+\frac{(y-\mu_Y)^2}{\sigma_Y^2}\right]} \Big/ 2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}$$

The conditional pdf of X given Y,

$$f_{X|Y}(x\,|\,y) = \frac{f_{XY}(x,y)}{f_Y(y)}, \text{ is } N(\mu_{X|Y}(y), \sigma_{X|Y}(y))$$

$$\mu_{X|Y}(y) = \mu_X + \rho\frac{\sigma_X}{\sigma_Y}(y-\mu_Y),\ \sigma_{X|Y}(y) = \sigma_X\sqrt{1-\rho^2}$$

If you divided

$$e^{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_X)^2}{\sigma_X^2}-2\rho\frac{(x-\mu_X)}{\sigma_X}\frac{(y-\mu_Y)}{\sigma_Y}+\frac{(y-\mu_Y)^2}{\sigma_Y^2}\right]} \Big/ 2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}$$

by $\dfrac{e^{-(y-\mu_Y)^2/2\sigma_Y^2}}{\sqrt{2\pi\sigma_Y^2}}$,

$$e^{-\left(x-\left[\mu_X+\rho\frac{\sigma_X}{\sigma_Y}(y-\mu_Y)\right]\right)^2 / 2[\sigma_X^2(1-\rho^2)]} \Big/ \sqrt{2\pi[\sigma_x^2(1-\rho^2)]}$$

The conditional pdf of X given Y,

$$f_{X|Y}(x \mid y) = \frac{f_{XY}(x,y)}{f_Y(y)}, \text{ is } N(\mu_{X|Y}(y), \sigma_{X|Y}(y))$$

$$\mu_{X|Y}(y) = \mu_X + \rho \frac{\sigma_X}{\sigma_Y}(y - \mu_Y), \quad \sigma_{X|Y}(y) = \sigma_X \sqrt{1-\rho^2}$$

After you learn what Y is,
if it's bigger than the
expected value μ_Y,
and if X and Y are positively correlated,
then you will expect X to be bigger
than its expected value μ_X.

Furthermore the additional information
reduces the uncertainty in X.

Summary of Important Equations for the Bivariate Gaussian

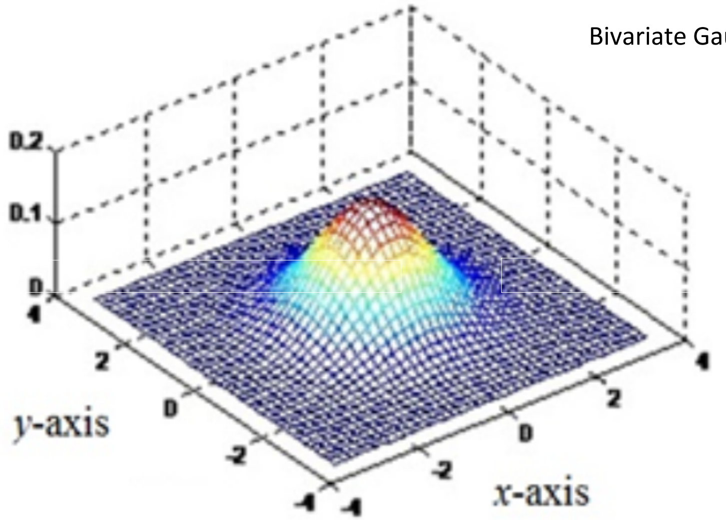
Bivariate Gaussian $f_{XY}(x,y) = e^{-\frac{1}{2(1-\rho^2)}[\frac{(x-\mu_X)^2}{\sigma_X^2} - 2\rho\frac{(x-\mu_X)}{\sigma_X}\frac{(y-\mu_Y)}{\sigma_Y} + \frac{(y-\mu_Y)^2}{\sigma_Y^2}]} \Big/ 2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}$

$$[\rho = E\{(X-\mu_X)(Y-\mu_Y)\}/(\sigma_X\,\sigma_Y)]$$

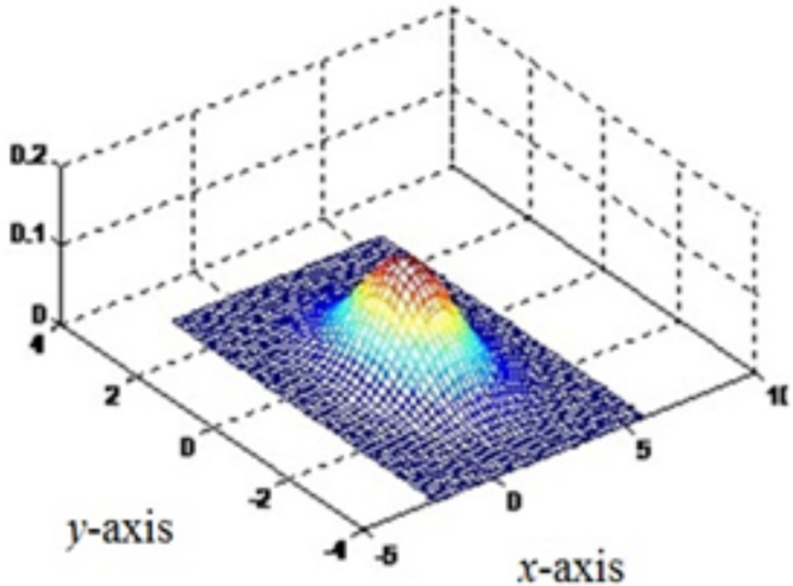
Marginal of bivariate Gaussian $f_X(x) = e^{-\frac{(x-\mu_X)^2}{2\sigma_X^2}} \Big/ \sqrt{2\pi\sigma_X^2}$

Conditional of bivariate Gaussian $f_{X|Y}(x\,|\,y) = N(\mu_X + \rho\frac{\sigma_X}{\sigma_Y}(y-\mu_Y),\ \sigma_X\sqrt{1-\rho^2})$

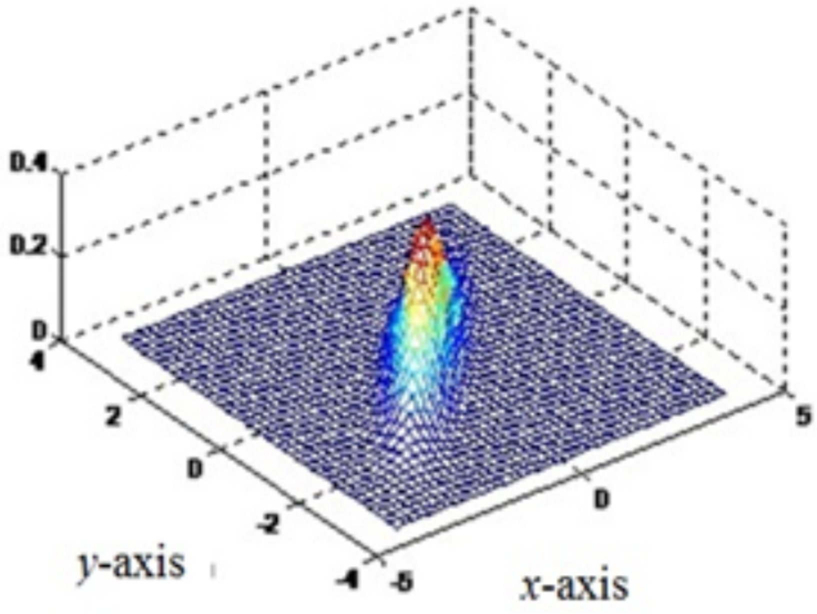
Bivariate Gaussian



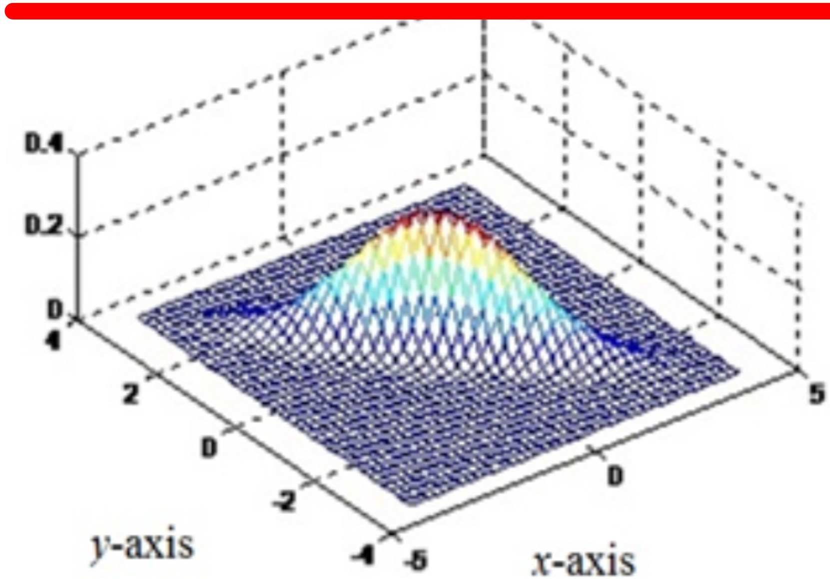
$$\mu_X = \mu_Y = 0,\ \sigma_X = \sigma_Y = 1,\ \rho = 0$$



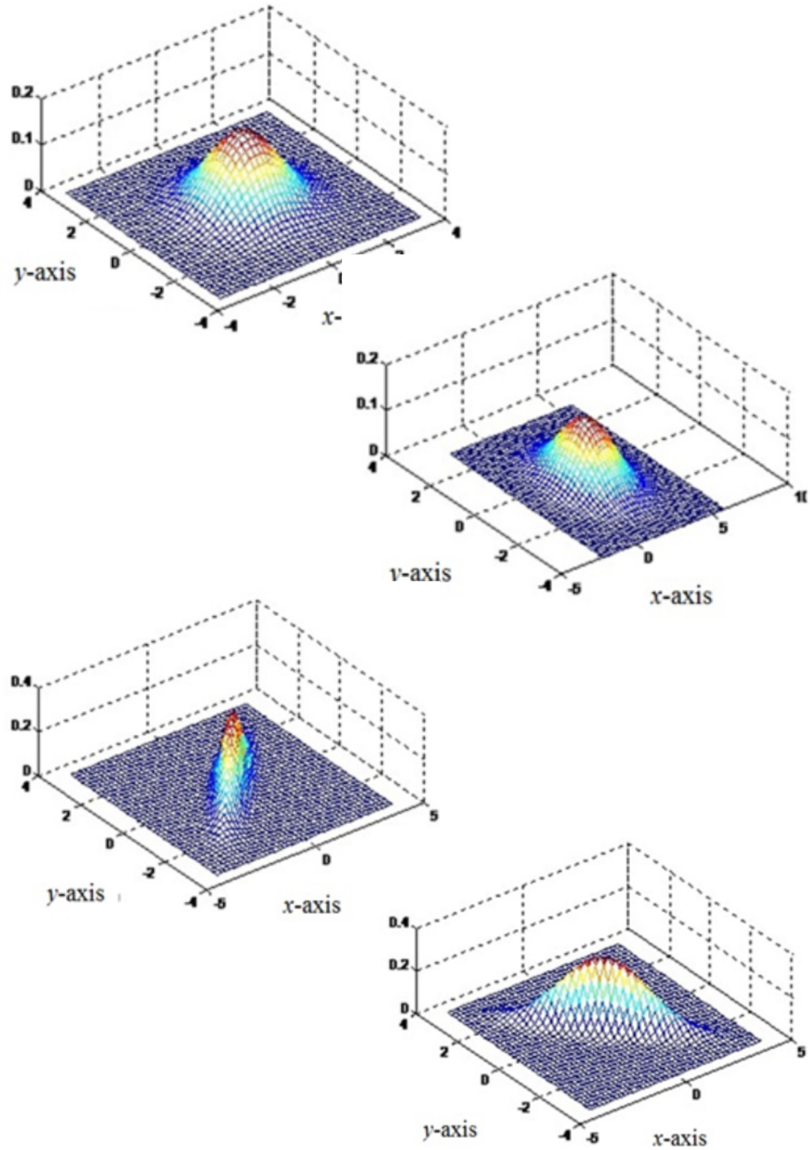
$$\mu_X = 1.8,\ \mu_Y = -0.5,\ \sigma_X = 1.1,\ \sigma_Y = 1,\ \rho = 0$$



$\mu_X = \mu_Y = 0,\ \sigma_X = 1.1,\ \sigma_Y = 1,\ \rho = 0.9$



$\mu_X = \mu_Y = 0,\ \sigma_X = 1.1,\ \sigma_Y = 1,\ \rho = -0.9$


y-axis
x-axis
y-axis
x-axis
y-axis
x-axis

Sums of Random Variables

Table 1.1 Probabilities

x	$p(X=x)$	y	$p(Y=y)$
0	0.2	2	0.3
1	0.3	3	0.4
2	0.4	4	0.1
3	0.1	5	0.2

If they are independent, what is
p(X+Y = 5)?

$$p(X=0 \ \& \ Y=5) = p(X=0)\, p(Y=5) = (0.2)\ (0.2)$$

$$p(X=1 \ \& \ Y=4) = p(X=1)\, p(Y=4) = (0.3)\ (0.1)$$

$$p(X=2 \ \& \ Y=3) = p(X=2)\, p(Y=3) = (0.4)\ (0.4)$$

$$p(X=3 \ \& \ Y=2) = p(X=3)\, p(Y=2) = (0.1)\ (0.3)$$

Table 1.1 Probabilities

x	$p(X=x)$	y	$p(Y=y)$
0	0.2	2	0.3
1	0.3	3	0.4
2	0.4	4	0.1
3	0.1	5	0.2

If they are independent, what is
p(X+Y = 5)?

$p(X=0 \ \& \ Y=5) = p(X=0)\,p(Y=5) = (0.2)\,(0.2)$

$p(X=1 \ \& \ Y=4) = p(X=1)\,p(Y=4) = (0.3)\,(0.1)$

$p(X=2 \ \& \ Y=3) = p(X=2)\,p(Y=3) = (0.4)\,(0.4)$

$p(X=3 \ \& \ Y=2) = p(X=3)\,p(Y=2) = (0.1)\,(0.3)$

$$p(X+Y=z) = \sum_{all\ x} p(X=x)\,p(Y=z-x)$$

If X and Y are independent,

$$p(X+Y=z) = \sum_{all\ x} p(X=x)\,p(Y=z-x)$$

If X and Y are independent and continuous,

then the pdf for the sum $Z = X + Y$ is

the *convolution*:

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x)\, f_Y(z-x)\, dx$$

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x)\, f_Y(z-x)\, dx \quad \text{(independent)}$$

Strategy. We'll get f_Z from f_{ZX}; we'll get f_{ZX} from $f_{Z|X}$.

Proof. Remember $Z = X + Y$.

$f_{Z|X}(z|x)\, dz =$

$p(z < Z < z + dz \mid X{=}x) = p(z < X{+}Y < z + dz \mid X{=}x)$

$= p(z - x < Y < z - x + dz \mid X{=}x)$

$= f_{Y|X}(z{-}x \mid x)\, dz$

independent $\rightarrow$ $= f_Y(z{-}x)\, dz$ so $f_{Z|X}(z|x) = f_Y(z{-}x)$.

So $f_{ZX}(z,x) = f_X(x)\, f_{Z|X}(z|x)$ and

$$f_Z(z) = \int_{-\infty}^{\infty} f_{ZX}(z,x)dx$$

$$= \int_{-\infty}^{\infty} f_X(x) f_Y(z-x)dx$$

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x)\, f_Y(z-x)\, dx \quad \text{(independent)}$$

Proof. Remember $Z = X + Y$.

$f_{Z|X}(z|x)\; dz =$

$p(z < Z < z + dz \mid X{=}x) = p(z < X{+}Y < z + dz \mid X{=}x)$

$= p(z - x < Y < z - x + dz \mid X{=}x)$

$= f_{Y|X}(z\text{-}x \mid x)\; dz$

independent → $= f_Y(z\text{-}x)\; dz$ so $f_{Z|X}(z|x) = f_Y(z\text{-}x)$.

So $f_{ZX}(z,x) = f_X(x)\, f_{Z|X}(z|x)$ and

$$f_Z(z) = \int_{-\infty}^{\infty} f_{ZX}(z,x)dx$$

$$= \int_{-\infty}^{\infty} f_X(x) f_Y(z-x)dx$$

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x)\, f_Y(z-x)\, dx \quad \text{(independent)}$$

Proof. Remember $Z = X + Y$.

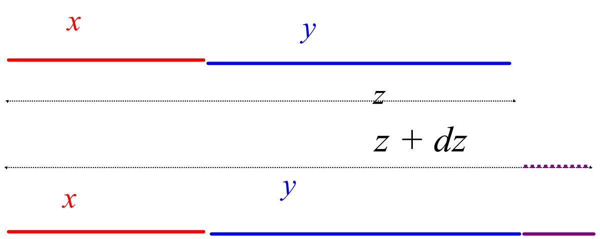
$f_{Z|X}(z|x)\ dz =$

$$p(z < Z < z + dz \mid X{=}x) = p(z < X{+}Y < z + dz \mid X{=}x)$$

$$= p(z - x < Y < z - x + dz \mid X{=}x)$$

$$= f_{Y|X}(z\text{-}x \mid x)\ dz$$

independent $\rightarrow$ $= f_Y(z\text{-}x)\ dz$ so $f_{Z|X}(z|x) = f_Y(z\text{-}x)$.



So $f_{ZX}(z,x) = f_X(x)\, f_{Z|X}(z|x)$ and

$$f_z(z) = \int_{-\infty}^{\infty} f_{ZX}(z,x)dx$$

$$= \int_{-\infty}^{\infty} f_X(x) f_Y(z-x)dx$$

$$f_{X+Y}(z) = \int_{-\infty}^{\infty} f_X(x)\, f_Y(z-x)\, dx$$

1. If X and Y are independent and Gaussian, then $X + Y$ is also Gaussian.

$$\mu_{X+Y} = \mu_X + \mu_Y$$

$$\sigma_{X+Y}^2 = \sigma_X^2 + \sigma_Y^2$$

2. If $X_1, X_2, X_3, \ldots, X_n$ are independent, not Gaussian, but have the same pdf, then

$$[X_1 + X_2 + X_3 + \ldots + X_n] \div n$$

has a pdf which approaches Gaussian as n becomes large.

$$\mu_{\sum X_n} = n\mu_X\,,\ \sigma_{\sum X_n}^2 = n\sigma_X^2$$

$$\mu_{\sum X_n/n} = \mu_X\,,\ \sigma_{\sum X_n/n}^2 = \frac{\sigma_X^2}{n}\,,\ \sigma_{\sum X_n/n} = \frac{\sigma_X}{\sqrt{n}}$$

1. If X and Y are independent and Gaussian, then $X + Y$ is also Gaussian.

$$\mu_{X+Y} = \mu_X + \mu_Y$$

$$\sigma_{X+Y}^2 = \sigma_X^2 + \sigma_Y^2$$

$$\mu_{X+Y} = \iint (x+y) f_{XY}(x,y)dx\,dy$$

$$= \iint x f_{XY}(x,y)dx\,dy + \iint y f_{XY}(x,y)dx\,dy$$

$$= \mu_X + \mu_Y$$

$$\sigma_{X+Y}^2 = \iint [(x + y - \mu_X - \mu_Y\,)^2] f_{XY}(x,y)dx\,dy$$

$$= \iint [(x - \mu_X\,)^2] f_{XY}(x,y)dx\,dy \;\leftarrow \sigma_X^2$$

$$+ \iint [(y - \mu_Y\,)^2] f_{XY}(x,y)dx\,dy \;\leftarrow \sigma_Y^2$$

$$+2 \iint (x - \mu_X\,)(y - \mu_Y\,) f_{XY}(x,y)dx\,dy$$

$$+2 \iint (x - \mu_X)(y - \mu_Y) f_{XY}(x,y) dx\, dy$$

$$2 \iint (x - \mu_X)(y - \mu_Y) f_{XY}(x,y) dx\, dy$$

$$2 \iint (x - \mu_X)(y - \mu_Y) f_X(x) f_Y(y) dx\, dy$$

$$2 \int (x - \mu_X) f_X(x) dx \int (y - \mu_Y) f_Y(y) dy$$

$$\mu_X - \mu_X \qquad\qquad \mu_Y - \mu_Y$$

But why is X+Y Gaussian?

$$f_{X+Y}(z) = \int_{-\infty}^{\infty} f_X(x)\, f_Y(z-x)\, dx$$

Fourier Convolution Theorem

$$\text{FT}\{f_{X+Y}\} = \text{FT}\{f_X\}\ \text{FT}\{f_Y\}$$

(characteristic functions)

$$= e^{\text{quadratic}}\, e^{\text{quadratic}}$$

$$\text{FT}\{f_{X+Y}\} = e^{\text{quadratic}}$$

So f_{X+Y} is Gaussian.

2. If X_1, X_2, X_3, ..., X_n are independent, not Gaussian, but have the same pdf, then

$$[X_1 + X_2 + X_3 + \ldots + X_n] \div n$$

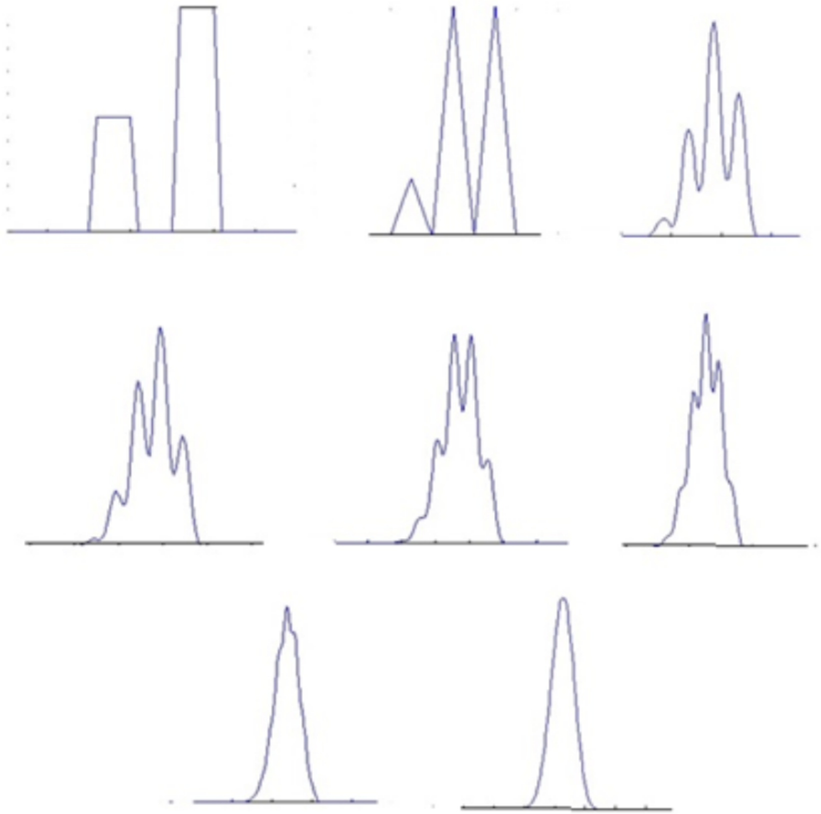
has a pdf which approaches Gaussian as n becomes large.

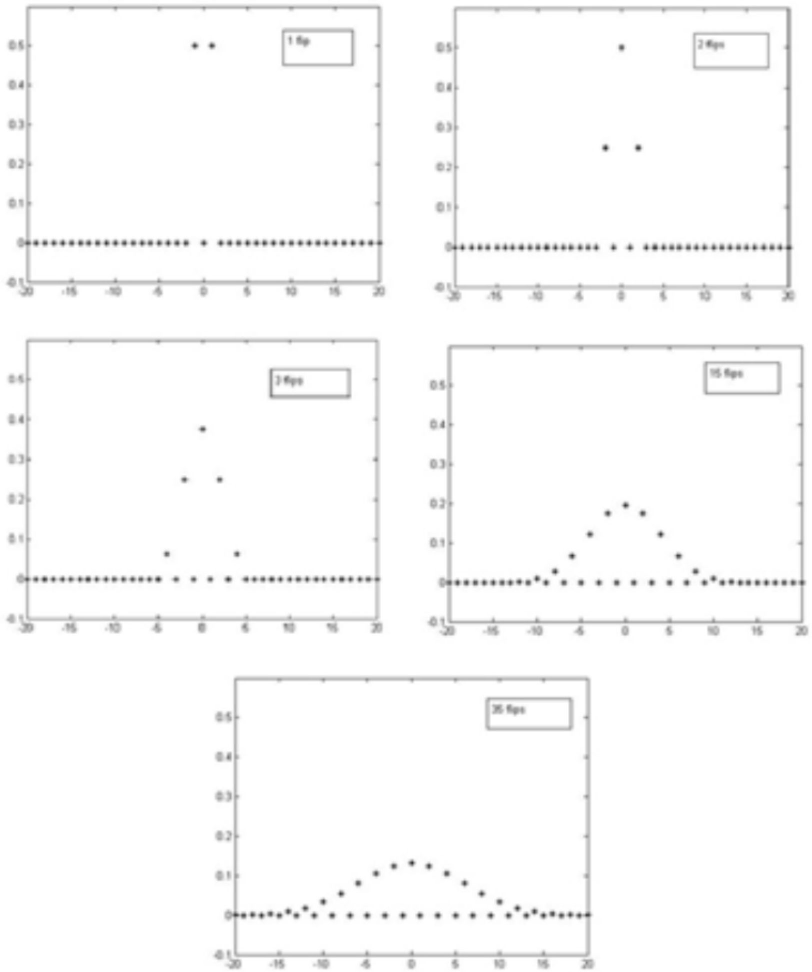
$$\mu_{\sum X_n} = n\mu_X \,, \quad \sigma^2_{\sum X_n} = n\sigma^2_X$$

$$\mu_{\sum X_n/n} = \mu_X \,, \quad \sigma^2_{\sum X_n/n} = \frac{\sigma^2_X}{n} \,, \quad \sigma_{\sum X_n/n} = \frac{\sigma_X}{\sqrt{n}}$$

The pdf of $X_1 + X_2 + X_3 + \ldots + X_n$ equals the pdf of X_1 convolved with the pdf of X_2, convolved with the pdf of X_3, and so on.

NO MATTER WHAT YOU START WITH, IF YOU KEEP CONVOLVING IT WITH ITSELF, YOU'LL EVENTUALLY GET A GAUSSIAN!




1 flip
2 flips
15 flips
35 flips

Equations for Sums of Random Variables

Sum $Z = X+Y$: if X & Y are independent,

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x)\, f_Y(z-x)\, dx$$

If X & Y are independent and Gaussian,

$$f_Z(z) = N(\mu_X + \mu_Y, \sqrt{\sigma_X^2 + \sigma_Y^2})$$

Central Limit Theorem: if $\{X_i\}$ are iid then

$$Z = \sum_{i=1}^{n} X_i \text{ implies } f_Z(z) \approx_{n\to\infty} N(n\mu_X, \sqrt{n}\sigma_X),$$

$$<X> = \frac{1}{n}\sum_{i=1}^{n} X_i \text{ implies } f_{<X>}(x) \approx_{n\to\infty} N(\mu_X, \sigma_X / \sqrt{n})$$

The Multivariate Gaussian

$$\mathbf{Cov} = \begin{bmatrix} E\{(X-\mu_X)^2\} & E\{(X-\mu_X)(Y-\mu_Y)\} \\ E\{(X-\mu_X)(Y-\mu_Y)\} & E\{(Y-\mu_Y)^2\} \end{bmatrix}$$

$$f_{\mathbf{X}}(\mathbf{x}) = e^{-\frac{1}{2}[\mathbf{x}-\mu][\mathbf{Cov}]^{-1}[\mathbf{x}-\mu]^T} \Big/ (2\pi)^{n/2}\sqrt{\det(\mathbf{Cov})}$$

WE ❤ THE GAUSSIAN BECAUSE

1. The Gaussian's bell-shaped curve is ubiquitous in nature.

2. The Gaussian's pdf is characterized by 2 parameters:

   mean and standard deviation.

   (But this is true of many other distributions as well.)

3. Any linear affine transformation ($aX+b$) of Gaussian variables is Gaussian. (So we only need to tabulate a standardized version using $[X-\mu]/\sigma$.)

4. Joint Gaussians have Gaussian marginals.

5. Joint Gaussians have Gaussian conditionals.

WE ❤ THE GAUSSIAN BECAUSE

6. Independence is equivalent to zero ~~correlation.~~ covariance

   (This is true of some other distributions as well.)

7. Any sum of Gaussian variables is Gaussian.

8. Sums of large numbers of independent identically distributed random variables have distributions that approach Gaussian, *even if the individual variables are not Gaussian*.

9. The Gaussian's characteristic function is Gaussian.

10. *Any* characteristic function whose logarithm is a polynomial is the characteristic of a Gaussian.

WE HATE the Gaussian because

you can't integrate e^{-x^2} in closed form.

Lecture 6
(2/1/2017)

Equations for Sums of Random Variables

Sum $Z = X+Y$: if X & Y are independent,

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x)\, f_Y(z-x)\, dx$$

If X & Y are independent and Gaussian,

$$f_Z(z) = N(\mu_X + \mu_Y, \sqrt{\sigma_X^2 + \sigma_Y^2})$$

Central Limit Theorem: if $\{X_i\}$ are iid then

$$Z = \sum_{i=1}^{n} X_i \text{ implies } f_Z(z) \approx_{n\to\infty} N(n\mu_X, \sqrt{n}\sigma_X),$$

$$\langle X \rangle = \frac{1}{n}\sum_{i=1}^{n} X_i \text{ implies } f_{\langle X \rangle}(x) \approx_{n\to\infty} N(\mu_X, \sigma_X / \sqrt{n})$$

!

Chapter 2 Random Processes
2.1 Examples of random processes
2.2 The Mathematical Characterization of Random Processes
2.3 Prediction: The Statistician's Task

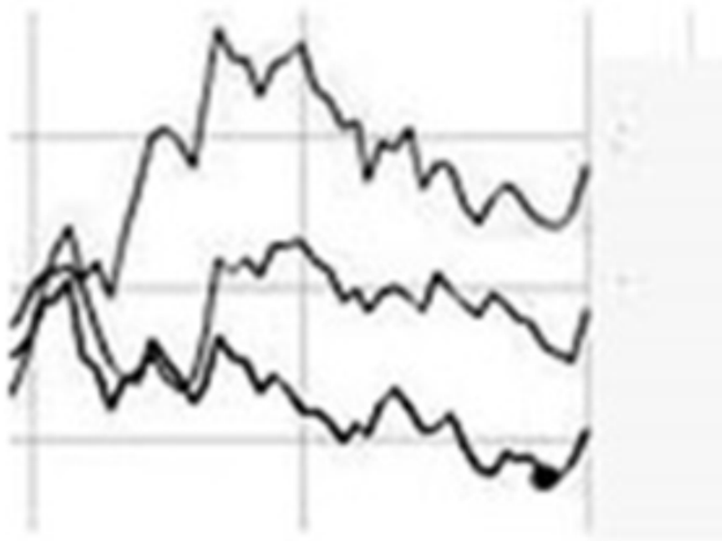


Figure 2.1 Stock market samples

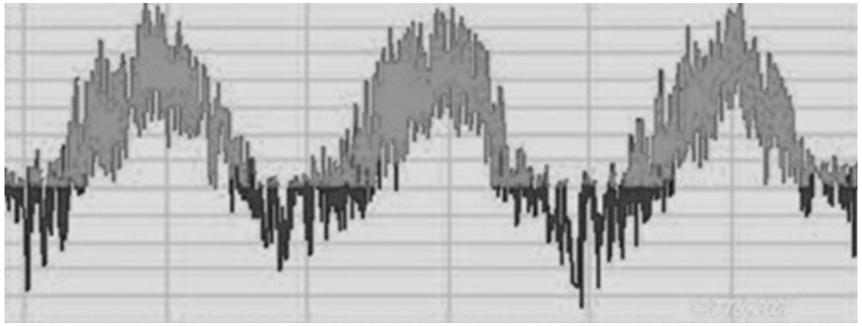


Figure 2.2　3-year temperature chart

Figure 2.3 Johnson noise

Figure 2.4 Shot noise

Figure 2.5 Popcorn noise

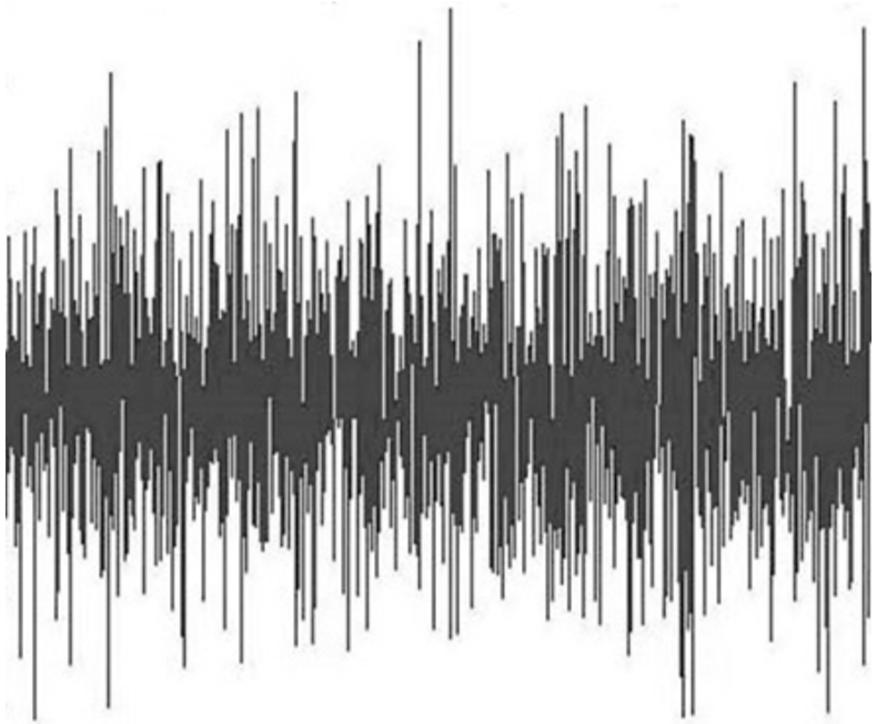


Figure 2.6 ARMA simulation

Autoregressive Moving Average

THTHHHTHTH…

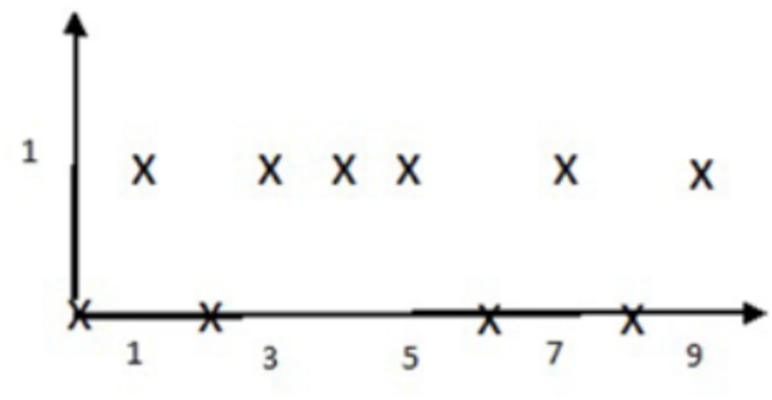


Figure 2.7 Bernoulli Process

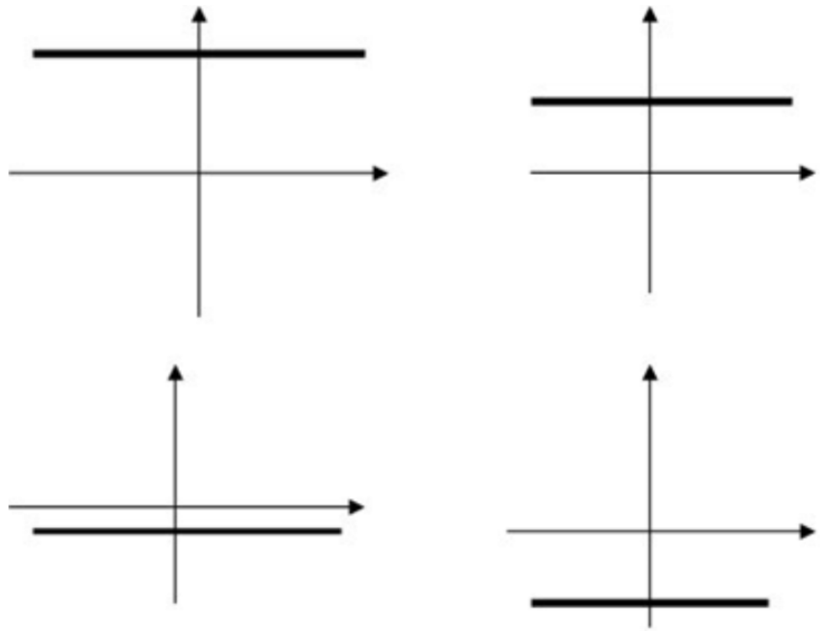


Figure 2.8   Random Settings for a DC Power Supply

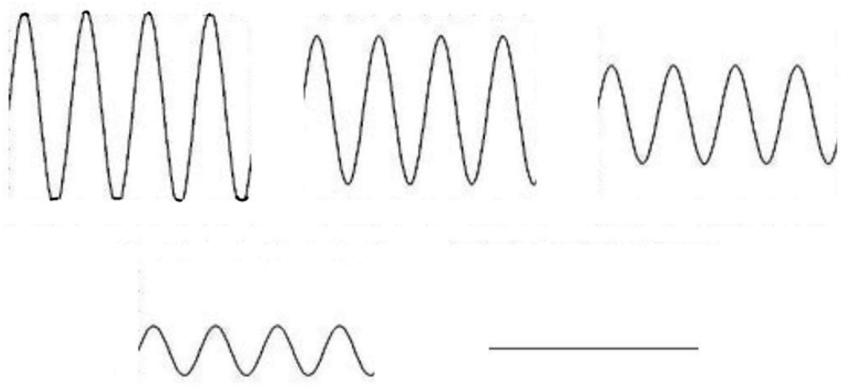


Figure 2.9   Random Settings for an AC Power Supply

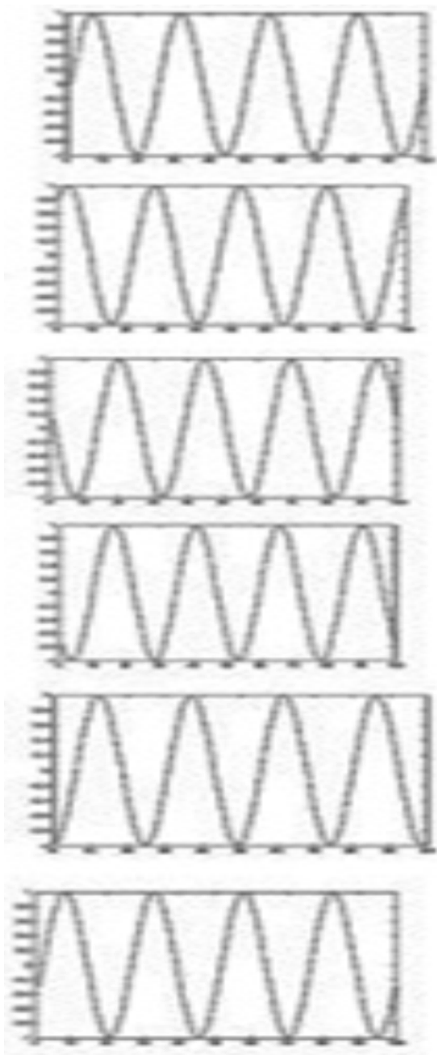


Figure 2.10 AC Power Supply Voltages
with Random Phase

the probability that the value of

X at time τ lies between a and b

$$= \int_{x=a}^{b} f_{X(\tau)}(x)\, dx$$

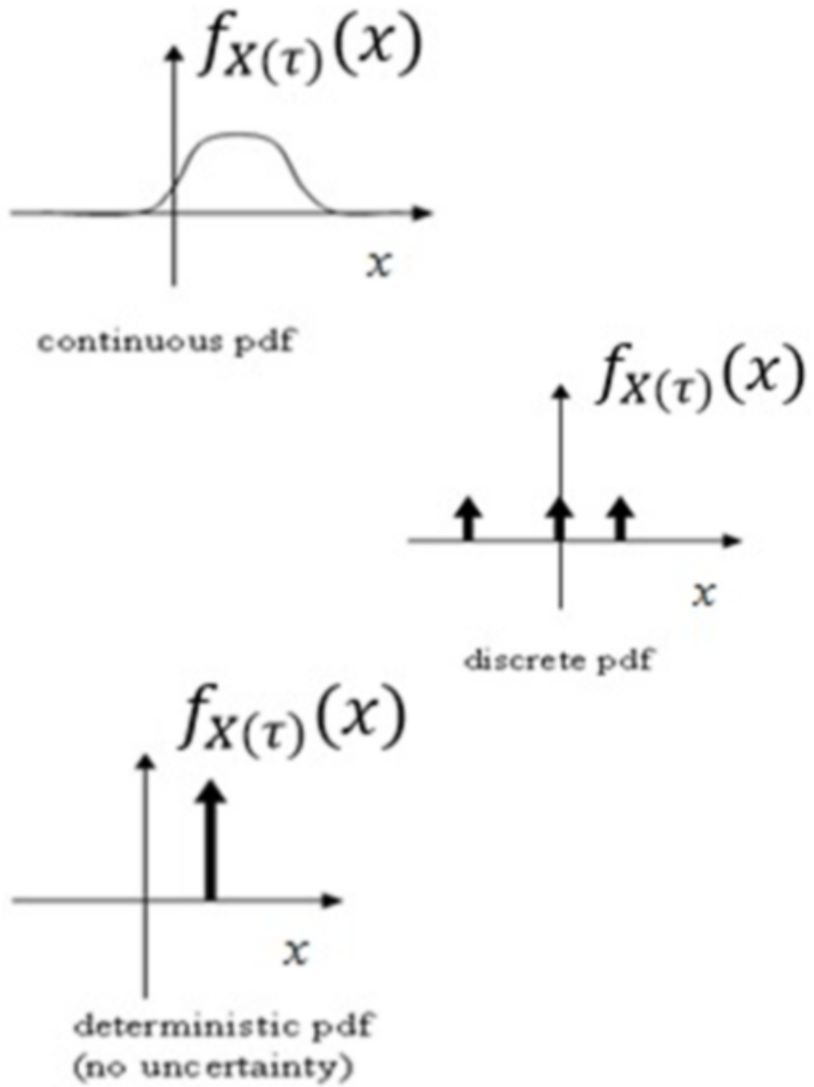


Figure 2.11 Continuous, Discrete, and Deterministic pdf's for a Random Process

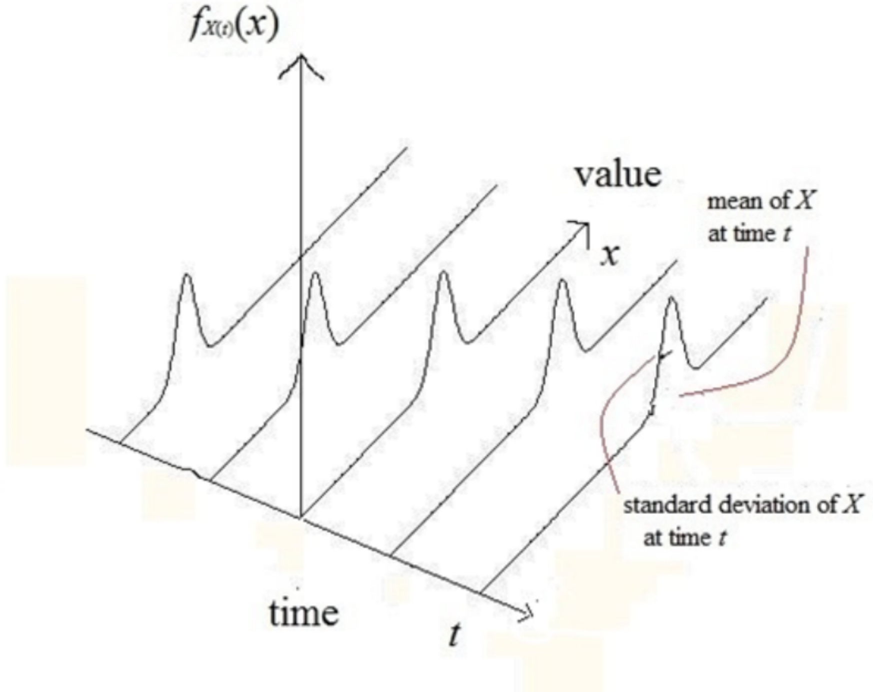
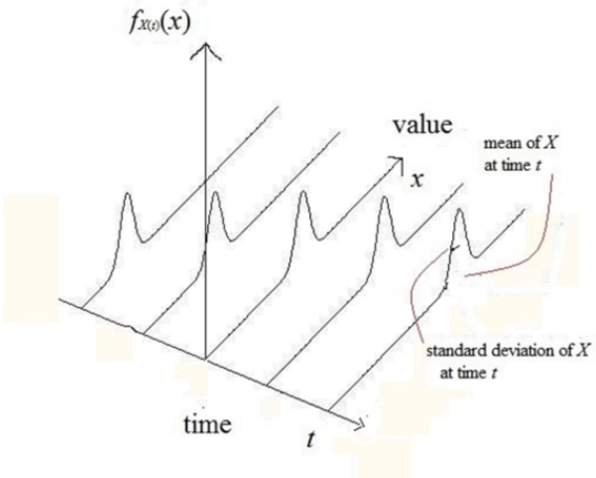
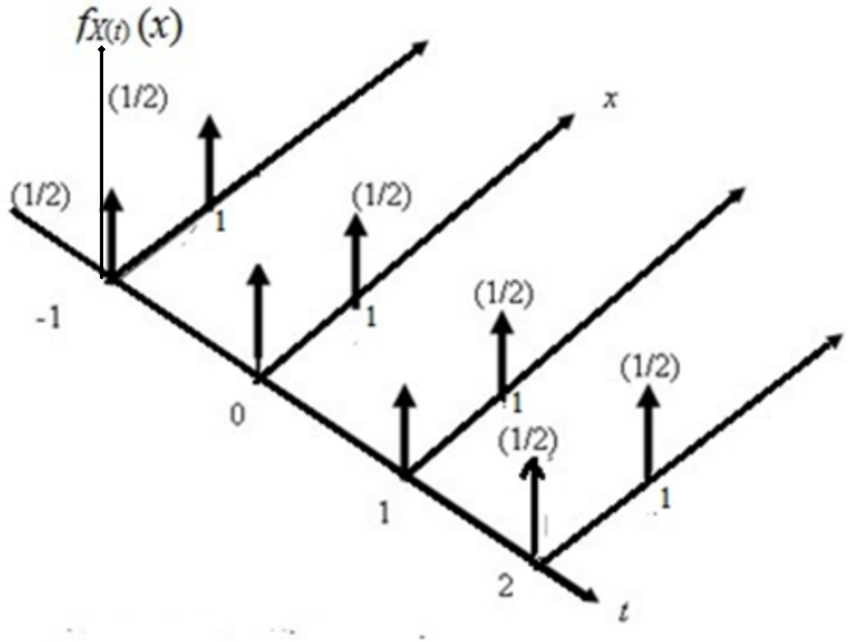


Figure 2.12   pdf's for X(t) at different times





$f_{X(t)}(x)$ for the Bernoulli Process

$$f_{X|Y}(x|y)$$

probability density for $X(t_2)$ taking the value x_2,

given that $X(t_1)$ took the value x_1.

$$f_{X(t_2)|X(t_1)}(x_2 \mid x_1)$$

joint pdf for $X(t_1)$ and $X(t_2)$

$$f_{X(t_1),X(t_2)}(x_1, x_2) = f_{X(t_1)}(x_1) f_{X(t_2)|X(t_1)}(x_2 \mid x_1)$$

$$f_{X(t_3)|X(t_1)\&X(t_2)}(x_3 \mid x_1, x_2)$$

$$f_{X(t_1),X(t_2),X(t_3)}(x_1, x_2, x_3)$$
$$= f_{X(t_1)}(x_1)\, f_{X(t_2)|X(t_1)}(x_2 \mid x_1)\, f_{X(t_3)|X(t_1)\&X(t_2)}(x_3 \mid x_1, x_2)$$

$$f_{X(t_1)}(x_1) = \int_{-\infty}^{\infty} f_{X(t_1),X(t_2)}(x_1, x_2)\, dx_2\ )$$

$$f_{X(t_3)|X(t_1)\&X(t_2)}(x_3 \mid x_1, x_2)$$

$$f_{X(t_1),X(t_2),X(t_3)}(x_1, x_2, x_3)$$
$$= f_{X(t_1)}(x_1)\, f_{X(t_2)|X(t_1)}(x_2 \mid x_1)\, f_{X(t_3)|X(t_1)\&X(t_2)}(x_3 \mid x_1, x_2)$$

$$f_{X(t_1)}(x_1) = \int_{-\infty}^{\infty} f_{X(t_1),X(t_2)}(x_1, x_2)\, dx_2\ )$$

Gaussian random processes.

All f's are normal.

$$f_{X(t_3)|X(t_1)\&X(t_2)}(x_3 \mid x_1, x_2)$$

$$f_{X(t_1),X(t_2),X(t_3)}(x_1, x_2, x_3)$$
$$= f_{X(t_1)}(x_1)\, f_{X(t_2)|X(t_1)}(x_2 \mid x_1)\, f_{X(t_3)|X(t_1)\&X(t_2)}(x_3 \mid x_1, x_2)$$

$$f_{X(t_1)}(x_1) = \int_{-\infty}^{\infty} f_{X(t_1),X(t_2)}(x_1, x_2)\, dx_2 \;)$$

Independent random processes

$$f_{X(t_3)|X(t_1)\&X(t_2)}(x_3 \mid x_1, x_2) = f_{X(t_3)}(x_3)$$

$$f_{X(t_1),X(t_2),X(t_3)}(x_1, x_2, x_3)$$
$$= f_{X(t_1)}(x_1)\, f_{X(t_2)}(x_2)\, f_{X(t_3)}(x_3)$$

$$\text{mean of } X(t) = E\{X(t)\} = \int_{-\infty}^{\infty} x\, f_{X(t)}(x)\, dx = \text{"}\mu_X(t)\text{"}$$

autocorrelation of $X(t)$ = correlation between $X(t_1)$ and $X(t_2)$

$$\begin{aligned} &= E\{X(t_1)X(t_2)\} \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1\, x_2\, f_{X(t_1),X(t_2)}(x_1, x_2)\, dx_1\, dx_2 \\ &= \text{"}R_X(t_1, t_2)\text{"} \end{aligned}$$

mean of $X(t) = E\{X(t)\} = \int_{-\infty}^{\infty} x\, f_{X(t)}(x)\, dx =$ "$\mu_X(t)$"

autocorrelation of $X(t)$ = correlation between $X(t_1)$ and $X(t_2)$

$$= E\{X(t_1)X(t_2)\}$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1\, x_2\, f_{X(t_1),X(t_2)}(x_1, x_2)\, dx_1\, dx_2$$

$$= \text{"}R_X(t_1, t_2)\text{"}$$

autocovariance of $X(t)$

= correlation between $[X(t_1) - \mu_X(t_1)]$ and $[X(t_2) - \mu_X(t_2)]$

$$= E\{[X(t_1) - \mu_X(t_1)][X(t_2) - \mu_X(t_2)]\}$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} [x_1 - \mu_X(t_1)][x_2 - \mu_X(t_2)]\, f_{X(t_1),X(t_2)}(x_1, x_2)\, dx_1\, dx_2$$

$$= \text{"}C_X(t_1, t_2)\text{"}$$

What is $C_X(t, t)$?

$$\text{mean of } X(t) = E\{X(t)\} = \int_{-\infty}^{\infty} x\, f_{X(t)}(x)\, dx = \text{"}\mu_X(t)\text{"}$$

autocorrelation of $X(t)$ = correlation between $X(t_1)$ and $X(t_2)$

$$= E\{X(t_1)X(t_2)\}$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1\, x_2\, f_{X(t_1),X(t_2)}(x_1, x_2)\, dx_1\, dx_2$$

$$= \text{"}R_X(t_1, t_2)\text{"}$$

autocovariance of $X(t)$

$= $ correlation between $[X(t_1) - \mu_X(t_1)]$ and $[X(t_2) - \mu_X(t_2)]$

$$= E\{[X(t_1) - \mu_X(t_1)][X(t_2) - \mu_X(t_2)]\}$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} [x_1 - \mu_X(t_1)][x_2 - \mu_X(t_2)]\, f_{X(t_1),X(t_2)}(x_1, x_2)\, dx_1\, dx_2$$

$$= \text{"}C_X(t_1, t_2)\text{"}$$

What is $C_X(t, t)$?

$\sigma_{X(t)}^{\;2}$

$$\text{mean of } X(t) = E\{X(t)\} = \int_{-\infty}^{\infty} x\, f_{X(t)}(x)\, dx = \text{"}\mu_X(t)\text{"}$$

autocorrelation of $X(t)$ = correlation between $X(t_1)$ and $X(t_2)$

$$= E\{X(t_1)X(t_2)\}$$
$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1 x_2\, f_{X(t_1),X(t_2)}(x_1, x_2)\, dx_1\, dx_2$$
$$= \text{"}R_X(t_1, t_2)\text{"}$$

autocovariance of $X(t)$

= correlation between $[X(t_1) - \mu_X(t_1)]$ and $[X(t_2) - \mu_X(t_2)]$

$$= E\{[X(t_1) - \mu_X(t_1)][X(t_2) - \mu_X(t_2)]\}$$
$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} [x_1 - \mu_X(t_1)][x_2 - \mu_X(t_2)]\, f_{X(t_1),X(t_2)}(x_1, x_2)\, dx_1\, dx_2$$
$$= \text{"}C_X(t_1, t_2)\text{"} \qquad [C_X(t, t) = \sigma_{X(t)}^{\ 2}]$$

correlation coefficient of $X(t)$

$$= \rho_X(t_1, t_2) = C_X(t_1, t_2) / \sqrt{C_X(t_1, t_1) C_X(t_2, t_2)}\ ,$$

$$-1 \le \rho_X(t_1, t_2) \le 1$$

$$\text{mean of } X(t) = E\{X(t)\} = \int_{-\infty}^{\infty} x\, f_{X(t)}(x)\, dx = "\mu_X(t)"$$

autocorrelation of $X(t)$ = correlation between $X(t_1)$ and $X(t_2)$

$$= E\{X(t_1)X(t_2)\}$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1\, x_2\, f_{X(t_1),X(t_2)}(x_1,x_2)\, dx_1\, dx_2$$

$$= "R_X(t_1,t_2)"$$

autocovariance of $X(t)$

= correlation between $[X(t_1)-\mu_X(t_1)]$ and $[X(t_2)-\mu_X(t_2)]$

$$= E\{[X(t_1)-\mu_X(t_1)][X(t_2)-\mu_X(t_2)]\}$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}[x_1-\mu_X(t_1)][x_2-\mu_X(t_2)]\, f_{X(t_1),X(t_2)}(x_1,x_2)\, dx_1\, dx_2$$

$$= "C_X(t_1,t_2)"$$

Second Moment Identity

$$R_X(t_1,t_2) = C_X(t_1,t_2) + \mu_X(t_1)\mu_X(t_2)$$

$$\text{mean of } X(t) = E\{X(t)\} = \int_{-\infty}^{\infty} x\, f_{X(t)}(x)\, dx = \text{"}\mu_X(t)\text{"}$$

autocorrelation of $X(t)$ = correlation between $X(t_1)$ and $X(t_2)$

$$= E\{X(t_1)X(t_2)\}$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1 x_2\, f_{X(t_1),X(t_2)}(x_1,x_2)\, dx_1\, dx_2$$

$$= \text{"}R_X(t_1,t_2)\text{"}$$

autocovariance of $X(t)$

= correlation between $[X(t_1)-\mu_X(t_1)]$ and $[X(t_2)-\mu_X(t_2)]$

$$= E\{[X(t_1)-\mu_X(t_1)][X(t_2)-\mu_X(t_2)]\}$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} [x_1-\mu_X(t_1)][x_2-\mu_X(t_2)]\, f_{X(t_1),X(t_2)}(x_1,x_2)\, dx_1\, dx_2$$

$$= \text{"}C_X(t_1,t_2)\text{"}$$

Second Moment Identity

$$R_X(t_1,t_2) = C_X(t_1,t_2) + \mu_X(t_1)\mu_X(t_2)$$

Two Random Processes $X(t)$, $Y(t)$

crosscorrelation

$$= \text{correlation between } X(t_1) \text{ and } Y(t_2)$$

$$= E\{X(t_1)Y(t_2)\}$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1\, y_2\, f_{X(t_1),Y(t_2)}(x_1, y_2)\, dx_1\, dy_2$$

$$= \text{"}R_{XY}(t_1, t_2)\text{"}$$

crosscovariance

$$= \text{correlation between } [X(t_1) - \mu_X(t_1)] \text{ and } [Y(t_2) - \mu_Y(t_2)]$$

$$= E\{[X(t_1) - \mu_X(t_1)][Y(t_2) - \mu_Y(t_2)]\}$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} [x_1 - \mu_X(t_1)][y_2 - \mu_Y(t_2)]\, f_{X(t_1),Y(t_2)}(x_1, y_2)\, dx_1\, dy_2$$

$$= \text{"}C_{XY}(t_1, t_2)\text{"}$$

Summary: The First and Second Moments of Random Processes

mean of $X(t) = E\{X(t)\} = \int_{-\infty}^{\infty} x\, f_{X(t)}(x)\,dx =$ "$\mu_X(t)$"

autocorrelation of $X(t)$ = correlation between $X(t_1)$ and $X(t_2)$

$$= E\{X(t_1)X(t_2)\}$$
$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1 x_2 f_{X(t_1),X(t_2)}(x_1,x_2)\,dx_1\,dx_2$$
$$= \text{"}R_X(t_1,t_2)\text{"}$$

autocovariance of $X(t)$ = correlation between $[X(t_1)-\mu_X(t_1)]$ and $[X(t_2)-\mu_X(t_2)]$

$$= E\{[X(t_1)-\mu_X(t_1)][X(t_2)-\mu_X(t_2)]\}$$
$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}[x_1-\mu_X(t_1)][x_2-\mu_X(t_2)]f_{X(t_1),X(t_2)}(x_1,x_2)\,dx_1\,dx_2$$
$$= \text{"}C_X(t_1,t_2)\text{"}$$

correlation coefficient of $X(t) = \rho_X(t_1,t_2) = C_X(t_1,t_2)/\sqrt{C_X(t_1,t_1)C_X(t_2,t_2)}$,

$$-1 \le \rho_X(t_1,t_2) \le 1$$

crosscorrelation of two random processes $X(t)$ and $Y(t)$

= correlation between $X(t_1)$ and $Y(t_2)$

$$= E\{X(t_1)Y(t_2)\}$$
$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1 y_2 f_{X(t_1),Y(t_2)}(x_1,y_2)\,dx_1\,dy_2$$
$$= \text{"}R_{XY}(t_1,t_2)\text{"}$$

$$R_X(t_1,t_2) = C_X(t_1,t_2) + \mu_X(t_1)\mu_X(t_2)$$
$$R_{XY}(t_1,t_2) = C_{XY}(t_1,t_2) + \mu_X(t_1)\mu_Y(t_2)$$

Example 1. For the Bernoulli process, the probability of 0 (or 1) on a given toss is 0.5, and the probability of 0 (or 1) on one toss and 0 (or 1) on another toss is $0.5^2 = 0.25$.

$$\mu_X(t) = (0.5) \times 0 + (0.5) \times 1$$

$$= 0.5 \text{ if } t = 0, \pm 1, \pm 2, \cdots \text{ (0 otherwise)}$$

$$\mu_X(t) = (0.5) \times 0 + (0.5) \times 1$$

$= 0.5$ if $t = 0, \pm 1, \pm 2, \cdots$ ~~(0 otherwise)~~

$$R_X(t_1, t_2) =$$

$$\frac{1}{4}(0 \times 0) + \frac{1}{4}(0 \times 1) + \frac{1}{4}(1 \times 0) + \frac{1}{4}(1 \times 1)$$

$= \frac{1}{4}$ if $t_1 \neq t_2$, both integers

$$R_X(t_1, t_2) = \frac{1}{2}(0 \times 0) + \frac{1}{2}(1 \times 1)$$

$= \frac{1}{2}$ if $t_1 = t_2$, both integers

~~$R_X(t_1, t_2) = 0$ otherwise~~ .

$$\mu_X(t) = (0.5) \times 0 + (0.5) \times 1$$

$$= 0.5 \text{ if } t = 0, \pm 1, \pm 2, \cdots \text{ (0 otherwise)}$$

$$R_X(t_1, t_2) =$$

$$\frac{1}{4}(0 \times 0) + \frac{1}{4}(0 \times 1) + \frac{1}{4}(1 \times 0) + \frac{1}{4}(1 \times 1)$$

$$= \frac{1}{4} \text{ if } t_1 \neq t_2, \text{ both integers}$$

$$R_X(t_1, t_2) = \frac{1}{2}(0 \times 0) + \frac{1}{2}(1 \times 1)$$

$$= \frac{1}{2} \text{ if } t_1 = t_2, \text{ both integers}$$

$$R_X(t_1, t_2) = 0 \text{ otherwise} \quad .$$

$$R_X(t_1, t_2) = C_X(t_1, t_2) + \mu_X(t_1)\mu_X(t_2)$$

$$C_X(t_1, t_2) = \frac{1}{4} - \left(\frac{1}{2}\right)^2 = 0 \text{ if } t_1 \neq t_2, \text{ both integers}$$

$$C_X(t_1, t_2) = \frac{1}{2} - \left(\frac{1}{2}\right)^2 = \frac{1}{4} \text{ if } t_1 = t_2, \text{ both integers}$$

$$R_X(t_1,t_2) = C_X(t_1,t_2) + \mu_X(t_1)\mu_X(t_2)$$

$$C_X(t_1,t_2) = \frac{1}{4} - \left(\frac{1}{2}\right)^2 = 0 \text{ if } t_1 \neq t_2 \text{, both integers}$$

$$C_X(t_1,t_2) = \frac{1}{2} - \left(\frac{1}{2}\right)^2 = \frac{1}{4} \text{ if } t_1 = t_2 \text{, both integers}$$

Remember $\rho \propto$ cov,

and $\rho = 0$ if they are independent !!!!

Example 2. Suppose $X(t)$ is a zero-mean

Gaussian process with autocorrelation function

$R_X(t_1, t_2) = 4e^{-|t_1 - t_2|}$.

Example 2. Suppose $X(t)$ is a zero-mean

Gaussian process with autocorrelation function

$R_X(t_1, t_2) = 4e^{-|t_1 - t_2|}$.

$$N(\mu, \sigma) = f_X(x) = e^{-\frac{(x-\mu)^2}{2\sigma^2}} \Big/ \sqrt{2\pi\sigma^2}$$

Example 2. Suppose $X(t)$ is a zero-mean

Gaussian process with autocorrelation function

$R_X(t_1, t_2) = 4e^{-|t_1 - t_2|}$.

$$\frac{e^{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_X)^2}{\sigma_X^2} - 2\rho\frac{(x-\mu_X)}{\sigma_X}\frac{(y-\mu_Y)}{\sigma_Y} + \frac{(y-\mu_Y)^2}{\sigma_Y^2}\right]}}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}}$$

X is $X(t_1)$ and Y is $X(t_2)$

ρ is the correlation coefficient

$$\rho = E\{(X-\mu_X)(Y-\mu_Y)\}/(\sigma_X\sigma_Y)$$

Example 2. Suppose $X(t)$ is a zero-mean

Gaussian process with autocorrelation function

$R_X(t_1, t_2) = 4e^{-|t_1 - t_2|}$.

0

$$\frac{e^{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_X)^2}{\sigma_X^2} - 2\rho\frac{(x-\mu_X)}{\sigma_X}\frac{(y-\mu_Y)}{\sigma_Y} + \frac{(y-\mu_Y)^2}{\sigma_Y^2}\right]}}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}}$$

ρ is the correlation coefficient

$\rho = E\{(X-\mu_X)(Y-\mu_Y)\}/(\sigma_X \sigma_Y)$

$$\text{mean of } X(t) = E\{X(t)\} = \int_{-\infty}^{\infty} x\, f_{X(t)}(x)\, dx = \text{"}\mu_X(t)\text{"}$$

autocorrelation of $X(t)$ = correlation between $X(t_1)$ and $X(t_2)$

$$= E\{X(t_1)X(t_2)\}$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1 x_2\, f_{X(t_1),X(t_2)}(x_1,x_2)\, dx_1\, dx_2$$

$$= \text{"}R_X(t_1,t_2)\text{"}$$

autocovariance of $X(t)$

= correlation between $[X(t_1)-\mu_X(t_1)]$ and $[X(t_2)-\mu_X(t_2)]$

$$= E\{[X(t_1)-\mu_X(t_1)][X(t_2)-\mu_X(t_2)]\}$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} [x_1-\mu_X(t_1)][x_2-\mu_X(t_2)]\, f_{X(t_1),X(t_2)}(x_1,x_2)\, dx_1\, dx_2$$

$$= \text{"}C_X(t_1,t_2)\text{"}$$

$[C_X(t, t) = \sigma_{X(t)}^2]$

correlation coefficient of $X(t)$

$$= \rho_X(t_1,t_2) = C_X(t_1,t_2)/\sqrt{C_X(t_1,t_1)C_X(t_2,t_2)}\ ,$$

$$-1 \le \rho_X(t_1,t_2) \le 1$$

mean of $X(t) = E\{X(t)\} = \int_{-\infty}^{\infty} x\, f_{X(t)}(x)\, dx = "\mu_X(t)"$

autocorrelation of $X(t)$ = correlation between $X(t_1)$ and $X(t_2)$

$$= E\{X(t_1)X(t_2)\}$$
$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1\, x_2\, f_{X(t_1),X(t_2)}(x_1,x_2)\, dx_1\, dx_2$$
$$= "R_X(t_1,t_2)"$$

autocovariance of $X(t)$

= correlation between $[X(t_1) - \mu_X(t_1)]$ and $[X(t_2) - \mu_X(t_2)]$

$$= E\{[X(t_1) - \mu_X(t_1)][X(t_2) - \mu_X(t_2)]\}$$
$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} [x_1 - \mu_X(t_1)][x_2 - \mu_X(t_2)]\, f_{X(t_1),X(t_2)}(x_1,x_2)\, dx_1\, dx_2$$
$$= "C_X(t_1,t_2)" \quad = R_X(t_1,t_2)$$

$[C_X(t, t) = \sigma_{X(t)}^2]$

correlation coefficient of $X(t)$

$$= \rho_X(t_1,t_2) = C_X(t_1,t_2) / \sqrt{C_X(t_1,t_1)\, C_X(t_2,t_2)}\ ,$$

$$-1 \le \rho_X(t_1,t_2) \le 1$$

Example 2. Suppose $X(t)$ is a zero-mean

Gaussian process with autocorrelation function

$R_X(t_1, t_2) = 4e^{-|t_1 - t_2|}$.

(i) What is the probability that the value of $X(7)$

lies between 6 and 7?

$X(7)$ is normal, so its pdf is normal

with mean and standard deviation

$$\mu_X(7) = 0, \sigma_X(7) = \sqrt{C_X(7,7)} = \sqrt{4e^{-|7-7|}} = 2$$

$$\int_6^7 N(0,2)\, dx \approx 0.0011$$

Example 2. Suppose $X(t)$ is a zero-mean

Gaussian process with autocorrelation function

$R_X(t_1, t_2) = 4e^{-|t_1 - t_2|}$.

(ii) If $X(6)$ is measured and found

to have the value 1, what is the

probability that the value of $X(7)$

lies between 6 and 7?

$$\rho = \mathrm{cov}(X, Y)/(\sigma_X \sigma_y)$$

$$\mu_{X|Y} = \mu_X + \rho \frac{\sigma_X}{\sigma_Y}(y - \mu_Y)$$

$$\sigma_{X|Y} = \sigma_X \sqrt{1 - \rho^2}$$

$$\rho = \text{cov}(X,Y)/(\sigma_X \sigma_y)$$

$$\mu_{X|Y} = \mu_X + \rho \frac{\sigma_X}{\sigma_Y}(y - \mu_Y)$$

$$\sigma_{X|Y} = \sigma_X \sqrt{1-\rho^2}$$

$$R_X(t_1, t_2) = 4e^{-|t_1 - t_2|}$$

(*ii*) If $X(6)$ is measured and found

to have the value 1

$$\mu_{X(7)} = \mu_{X(6)} = 0, \sigma_{X(7)} = \sigma_{X(6)} = \sqrt{4e^0} = 2,$$

$$\rho_{X(7)X(6)} = \frac{4e^{-|7-6|}}{2 \times 2} = e^{-1} \approx 0.3769,$$

$$\mu_{X(7)|X(6)} = 0 + e^{-1}\frac{2}{2}(1-0) = e^{-1} \approx 0.3769,$$

$$\sigma_{X(7)|X(6)} = 2\sqrt{1-e^{-2}} \approx 1.8597 \text{ .}$$

$$\rho = \text{cov}(X,Y)/(\sigma_X \sigma_y)$$

$$\mu_{X|Y} = \mu_X + \rho \frac{\sigma_X}{\sigma_Y}(y - \mu_Y)$$

$$\sigma_{X|Y} = \sigma_X \sqrt{1-\rho^2}$$

$$R_X(t_1,t_2) = 4e^{-|t_1-t_2|}$$

(*ii*) If $X(6)$ is measured and found

to have the value 1, what is the

$$\mu_{X(7)} = \mu_{X(6)} = 0, \sigma_{X(7)} = \sigma_{X(6)} = \sqrt{4e^0} = 2,$$

$$\rho_{X(7)X(6)} = \frac{4e^{-|7-6|}}{2 \times 2} = e^{-1} \approx 0.3769,$$

$$\mu_{X(7)|X(6)} = 0 + e^{-1}\frac{2}{2}(1-0) = e^{-1} \approx 0.3769,$$

$$\sigma_{X(7)|X(6)} = 2\sqrt{1-e^{-2}} \approx 1.8597\ .$$

$$\int_6^7 N(0.3769, 1.8597)dx \approx 0.0010$$

Example 2. Suppose $X(t)$ is a zero-mean

Gaussian process with autocorrelation function

$R_X(t_1, t_2) = 4e^{-|t_1 - t_2|}$.

(iii) What is the probability that $X(6)$

and $X(7)$ both lie between 0 and 7?

$$\frac{e^{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_X)^2}{\sigma_X^2} - 2\rho\frac{(x-\mu_X)}{\sigma_X}\frac{(y-\mu_Y)}{\sigma_Y} + \frac{(y-\mu_Y)^2}{\sigma_Y^2}\right]}}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}}$$

$$\mu_{X(7)} = \mu_{X(6)} = 0, \sigma_{X(7)} = \sigma_{X(6)} = \sqrt{4e^0} = 2,$$

$$\rho_{X(7)X(6)} = \frac{4e^{-|7-6|}}{2 \times 2} = e^{-1} \approx 0.3769$$

$$\int_0^7 \int_0^7 (\cdots) dx\, dy \approx 0.3095$$

Example 3. Describe the mean and autocorrelation of a switch that is turned on at a random time between $t=0$ and $t=1$. That is, the process $X(t)$ is zero for $t<c$ and one for $t>c$, where c is a random variable uniformly distributed in [0,1].

Example 3. Describe the mean and autocorrelation of a switch that is turned on at a random time between t=0 and t=1. That is, the process $X(t)$ is zero for $t<c$ and one for $t>c$, where c is a random variable uniformly distributed in [0,1].

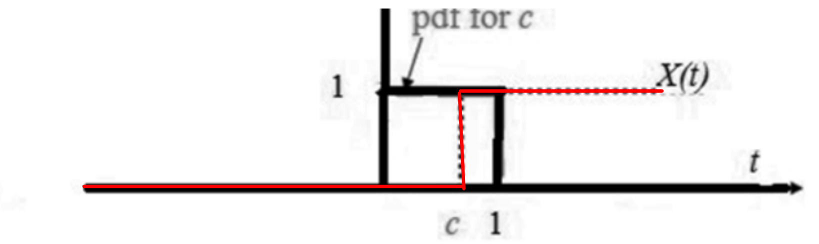


Figure 2.14 Random switching function

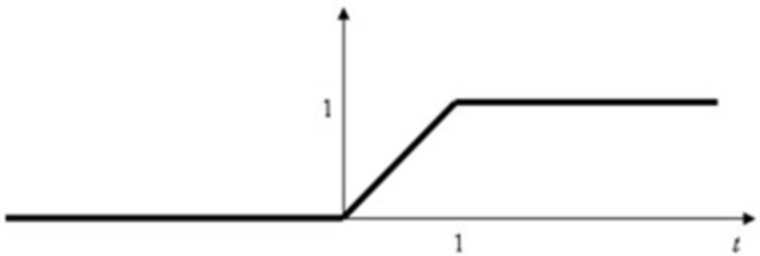
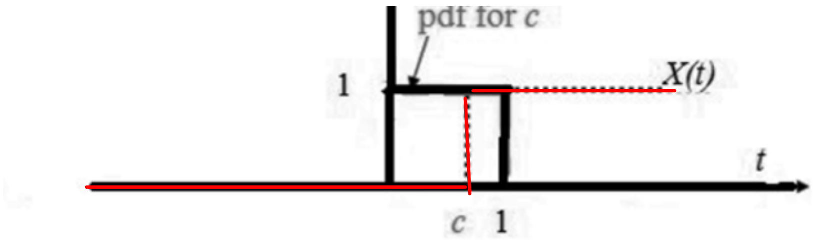


Figure 2.15 Probability that X(t) = 1 (and, the mean of X(t))

$E\{X(t)\} = 0 \times p[X(t)=0] + 1 \times p[X(t) = 1]$

Example 3. Describe the mean and autocorrelation of a switch that is turned on at a random time between t=0 and t=1. That is, the process $X(t)$ is zero for t<c and one for t>c, where c is a random variable uniformly distributed in [0,1].



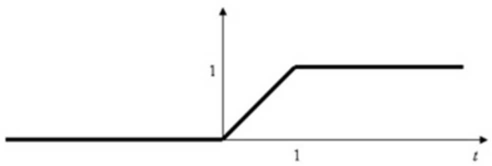


Figure 2.15 Probability that X(t) = 1 (and, the mean of X(t))

$X(t_1)X(t_2)$ is either 0 or 1.

E{$(X(t_1)X(t_2))$}:

if $\min(t_1, t_2) < 0$ certainly $X(t_1)X(t_2) = 0$;

if $\min(t_1, t_2) > 1$ certainly $X(t_1)X(t_2) = 1$;

otherwise $p[X(t_1)X(t_2) = 1] = \min(t_1, t_2)$.

Lecture 7
(2/6/2017)

Example 2. Suppose $X(t)$ is a zero-mean

Gaussian process with autocorrelation function

$R_X(t_1, t_2) = 4e^{-|t_1 - t_2|}$.

(*iii*) What is the probability that $X(6)$

and $X(7)$ both lie between 0 and 7?

$$\frac{e^{-\frac{1}{2(1-\rho^2)}[\frac{(x-\mu_X)^2}{\sigma_X^2} - 2\rho\frac{(x-\mu_X)}{\sigma_X}\frac{(y-\mu_Y)}{\sigma_Y} + \frac{(y-\mu_Y)^2}{\sigma_Y^2}]}}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}}$$

$$\mu_{X(7)} = \mu_{X(6)} = 0, \sigma_{X(7)} = \sigma_{X(6)} = \sqrt{4e^0} = 2,$$

$$\rho_{X(7)X(6)} = \frac{4e^{-|7-6|}}{2 \times 2} = e^{-1} \approx 0.3769$$

$$\int_0^7 \int_0^7 (\cdots) dx\, dy \approx 0.3095$$

Example 3. Describe the mean and autocorrelation of a switch that is turned on at a random time between t=0 and t=1. That is, the process $X(t)$ is zero for $t<c$ and one for $t>c$, where c is a random variable uniformly distributed in [0,1]. c is the turn-on time.

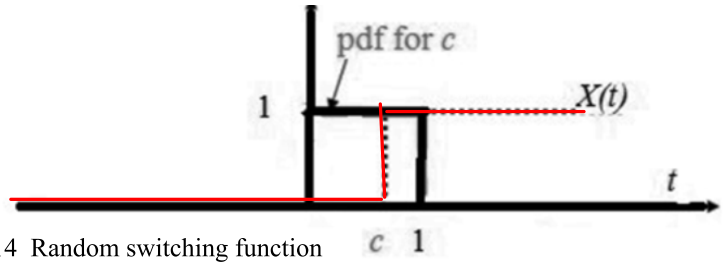


Figure 2.14 Random switching function

$$p\{X(t)=1\} = p\{\text{turned on before } t\}$$

$$= p\{c<t\}$$

$$= 0 \text{ if } t<0 , \quad = 1 \text{ if } t>1 ,$$

$$= \int_0^t (1)dt = t \quad if \;\; 0 < t < 1$$

Example 3. Describe the mean and autocorrelation of a switch that is turned on at a random time between t=0 and t=1. That is, the process $X(t)$ is zero for t<c and one for t>c, where c is a random variable uniformly distributed in [0,1].

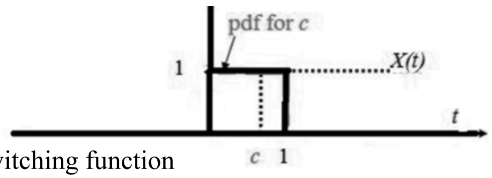


Figure 2.14 Random switching function

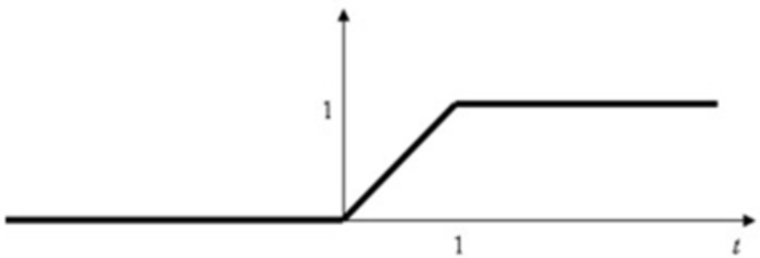


Figure 2.15 Probability that X(t) = 1 (and, the mean of X(t))

$$E\{X(t)\} = 0 \times p[X(t)=0] + 1 \times p[X(t) = 1]$$

$$= p[X(t) = 1]$$

Example 3. Describe the mean and autocorrelation of a switch that is turned on at a random time between t=0 and t=1. That is, the process $X(t)$ is zero for t<c and one for t>c, where c is a random variable uniformly distributed in [0,1].

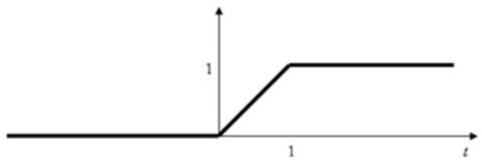


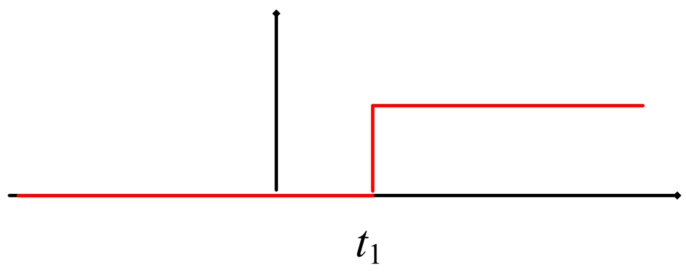
Figure 2.15 Probability that X(t) = 1 (and, the mean of X(t))

$E\{(X(t_1)X(t_2)\}$:

$X(t_1)X(t_2)$ is either 0 or 1.

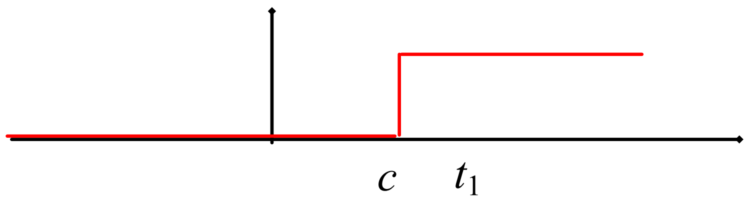
The turn-on time is c.

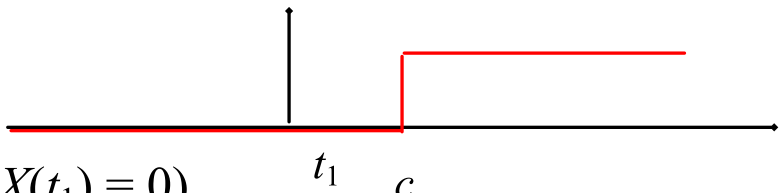
Remember: t_1 is NOT the turn-on time for X. It is a time when X is observed. c is the turn-on time. Same for t_2.



THIS IS NOT A GRAPH OF $X(t_1)$!!

THESE are possible graphs of $X(t_1)$:





$(X(t_1) = 0)$

Example 3. Describe the mean and autocorrelation of a switch that is turned on at a random time between t=0 and t=1. That is, the process $X(t)$ is zero for t<c and one for t>c, where c is a random variable uniformly distributed in [0,1].

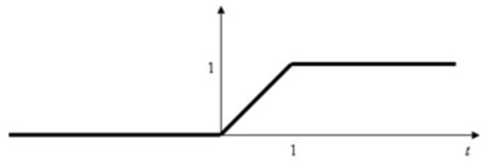
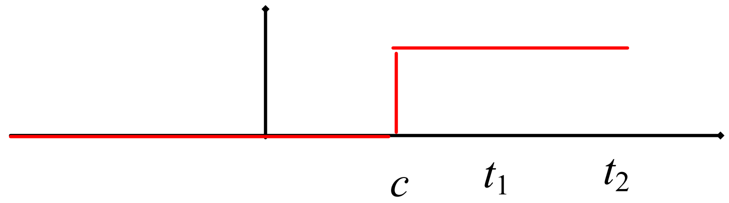


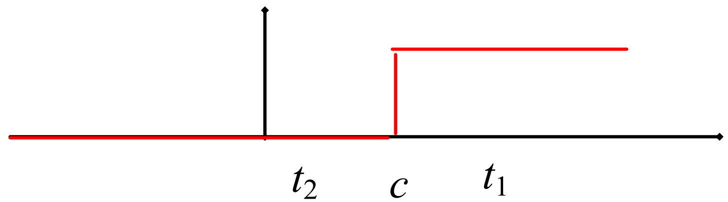
Figure 2.15 Probability that X(t) = 1 (and, the mean of X(t))

$X(t_1)X(t_2)$ is either 0 or 1.

The turn-on time is c.

E{($X(t_1)X(t_2)$}:





Example 3. Describe the mean and autocorrelation of a switch that is turned on at a random time between t=0 and t=1. That is, the process $X(t)$ is zero for $t<c$ and one for $t>c$, where c is a random variable uniformly distributed in [0,1].

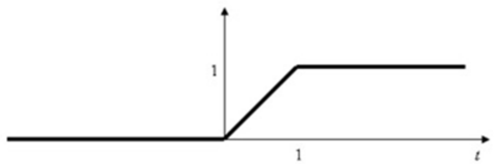


Figure 2.15 Probability that X(t) = 1 (and, the mean of X(t))

$X(t_1)X(t_2)$ is either 0 or 1.

The turn-on time is c.

E{$(X(t_1)X(t_2)$}: **let t** = min(t_1, t_2)

if t < 0 certainly $X(t_1)X(t_2) = 0$;

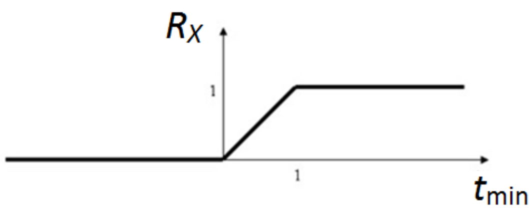
if t > 1 certainly $X(t_1)X(t_2) = 1$;

if 0 < t < 1, then $X(t_1)X(t_2)$ = 0 if c > t

and $X(t_1)X(t_2)$ = 1 if c < t.

E{$(X(t_1)X(t_2)$} = 0 x p{c>t} + 1 x p(c<t)

= p(c<t) = t



PREDICTION

B = Boston temperature

N = New York temperature

Goal: get an estimator of B from N

$$\hat{B} = \alpha + \beta N$$

How should we choose α and β ?

Mean squared error:

$$\text{error } B - \hat{B} = B - (\alpha + \beta N)$$

$$\text{MSE} \equiv E\{(B - \hat{B})^2\} = E\{[B - (\alpha + \beta N)]^2\}$$

The optimal predictor will have the least mean squared error.

$$\text{MSE} \equiv E\{(B-\hat{B})^2\} = E\{[B-(\alpha+\beta N)]^2\}$$

$$E\{[B-(\alpha+\beta N)]^2\}$$
$$= E\{B^2+\alpha^2+\beta^2 N^2-2\alpha B-2\beta BN+2\alpha\beta N\}$$
$$= E\{B^2\}+\alpha^2+\beta^2 E\{N^2\}-2\alpha E\{B\}-2\beta E\{BN\}+2\alpha\beta E\{N\}$$

The MSE depends on α, β. To minimize,

$$\frac{\partial}{\partial\alpha}E\{[B-(\alpha+\beta N]^2\}$$
$$=2\alpha-2E\{B\}+2\beta E\{N\}=0$$

$$\frac{\partial}{\partial\beta}E\{[B-(\alpha+\beta N]^2\}$$
$$=2\beta E\{N^2\}-2E\{BN\}+2\alpha E\{N\}=0$$

$$\frac{\partial}{\partial \alpha} E\{[B-(\alpha+\beta N]^2\}$$

$$= 2\alpha - 2E\{B\} + 2\beta E\{N\} = 0$$

$$\frac{\partial}{\partial \beta} E\{[B-(\alpha+\beta N]^2\}$$

$$= 2\beta E\{N^2\} - 2E\{BN\} + 2\alpha E\{N\} = 0$$

$$\begin{bmatrix} & \\ & \end{bmatrix} \begin{bmatrix} \\ \end{bmatrix} = \begin{bmatrix} \\ \end{bmatrix}$$

$$\alpha = \frac{E\{B\}E\{N^2\} - E\{N\}E\{BN\}}{E\{N^2\} - E\{N\}^2} = \mu_B - \rho \frac{\sigma_B}{\sigma_N} \mu_N$$

$$\beta = \frac{E\{BN\} - E\{N\}E\{B\}}{E\{N^2\} - E\{N\}^2} = \rho \frac{\sigma_B}{\sigma_N}$$

$$\hat{B} = \alpha + \beta N$$

$$\alpha = \mu_B - \rho \frac{\sigma_B}{\sigma_N} \mu_N$$

$$\beta = \rho \frac{\sigma_B}{\sigma_N}$$

(What *is* the least possible mean squared error?)

$$E\{[B - (\alpha + \beta N)]^2\} = (1 - \rho^2)\sigma_B^2$$

$$= 0 \text{ if } \rho = \pm 1$$

$$= \sigma_B^2 \text{ if } \rho = 0$$

Chapter 3 Analysis of Raw Data: Spectral Methods
3.1 Stationarity and Ergodicity
3.2 The Limit Concept in Random Processes
3.3 Spectral Methods for Obtaining Autocorrelations
3.4 Interpretation of the Discrete Time Fourier Transform
3.5 The Power Spectral Density
3.6 Interpretation of the Power Spectral Density
3.7 Engineering the Power Spectral Density
3.8 Back to Estimating the Autocorrelation
3.9 The Secret of Bartlett's Method
3.10 Spectral Analysis for Continuous Random Processes

STATIONARITY and ERGODICITY

Samples of $X(t)$

$$X(\Delta t), X(2\Delta t), \cdots, X(N\Delta t)$$

$$X(1), X(2), \ldots, X(N)$$

Estimate the mean

$$E\{X(N+1)\} \approx \frac{X(1)+X(2)+\ldots+X(N)}{N}$$

Estimate the mean

$$E\{X(N+1)\} \approx \frac{X(1)+X(2)+\ldots+X(N)}{N}$$

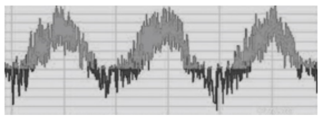
PRESUMPTION:

$$E\{X(N+1)\} = E\{X(N+2)\}$$

$$E\{X(m)\} = E\{X(n)\} \text{ for all } m,n$$

$$E\{X(t_1)\} = E\{X(t_2)\}$$

Yes



No?

Process is STATIONARY

$$E\{X(N+1)\} = E\{X(N+2)\}$$

$$E\{X(m)\} = E\{X(n)\} \text{ for all } m,n$$

$$E\{X(t_1)\} = E\{X(t_2)\}$$

$$= \mu_X \text{ (independent of time)}$$

"Stationary *in the mean*"

weak-sense stationarity

wide-sense stationarity (WSS)

covariance stationarity

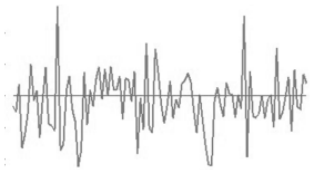
second-order stationarity

strict stationarity

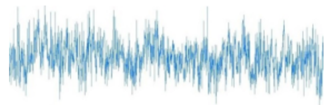
strong-sense stationarity

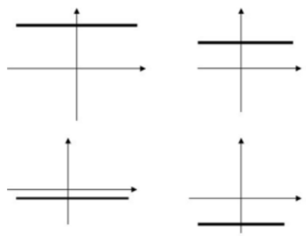
Estimate the mean

$$E\{X(N+1)\} \approx \frac{X(1)+X(2)+...+X(N)}{N}$$

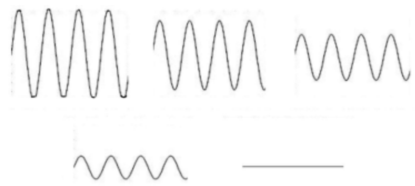


Yes





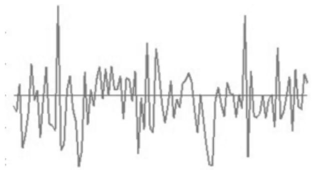
No



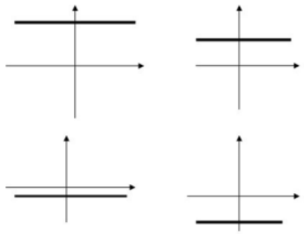
PRESUMPTION:

Averages over all samples can be estimated by taking time averages for any particular sample.

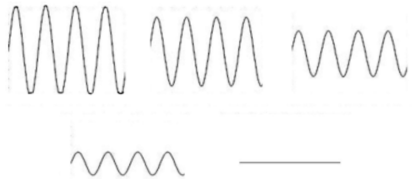
A process is ERGODIC if ensemble means equal time averages for a particular sample.



Yes



No



"Ergodic *in the mean*"

weak-sense ergodicity

wide-sense ergodicity

covariance ergodicity

second-order ergodicity

strict ergodicity

strong-sense ergodicity

$$\text{mean of } X(t) = E\{X(t)\} = \text{"}\mu_X(t)\text{"}$$

$$= \int_{-\infty}^{\infty} x\, f_{X(t)}(x)\, dx$$

STATIONARY

$$\mu_X(t) = \mu_X$$

ERGODIC

$$\mu_X \approx \frac{X(1) + X(2) + \ldots + X(N)}{N}$$

$$\approx \frac{1}{2T} \int_{-T}^{T} X(t)\, dt$$

Does ergodicity presume stationarity?

A process is ERGODIC if ensemble means equal time averages for a particular sample.

If we can't predict the future from the past, what can we predict it from?

A process is ERGODIC if ensemble means equal time averages for a particular sample.

"Those who do not learn history

are doomed to repeat it."

George Santayana

Lecture 8
(2/8/2017)

$$\hat{B} = \alpha + \beta N$$

$$\alpha = \mu_B - \rho \frac{\sigma_B}{\sigma_N} \mu_N$$

$$\beta = \rho \frac{\sigma_B}{\sigma_N}$$

You need to know the mean and standard deviations of both *B* and *N*, and their correlation.

CONVERGENCE; LIMITS

$$\frac{dx}{dt} = \lim_{\Delta t \to 0} \frac{x(t + \Delta t) - x(t)}{\Delta t}$$

$$Y_n(t) \Rightarrow Y(t) \qquad ?????$$

1. Convergence in probability

2. Convergence with probability 1

3. Convergence in distribution

4. Mean-Square convergence:

$$\mathrm{E}\{[Y_n(t) - Y(t)]^2\} \Rightarrow 0$$

Section 3.2

Process is STATIONARY

$$E\{X(N+1)\} = E\{X(N+2)\}$$

$E\{X(m)\} = E\{X(n)\}$ for all m,n

$$E\{X(t_1)\} = E\{X(t_2)\}$$

$= \mu_X$ (independent of time)

$$E\{X(N+1)\} \approx \frac{X(1) + X(2) + \ldots + X(N)}{N}$$

$\approx E\{X\}$ (for all N)

STATIONARY

$$R_X(1002,1001) \approx \frac{X(2)X(1) + X(3)X(2) + \ldots + X(1000)X(999)}{999}$$

$$\approx R_X(1002,\ 1003)$$

$$\approx R_X(1023,\ 1024)$$

$$\approx R_X(N,\ N+1) \approx R_X(N,\ N-1)$$

$R_X(m,n)$ depends only on $|m-n|$.

Rewrite it as $R_X(|m-n|)$ or $R_X(p)$

$$E\{X(N+1)\} \approx \frac{X(1)+X(2)+\ldots+X(N)}{N}$$

easy to calculate

$R_X(1002,1001) \approx$ $R_X(1) \approx$

$$\frac{X(2)X(1)+X(3)X(2)+\ldots+X(1000)X(999)}{999}$$

also have $R_X(0)$, $R_X(2)$, ...

$$\hat{R}_X(m) = \frac{1}{\text{number of terms}} \sum_n X(n)X(n-m)$$

looks like convolution:

$$X \circ Y(m) \equiv \sum_{n=-\infty}^{\infty} X(n)Y(m-n)$$

(Let $Y = X_{\text{time-reversed}}$;

truncate to finite number of terms ;

divide by number of terms.)

$$\hat{R}_X(m) = \frac{1}{\text{number of terms}} \sum_n X(n)X(n-m)$$

looks like convolution:

$$X \circ Y(m) \equiv \sum_{n=-\infty}^{\infty} X(n)Y(m-n)$$

FOURIER says
Fourier Transform of convolution
=
(FT of first) x (FT of second)

USE Fast FT to get FT$\{X(n)\}$.

J.W. Cooley and John Tukey
1965

$$\hat{R}_X(m) = \frac{1}{\text{number of terms}} \sum_n X(n)X(n-m)$$

looks like convolution:

$$X \circ Y(m) \equiv \sum_{n=-\infty}^{\infty} X(n)Y(m-n)$$

FOURIER says
Fourier Transform of convolution
=
(FT of first) x (FT of second)

USE Fast FT to get $\mathrm{FT}\{X(n)\}$.

$Y = X_{\text{time-reversed}} = X(-n)$

FOURIER says
Fourier transform of $X(-n)$ is just the complex conjugate of $\mathrm{F}\{X(n)\}$.

$$\hat{R}_X(m) = \frac{1}{\text{number of terms}} \sum_n X(n)X(n-m)$$

looks like convolution:

$$X \circ Y(m) \equiv \sum_{n=-\infty}^{\infty} X(n)Y(m-n)$$

FOURIER says
Fourier Transform of convolution
=
(FT of first) x (FT of second)

USE Fast FT to get FT$\{X(n)\}$.

$Y = X_{\text{time-reversed}} = X(-n)$

FOURIER says
Fourier transform of $X(-n)$ is just the complex conjugate of F$\{X(n)\}$.

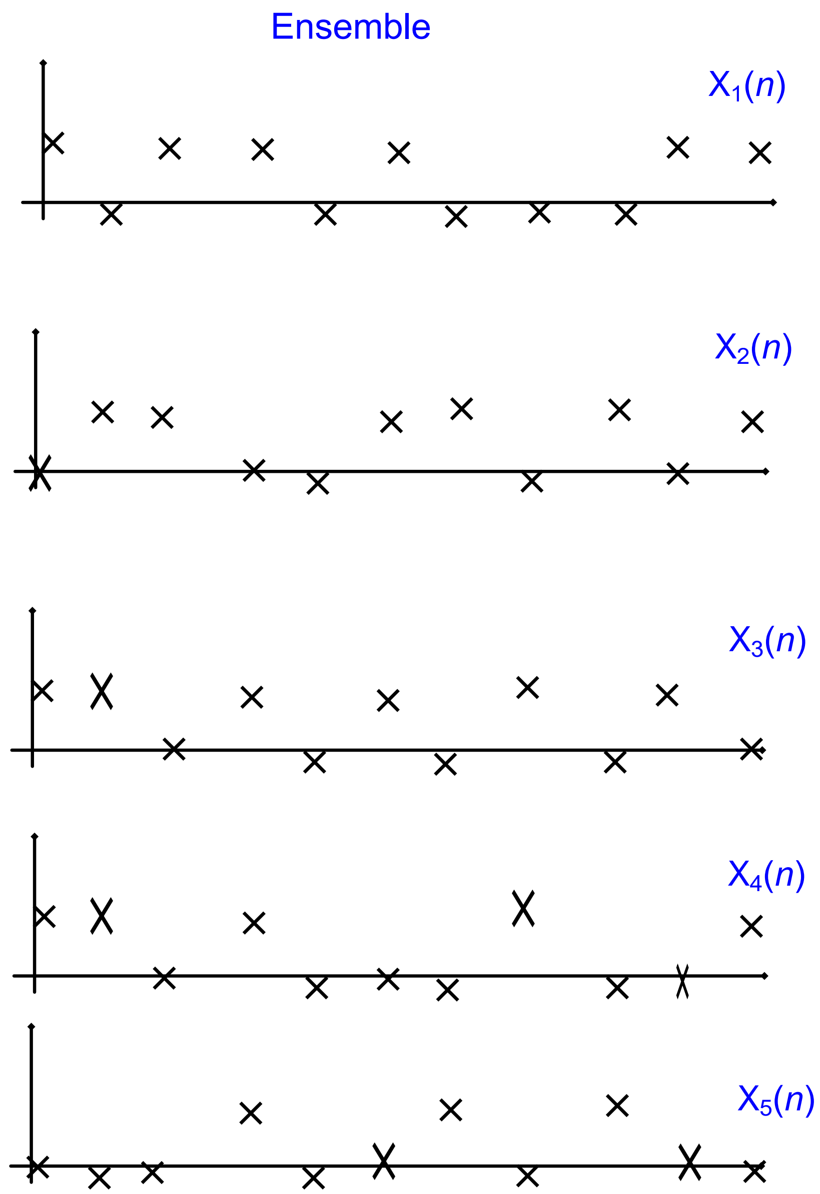
To estimate $R_X(m)$:
1. Take FFT of $\{X(n)\}$;
2. Multiply by complex conjugate;
3. Take inverse FFT;
4. Divide by number of terms.

$\Rightarrow \quad R_X(0), R_X(1), R_X(2), \ldots$

What can go wrong?

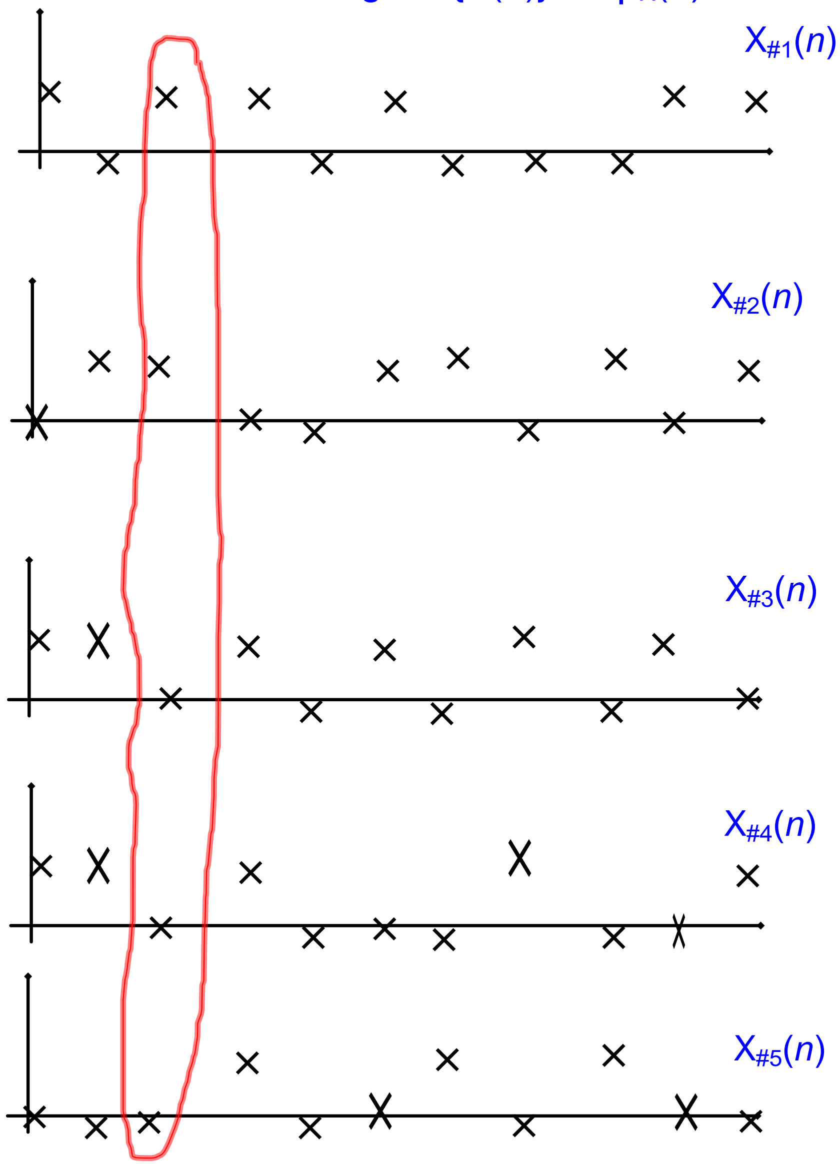
Lecture 9
(Feb. 15, 2017)

Bernoulli trials: Binomial Process



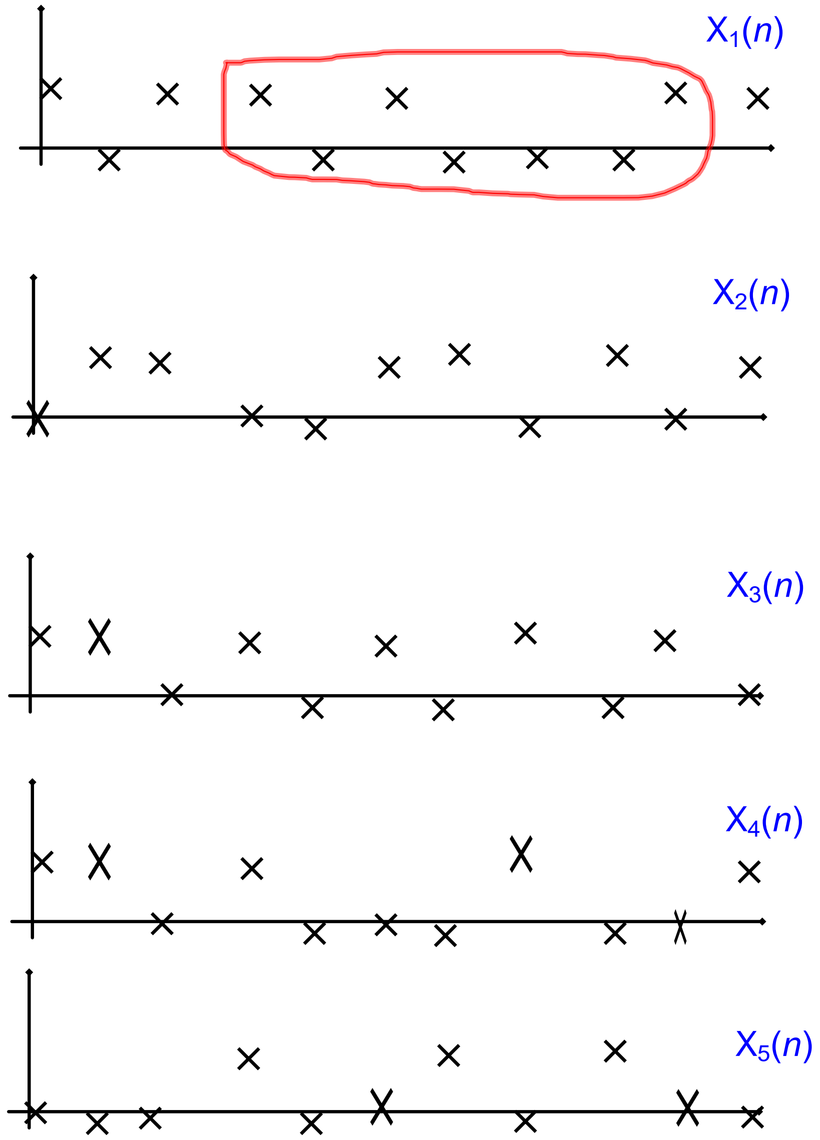
Bernoulli trials: Binomial Process

Ensemble Average: $E\{X(2)\} = \mu_X(2)$



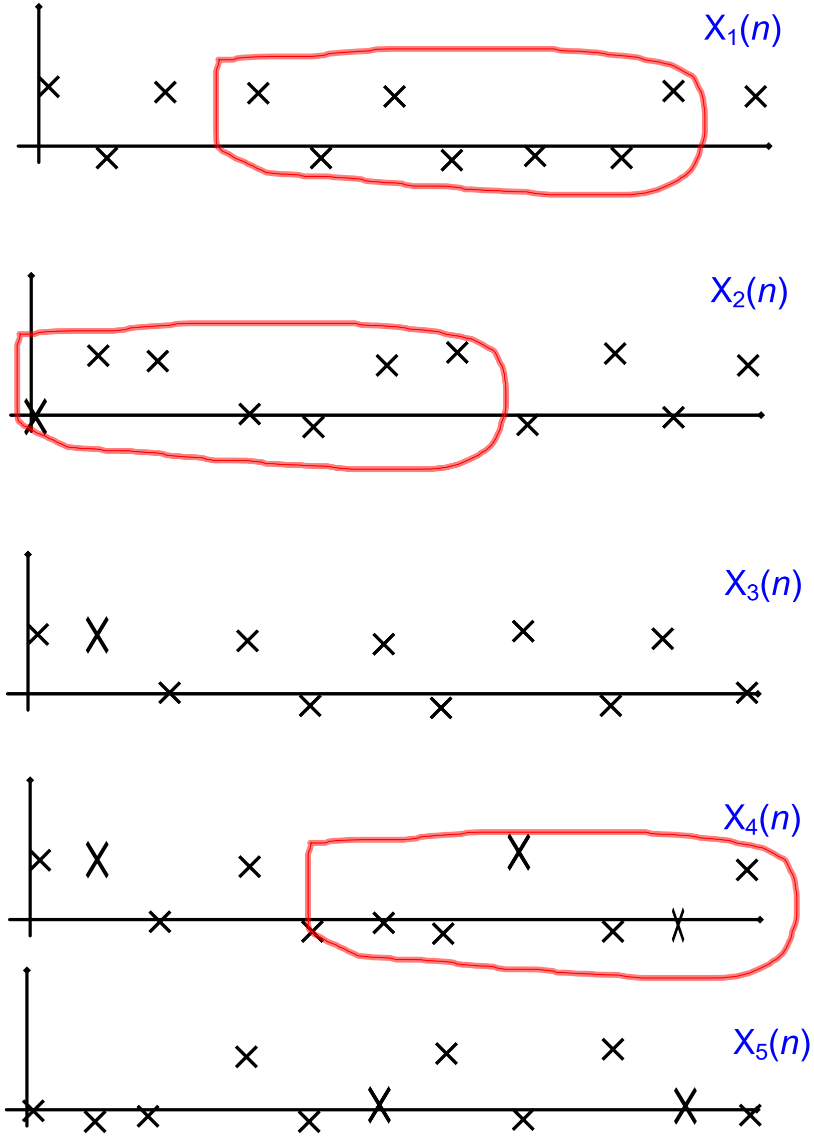
Bernoulli trials: Binomial Process

Time Average



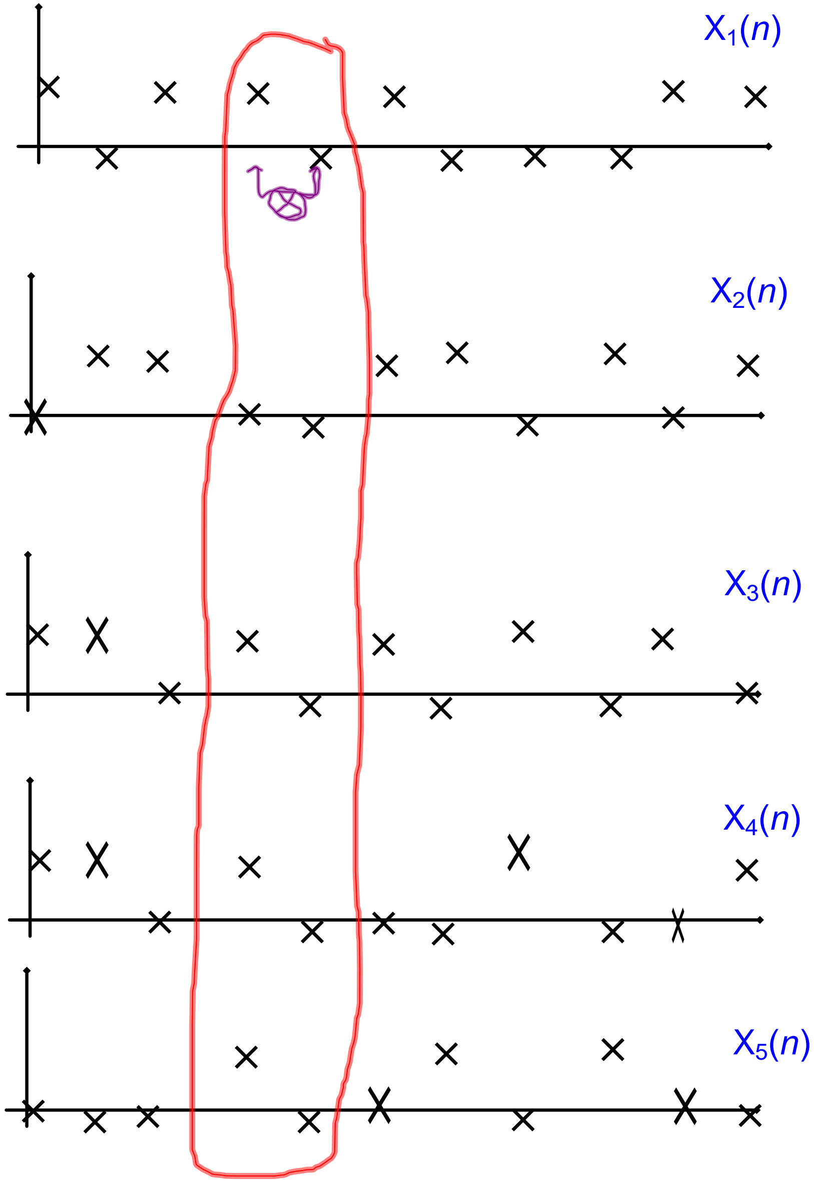
Bernoulli trials: Binomial Process

Time Average <X>



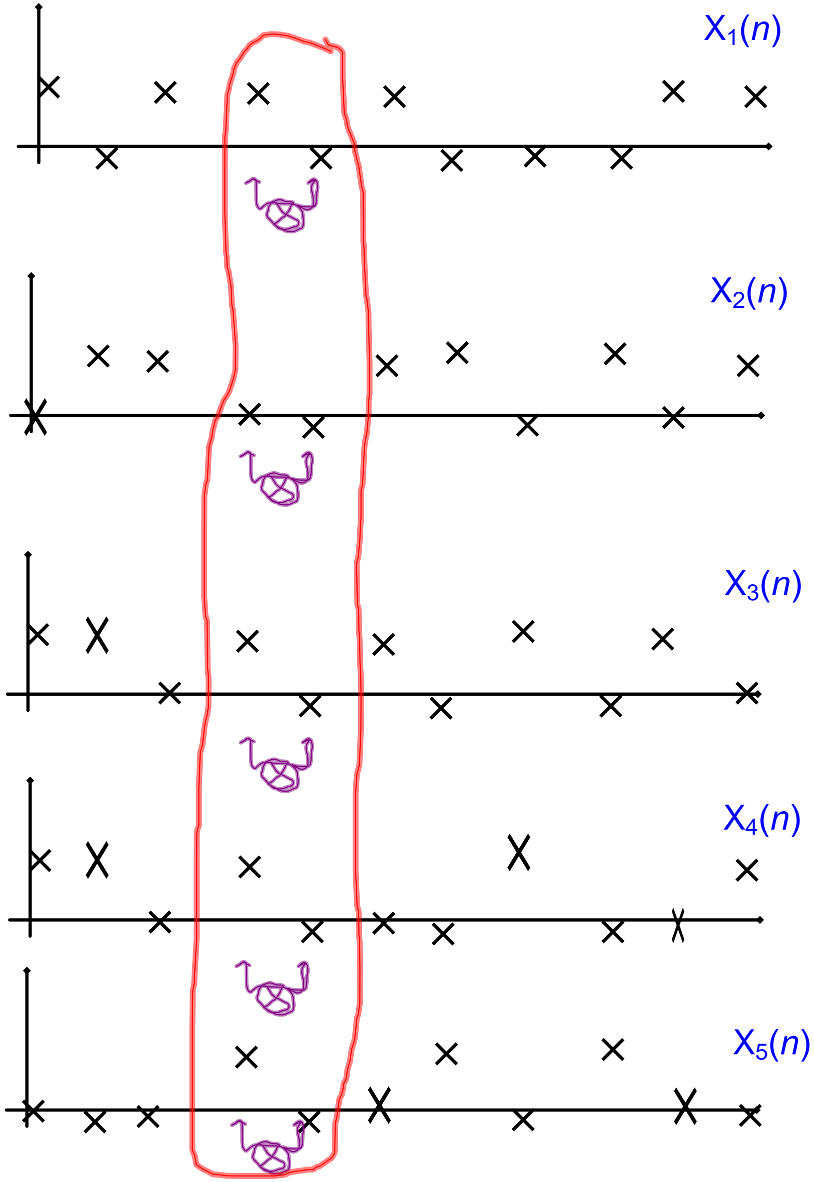
Bernoulli trials: Binomial Process

Ensemble Average $R_X(3,4)$



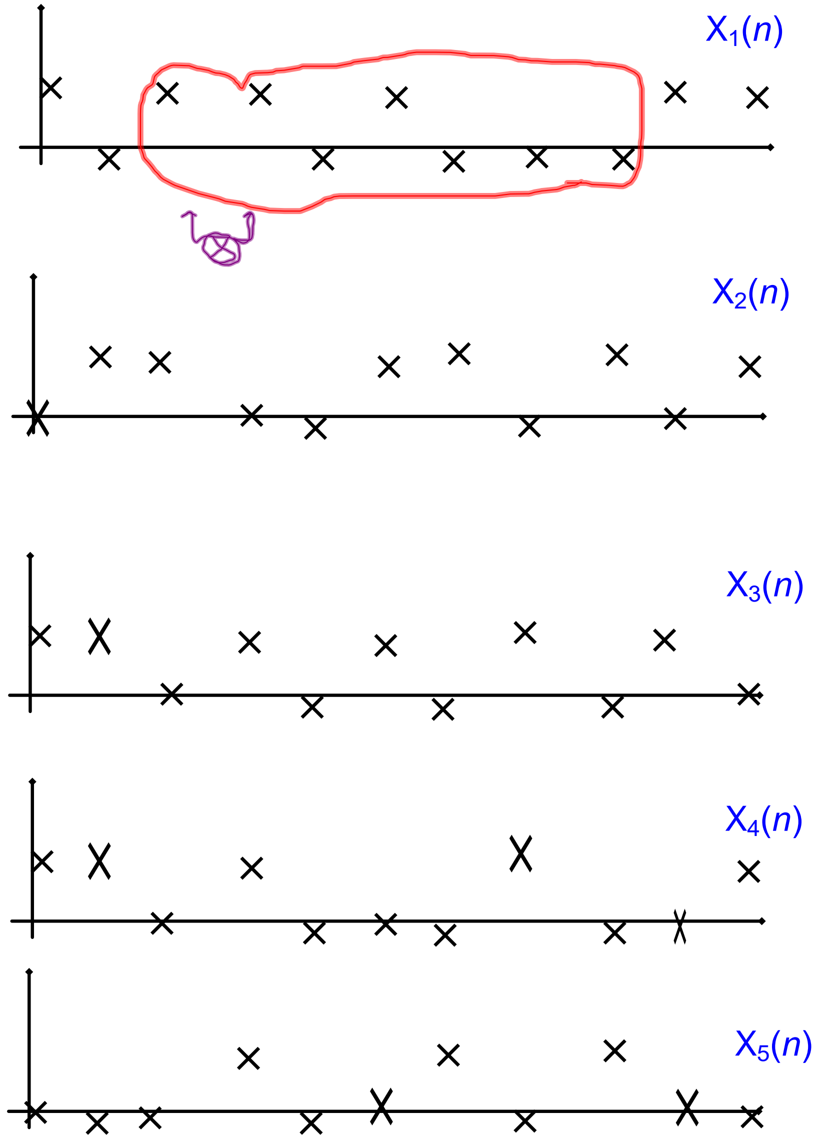
Bernoulli trials: Binomial Process

Ensemble Average $R_X(3,4)$



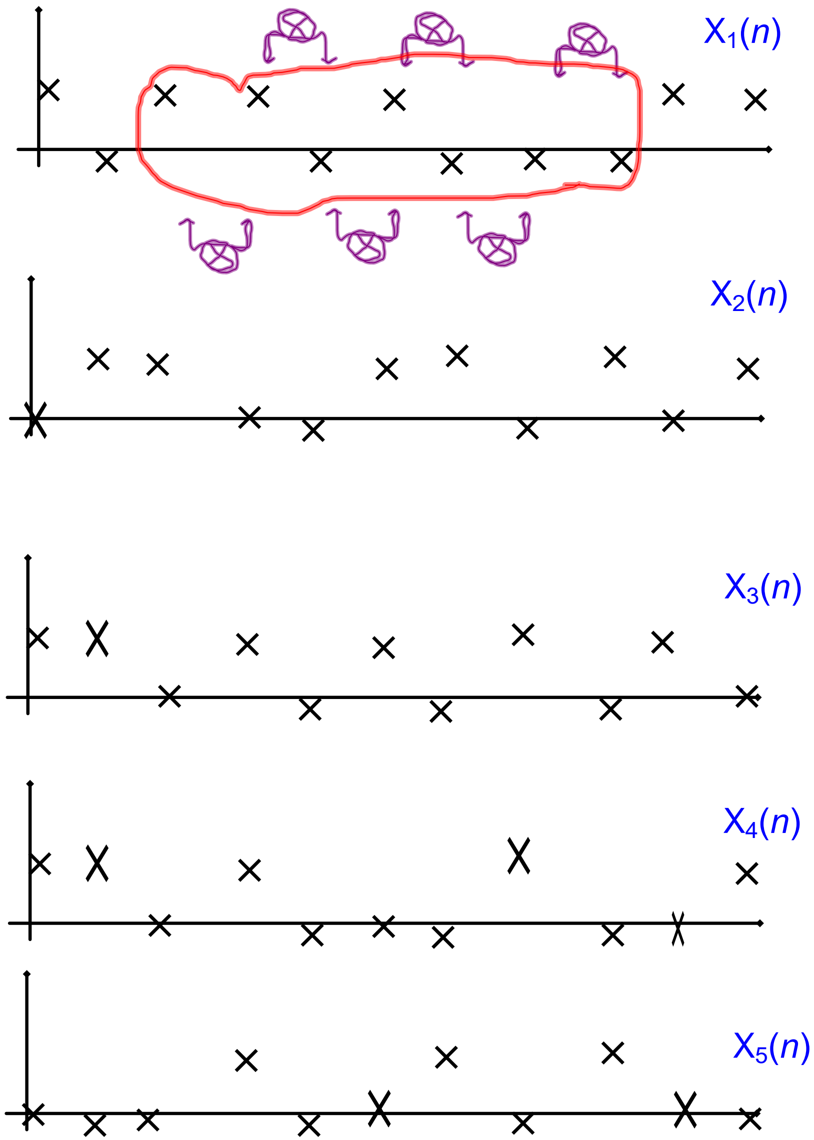
Bernoulli trials: Binomial Process

Time average $R_X(3,4) = R_X(1)$

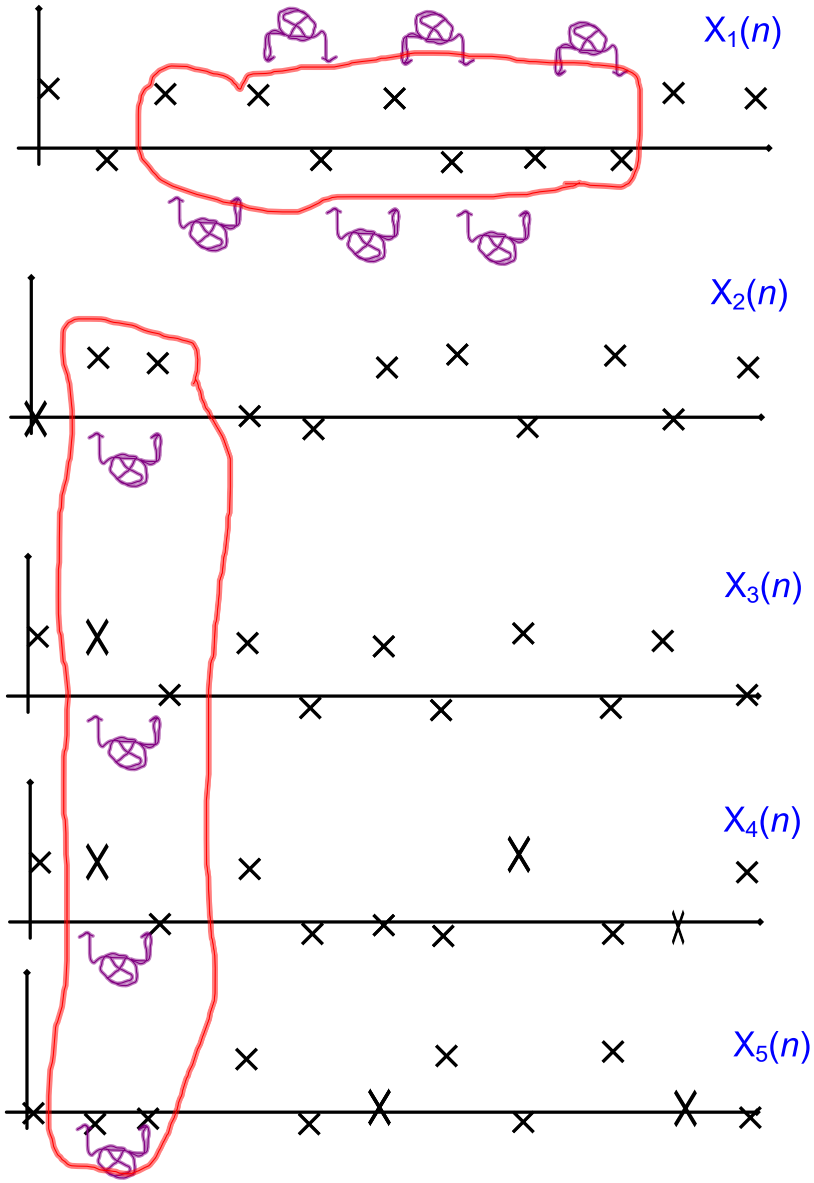


Bernoulli trials: Binomial Process

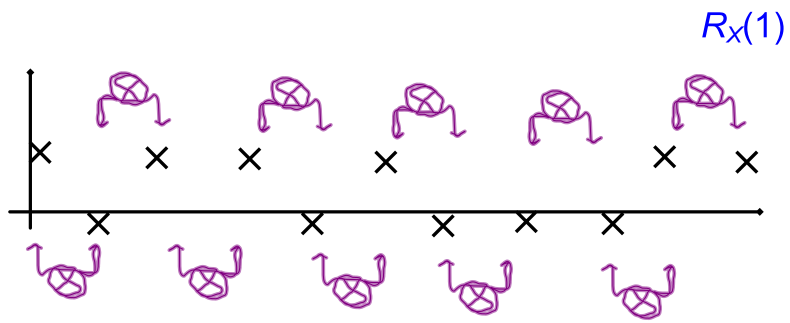
Time average $R_X(3,4) = R_X(1)$

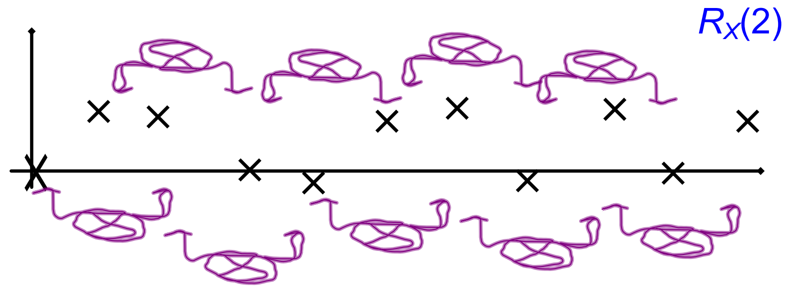


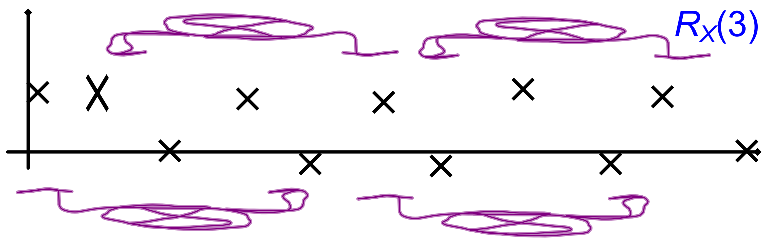
If a process is *ergodic* then ensemble averages ("means") equal time averages

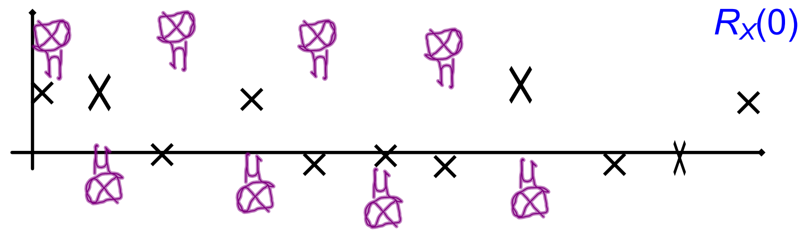


Autocorrelations

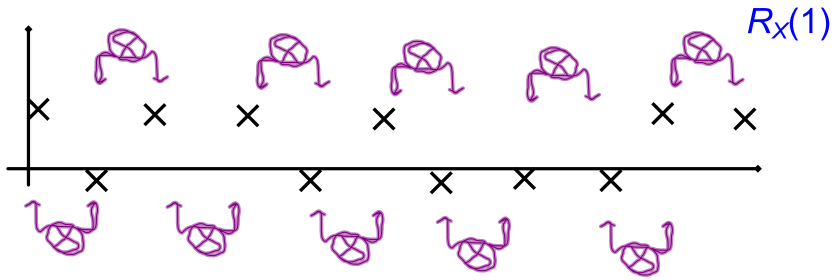


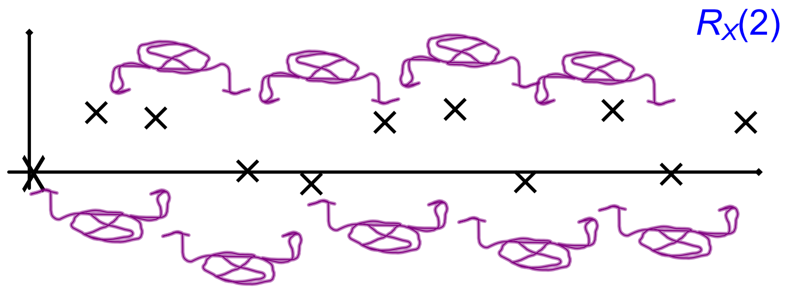


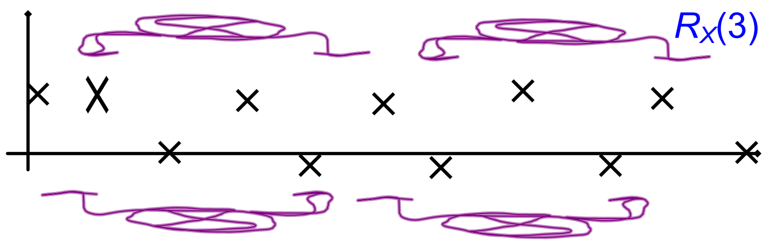


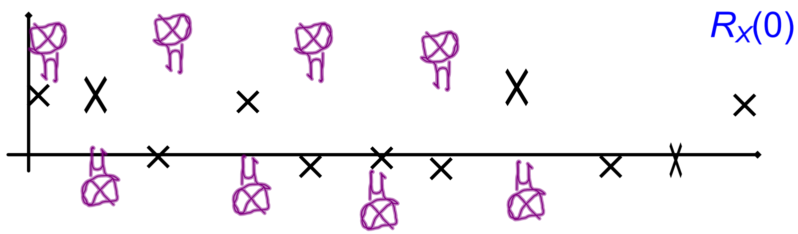


$$\hat{R}_X(m) = \frac{1}{\text{number of terms}} \sum_n X(n)X(n-m)$$

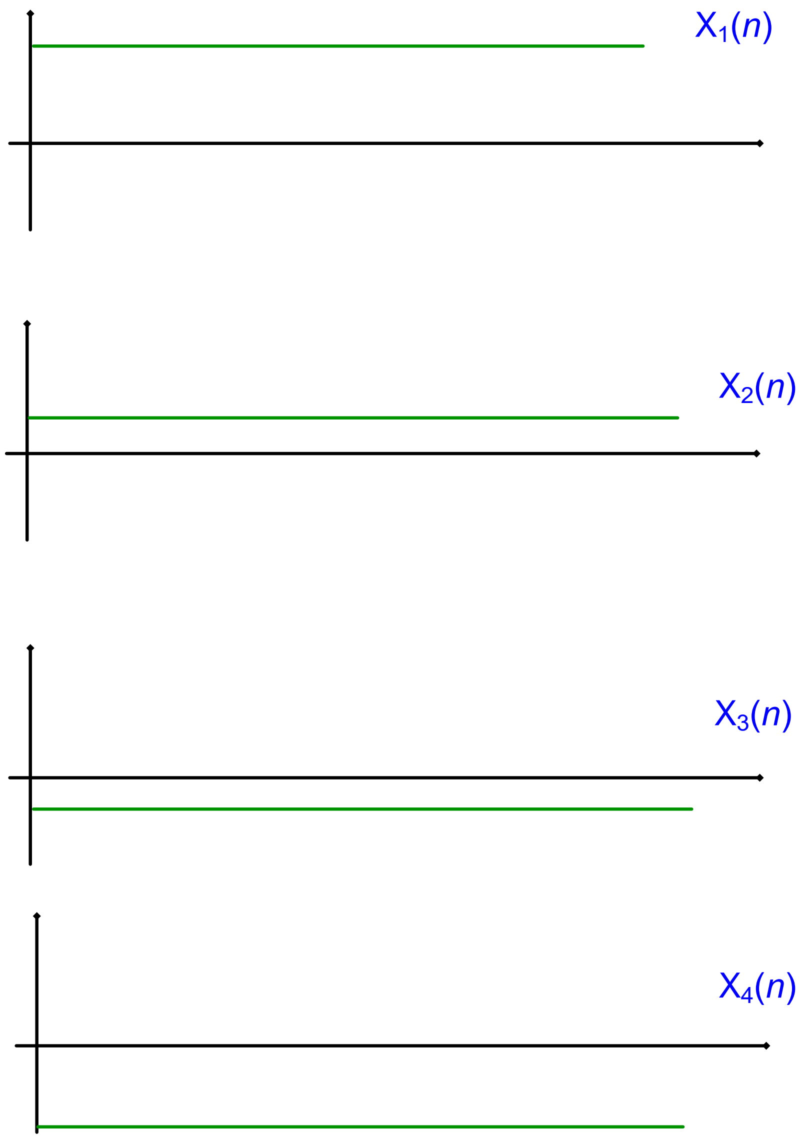




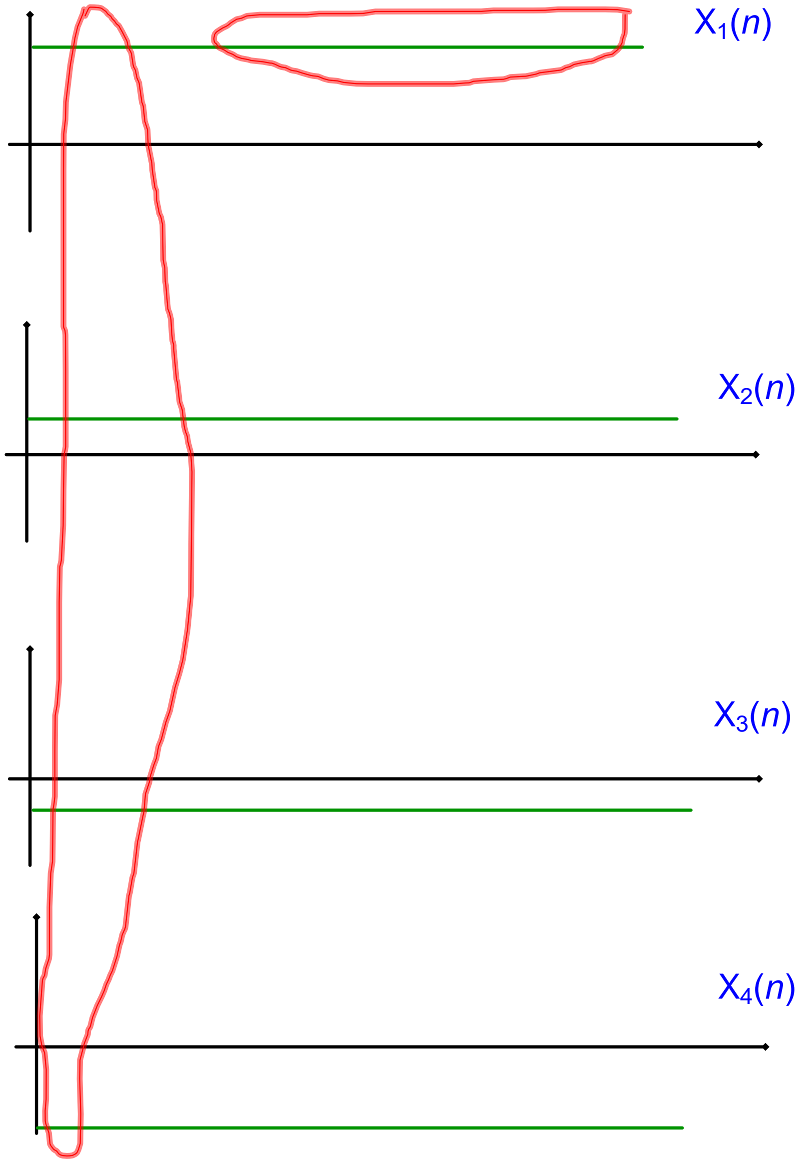




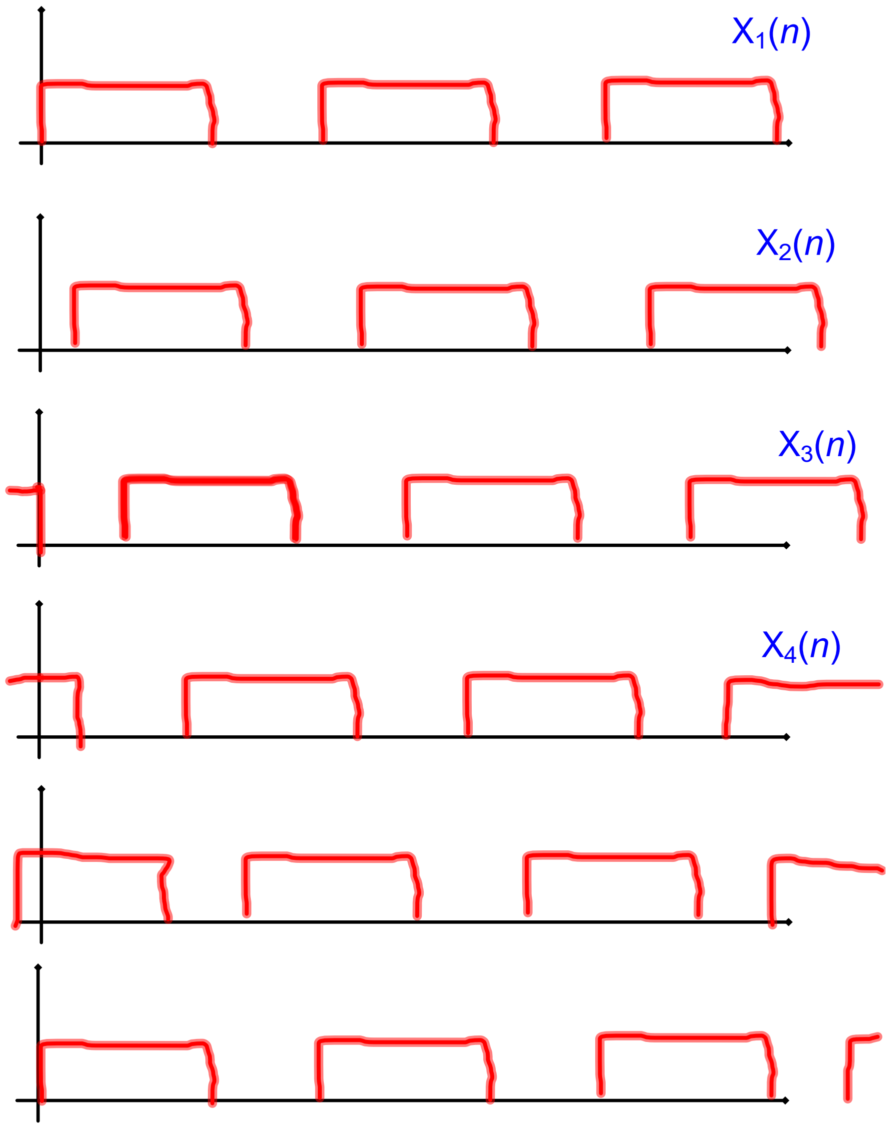
The randomly-set DC voltage source is NOT ergodic



The randomly-set DC voltage source is NOT ergodic

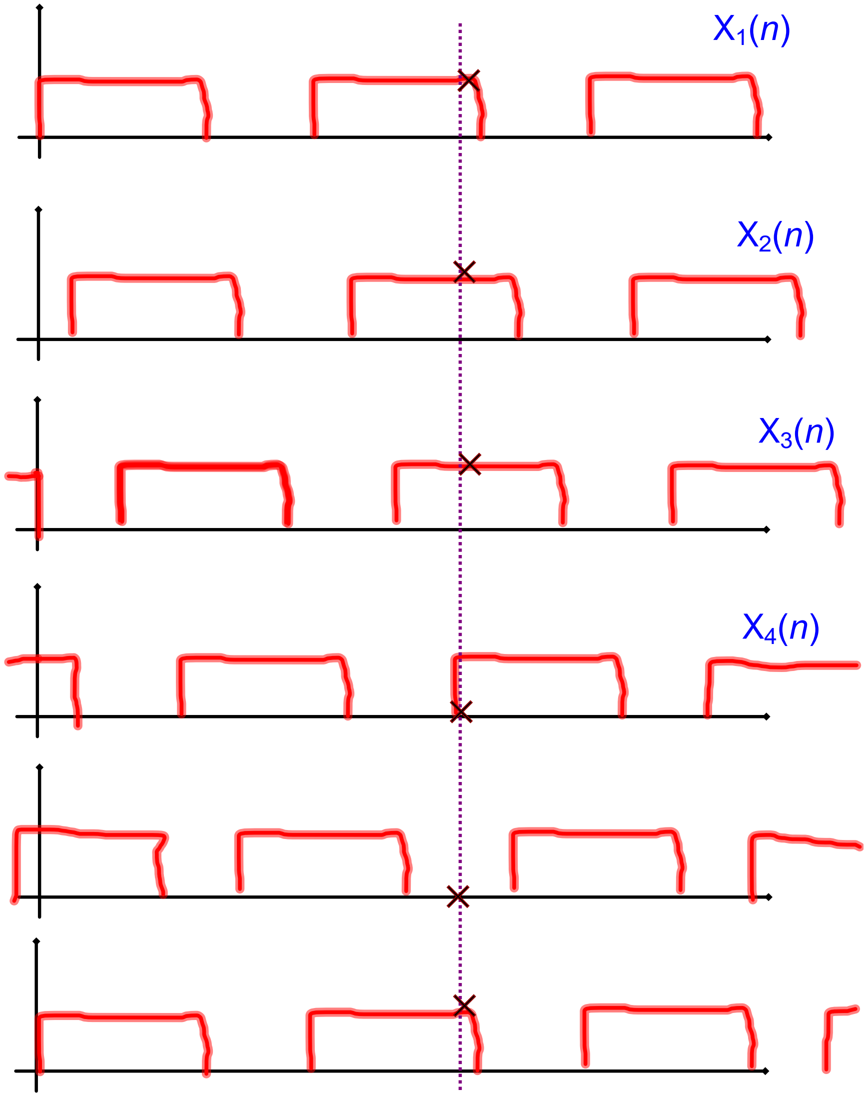


A random-shifted periodic function

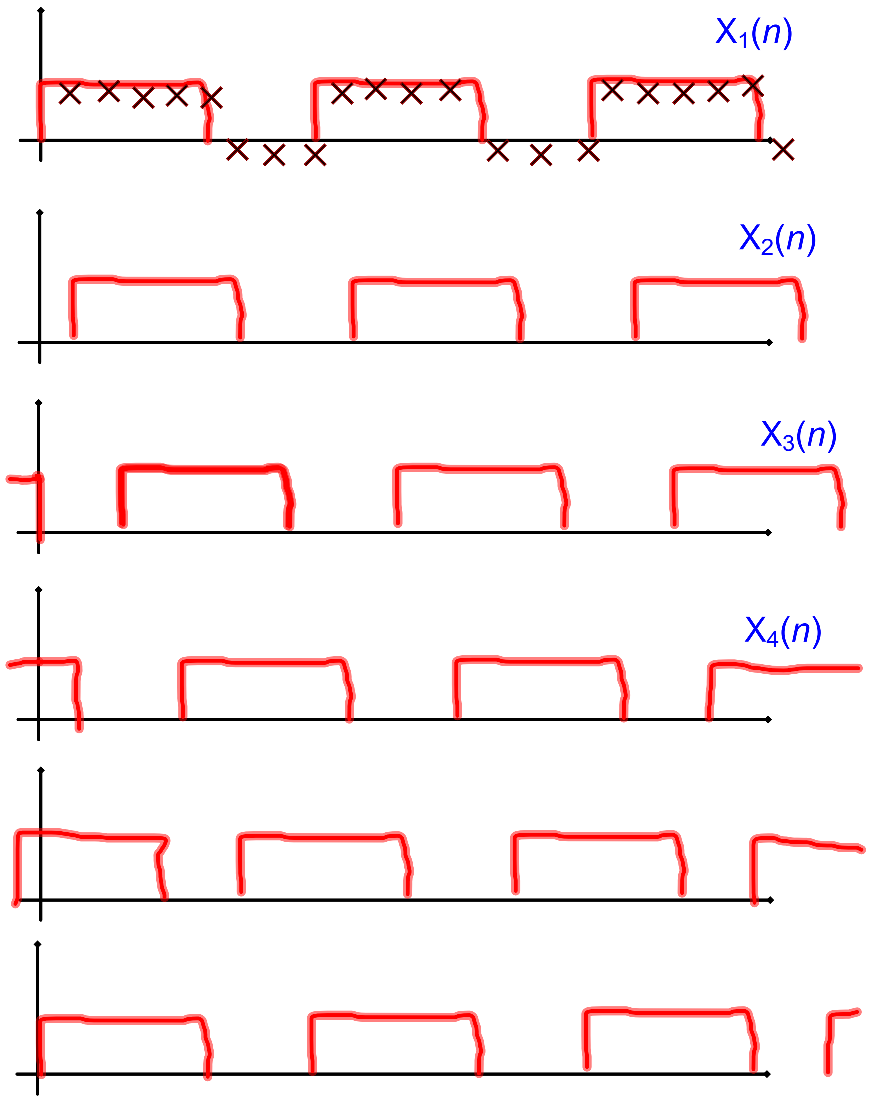


A random-shifted periodic function

Ensemble mean

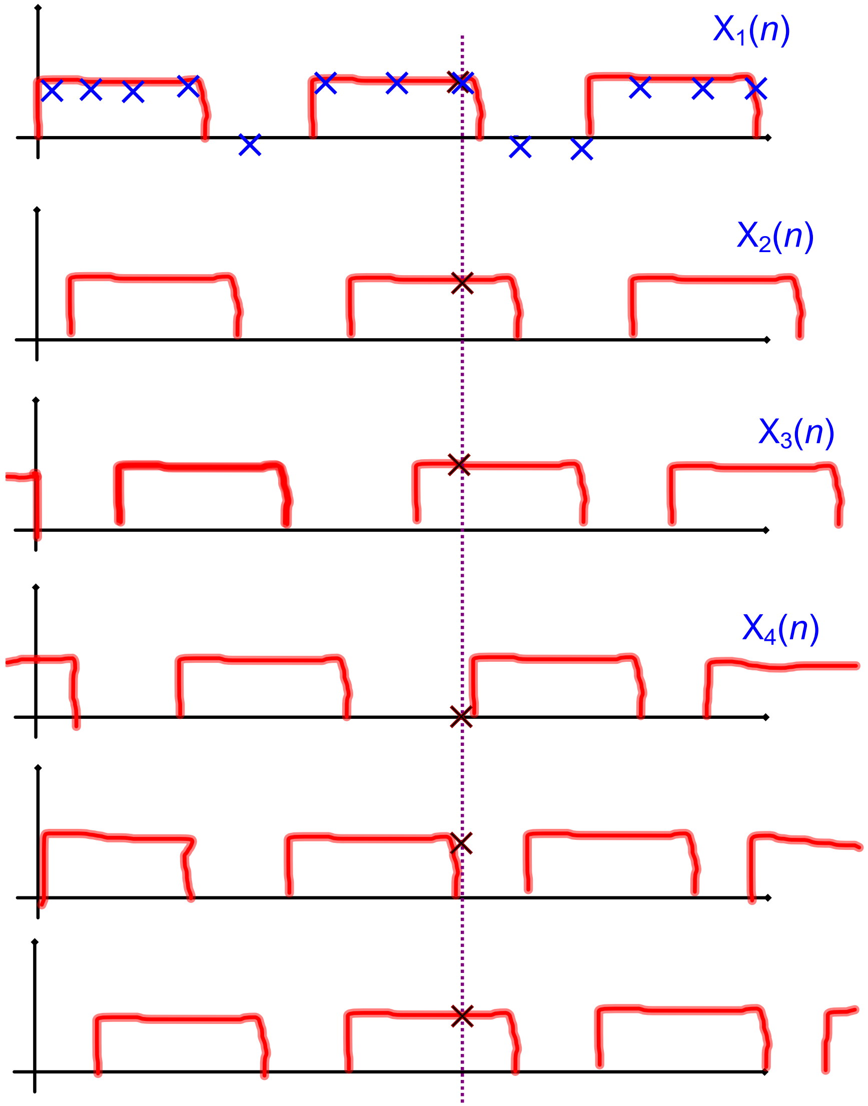


Time-averaged mean

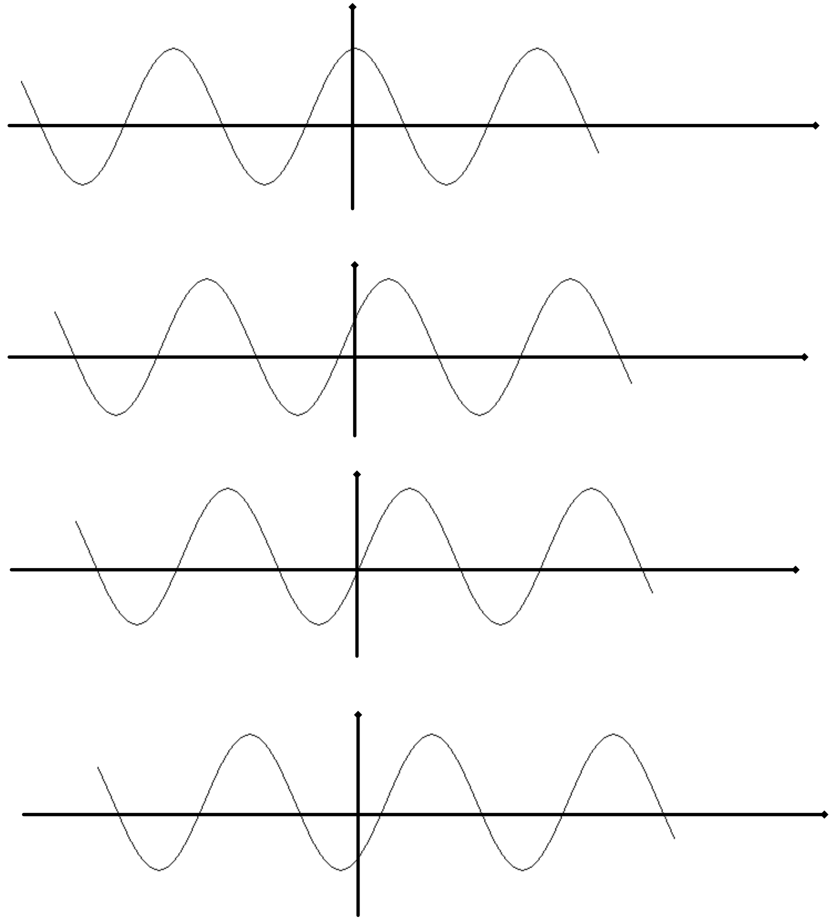


A random-shifted periodic function

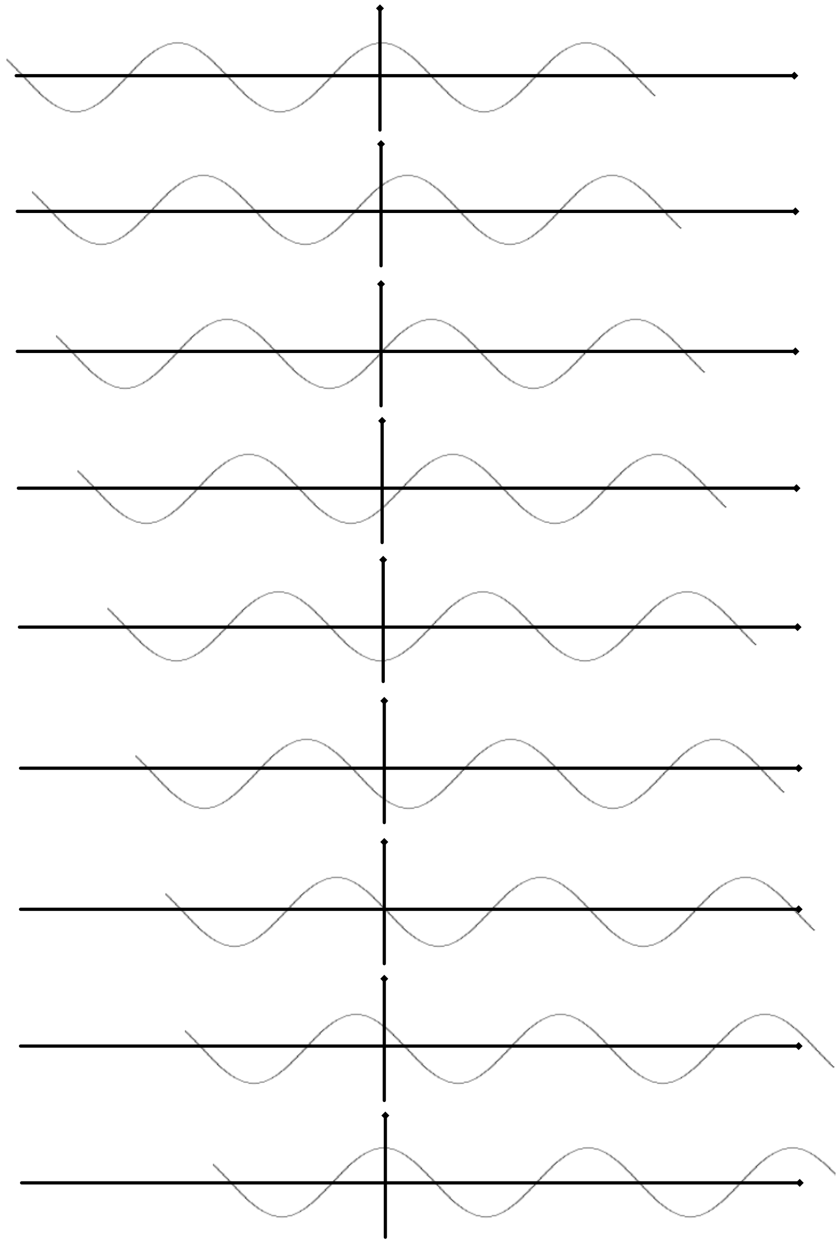
Ensemble mean = Time average



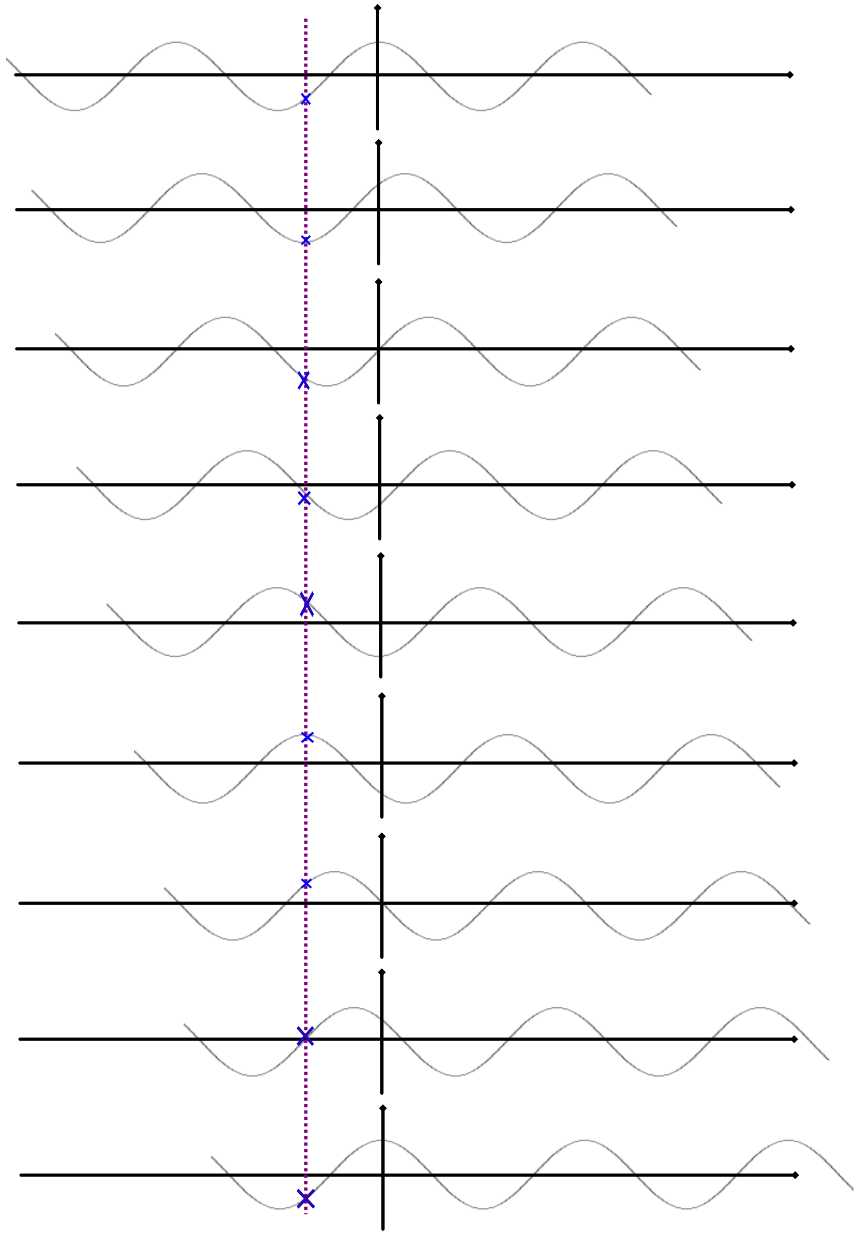
Random-Phase Sine Wave



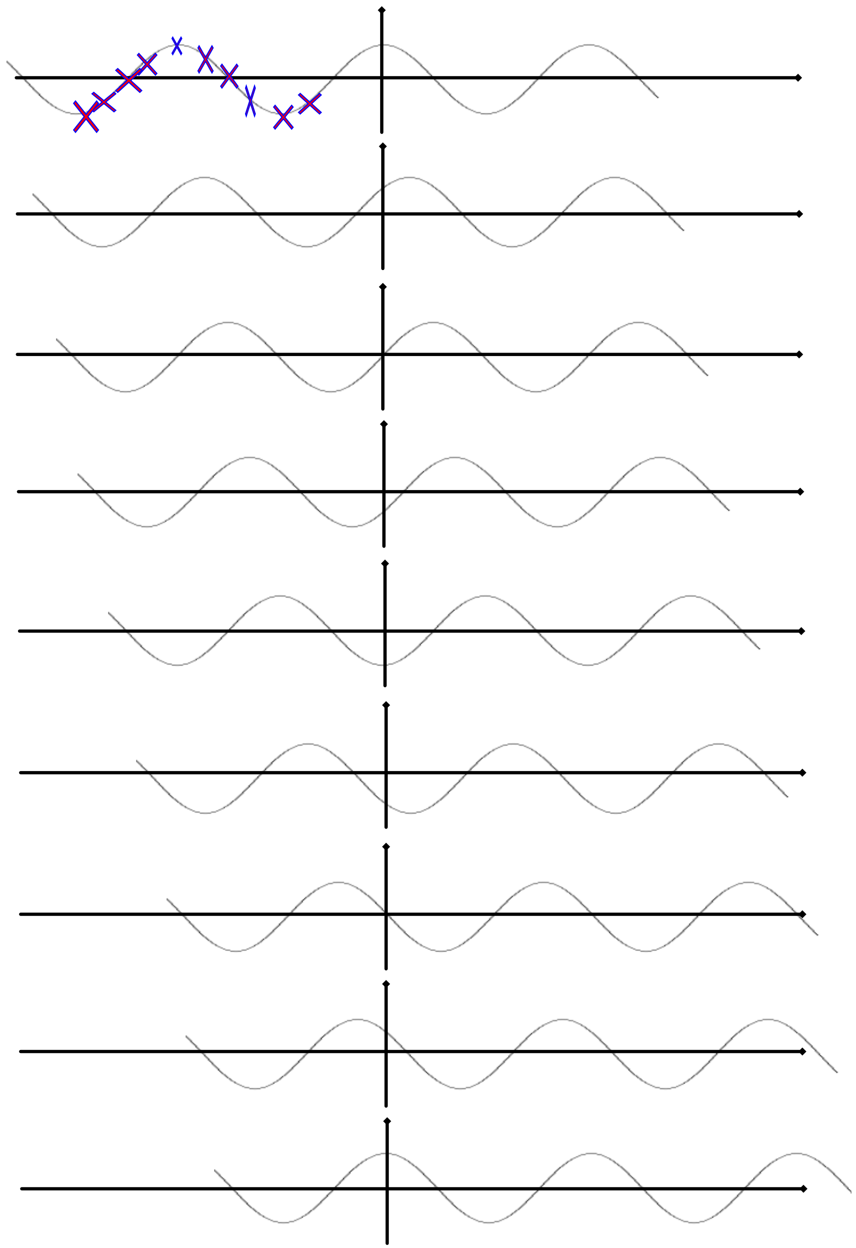
Random-Phase Sine Wave



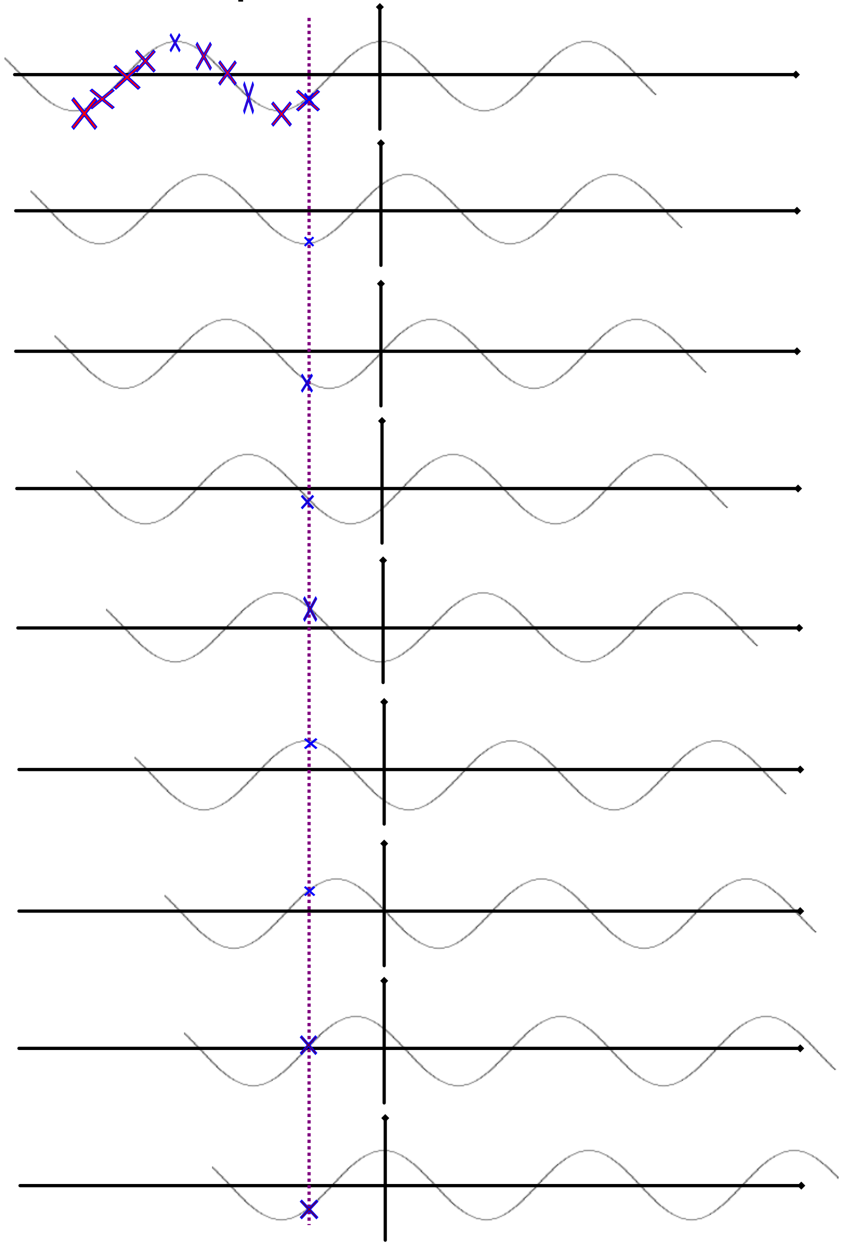
E{$X(t)$}, ensemble average



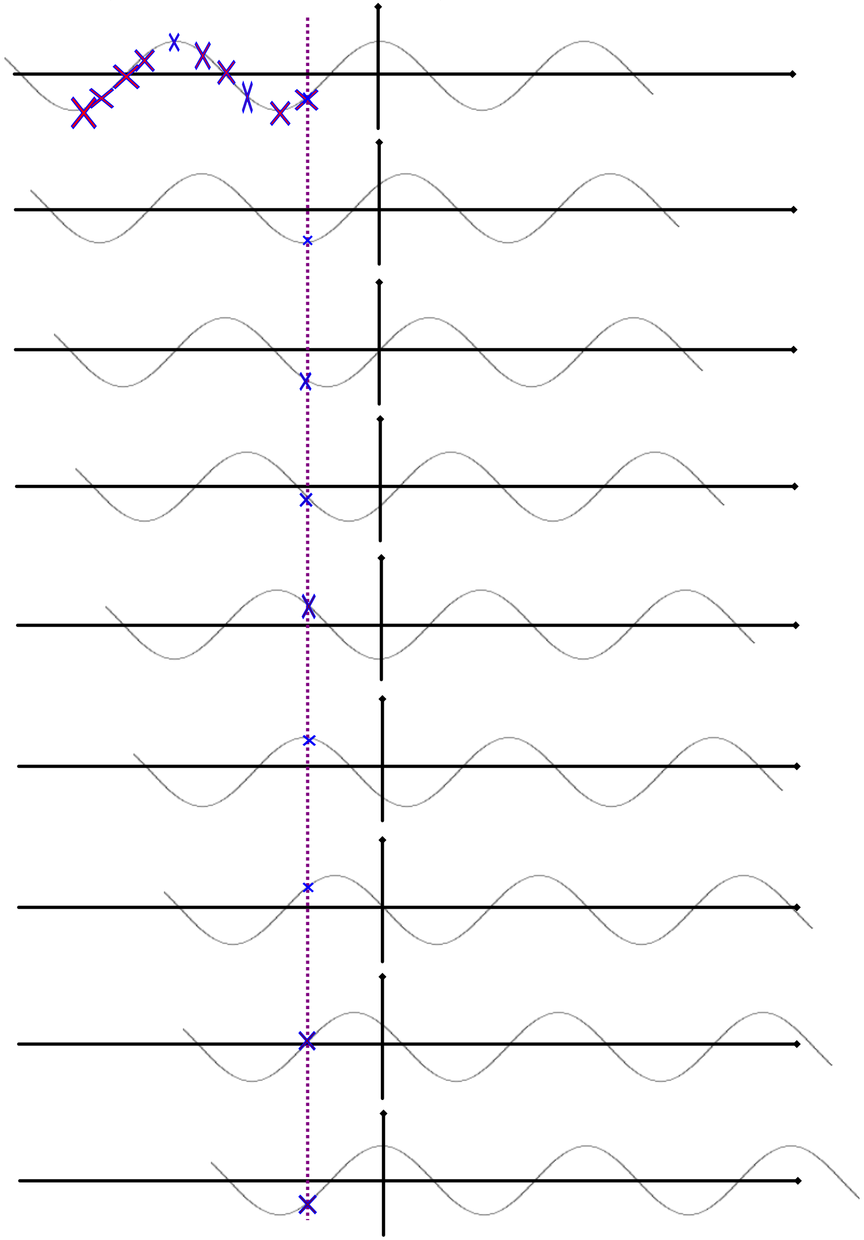
<*X*>, time average



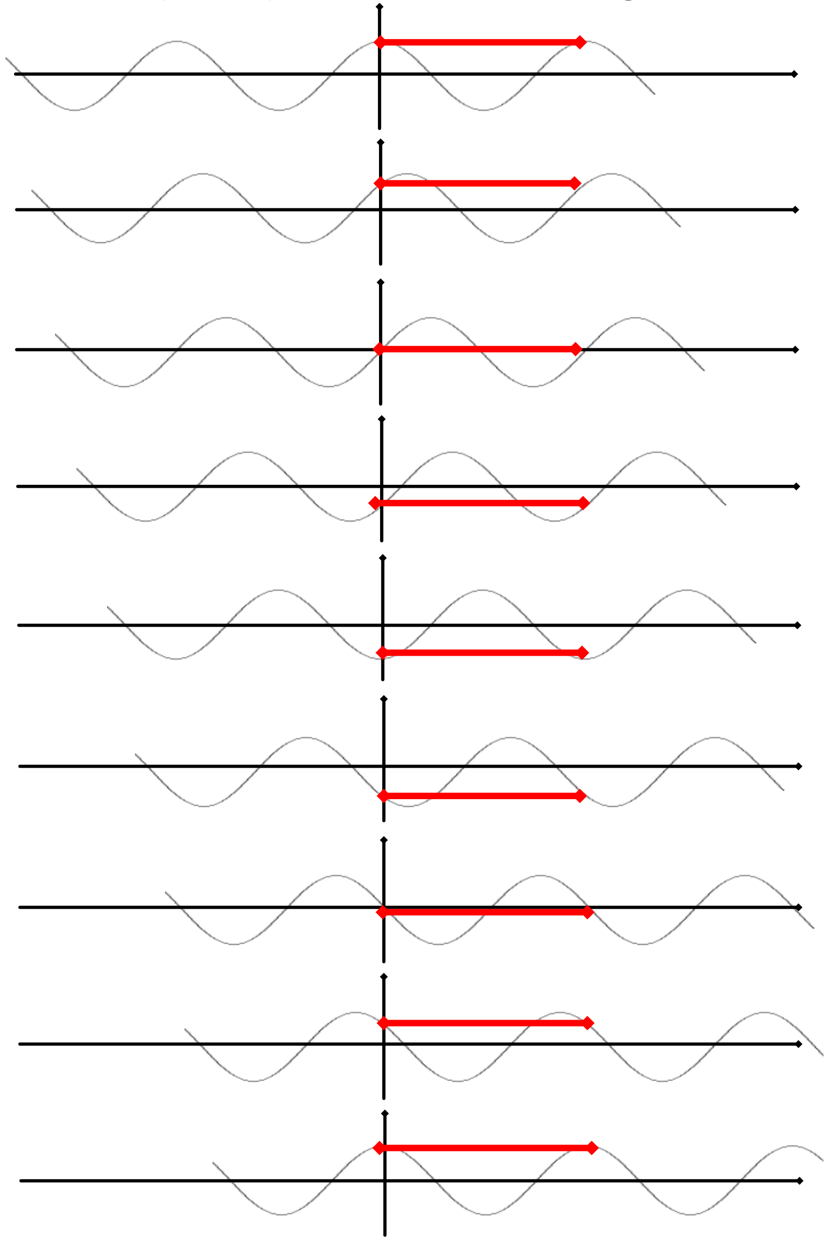
$<X> = \mu_X$



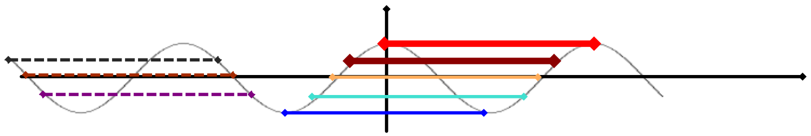
(and $<X^2> = \sigma_X{}^2$)

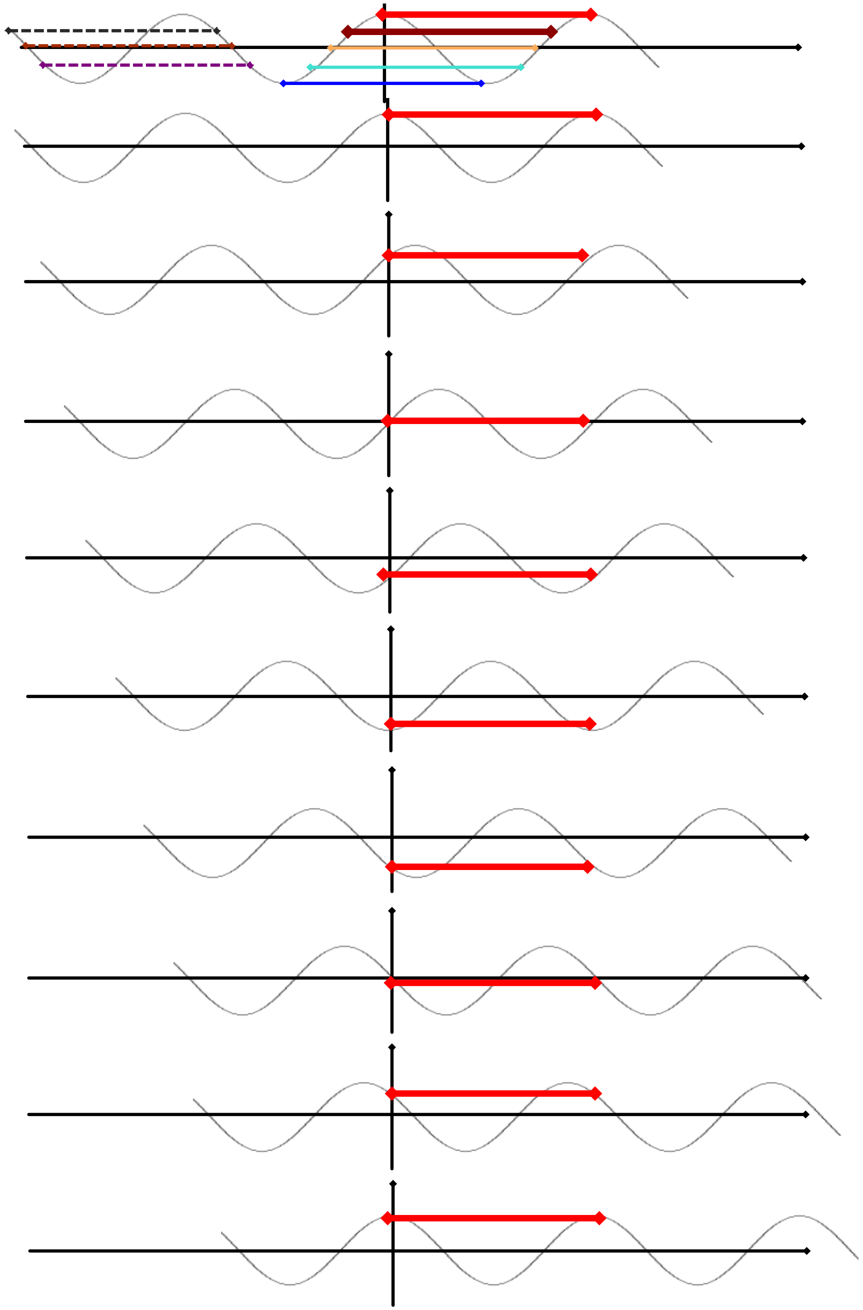


$R_X(0,2\pi)$, ensemble average

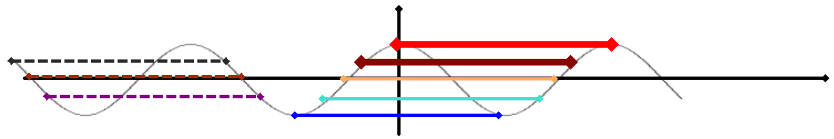


$R_X(2\pi)$, time average



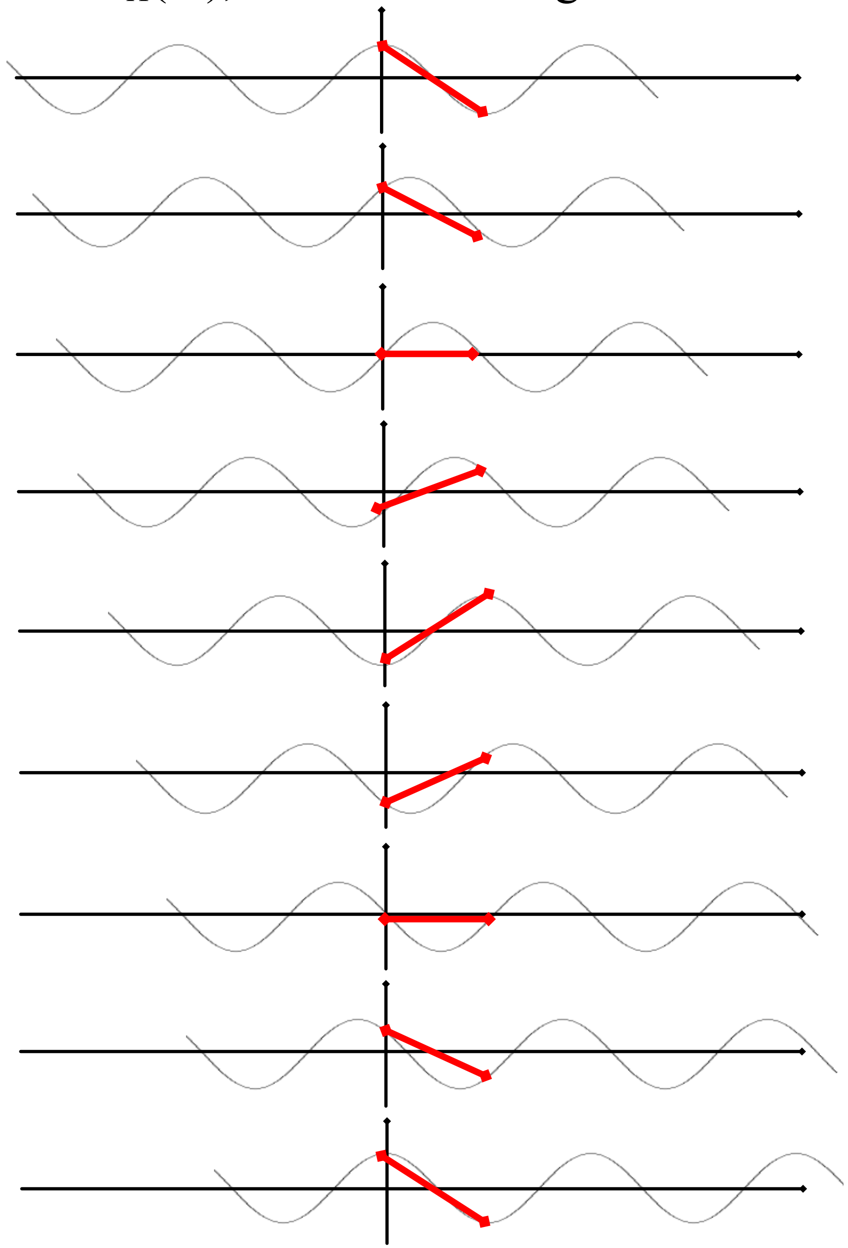


$R_X(2\pi)$, time average

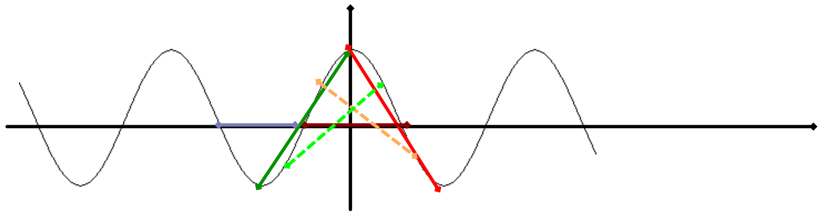


$R_X(2\pi)$ is large positive
[same as $R_X(0)$]

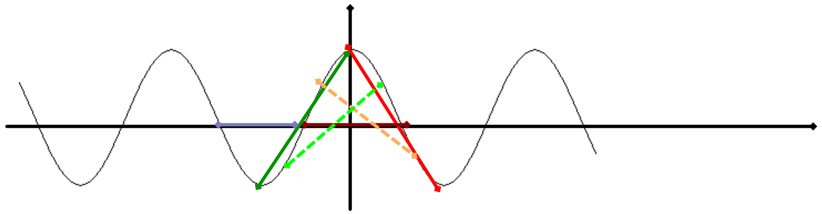
$R_X(\pi)$, ensemble average



$R_X(\tau)$, time average

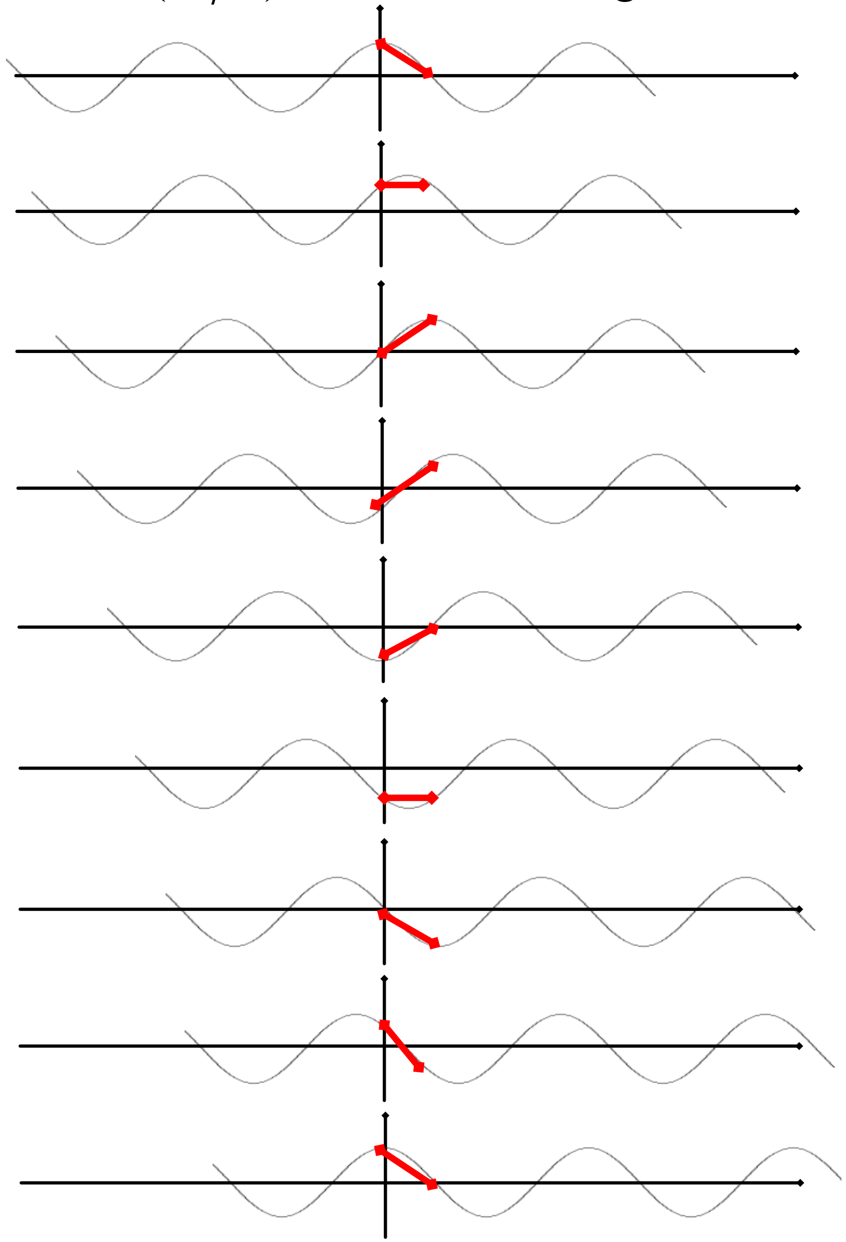


$R_X(\pi)$, time average

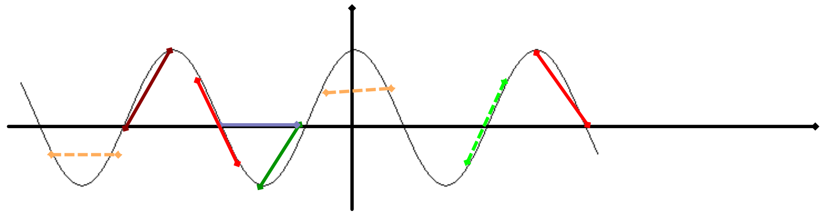


$R_X(\pi)$ is large negative

$R_X(\pi/2)$, ensemble average

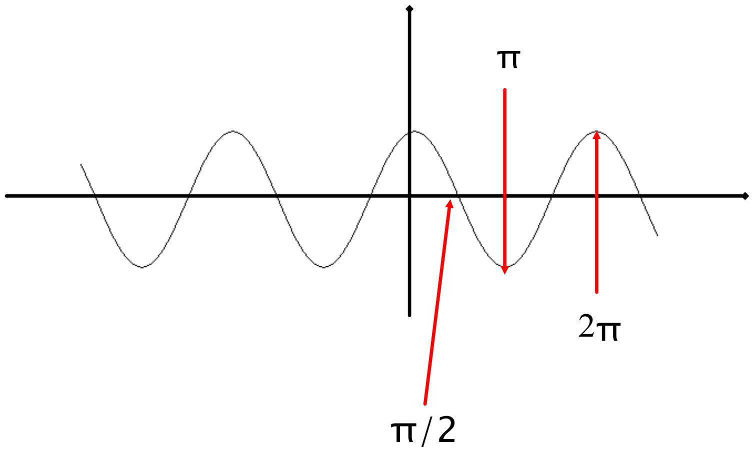


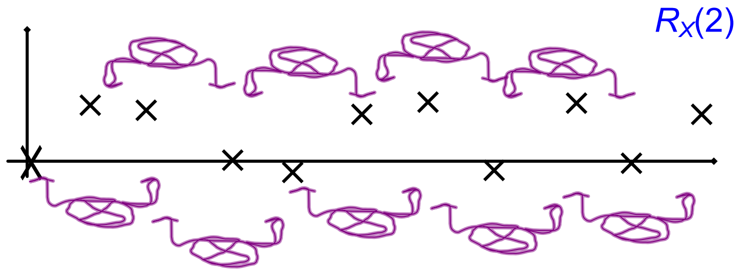
$R_X(\pi/2)$, time average



$$R_X(\pi/2) = 0$$

$R_X(t)$ is the (unshifted) cosine !



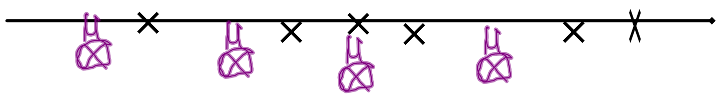


$$\hat{R}_X(m) = \frac{1}{\text{number of terms}} \sum_n X(n)X(n-m)$$

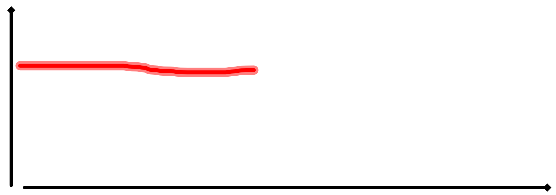
looks like convolution:

$$X \circ Y(m) \equiv \sum_{n=-\infty}^{\infty} X(n)Y(m-n)$$

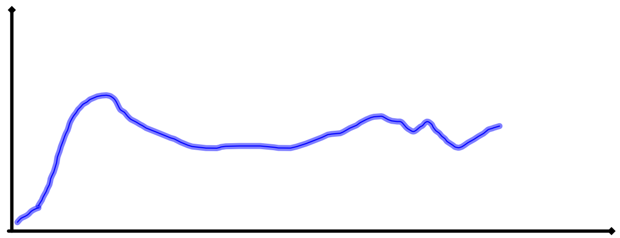
Why?



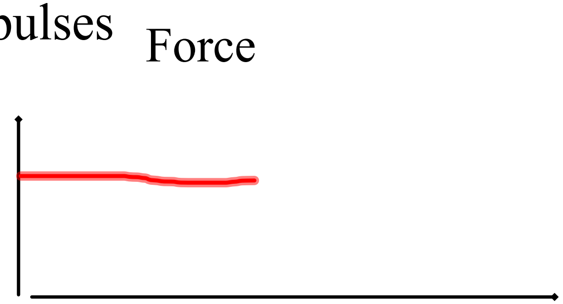
Force

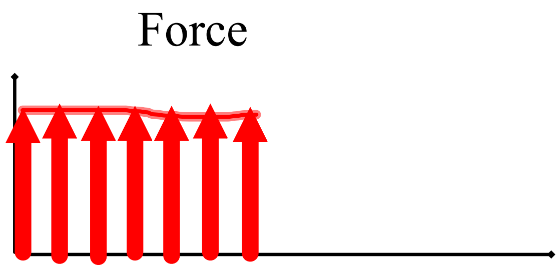


Response

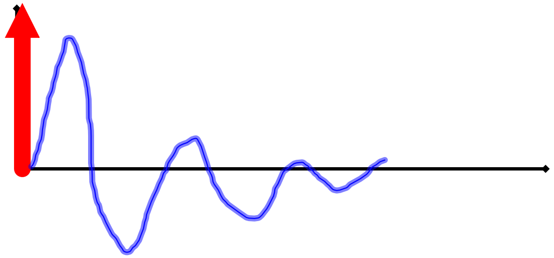


Model the force as a superposition of impulses

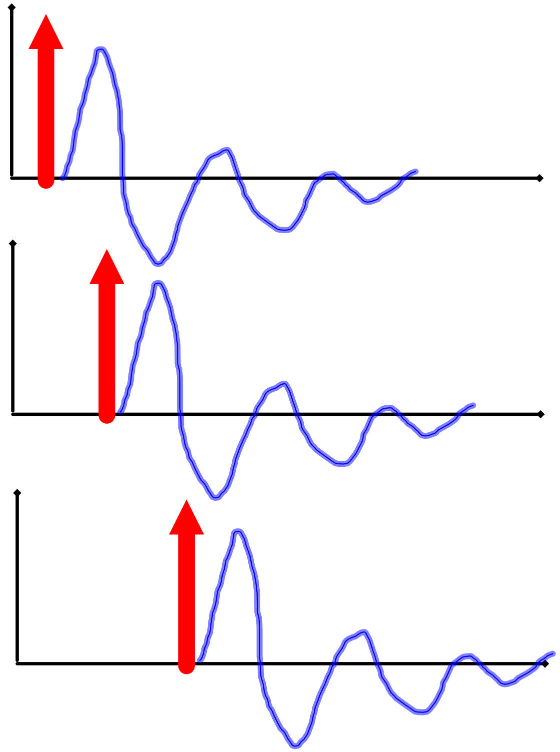




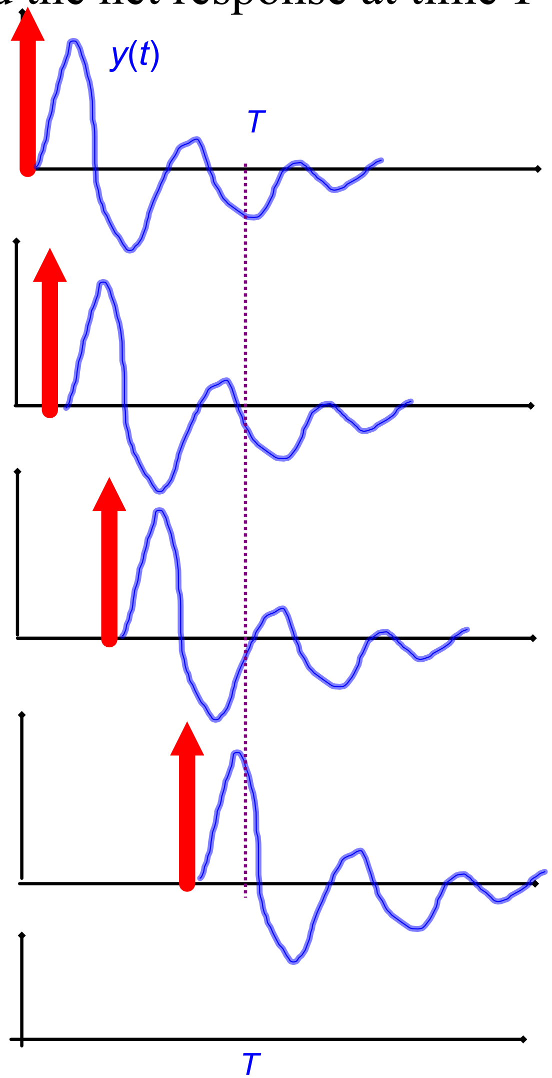
If the response to a single impulse looks like this:



then the response to each of the impulses looks like

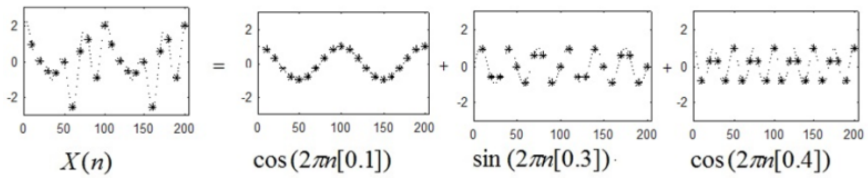


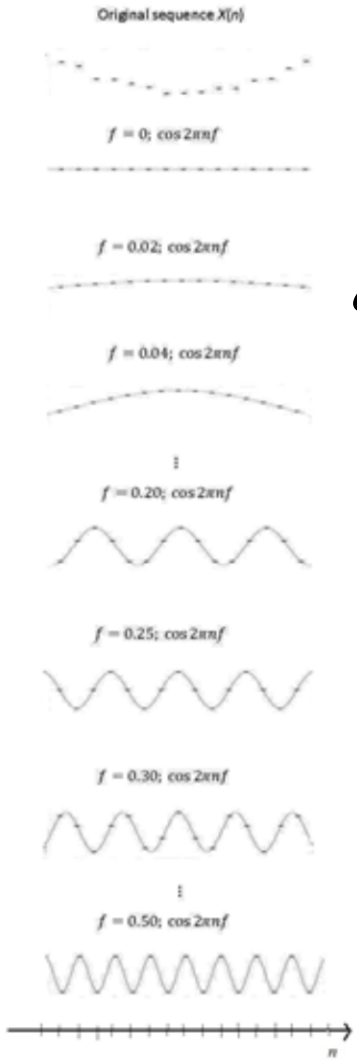
and the net response at time T equals



$F(0)y(T) + F(\tau)y(T-\tau) + F(2\tau)y(T-2\tau)$
$+ F(3\tau)y(T-3\tau) + \ldots$
(CONVOLUTION!)

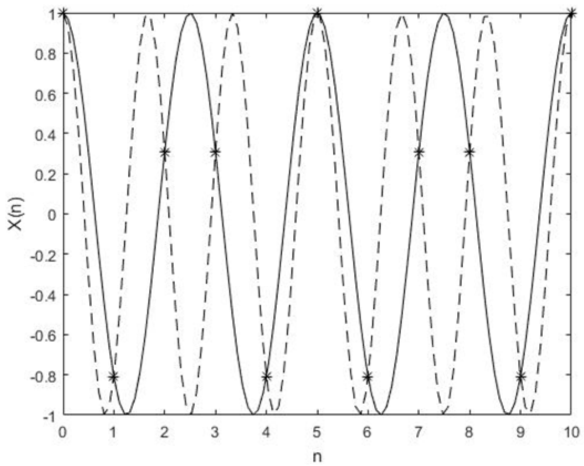
Original sequence $X(n)$


X(n)
=
cos(2πn[0.1])
+
sin(2πn[0.3])
+
cos(2πn[0.4])



$$X(n) = \int_{-1/2}^{1/2} \tilde{X}(f)\, e^{j2\pi nf}\, df$$

$$e^{j\theta} = \cos\theta + j\sin\theta$$



Aliasing

Don't need frequencies higher than f = 1/2

$$X(n) = \int_{-1/2}^{1/2} \tilde{X}(f) e^{j2\pi nf}\, df$$

$$X(n) = \int_{-1/2}^{1/2} \tilde{X}(f) e^{j2\pi nf}\, df$$

$$\tilde{X}(f) = \sum_{n=-\infty}^{\infty} X(n) e^{-j2\pi nf}, \; -\frac{1}{2} < f < \frac{1}{2}$$

$\tilde{X}(f)$ = Discrete Time Fourier Transform of $X(n)$

{$X(n)$ = *Inverse* DTFT of $\tilde{X}(f)$}

$$X(n) = \int_{-1/2}^{1/2} \tilde{X}(f) e^{j2\pi nf} \, df$$

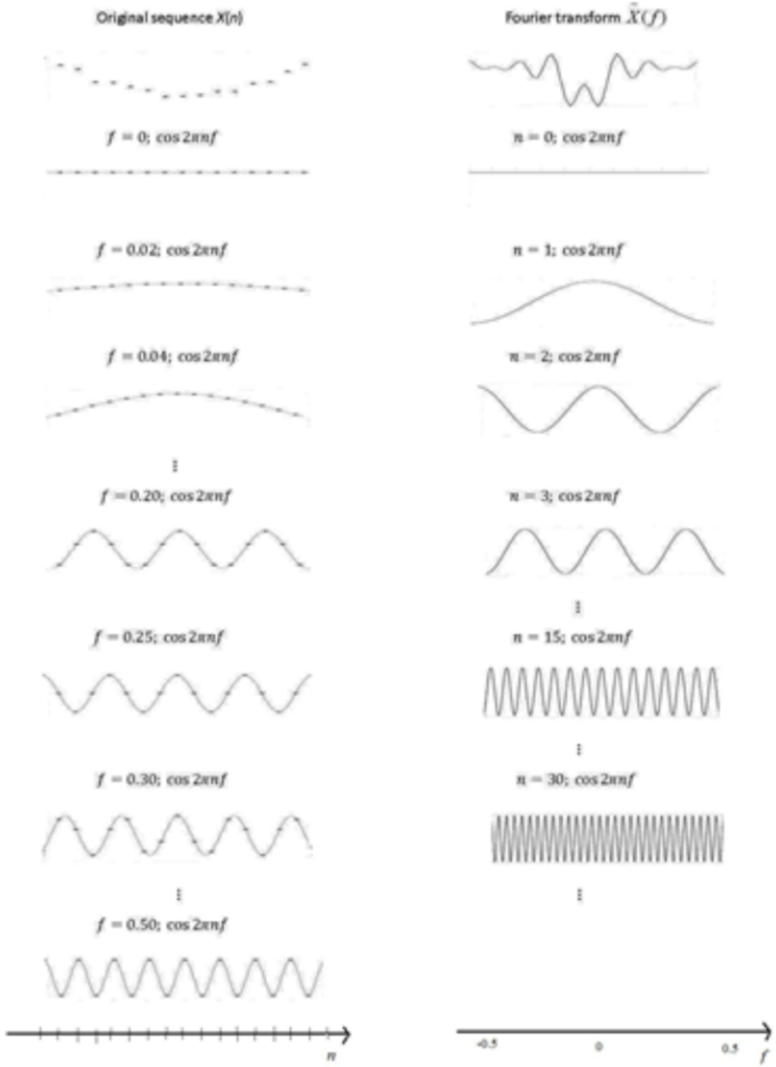
$$\tilde{X}(f) = \sum_{n=-\infty}^{\infty} X(n) e^{-j2\pi nf}, \quad -\frac{1}{2} < f < \frac{1}{2}$$

$\tilde{X}(f)$ = Discrete Time Fourier Transform of $X(n)$

{$X(n)$ = *Inverse* DTFT of $\tilde{X}(f)$}

Compute them both using the Fast Fourier Transform

$$X(n) = \int_{-1/2}^{1/2} \tilde{X}(f)\, e^{j2\pi nf}\, df \qquad \tilde{X}(f) = \sum_{n=-\infty}^{\infty} X(n)\, e^{-j2\pi nf}, \; -\frac{1}{2} < f < \frac{1}{2}$$



If $\tilde{X}(f) = \sum_{n=-\infty}^{\infty} X(n) e^{-j2\pi nf}$ and $\tilde{Y}(f) = \sum_{n=-\infty}^{\infty} Y(n) e^{-j2\pi nf}$ then

$$\tilde{X}(f)\tilde{Y}(f) = \sum_{m=-\infty}^{\infty} Z(m) e^{-i2\pi mf} \text{ where}$$

$$Z(m) = \sum_{n=-\infty}^{\infty} X(n) Y(m-n) \ .$$

In short, $\tilde{X}\tilde{Y} = \widetilde{X \circ Y}$.

If $\tilde{X}(f) = \sum_{n=-\infty}^{\infty} X(n) e^{-j2\pi nf}$ and $\tilde{Y}(f) = \sum_{n=-\infty}^{\infty} Y(n) e^{-j2\pi nf}$ then

$$\tilde{X}(f)\tilde{Y}(f) = \sum_{m=-\infty}^{\infty} Z(m) e^{-i2\pi mf} \text{ where}$$

$$Z(m) = \sum_{n=-\infty}^{\infty} X(n)Y(m-n) \; .$$

In short, $\tilde{X}\tilde{Y} = \widetilde{X \circ Y}$.

Fourier Convolution Theorem

$$\hat{R}_X(m) = \frac{1}{\text{number of terms}} \sum_n X(n)X(n-m)$$

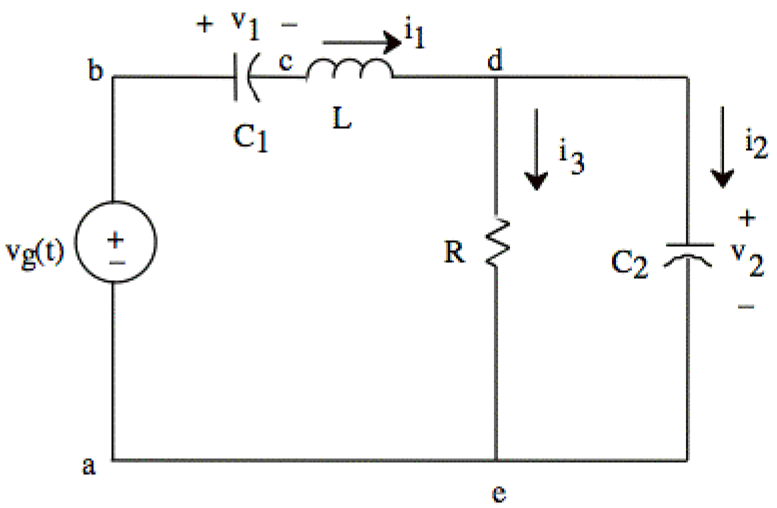
looks like convolution:

$$X \circ Y(m) \equiv \sum_{n=-\infty}^{\infty} X(n)Y(m-n)$$

Students: This is an elaboration of of the next lecture - lecture notes in the usual format will follow at the end.

Fourier Analysis in One Hour

For an engineer, a linear time-invariant system (like an RLC network) is characterized by the condition that no one is twiddling with the dials; the capacitances, e.g., are constant. To a mathematician it is a linear system of differential equations with constant coefficients.



$$di_1/dt = -\, v_1/L - v_2/L + v_g(t)/L$$

$$-v_g(t) + v_1 + L\, di_1/dt + v_2 = 0$$

$$dv_2/dt = i_1/C_2 - v_2/RC_2$$

(http://www.ece.utah.edu/eceCTools/StateSpace/Circuits/Matlab/StateSpaceCircTutor.htm)

It is well known that when such a system is driven by a sinusoid, it eventually responds with a sinusoid of the same frequency. (We ignore resonance for the moment.) The engineer says the system responds "in sync" with the driver; the mathematician says the equations have a "particular solution" of the same mathematical form.

From Euler's equations $\cos 2\pi f t = \frac{e^{j2\pi f t} + e^{-j2\pi f t}}{2}$ and $\sin 2\pi f t = \frac{e^{j2\pi f t} - e^{-j2\pi f t}}{2j}$, we know that we can generically represent a sinusoid by $e^{j2\pi f t}$ if we allow for negative frequencies. For sinusoidal drivers

$$v_g(t) \leftarrow \mathrm{v_g} e^{j2\pi f t}$$

and sinusoidal responses

$$i_1(t) \leftarrow \mathrm{i_1} e^{j2\pi f t},\ v_1(t) \leftarrow \mathrm{v_1} e^{j2\pi f t}, v_2(t) \leftarrow \mathrm{v_2} e^{j2\pi f t};$$

time differentiation in the differential equations becomes replaced by mere multiplication:

$$\boldsymbol{j2\pi f} e^{j2\pi f t}\mathrm{i_1} = -\, \mathrm{v_1} e^{j2\pi f t}/L - \mathrm{v_2} e^{j2\pi f t}/L + \mathrm{v_g} e^{j2\pi f t}/L$$

$$- v_g e^{j2\pi ft} + v_1 e^{j2\pi ft} + L j2\pi f e^{j2\pi ft} i_1 + v_2 e^{j2\pi ft} = 0$$

$$j2\pi f \; e^{j2\pi ft} v_2 = v_1 e^{j2\pi ft}/C_2 - v_2 e^{j2\pi ft}/RC_2 \; ;$$

and the time-dependent differential equations are replaced by time-independent algebraic equations:

$$j2\pi f i_1 = - v_1/L - v_2/L + v_g/L$$
$$- v_g + v_1 + L j2\pi f i_1 + v_2 = 0$$

$$j2\pi f v_2 = v_1/C_2 - v_2/RC_2 \, ,$$

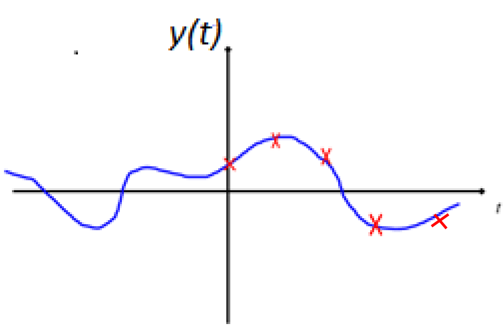
or
$$\begin{bmatrix} j2\pi f & 1/L & 1/L \\ Lj2\pi f & 1 & 1 \\ 0 & -1/C_2 & j2\pi f + 1/RC_2 \end{bmatrix} \begin{bmatrix} i_1 \\ v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} v_g/L \\ v_g \\ 0 \end{bmatrix}.$$

Engineering interpretation: the capacitors and inductors are replaced by resistors ("impedances") and the dynamic system becomes static. The task is recast in the *frequency domain*.

So the analysis of a system driven by sinusoids is much simpler than for arbitrary inputs.

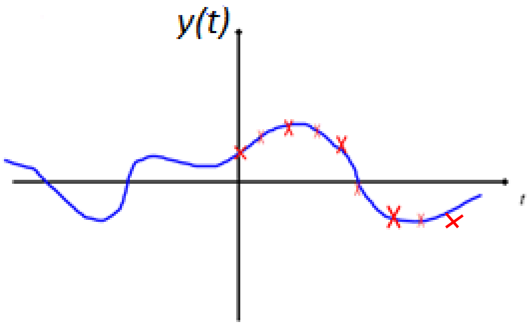
The contribution of Fourier and his disciples is the recognition that *practically every function can be written as a superposition of sinusoids*. Therefore - by superposition - the analysis of every linear time invariant system, for arbitrary inputs, enjoys this same simplification, reducing dynamics to statics and replacing differential equations by algebraic ones.

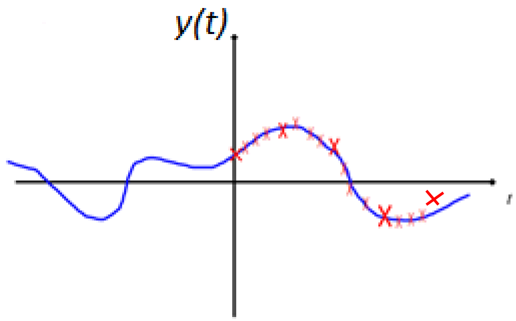
Let's see how this superposition comes about. We start by considering uniformly spaced measurements of a function $y(t)$:



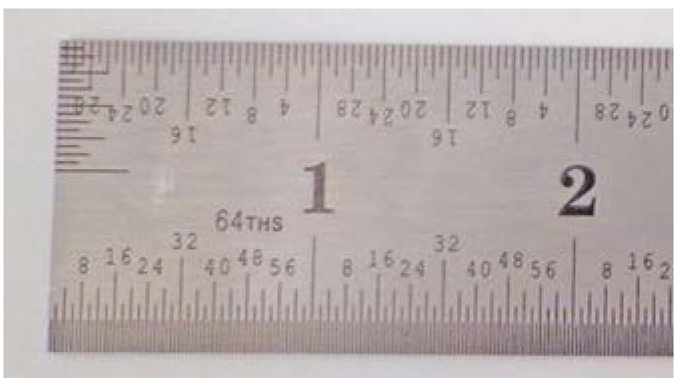
$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} y(t_1) \\ y(t_2) \\ y(t_3) \\ \vdots \\ y(t_n) \end{bmatrix}.$$

We envision a scheme wherein we refine our measurements by successively incorporating the midpoints:





Specifically, the measurements are taken at values of t ($0 \le t < 1$; we'll rescale later) which are the *dyadic fractions* ("foot-ruler markings"): $\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_{2^N} \end{bmatrix} = \begin{bmatrix} y(0) \\ y(\frac{1}{2^N}) \\ y(\frac{2}{2^N}) \\ \vdots \\ y(\frac{2^N-1}{2^N}) \end{bmatrix}$,



and we exclude the final point t=1 to make the numbering come out right.

Now consider a matrix-transformed version of *y*,

(1) $$\hat{y} = My\,, \quad \begin{bmatrix} \widehat{y_1} \\ \widehat{y_2} \\ \widehat{y_3} \\ \vdots \\ \widehat{y_{2^N}} \end{bmatrix} = \begin{bmatrix} & M & \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_{2^N} \end{bmatrix},$$

(2) $$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_{2^N} \end{bmatrix} = \begin{bmatrix} & M^{-1} & \end{bmatrix} \begin{bmatrix} \widehat{y_1} \\ \widehat{y_2} \\ \widehat{y_3} \\ \vdots \\ \widehat{y_{2^N}} \end{bmatrix}.$$

The matrix is $M = \begin{bmatrix} & M_{mn} & \end{bmatrix}$ with M_{mn} given by $\frac{1}{2^N} e^{-j2\pi(m-1-2^{N-1})(n-1)/2^N}$. Much more transparently, if we define $w \equiv e^{-j2\pi/2^N}$, then for N=2 (2^N=4 measured values, w = -j) the pattern is

$$M = \tfrac{1}{4}\begin{bmatrix} 1 & -1 & 1 & -1 \\ 1 & -w & w^2 & -w^3 \\ 1 & -w^2 & w^4 & -w^6 \\ 1 & -w^3 & w^6 & -w^9 \end{bmatrix}.$$

M is easy to invert; M^{-1} is the transposed complex conjugate ("Hermitian"), rescaled by the factor 2^N . ($M^{-1} = 2^N M^{\mathrm{H}}$.)

$$M^{-1} = 4M^{\mathrm{H}} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ -1 & -\overline{w} & -\overline{w}^2 & -\overline{w}^3 \\ 1 & \overline{w}^2 & \overline{w}^4 & \overline{w}^6 \\ -1 & -\overline{w}^3 & -\overline{w}^6 & -\overline{w}^9 \end{bmatrix}.$$

You can almost verify this mentally. The diagonal terms of $M^{\mathrm{H}}M$ entail adding 4 copies of ones ($\overline{w}w = 1$). The offdiagonal terms each take the form $1 + r + r^2 + r^4$, so the familiar formula $1 + r + r^2 + \cdots + r^n = \frac{1-r^{n+1}}{1-r}$ gives zero when you identify the "r".

$\hat{y}$ in (1) is called the ***Discrete Fourier Transform*** of y. It's the Fourier transform that we use whenever we are doing simulations on a computer. (2) expresses the inverse transform. To compute a single coefficient $\hat{y}_n$ by (1) requires 2^N complex multiplications and additions. To compute all 2^N coefficients would take 2^{2N} multiplies. For large data sets, that’s prohibitive. But the sinusoidal factors in these products are very redundant. For example when N=4 (2^N = 16), the data in the rows of M are given by

n	$e^{j2\pi n/16}$
-8	-1
-7	-.924 - .383j
-6	-.707 - .707j
-5	-.383 - .924j
-4	-j
-3	.383 - .924j
-2	.707 - .707j
-1	.924 - .383j
0	1
1	.924 + .383j
2	.707 + .707j
3	.383 + .924j
4	j
5	-.383 + .924j
6	-.707 + .707j
7	-.924 + .383j

so there are only 3 different nontrivial factors. With appropriate grouping only 3 multiplications are required for each $\widehat{y_n}$. The *Fast Fourier Transform* exploits these redundancies and computes all 2^N coefficients $\hat{y}$ using only $2^N(1+N/2)$ multiplies. The FFT has rendered data analysis feasible for an enormous number of applications. (There are versions of the algorithm that apply when the number of data is *not* a power of 2.)

This form of the Discrete Fourier Transform is good for visualizing the matrix algebra (and for computer coding), but the physical meaning is obscure; it will become more transparent when we change the notation. We identify $(n-1)/2^N$ as the n^{th} measurement time, after N refinements (2^N measurements), since we count measurements starting from n=1 but time starts from t=0.

(2) $$M_{mn} = \frac{1}{2^N} e^{-j2\pi(m-1-2^{N-1})(n-1)/2^N} = \frac{1}{2^N} e^{-j2\pi(m-1-2^{N-1})t} \text{ for } t = (n-1)/2^N .$$

This is the formula for a sinusoid, oscillating at frequency $m-1-2^{N-1}$ cycles per second, for $m = 1, 2, \cdots, 2^N$. If we rename $m-1-2^{N-1}$ as f, running from $f = -2^{N-1}$ to $2^{N-1}-1$, then

$$M_{mn} = \frac{1}{2^N} e^{-j2\pi f t} \text{ with } t = (n-1)/2^N \text{ and } f = m-1-2^{N-1}.$$

Now the expression (2) takes the familiar form we've been looking for:

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_{2^N} \end{bmatrix} \equiv \begin{bmatrix} y(t_1) \\ y(t_2) \\ y(t_3) \\ \vdots \\ y(t_{2^N}) \end{bmatrix} = \begin{bmatrix} & & \\ & (e^{+j2\pi ft}) & \\ & & \end{bmatrix} \begin{bmatrix} \widehat{y_1} \\ \widehat{y_2} \\ \widehat{y_3} \\ \vdots \\ \widehat{y_{2^N}} \end{bmatrix},$$

or (row by row)

(3) $$y(t) = \sum_{f=-2^{N-1}}^{(2^{N-1}-1)} \widehat{y(f)}\, e^{j2\pi ft} \text{ for } \; t = \frac{n-1}{2^N}$$

(where we have written $\widehat{y(f)}$ for $y_{\widehat{m-1-2^{N-1}}}$). So now *y*(*t*) is expressed as a superposition of 2^N *sinusoids* - at least, for the dyadic fractional times between 0 and 1 - with frequencies

$f = -2^{N-1}, -2^{N-1}, \cdots, 2^{N-1} - 1$ Hz.

(3) is *almost* the Fourier series for *y*(*t*); we'll apply the finishing touch soon. But let's jump the gun and make some (premature) observations now.

(3) is *exact*, not an approximation, at every *dyadic* time value - eventually; once *N* is large enough that the dyadic *t* can be expressed as $(n\text{-}1)/2^N$, (3) holds. Therefore,

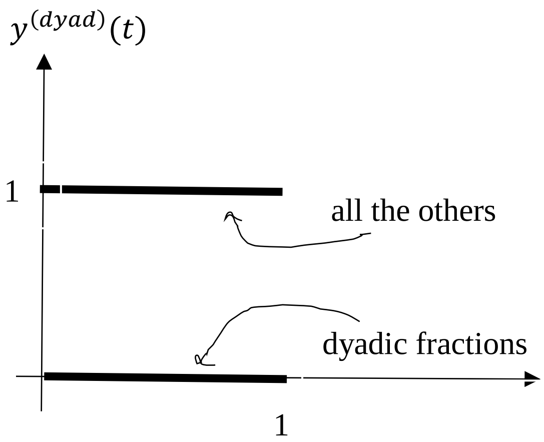
(4) $$y(t) = \lim_{N\to\infty} \sum_{f=-2^{N-1}}^{(2^{N-1}-1)} \widehat{y(f)}\, e^{j2\pi ft} \text{ for every dyadic fraction time } \; t,\; 0 \le t < 1,$$

$$\text{when } \widehat{y(f)} \text{ is given by } \frac{1}{2^N}\sum_t y(t)e^{-j2\pi ft} \quad \left(t = 0, \frac{1}{2^N}, \frac{2}{2^N}, \cdots, \frac{2^N-1}{2^N}\right).$$

We ask: Is (4) valid for the *other* values of *t*? Can we use it to interpolate y at the nondyadic points?

Remember that the dyadic fractions are infinitely dense in the interval [0,1). Now if a function's value at a *non*dyadic point, like *t* = 1/3, is radically different from its values at all the neighboring dyadic fractions … - well, that function is weird. You can take an infinite number of measurements of the function on an infinitely dense set of points and never learn anything about its value at 1/3. *You cannot simulate it on any commercial computer* (because binary numbers are dyadic fractions). If we were to define a "measureable" function as a function which can be completely determined by *measuring* it on any dense point set, then (4) is a Fourier representation for all "measureable" functions. Those of you whose mathematical curiosity extends no further than computer simulations can stop reading. (You learned Fourier analysis in ~~one hour~~ 15 minutes!)

In formulating his extension of the notion of Riemann integration, one of the weird functions considered by the mathematician Lebesque was the function defined to be 0 on the dyadic fractions, and 1 on all the other numbers; call it $y^{(dyad)}(t)$. The graph of this function is deceptive,



because neither of the segments is connected; they are perforated infinitely many times. (4) is certainly false for $y^{(dyad)}(t)$; all of the coefficients $\widehat{y(f)}$ are zero, so *the sum in* (4) *is zero for every t*, including *t* = 1/3 where $y^{(dyad)}(t)$ equals 1. The set of dyadic fractions is a subset of the set of rational numbers; it is countable. But the set of all numbers in [0,1) is uncountable. So the number of instances where $y^{(dyad)}(t)$ = 1 greatly "outweighs" the number of instances where it is zero, and (4) is wrong most of the time!

> (Lebesque had his own notion of what "measureable" means, and he considered $y^{(dyad)}(t)$ to be a measureable function - even though the true nature of the function is not revealed by an infinite number of "point" measurements (at the dyadic fractions). Lebesque felt that most physical measurements are performed by devices that "average" properties over small connected sets - like a thermometer, which does not record tempertures at a point but rather "smears" them over the surface of its bulb. A smeared measurement of $y^{(dyad)}(t)$ would report its value as 1. Indeed, Lebesque states $y^{(dyad)}(t) = 1$ "almost everywhere.")

But since (4) *is* valid at every dyadic fraction, and the dyadic fractions are infinitely dense, we have good reason to believe that (4) should converge to *y*(*t*) for, at least, the continuous functions. Before we delve too deeply into the convergence properties, let's clean up one last detail and give the accurate definition of the Fourier transform.

Again, we change the notation to get more insight. In the formula for the coefficient

$\widehat{y(f)} = \frac{1}{2^N}\sum_t y(t)e^{-j2\pi ft}$ $\left(t = 0, \frac{1}{2^N}, \frac{2}{2^N}, \cdots, \frac{2^N-1}{2^N}\right)$, we note that the spacing between the dyadic times Δt equals $\frac{1}{2^N}$, so we can write

$$\widehat{y(f)} = \sum_t y(t)e^{-j2\pi ft}\ \Delta t,$$

which is a Riemann sum for the integral $\int_0^1 y(t)\, e^{-j2\pi ft} dt$. Therefore in the *limit* as $N \to \infty$ we get $\widehat{y(f)} \to \int_0^1 y(t)\, e^{-j2\pi ft} dt$. The (legitimate) *Finite Fourier Transform* of *y*(*t*) is defined as this limit, and the *Fourier Series* for *y*(*t*) is $\sum_{f=-\infty}^{\infty} \widehat{y(f)}\, e^{j2\pi ft}$ (with $\widehat{y(f)}$ given by the integral, not the sum). The question that mathematicians asked, then, is

I*s y*(*t*) *given by the sum of its Fourier Series for all t in* [0, 1):

(5) $$\widehat{y(f)} = \int_0^1 y(t)\, e^{-j2\pi ft} dt \;\; ; \;\; y(t) \stackrel{?}{=} \sum_{f=-\infty}^{\infty} \widehat{y(f)}\, e^{j2\pi ft} \; .$$

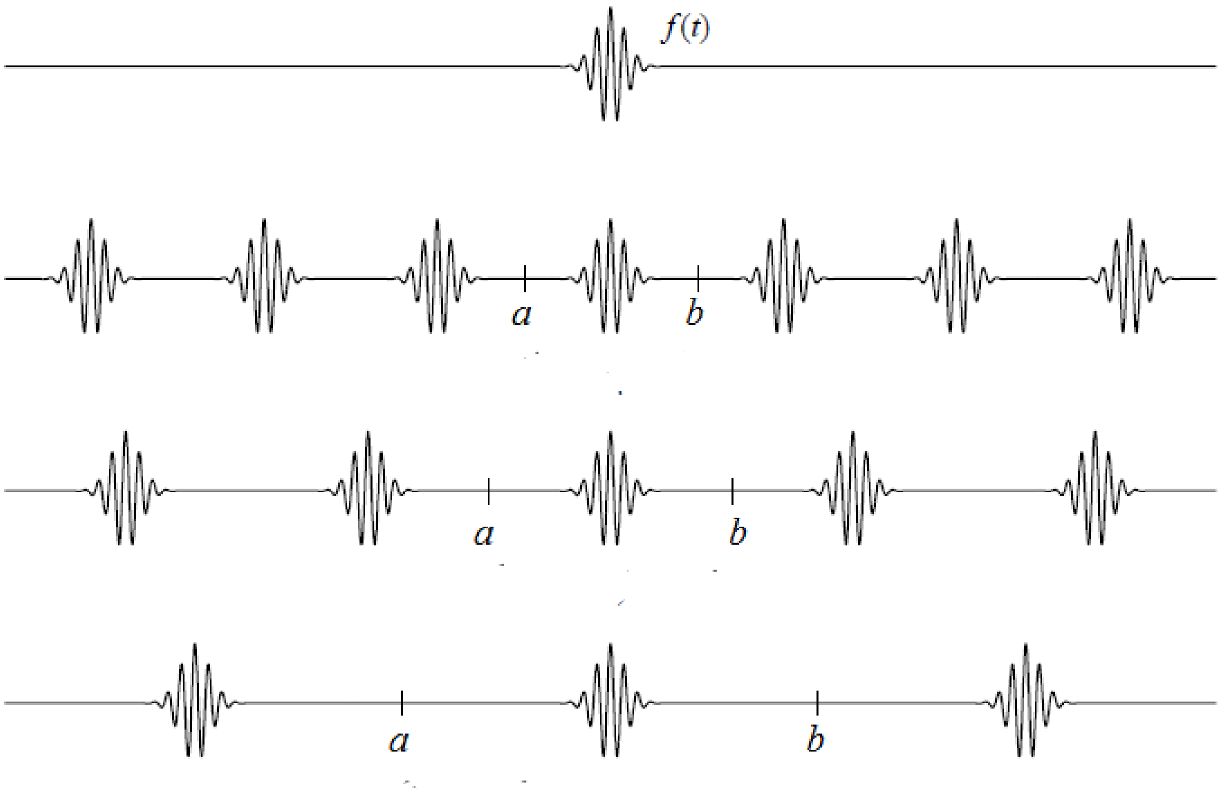
A few remarks before we address this question:

(*i*) The representation can be moved from $0 \le t < 1$ to any other interval $a \le \tau < b$ by a simple change of variable $t = \frac{\tau - a}{b - a}$. With a little juggling the ***Finite Fourier Transform*** can then be expressed

(6) $$\widehat{Y(f)} = \frac{1}{b-a} \int_a^b Y(t)\, e^{-j2\pi ft} dt \;\; ; \;\; Y(t) \stackrel{?}{=} \sum_{p=-\infty}^{\infty} \widehat{Y(f)}\, e^{j2\pi ft} \, , \;\; f = \frac{p}{b-a} \, .$$

The frequencies are now spaced by $\Delta f = \frac{1}{b-a}$.

(*ii*) We do not expect the Fourier series in (6) to converge to *Y*(*t*) *outside* [*a*,*b*), because the sinusoids in the sum are *periodic*; the Fourier series is going to replace *Y*(*t*) by periodic replicas outside the interval. In particular, even if the series converged to *Y*(*a*) at *t*=*a*, it will not converge to *Y*(*b*) unless *Y*(*b*) = *Y*(*a*). The figure below displays what the various Fourier series for a particular function look like if the approximation interval $[a, b)$ is changed.



Fourier expansions over increasing intervals

(*iii*) The figure suggests that if let $a \to -\infty$ and $b \to +\infty$, the periodic replicas would be banished and we would have a sinusoidal representation of *Y*(*t*) for all *t*. The following mutation of (5) -

$$\widehat{Y(f)} = \int_a^b \frac{Y(t)}{b-a}\, e^{-j2\pi ft} dt \;\; ; \;\; \frac{Y(t)}{b-a} = \sum_{f=0,\ \frac{\pm 1}{b-a},\ \frac{\pm 2}{b-a},\ \ldots} \widehat{Y(f)}\, e^{j2\pi ft}\, \frac{1}{b-a}$$

- together with the identification of the frequency spacing as $\Delta f = \frac{1}{b-a}$ - warrants the conjecture

$$\widehat{F(f)} = \int_{-\infty}^{\infty} F(t)\, e^{-j2\pi ft} dt \;\; ; \;\; F(t) = \int_{-\infty}^{\infty} \widehat{F(f)}\, e^{j2\pi ft} df \;\; .$$

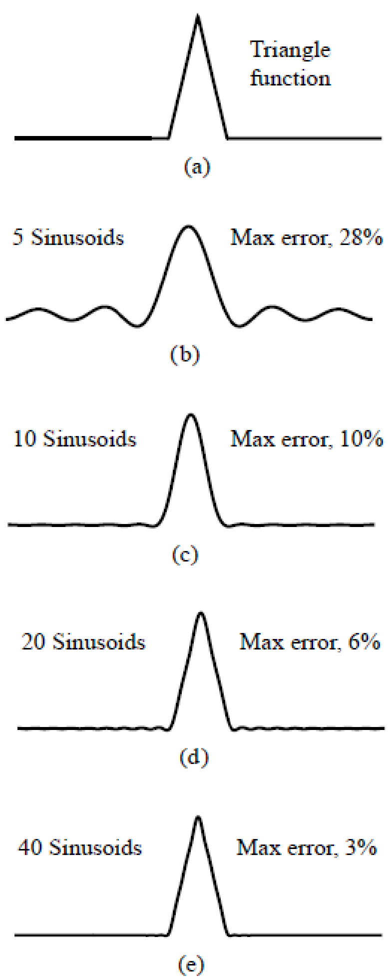
$\widehat{F(f)}$ is called the *Fourier Transform* of *F*(*t*) and the inverse transform tenders the expression of *F*(*t*) as a superposition of sinusoids for *all t in* $(-\infty, \infty)$.

(*iv*) Although the correctness of (4) at the dyadic fraction points in the interval is quite compelling, we cannot make the same claim for (5), since we have replaced the *sums* $\widehat{y(f)}$ by the *integrals* $\widehat{Y(f)}$.

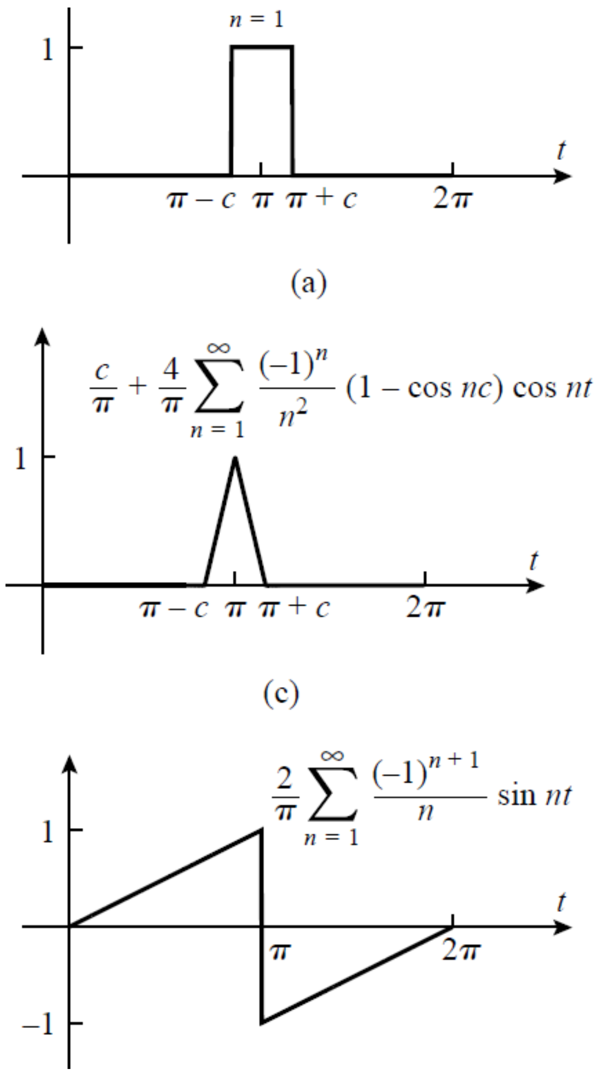
Investigation into the convergence question has spawned much of modern analysis. Surprisingly, continuity alone is not enough to guarantee (5); but if *y* is differentiable at *t*, the Fourier series converges there. Twentieth century mathematics is replete with examples demonstrating how "weird" functions can avoid convergence of their Fourier series. However in practical applications the Fourier series of a smooth function converges to the proper value; while if the function has jump discontinuities (like a switching function), the series converges to the midpoint of the break.

Even if the series fails to converge at various points, it may "converge" in a weaker sense of the word. For example, the area between the function's graph and the graph of its Fourier series may go to zero. Here are some graphics illustrating various types of convergence.

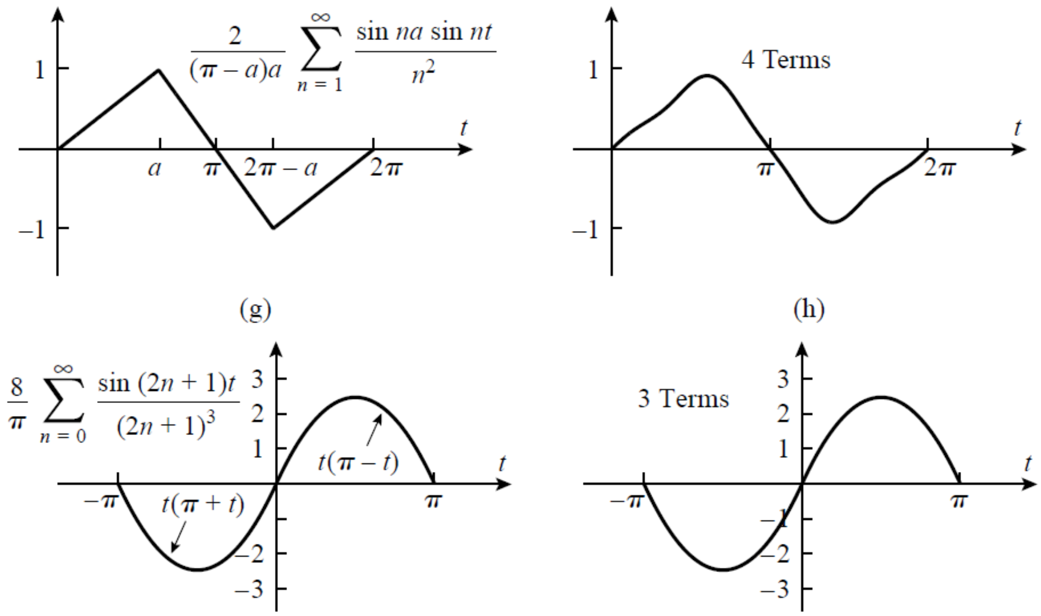
Snider, A. D., ***Partial Differential Equations - Sources and Solutions***, Prentice-Hall, 1999; Dover Pub., 2006. (ISBN 0-486-45340-5)



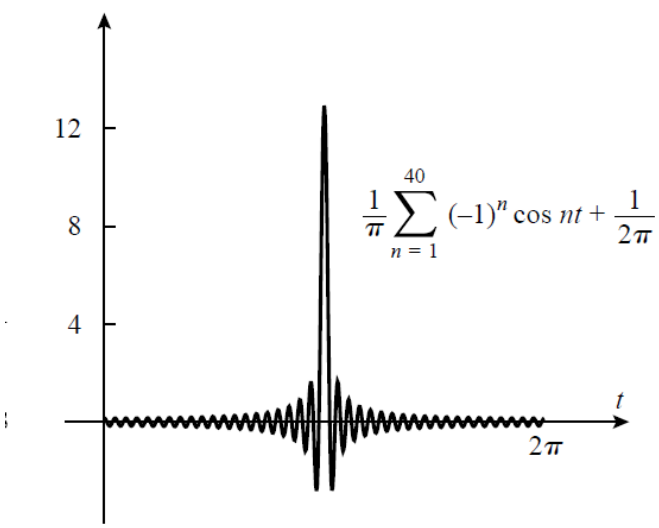
(Continuous function, differentiable everywhere except for three points)



Jump discontinuities; discontinuities in the derivative; jump discontinuity.

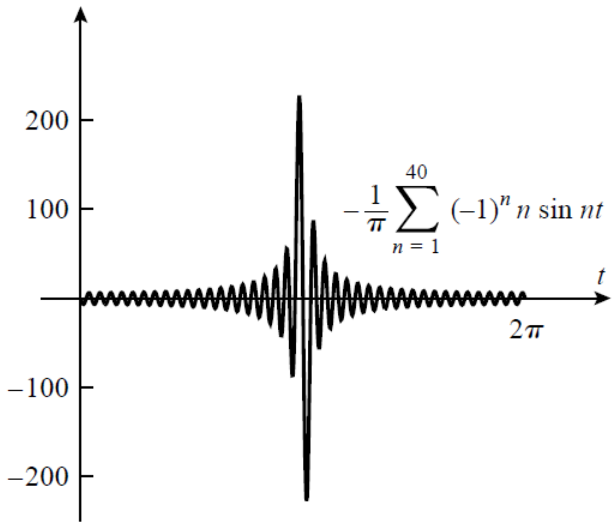


Continuous sawtooth; two parabolas joined smoothly.



Truncated Fourier series for $\delta(t - \pi)$

Wildly discontinuous delta function. The Fourier approximation's peak is so high that the central lobe contains unit area. The other lobes oscillate so fast that they cancel each other when integrated.



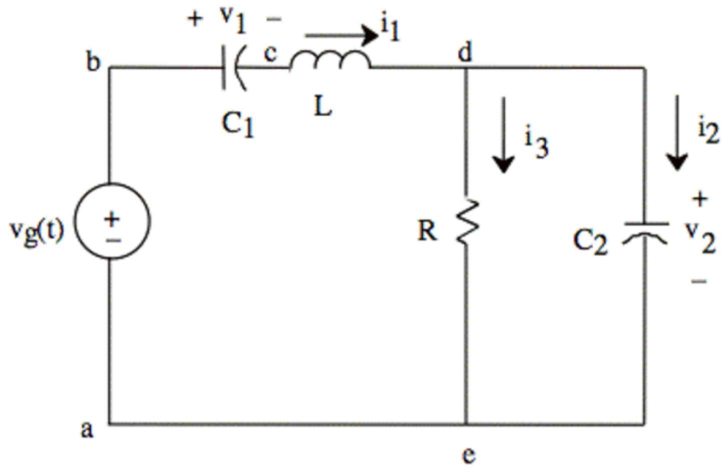
Truncated Fourier series for $\delta'(t-\pi)$

Even wilder - the derivative of the delta function can be interpreted as a dipole or a doublet.

Lecture 10
Feb. 20, 2017

Fourier Analysis in One Hour

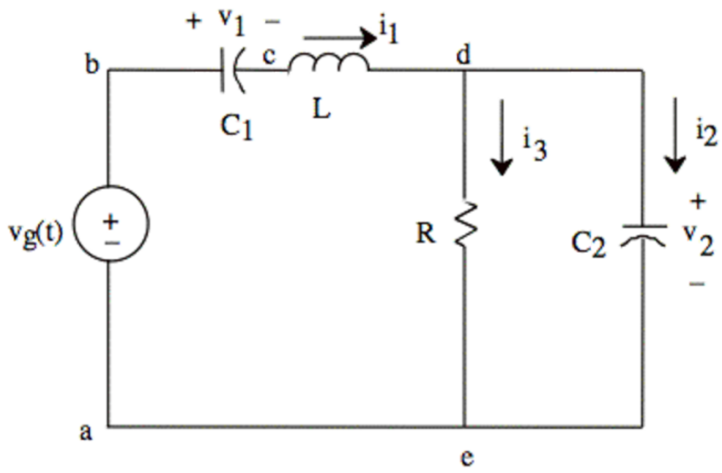
Linear time-invariant Systems



$di_1/dt = -v_1/L - v_2/L + v_g(t)/L$

$-v_g(t) + v_1 + L\, di_1/dt + v_2 = 0$

$dv_2/dt = i_1/C_2 - v_2/RC_2$



$$di_1/dt = -\,v_1/L - v_2/L + v_g(t)/L$$

$$-v_g(t) + v_1 + L\,di_1/dt + v_2 = 0$$

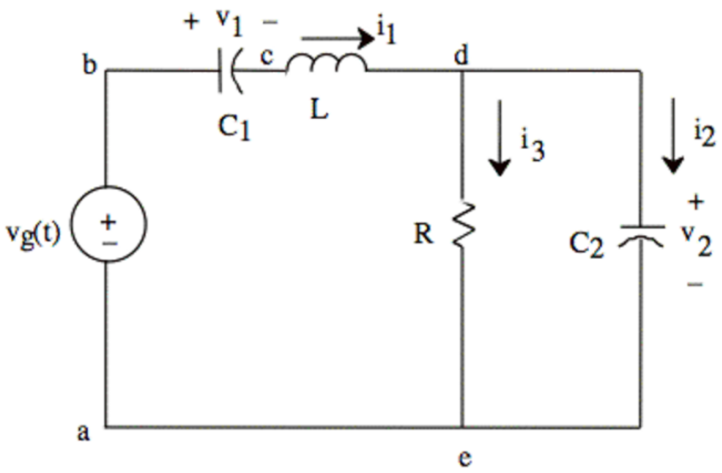
$$dv_2/dt = i_1/C_2 - v_2/RC_2$$

When such a system is driven by a sinusoid, it eventually responds with a sinusoid of the same frequency.

Engineer: "in sync"

Mathematician: "Method of Undetermined Coefficients"

Trial solution: sinusoid



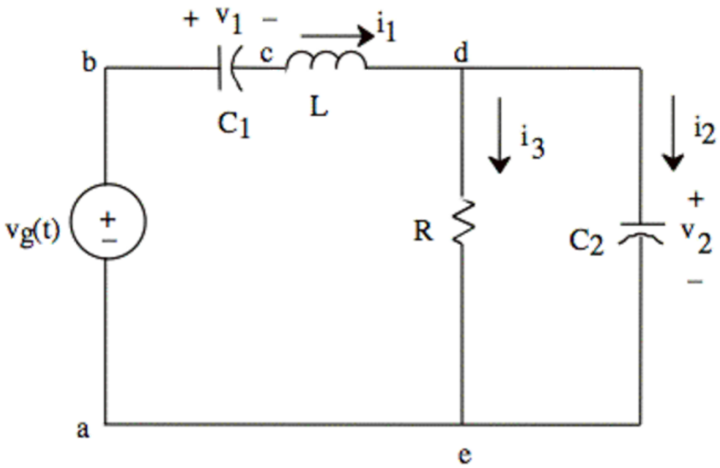
$di_1/dt = -v_1/L - v_2/L + v_g(t)/L$

$-v_g(t) + v_1 + L\, di_1/dt + v_2 = 0$

$dv_2/dt = i_1/C_2 - v_2/RC_2$

$$\cos 2\pi ft = \frac{e^{j2\pi ft} + e^{-j2\pi ft}}{2} \text{ and } \sin 2\pi ft = \frac{e^{j2\pi ft} - e^{-j2\pi ft}}{2j}$$

generically represent a sinusoid by $e^{j2\pi ft}$



$di_1/dt = -v_1/L - v_2/L + v_g(t)/L$

$-v_g(t) + v_1 + L\, di_1/dt + v_2 = 0$

$dv_2/dt = i_1/C_2 - v_2/RC_2$

$$v_g(t) \leftarrow \mathrm{v_g} e^{j2\pi ft}$$

$$i_1(t) \leftarrow \mathrm{i_1} e^{j2\pi ft},\ v_1(t) \leftarrow \mathrm{v_1} e^{j2\pi ft}, v_2(t) \leftarrow \mathrm{v_2} e^{j2\pi ft}$$

$$\boldsymbol{j2\pi f} e^{j2\pi ft}\mathrm{i_1} = -\mathrm{v_1}e^{j2\pi ft}/L - \mathrm{v_2}e^{j2\pi ft}/L + \mathrm{v_g}e^{j2\pi ft}/L$$

$$-\mathrm{v_g}e^{j2\pi ft} + \mathrm{v_1}e^{j2\pi ft} + L\boldsymbol{j2\pi f}e^{j2\pi ft}\mathrm{i_1} + \mathrm{v_2}e^{j2\pi ft} = 0$$

$$\boldsymbol{j2\pi f}\ e^{j2\pi ft}v_2 = \mathrm{v_1}e^{j2\pi ft}/C_2 - \mathrm{v_2}e^{j2\pi ft}/RC_2\ ;$$

$$\boldsymbol{j2\pi f} e^{j2\pi ft} \mathrm{i}_1 = -\, \mathrm{v}_1 e^{j2\pi ft}/L - \mathrm{v}_2 e^{j2\pi ft}/L + \mathrm{v}_g e^{j2\pi ft}/L$$

$$- \mathrm{v}_g e^{j2\pi ft} + \mathrm{v}_1 e^{j2\pi ft} + L\boldsymbol{j2\pi f} e^{j2\pi ft} \mathrm{i}_1 + \mathrm{v}_2 e^{j2\pi ft} = 0$$

$$\boldsymbol{j2\pi f}\ e^{j2\pi ft} v_2 = \mathrm{v}_1 e^{j2\pi ft}/C_2 - \mathrm{v}_2 e^{j2\pi ft}/RC_2\ ;$$

$$j2\pi f \mathrm{i}_1 = -\, \mathrm{v}_1/L - \mathrm{v}_2/L + \mathrm{v}_g/L$$

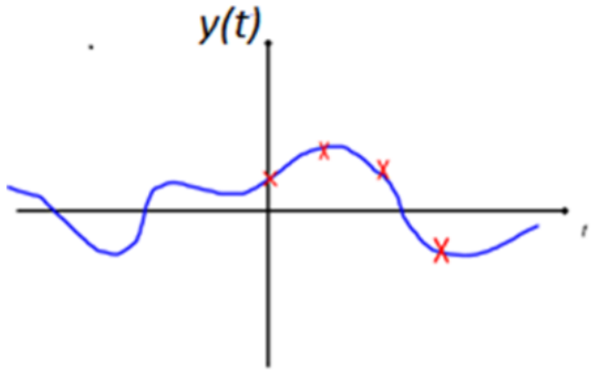
$$- \mathrm{v}_g + \mathrm{v}_1 + Lj2\pi f \mathrm{i}_1 + \mathrm{v}_2 = 0$$

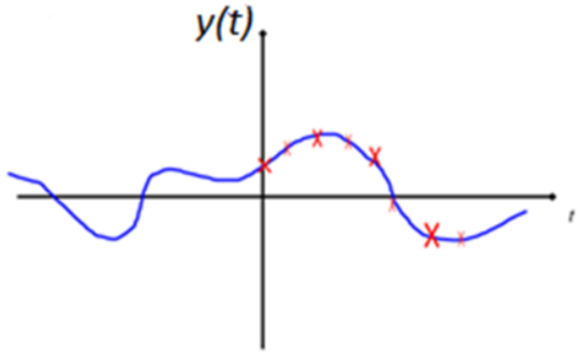
$$j2\pi f v_2 = \mathrm{v}_1/C_2 - \mathrm{v}_2/RC_2\ ,$$

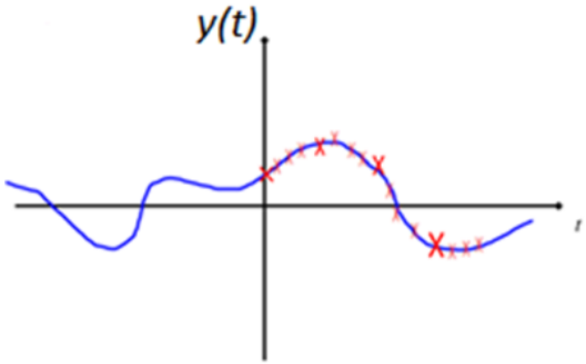
$$\begin{bmatrix} j2\pi f & 1/L & 1/L \\ Lj2\pi f & 1 & 1 \\ 0 & -1/C_2 & j2\pi f + 1/RC_2 \end{bmatrix} \begin{bmatrix} i_1 \\ v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} v_g/L \\ v_g \\ 0 \end{bmatrix}$$

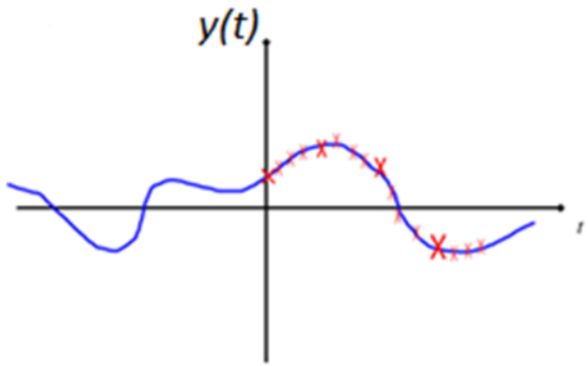
Dynamics ↓ Statics

Capacitors, Inductors ↓ Resistors (Impedances)

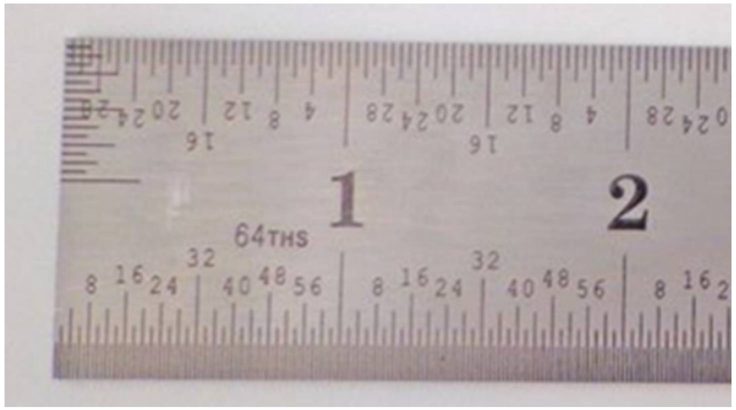

y(t)
t


y(t)
t


y(t)
t



$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_{2^N} \end{bmatrix} = \begin{bmatrix} y(0) \\ y(\frac{1}{2^N}) \\ y(\frac{2}{2^N}) \\ \vdots \\ y(\frac{2^N-1}{2^N}) \end{bmatrix}$$



$$\hat{y} = My\ , \quad \begin{bmatrix} \widehat{y_1} \\ \widehat{y_2} \\ \widehat{y_3} \\ \vdots \\ \widehat{y_{2^N}} \end{bmatrix} = \begin{bmatrix} \square & \square & \square \\ \square & M & \square \\ \square & \square & \square \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_{2^N} \end{bmatrix}$$

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_{2^N} \end{bmatrix} = \begin{bmatrix} \square & \square & \square \\ \square & M^{-1} & \square \\ \square & \square & \square \end{bmatrix} \begin{bmatrix} \widehat{y_1} \\ \widehat{y_2} \\ \widehat{y_3} \\ \vdots \\ \widehat{y_{2^N}} \end{bmatrix}$$

$$M_{mn} = \frac{1}{2^N} e^{-j2\pi(m-1-2^{N-1})(n-1)/2^N}$$

$$M = \frac{1}{4} \begin{bmatrix} 1 & -1 & 1 & -1 \\ 1 & -w & w^2 & -w^3 \\ 1 & -w^2 & w^4 & -w^6 \\ 1 & -w^3 & w^6 & -w^9 \end{bmatrix}$$

$$w = -j$$

$$M = \frac{1}{4}\begin{bmatrix} 1 & -1 & 1 & -1 \\ 1 & -w & w^2 & -w^3 \\ 1 & -w^2 & w^4 & -w^6 \\ 1 & -w^3 & w^6 & -w^9 \end{bmatrix}$$

$$M^{-1} = 4M^{\mathrm{H}} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ -1 & -\bar{w} & -\bar{w}^2 & -\bar{w}^3 \\ 1 & \bar{w}^2 & \bar{w}^4 & \bar{w}^6 \\ -1 & -\bar{w}^3 & -\bar{w}^6 & -\bar{w}^9 \end{bmatrix}$$

$$(\bar{w}w = 1)$$

$$\hat{y} = My\ , \quad \begin{bmatrix} \widehat{y_1} \\ \widehat{y_2} \\ \widehat{y_3} \\ \vdots \\ \widehat{y_{2^N}} \end{bmatrix} = \begin{bmatrix} \square & \square & \square \\ \square & M & \square \\ \square & \square & \square \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_{2^N} \end{bmatrix}$$

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_{2^N} \end{bmatrix} = \begin{bmatrix} \square & \square & \square \\ \square & M^{-1} & \square \\ \square & \square & \square \end{bmatrix} \begin{bmatrix} \widehat{y_1} \\ \widehat{y_2} \\ \widehat{y_3} \\ \vdots \\ \widehat{y_{2^N}} \end{bmatrix}$$

$$M_{mn} = \frac{1}{2^N} e^{-j2\pi(m-1-2^{N-1})(n-1)/2^N}$$

Discrete Fourier Transform

$$\hat{y} = My\,, \quad \begin{bmatrix} \widehat{y_1} \\ \widehat{y_2} \\ \widehat{y_3} \\ \vdots \\ \widehat{y_{2^N}} \end{bmatrix} = \begin{bmatrix} \square & \square & \square \\ \square & M & \square \\ \square & \square & \square \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_{2^N} \end{bmatrix}$$

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_{2^N} \end{bmatrix} = \begin{bmatrix} \square & \square & \square \\ \square & M^{-1} & \square \\ \square & \square & \square \end{bmatrix} \begin{bmatrix} \widehat{y_1} \\ \widehat{y_2} \\ \widehat{y_3} \\ \vdots \\ \widehat{y_{2^N}} \end{bmatrix}$$

$$M_{mn} = \frac{1}{2^N} e^{-j2\pi(m-1-2^{N-1})(n-1)/2^N}$$

Discrete Fourier Transform

2^N multiplies per term,
2^N terms,
2^{2N} multiplies

$$\hat{y} = My\ , \quad \begin{bmatrix} \widehat{y_1} \\ \widehat{y_2} \\ \widehat{y_3} \\ \vdots \\ \widehat{y_{2^N}} \end{bmatrix} = \begin{bmatrix} \square & \square & \square \\ \square & M & \square \\ \square & \square & \square \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_{2^N} \end{bmatrix}$$

n	$e^{j2\pi n/16}$
-8	-1
-7	-.924 - .383j
-6	-.707 - .707j
-5	-.383 - .924j
-4	-j
-3	.383 - .924j
-2	.707 - .707j
-1	.924 - .383j
0	1
1	.924 + .383j
2	.707 + .707j
3	.383 + .924j
4	j
5	-.383 + .924j
6	-.707 + .707j
7	-.924 + .383j

$2^4 = 16$

Only need 3 multiplies per term.
48 multiplies instead of 256

$$\hat{y} = My \,, \quad \begin{bmatrix} \widehat{y_1} \\ \widehat{y_2} \\ \widehat{y_3} \\ \vdots \\ \widehat{y_{2^N}} \end{bmatrix} = \begin{bmatrix} \square & \square & \square \\ \square & M & \square \\ \square & \square & \square \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_{2^N} \end{bmatrix}$$

Discrete Fourier Transform

2^{2N} multiplies

Fast Fourier Transform

$2^N(1 + N/2)$ multiplies

$$\hat{y} = My\,, \quad \begin{bmatrix} \widehat{y_1} \\ \widehat{y_2} \\ \widehat{y_3} \\ \vdots \\ \widehat{y_{2^N}} \end{bmatrix} = \begin{bmatrix} \square & \square & \square \\ \square & M & \square \\ \square & \square & \square \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_{2^N} \end{bmatrix}$$

$$M_{mn} = \frac{1}{2^N} e^{-j2\pi\left(m-1-2^{N-1}\right)(n-1)/2^N}$$

$$= \frac{1}{2^N} e^{-j2\pi(m-1-2^{N-1})t} \text{ for } t = (n-1)/2^N$$

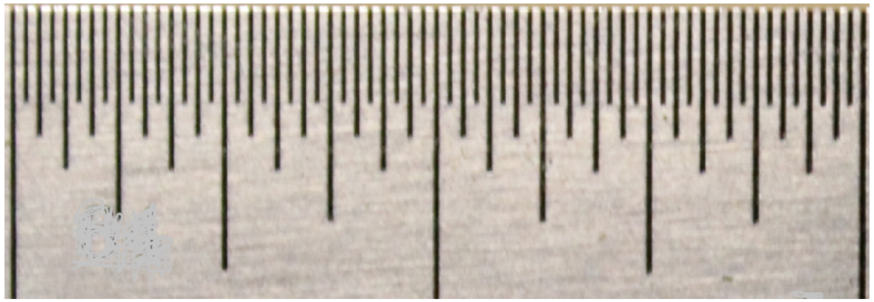
$$M_{mn} = \frac{1}{2^N} e^{-j2\pi f t} \text{ with } t = (n-1)/2^N$$

and $f = m - 1 - 2^{N-1}$.

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_{2^N} \end{bmatrix} \equiv \begin{bmatrix} y(t_1) \\ y(t_2) \\ y(t_3) \\ \vdots \\ y(t_{2^N}) \end{bmatrix} = \begin{bmatrix} \square & \square & \square \\ \square & (e^{+j2\pi f t}) & \square \\ \square & \square & \square \end{bmatrix} \begin{bmatrix} \widehat{y_1} \\ \widehat{y_2} \\ \widehat{y_3} \\ \vdots \\ \widehat{y_{2^N}} \end{bmatrix}$$

$$y(t) = \sum_{f=-2^{N-1}}^{(2^{N-1}-1)} \widehat{y(f)}\, e^{j2\pi f t} \text{ for } t = \frac{n-1}{2^N}$$

$$f = -2^{N-1}, -2^{N-1}, \cdots, 2^{N-1} - 1 \text{ Hz}$$



$t = \frac{n-1}{2^N}$

Let N -> ∞

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_{2^N} \end{bmatrix} \equiv \begin{bmatrix} y(t_1) \\ y(t_2) \\ y(t_3) \\ \vdots \\ y(t_{2^N}) \end{bmatrix} = \begin{bmatrix} \square & \square & \square \\ \square & (e^{+j2\pi ft}) & \square \\ \square & \square & \square \end{bmatrix} \begin{bmatrix} \widehat{y_1} \\ \widehat{y_2} \\ \widehat{y_3} \\ \vdots \\ \widehat{y_{2^N}} \end{bmatrix}$$

$$f = -2^{N-1}, -2^{N-1}, \cdots, 2^{N-1} - 1 \text{ Hz}$$

$$y(t) = \sum_{f=-2^{N-1}}^{(2^{N-1}-1)} \widehat{y(f)}\, e^{j2\pi ft} \text{ for } \ t = \frac{n-1}{2^N}$$

$$\begin{bmatrix} \widehat{y_1} \\ \widehat{y_2} \\ \widehat{y_3} \\ \vdots \\ \widehat{y_{2^N}} \end{bmatrix} = \begin{bmatrix} \square & \square & \square \\ \square & M & \square \\ \square & \square & \square \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_{2^N} \end{bmatrix}$$

$$\frac{1}{2^N} \sum_t y(t) e^{-j2\pi ft}$$

$$\left(t = 0, \frac{1}{2^N}, \frac{2}{2^N}, \cdots, \frac{2^N-1}{2^N}\right)$$

$$\begin{bmatrix} \widehat{y_1} \\ \widehat{y_2} \\ \widehat{y_3} \\ \vdots \\ \widehat{y_{2^N}} \end{bmatrix} = \begin{bmatrix} \square & \square & \square \\ \square & M & \square \\ \square & \square & \square \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_{2^N} \end{bmatrix}$$

$$= \frac{1}{2^N} \sum_t y(t) e^{-j2\pi f t} \qquad \Delta t \text{ equals } \frac{1}{2^N}$$

$$(t = 0, \frac{1}{2^N}, \frac{2}{2^N}, \cdots, \frac{2^N - 1}{2^N})$$

$$\widehat{y(f)} = \sum_t y(t) e^{-j2\pi f t} \ \Delta t$$

$$\rightarrow \int_0^1 y(t)\, e^{-j2\pi f t} dt$$

$$f = -2^{N-1}, -2^{N-1}, \cdots, 2^{N-1} - 1 \text{ Hz}$$

Finite Fourier Transform

$$\widehat{y(f)} \rightarrow \int_0^1 y(t)\, e^{-j2\pi f t} dt$$

Fourier Series

$$y(t) = \sum_{f=-\infty}^{\infty} \widehat{y(f)}\, e^{j2\pi f t}$$

for all t in [0, 1)

$$t = \frac{\tau - a}{b - a}$$

Finite Fourier Transform

$$\widehat{Y(f)} = \frac{1}{b-a} \int_a^b Y(t)\, e^{-j2\pi f t} dt$$

Fourier Series

$$Y(t) = \sum_{p=-\infty}^{\infty} \widehat{Y(f)}\, e^{j2\pi f t}$$

$$f = \frac{p}{b-a} \qquad \Delta f = \frac{1}{b-a}$$

$$a \le t < b$$

Fourier Series

$$f = \frac{p}{b-a}$$

$$Y(t) = \sum_{p=-\infty}^{\infty} \widehat{Y(f)}\, e^{j2\pi f t}$$

$$a \leq t < b$$

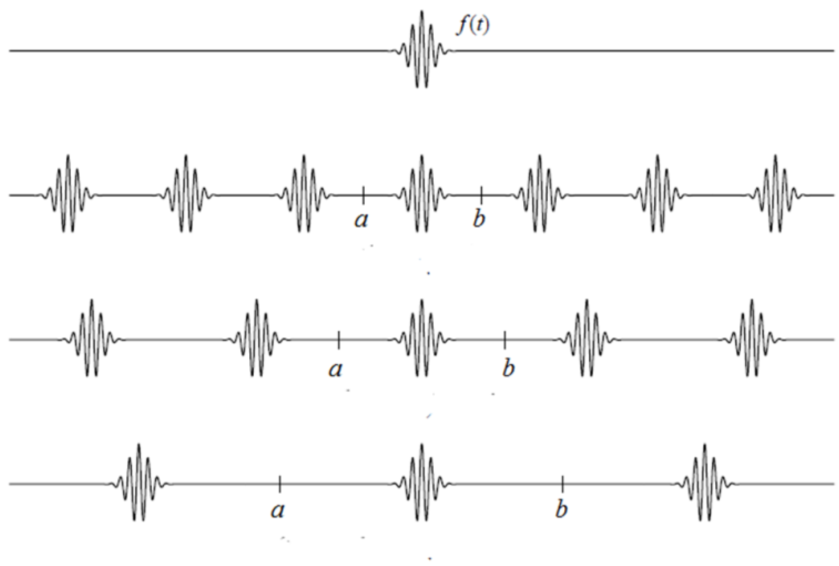
Periodic

Fourier Series

$$Y(t) = \sum_{p=-\infty}^{\infty} \widehat{Y(f)}\, e^{j2\pi f t}$$

$$a \leq t < b$$

Periodic

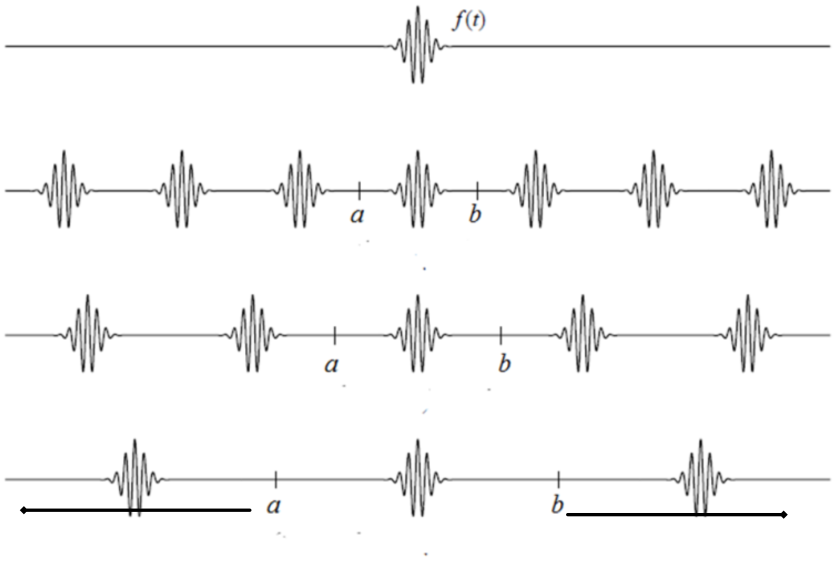


Fourier expansions over increasing intervals

Fourier Series

$$Y(t) = \sum_{p=-\infty}^{\infty} \widehat{Y(f)}\, e^{j2\pi f t}$$

$$\widehat{Y(f)} = \frac{1}{b-a} \int_a^b Y(t)\, e^{-j2\pi f t} dt$$



Fourier expansions over increasing intervals

$$F(t) = \int_{-\infty}^{\infty} \widehat{F(f)}\, e^{j2\pi f t} df$$

$$\widehat{F(f)} = \int_{-\infty}^{\infty} F(t)\, e^{-j2\pi f t} dt$$

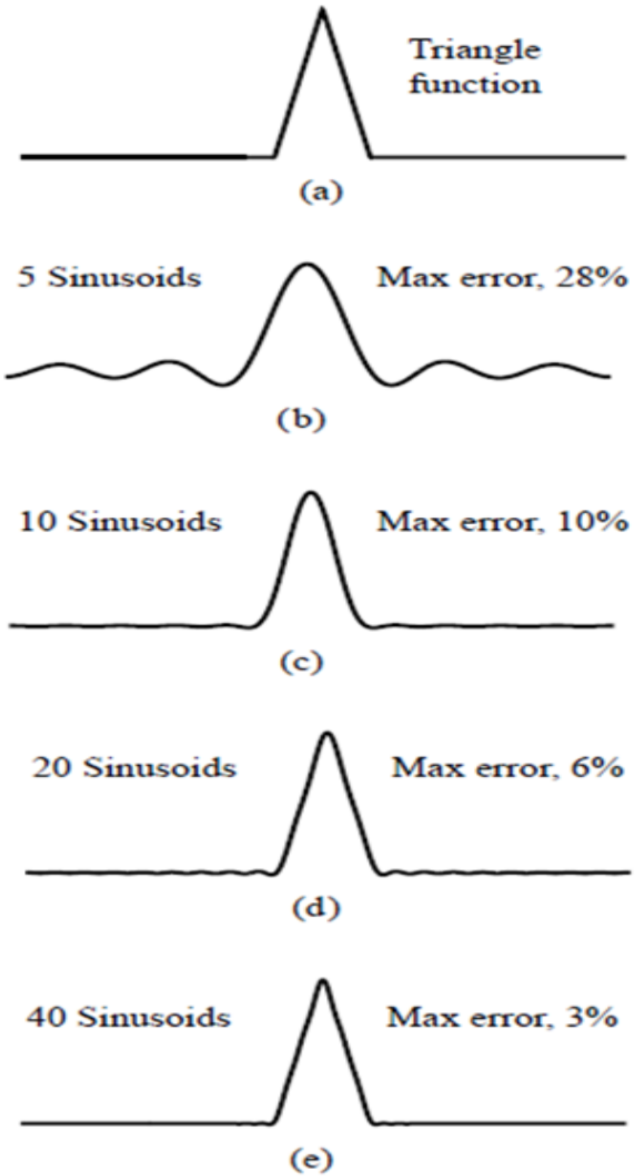
Fourier Transform

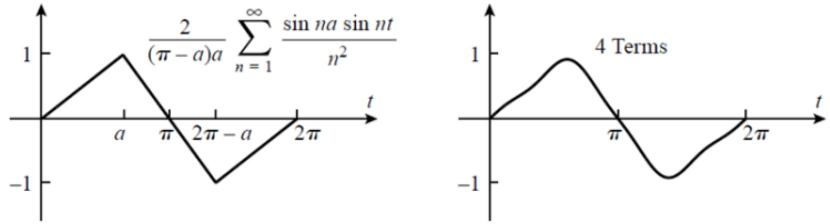
$$\widehat{F(f)} = \int_{-\infty}^{\infty} F(t)\, e^{-j2\pi f t} dt$$

Fourier Integral Representation

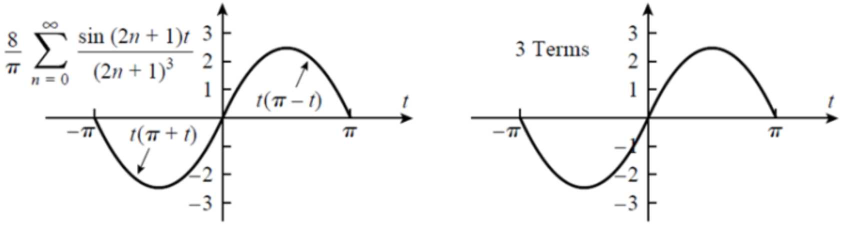
$$F(t) = \int_{-\infty}^{\infty} \widehat{F(f)}\, e^{j2\pi f t} df$$

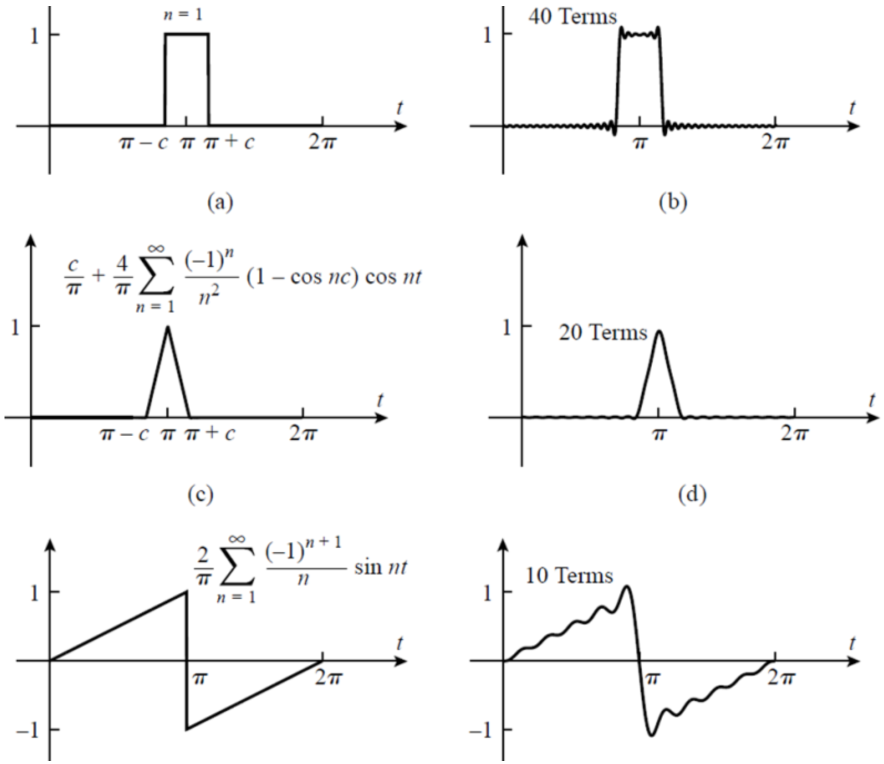
$$\text{all } t \text{ } in \ (-\infty, \infty)$$

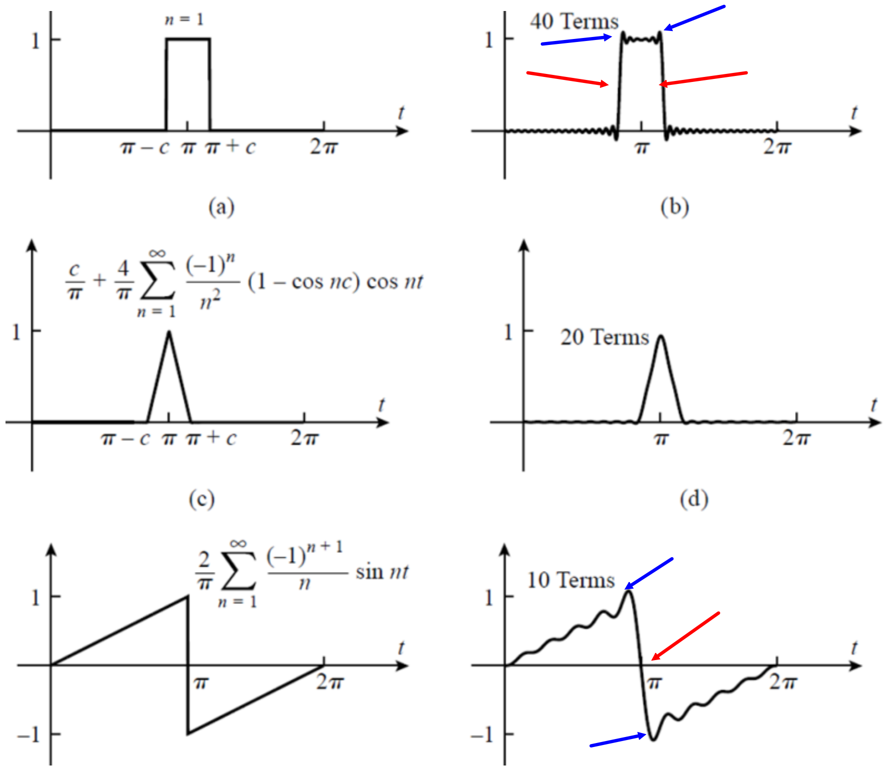

Triangle function
(a)
5 Sinusoids
Max error, 28%
(b)
10 Sinusoids
Max error, 10%
(c)
20 Sinusoids
Max error, 6%
(d)
40 Sinusoids
Max error, 3%
(e)


$\frac{2}{(\pi - a)a} \sum_{n=1}^{\infty} \frac{\sin na \sin nt}{n^2}$
1
−1
t
a
π
2π − a
2π
4 Terms

(g) (h)


$\frac{8}{\pi} \sum_{n=0}^{\infty} \frac{\sin (2n+1)t}{(2n+1)^3}$
3
2
1
−2
−3
t
−π
π
t(π − t)
t(π + t)
3 Terms


n = 1
1
t
π − c π π + c 2π
(a)
40 Terms
1
t
π 2π
(b)
$\frac{c}{\pi} + \frac{4}{\pi}\sum_{n=1}^{\infty}\frac{(-1)^n}{n^2}(1-\cos nc)\cos nt$
1
t
π − c π π + c 2π
(c)
1
20 Terms
t
π 2π
(d)
$\frac{2}{\pi}\sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{n}\sin nt$
1
t
π 2π
−1
10 Terms
1
t
π 2π
−1



n = 1
1
t
π − c π π + c 2π
(a)
40 Terms
1
t
π
2π
(b)
$\frac{c}{\pi} + \frac{4}{\pi}\sum_{n=1}^{\infty}\frac{(-1)^n}{n^2}(1-\cos nc)\cos nt$
1
t
π − c π π + c 2π
(c)
20 Terms
1
t
π 2π
(d)
$\frac{2}{\pi}\sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{n}\sin nt$
1
−1
t
π 2π
10 Terms
1
−1
t
π 2π

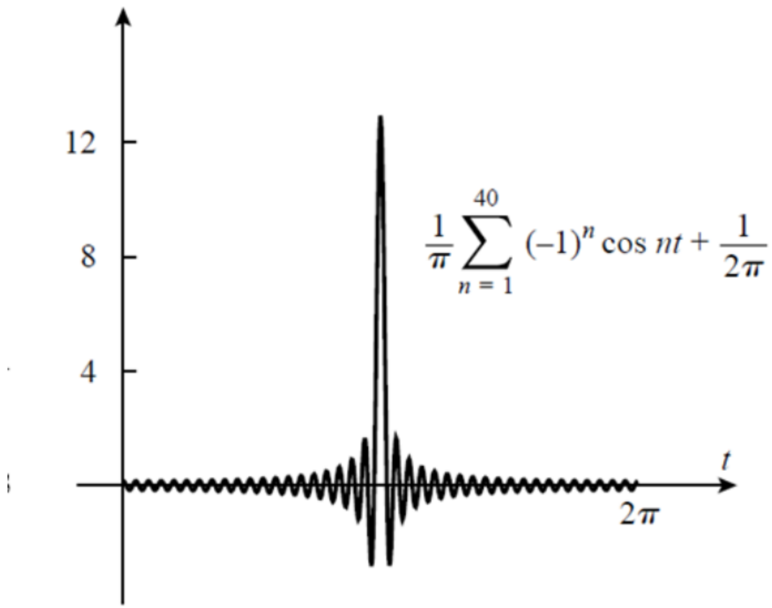
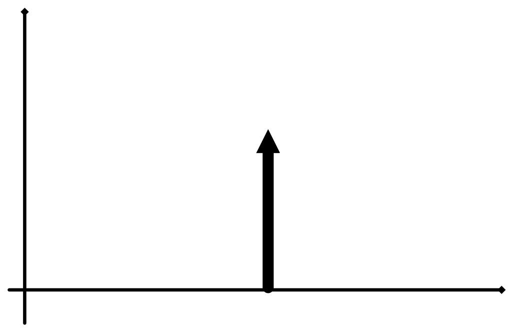


Figure 1 *Truncated Fourier series for* $\delta(t-\pi)$



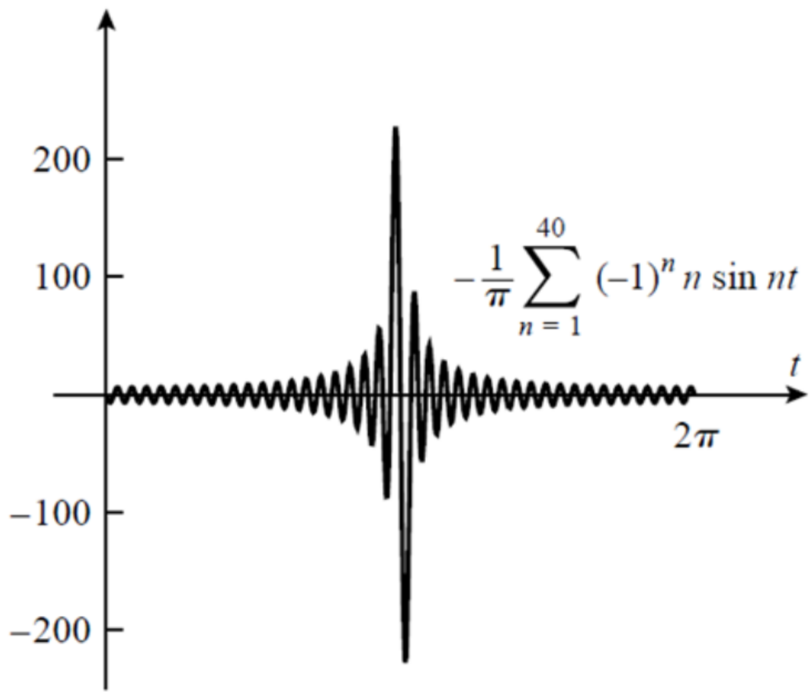
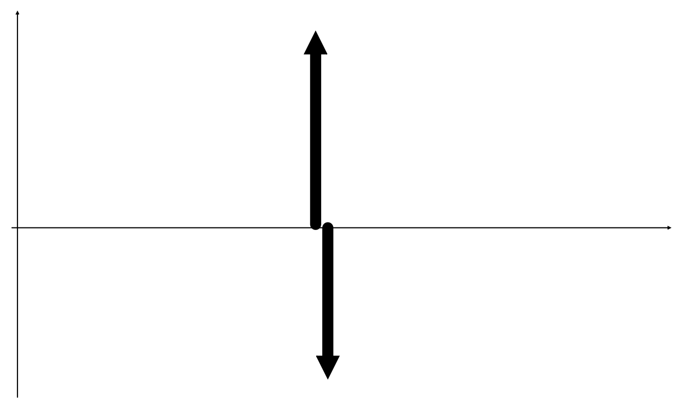


Figure 3 *Truncated Fourier series for* $\delta'(t-\pi)$



For physics and engineering

Finite Fourier Transform

$$\widehat{y(f)} \to \int_0^1 y(t)\, e^{-j2\pi f t} dt$$

Fourier Series *for all t in* [0, 1)

$$y(t) = \sum_{f=-\infty}^{\infty} \widehat{y(f)}\, e^{j2\pi f t}$$

For signal processing and random processes, switch the roles of *f* and *t*

$$\tilde{X}(f) = \sum_{n=-\infty}^{\infty} X(n)\, e^{-j2\pi n f}, \; -\frac{1}{2} < f < \frac{1}{2}$$

$$X(n) = \int_{-1/2}^{1/2} \tilde{X}(f)\, e^{j2\pi n f}\, df$$

(next slide)

For physics and engineering

Finite Fourier Transform

$$\widehat{y(f)} \rightarrow \int_0^1 y(t)\, e^{-j2\pi f t} dt$$

Fourier Series *for all t in* [0, 1)

$$y(t) \doteq \sum_{f=-\infty}^{\infty} \widehat{y(f)}\, e^{j2\pi f t}$$

For signal processing and random processes, switch the roles of *f* and ***t***

Times are discrete, *t* = *n*, (-∞ to ∞)

$$y(t) \doteq \sum_{f=-\infty}^{\infty} \widehat{y(f)}\, e^{j2\pi f t}$$

$$\tilde{X}(f) = \sum_{n=-\infty}^{\infty} X(n) e^{-j2\pi n f}, \; -\frac{1}{2} < f < \frac{1}{2}$$

The old *t* is reinterpreted as frequency and shifted from (0,1) to (-1/2, 1/2)

$$X(n) = \int_{-1/2}^{1/2} \tilde{X}(f) e^{j2\pi n f}\, df$$

For physics and engineering

Finite Fourier Transform

$$\widehat{y(f)} \to \int_0^1 y(t)\, e^{-j2\pi f t} dt$$

Fourier Series *for all t in* [0, 1)

$$y(t) = \sum_{f=-\infty}^{\infty} \widehat{y(f)}\, e^{j2\pi f t}$$

For signal processing and random processes, switch the roles of *f* and ***t***

Times are discrete, *t* = *n*, (-∞ to ∞)

$$y(t) = \sum_{f=-\infty}^{\infty} \widehat{y(f)}\, e^{j2\pi f t}$$

$$\tilde{X}(f) = \sum_{n=-\infty}^{\infty} X(n) e^{-j2\pi n f}, \quad -\frac{1}{2} < f < \frac{1}{2}$$

The old *t* is reinterpreted as frequency and shifted from (0,1) to (-1/2, 1/2)

$$X(n) = \int_{-1/2}^{1/2} \tilde{X}(f) e^{j2\pi n f}\, df$$

Discrete Time Fourier Transform

$$\tilde{X}(f) = \sum_{n=-\infty}^{\infty} X(n)\, e^{-j2\pi nf}, \quad -\frac{1}{2} < f < \frac{1}{2}$$

$$X(n) = \int_{-1/2}^{1/2} \tilde{X}(f)\, e^{j2\pi nf}\, df$$

(But for computer simulations,
n only runs from 1 to 2^N and
f is discretized and rescaled to
run from 1 to 2^N and
you use the Fast Fourier Transform.)

Discrete Time Fourier Transform

$$\tilde{X}(f) = \sum_{n=-\infty}^{\infty} X(n)\, e^{-j2\pi nf}, \quad -\frac{1}{2} < f < \frac{1}{2}$$

$$X(n) = \int_{-1/2}^{1/2} \tilde{X}(f)\, e^{j2\pi nf}\, df$$

The DTFT of time-reversed *X*(-*n*) is the complex conjugate of the DTFT of *X*(*n*):

$$\tilde{X}(f) = \sum_{n=-\infty}^{\infty} X(n)\, e^{-j2\pi nf}, \quad -\frac{1}{2} < f < \frac{1}{2}$$

$$\tilde{X}(f) = \sum_{n=-\infty}^{\infty} X(n)\, e^{-j2\pi nf}, \quad -\frac{1}{2} < f < \frac{1}{2}$$

Discrete Time Fourier Transform

$$\tilde{X}(f) = \sum_{n=-\infty}^{\infty} X(n)\, e^{-j2\pi nf}, \quad -\frac{1}{2} < f < \frac{1}{2}$$

$$X(n) = \int_{-1/2}^{1/2} \tilde{X}(f)\, e^{j2\pi nf}\, df$$

The DTFT of time-reversed *X*(-*n*) is the complex conjugate of the DTFT of *X*(*n*):

$$\tilde{X}(f) = \sum_{n=-\infty}^{\infty} X(-n)\, e^{-j2\pi nf}, \quad -\frac{1}{2} < f < \frac{1}{2}$$

$$\tilde{X}(f) = \sum_{n=-\infty}^{\infty} X(n)\, e^{-j2\pi nf}, \quad -\frac{1}{2} < f < \frac{1}{2}$$

$$\tilde{X}(f) = \sum_{n=-\infty}^{\infty} X(n)\, e^{-j2\pi nf}\ ,\ -\frac{1}{2} < f < \frac{1}{2}$$

$$X(n) = \int_{-1/2}^{1/2} \tilde{X}(f)\, e^{j2\pi nf}\, df$$

Parseval's Theorem

$$\sum_{n=-\infty}^{\infty} X(n)^2 = \int_{-1/2}^{1/2} |\tilde{X}(f)|^2\, df$$

ENERGY

$$\frac{1}{2} m v^2 \qquad \frac{1}{2} k x^2$$

$$\iiint \frac{1}{2} \rho v^2\, dx\, dy\, dz$$

$$\frac{1}{2} \iiint \{\varepsilon\, |\mathbf{E}|^2 + \mu\, |\mathbf{H}|^2\}\, dx\, dy\, dz$$

$$X(n) = \int_{-1/2}^{1/2} \tilde{X}(f) e^{j2\pi nf} \, df$$

Parseval's Theorem

$$\sum_{n=-\infty}^{\infty} X(n)^2 = \int_{-1/2}^{1/2} |\tilde{X}(f)|^2 \, df$$

Energy Spectral Density

$$X(n) = \int_{-1/2}^{1/2} \tilde{X}(f) e^{j2\pi nf} \, df$$

Parseval's Theorem

$$\sum_{n=-\infty}^{\infty} X(n)^2 = \int_{-1/2}^{1/2} |\tilde{X}(f)|^2 \, df$$

What if *X*(*n*) is a stationary random process?

(mean zero)

$$E\{\sum_{n=-\infty}^{\infty} X(n)^2\} = \sum_{n=-\infty}^{\infty} R_X(0) = \sum_{n=-\infty}^{\infty} \sigma^2 = \infty$$

Lecture 11

Feb. 22, 2017

$$\hat{R}_X(m) = \frac{1}{\text{number of terms}} \sum_n X(n)X(n-m)$$

Takes too long

---

looks like convolution:

$$X \circ Y(m) \equiv \sum_{n=-\infty}^{\infty} X(n)Y(m-n)$$

If $\tilde{X}(f) = \sum_{n=-\infty}^{\infty} X(n)e^{-j2\pi nf}$ and $\tilde{Y}(f) = \sum_{n=-\infty}^{\infty} Y(n)e^{-j2\pi nf}$ then

$$\tilde{X}(f)\tilde{Y}(f) = \sum_{m=-\infty}^{\infty} Z(m)e^{-i2\pi mf} \quad \text{where}$$

$$Z(m) = \sum_{n=-\infty}^{\infty} X(n)Y(m-n) \ .$$

Fourier Convolution Theorem

In short, $\tilde{X}\tilde{Y} = \widetilde{X \circ Y}$.

$$\hat{R}_X(m) = \frac{1}{\text{number of terms}} \sum_n X(n)X(n-m)$$

Takes too long

looks like convolution:

$$X \circ Y(m) \equiv \sum_{n=-\infty}^{\infty} X(n)Y(m-n)$$

If $\tilde{X}(f) = \sum_{n=-\infty}^{\infty} X(n)e^{-j2\pi nf}$ and $\tilde{Y}(f) = \sum_{n=-\infty}^{\infty} Y(n)e^{-j2\pi nf}$ then

$$\tilde{X}(f)\tilde{Y}(f) = \sum_{m=-\infty}^{\infty} Z(m)e^{-i2\pi mf}$$

Fourier Convolution Theorem

$$Z(m) = \sum_{n=-\infty}^{\infty} X(n)Y(m-n) \ .$$

The DTFT of time-reversed *X*(-*n*) is the complex conjugate of the DTFT of *X*(*n*):

So..., to find

$$\sum_n X(n)X(n-m)$$

you FFT the data $X(n) \Rightarrow \tilde{X}(f)$

take its conjugate (FFT of time-reversed) ,

multiply them, and take inverse FFT.

So..., to find

$$\sum_n X(n)X(n-m)$$

you FFT the data $X(n) \Rightarrow \tilde{X}(f)$

take its conjugate (FFT of time-reversed) ,

multiply them, and take inverse FFT.

Parseval's Theorem

$$\sum_{n=-\infty}^{\infty} X(n)^2 = \int_{-1/2}^{1/2} |\tilde{X}(f)|^2 \, df$$

What if $X(n)$ is a stationary random process?

(mean zero)

$$E\{\sum_{n=-\infty}^{\infty} X(n)^2\} = \sum_{n=-\infty}^{\infty} R_X(0) = \sum_{n=-\infty}^{\infty} \sigma^2 = \infty$$

Wiener-Khintchine Theory

Replace

$$E\{\sum_{n=-\infty}^{\infty} X(n)^2\} = \lim_{N\to\infty} E\{\sum_{n=-N}^{N-1} X(n)^2\}$$

by

$$\lim_{N\to\infty} E\{\frac{1}{2N}\sum_{n=-N}^{N-1} X(n)^2\}$$

$$= \lim_{N\to\infty} \frac{1}{2N}\sum_{n=-N}^{N-1} R_X(0)$$

$$= \frac{2N}{2N} R_X(0) = R_X(0)$$

$$= E\{X(n)^2\} = (\mu_X^2 +)\sigma_X^2$$

$2N$ is the elapsed time,
energy/time = POWER

The Scheme

Truncate

$X_N(n) = \{.., 0, 0, 0, 0, X(-N), X(-N+1), \ldots, X(N-1), 0, 0, \ldots\}$

Take the DTFT

$$\tilde{X}_N(f) = \sum_{n=--\infty}^{\infty} X_N(n) e^{-j2\pi nf} = \sum_{n=-N}^{N-1} X(n) e^{-j2\pi nf}$$

Square the magnitude

$$|\tilde{X}_N(f)|^2 = \tilde{X}_N(f)\overline{\tilde{X}_N(f)}$$

$$= \{\sum_{n=-N}^{N-1} X(n) e^{-j2\pi nf}\}\{\sum_{p=-N}^{N-1} X(p) e^{+j2\pi pf}\}$$

Divide by 2N

Example: Suppose N=2 .

$$|\tilde{X}_N(f)|^2 = \{\sum_{n=-N}^{N-1} X(n)e^{-j2\pi nf}\}\{\sum_{p=-N}^{N-1} X(p)e^{+j2\pi pf}\}$$

$$\{X(-2)e^{+j2\pi 2f} + X(-1)e^{+j2\pi 1f} + X(0)e^{0} + X(1)e^{-j2\pi 1f}\}$$

$$\{X(-2)e^{-j2\pi 2f} + X(-1)e^{-j2\pi 1f} + X(0)e^{0} + X(1)e^{+j2\pi 1f}\}$$

$$\{X(-2)^2 + X(-1)^2 + X(0)^2 + X(1)^2\}$$

$$+\{X(-2)X(-1) + X(-1)X(0) + X(0)X(1)\}e^{-j2\pi 1f}$$

$$+\{X(-2)X(0) + X(-1)X(1)\}e^{-j2\pi 2f}$$

$$+\{X(-2)X(1)\}e^{-j2\pi 3f} \qquad + \{X(-2)X(1)\}e^{+j2\pi 3f}$$

$$+\{X(-2)X(0) + X(-1)X(1)\}e^{+j2\pi 2f}$$

$$+\{X(-2)X(-1) + X(-1)X(0) + X(0)X(1)\}e^{+j2\pi 1f}$$

$X(-3)\quad X(-2)\quad X(-1)\quad X(0)\quad X(1)\quad X(2)\quad X(3)$

$X(-3)\quad X(-2)\quad X(-1)\quad X(0)\quad X(1)\quad X(2)\quad X(3)$

Now take expected values. N=2

$$|\tilde{X}_N(f)|^2 = \{\sum_{n=-N}^{N-1} X(n)e^{-j2\pi nf}\}\{\sum_{p=-N}^{N-1} X(p)e^{+j2\pi pf}\}$$

$$\{X(-2)e^{+j2\pi 2f} + X(-1)e^{+j2\pi 1f} + X(0)e^{0} + X(1)e^{-j2\pi 1f}\}$$

$$\{X(-2)e^{-j2\pi 2f} + X(-1)e^{-j2\pi 1f} + X(0)e^{0} + X(1)e^{+j2\pi 1f}\}$$

$R_X(0)$ $\{X(-2)^2 + X(-1)^2 + X(0)^2 + X(1)^2\}$

$R_X(1)$ $+\{X(-2)X(-1) + X(-1)X(0) + X(0)X(1)\}e^{-j2\pi 1f}$

$R_X(2)$ $+\{X(-2)X(0) + X(-1)X(1)\}e^{-j2\pi 2f}$

$R_X(3)$ $+\{X(-2)X(1)\}e^{-j2\pi 3f}$ $+\{X(-2)X(1)\}e^{+j2\pi 3f}$ $R_X(3)$

$+\{X(-2)X(0) + X(-1)X(1)\}e^{+j2\pi 2f}$ $R_X(2)$

$+\{X(-2)X(-1) + X(-1)X(0) + X(0)X(1)\}e^{+j2\pi 1f}$ $R_X(1)$

$$4R_X(0) + 3R_X(1)e^{-j2\pi 1f} + 3R_X(1)e^{+j2\pi 1f}$$

$$+2R_X(2)e^{-j2\pi 2f} + 2R_X(2)e^{+j2\pi 2f}$$

$$+1R_X(3)e^{-j2\pi 3f} + 1R_X(3)e^{+j2\pi 3f}$$

Divide by $2N$. N=2

$$4R_X(0) + 3R_X(1)e^{-j2\pi 1f} + 3R_X(1)e^{+j2\pi 1f}$$

$$+2R_X(2)e^{-j2\pi 2f} + 2R_X(2)e^{+j2\pi 2f}$$

$$+1R_X(3)e^{-j2\pi 3f} + 1R_X(3)e^{+j2\pi 3f}$$

$$R_X(0) + \frac{3}{4}R_X(1)e^{-j2\pi 1f} + \frac{3}{4}R_X(1)e^{+j2\pi 1f}$$

$$+\frac{2}{4}R_X(2)e^{-j2\pi 2f} + \frac{2}{4}R_X(2)e^{+j2\pi 2f}$$

$$+\frac{1}{4}R_X(3)e^{-j2\pi 3f} + \frac{1}{4}R_X(3)e^{+j2\pi 3f}$$

If you had used 4000 points,

4000

$$4R_X(0) + 3R_X(1)e^{-j2\pi 1f} + 3R_X(1)e^{+j2\pi 1f}$$

$$+2R_X(2)e^{-j2\pi 2f} + 2R_X(2)e^{+j2\pi 2f}$$

$$+1R_X(3)e^{-j2\pi 3f} + 1R_X(3)e^{+j2\pi 3f}$$

would be

$$4R_X(0) + 3R_X(1)e^{-j2\pi 1f} + 3R_X(1)e^{+j2\pi 1f}$$

4000 3999 3999

$$+2R_X(2)e^{-j2\pi 2f} + 2R_X(2)e^{+j2\pi 2f}$$

3998 3998

+ . . .

$$+1R_X(3)e^{-j2\pi 3f} + 1R_X(3)e^{+j2\pi 3f}$$

3999 3999

$$R_X(0) + \frac{3}{4}R_X(1)e^{-j2\pi 1f} + \frac{3}{4}R_X(1)e^{+j2\pi 1f}$$

3999/4000

$$+\frac{2}{4}R_X(2)e^{-j2\pi 2f} + \frac{2}{4}R_X(2)e^{+j2\pi 2f}$$

3998/4000

3999

$$+\frac{1}{4}R_X(3)e^{-j2\pi 3f} + \frac{1}{4}R_X(3)e^{+j2\pi 3f}$$

1/4000

$$R_X(0) + \frac{3}{4}R_X(1)e^{-j2\pi 1f} + \frac{3}{4}R_X(1)e^{+j2\pi 1f}$$

$$+\frac{2}{4}R_X(2)e^{-j2\pi 2f} + \frac{2}{4}R_X(2)e^{+j2\pi 2f}$$

$$+\frac{1}{4}R_X(3)e^{-j2\pi 3f} + \frac{1}{4}R_X(3)e^{+j2\pi 3f}$$

if $N \approx \infty$ and you were very forgiving

$$R_X(0) + R_X(1)e^{-j2\pi 1f} + R_X(1)e^{+j2\pi 1f}$$

$$+R_X(2)e^{-j2\pi 2f} + R_X(2)e^{+j2\pi 2f}$$

$$+R_X(3)e^{-j2\pi 3f} + R_X(3)e^{+j2\pi 3f}$$

Wiener-Khintchine justified this

To sum up:
Take the DTFT of your (finite) data set:

$$\tilde{X}_N(f) = \sum_{n=-\infty}^{\infty} X_N(n) e^{-j2\pi nf} = \sum_{n=-N}^{N-1} X(n) e^{-j2\pi nf}$$

Square the magnitude

$$|\tilde{X}_N(f)|^2 = \tilde{X}_N(f)\overline{\tilde{X}_N(f)}$$

Divide by $2N$

And you'll have a time-average estimate of the DTFT of the autocorrelation $R_X(n)$ - $|\tilde{X}^2(f)|$

Take the inverse DTFT and you'll have R_X !!!!!

$$R_X(0) + R_X(1)e^{-j2\pi 1f} + R_X(1)e^{+j2\pi 1f}$$

$$+R_X(2)e^{-j2\pi 2f} + R_X(2)e^{+j2\pi 2f}$$

$$+R_X(3)e^{-j2\pi 3f} + R_X(3)e^{+j2\pi 3f}$$

To sum up:
Take the DTFT of your (finite) data set:

$$\tilde{X}_N(f) = \cancel{\sum_{n=-\infty}^{\infty} \tilde{X}_N(n) e^{-j2\pi nf}} = \sum_{n=-N}^{N-1} X(n) e^{-j2\pi nf}$$

$$\hat{y} = My\,, \quad \begin{bmatrix} \widehat{y_1} \\ \widehat{y_2} \\ \widehat{y_3} \\ \vdots \\ \widehat{y_{2N}} \end{bmatrix} = \begin{bmatrix} \square & \square & \square \\ \square & M & \square \\ \square & \square & \square \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_{2N} \end{bmatrix}$$

Square the magnitude and divide by $2N$

$$\hat{y} \leftarrow |\hat{y}|^2/2N$$

And you'll have a time-average estimate of the DTFT of the autocorrelation $R_X(n)$.

Take the inverse DTFT and you'll have R_X !!!!!

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_{2N} \end{bmatrix} = \begin{bmatrix} \square & \square & \square \\ \square & M^{-1} & \square \\ \square & \square & \square \end{bmatrix} \begin{bmatrix} \widehat{y_1} \\ \widehat{y_2} \\ \widehat{y_3} \\ \vdots \\ \widehat{y_{2N}} \end{bmatrix}$$

To sum up:
Take the DTFT of your (finite) data set:

Square the magnitude

Divide by $2N$

And you'll have a time-average estimate of the DTFT of the autocorrelation $R_X(n)$.

Take the inverse DTFT and you'll have R_X !!!!!

The estimates will be better for the lower values of $R_X(n)$.

WHY?

$$|\tilde{X}_N(f)|^2 = \{\sum_{n=-N}^{N-1} X(n)e^{-j2\pi nf}\}\{\sum_{p=-N}^{N-1} X(p)e^{+j2\pi pf}\}$$

You have 4 estimates of $R_X(0)$, 3 of $R_X(1)$, 2 of .....

$$\{X(-2)^2 + X(-1)^2 + X(0)^2 + X(1)^2\}$$

$$+\{X(-2)X(-1) + X(-1)X(0) + X(0)X(1)\}e^{-j2\pi 1f}$$

$$+\{X(-2)X(0) + X(-1)X(1)\}e^{-j2\pi 2f}$$

$$+\{X(-2)X(1)\}e^{-j2\pi 3f} \qquad + \{X(-2)X(1)\}e^{+j2\pi 3f}$$

$$+\{X(-2)X(0) + X(-1)X(1)\}e^{+j2\pi 2f}$$

$$+\{X(-2)X(-1) + X(-1)X(0) + X(0)X(1)\}e^{+j2\pi 1f}$$

and you divide them all by 4.

In practice, you'll have 4096 estimates of $R_X(0)$, 4095 estimates of $R_X(1)$, ..., and 1 estimate of $R_X(4095)$.

And you'll divide them all by 4096.

(Two crimes)

To sum up:
Take the DTFT of your (finite) data set:

Square the magnitude

Divide by $2N$

And you'll have a time-average estimate of the DTFT of the autocorrelation $R_X(n)$.

Take the inverse DTFT and you'll have R_X !!!!!

The estimates will be better for the lower values of $R_X(n)$.

$$\lim_{N\to\infty} \frac{|\tilde{X}_N(f)|^2}{2N} = \sum_{m=-\infty}^{\infty} R_X(m)\, e^{-j2\pi mf}$$

$$= \text{DTFT of } R_X \equiv S_X(f)$$

Note that

$$R_X(0) = \int_{-1/2}^{1/2} S_X(f)\, df$$

so $S_X(f)$ is called the Power Spectral Density.

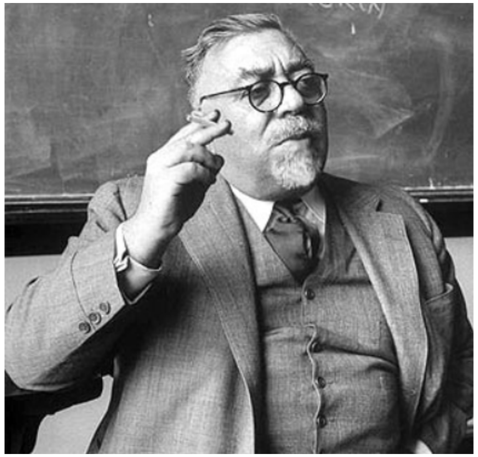
$$\lim_{N\to\infty} \frac{|\tilde{X}_N(f)|^2}{2N} = \sum_{m=-\infty}^{\infty} R_X(m)\, e^{-j2\pi mf}$$

$$= \text{DTFT of } R_X \equiv S_X(f)$$

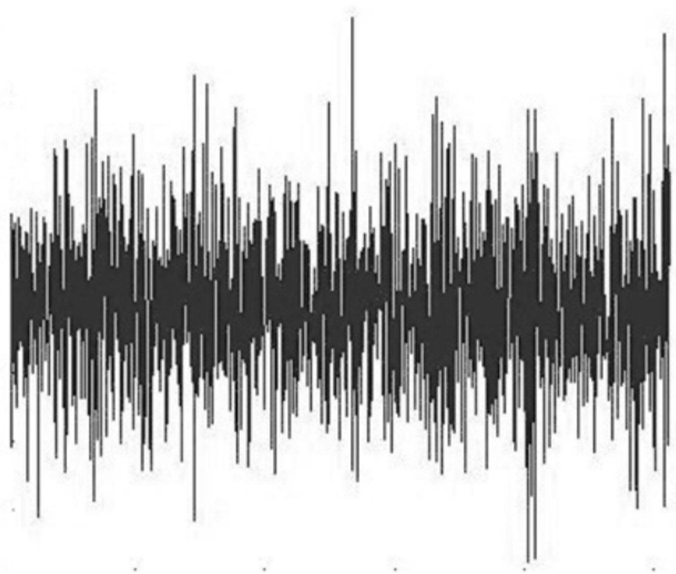
$$R_X(0) = \int_{-1/2}^{1/2} S_X(f)\, df$$

Wiener Khintchine Theorem (?)

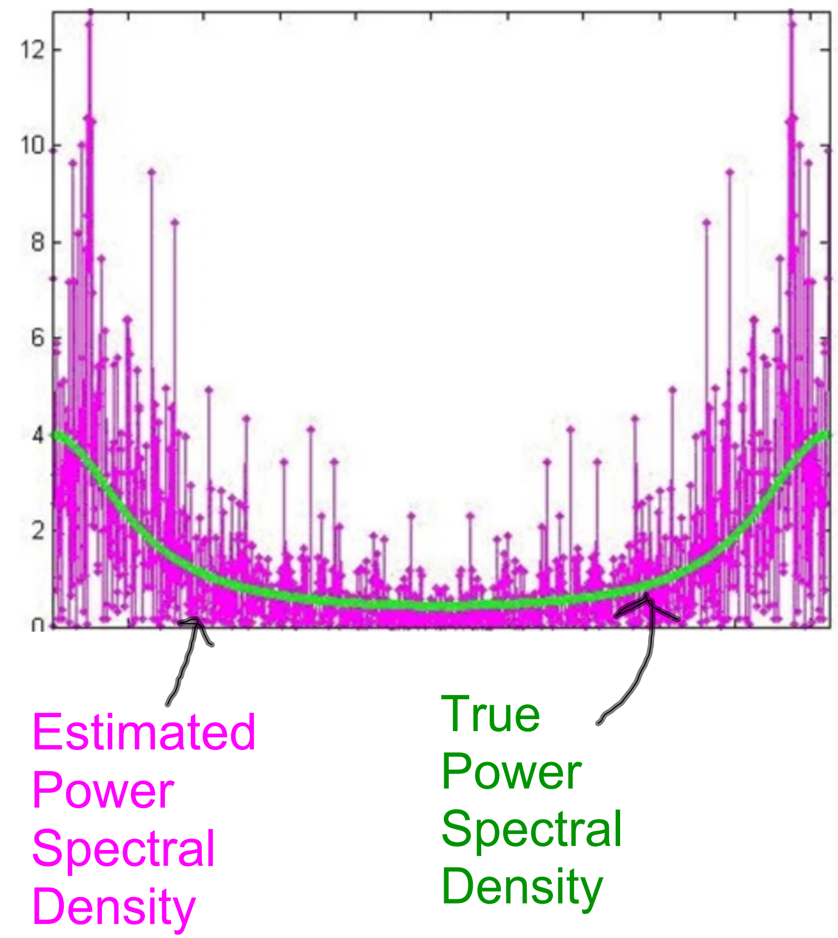
The Power Spectral Density is the DTFT of the autocorrelation



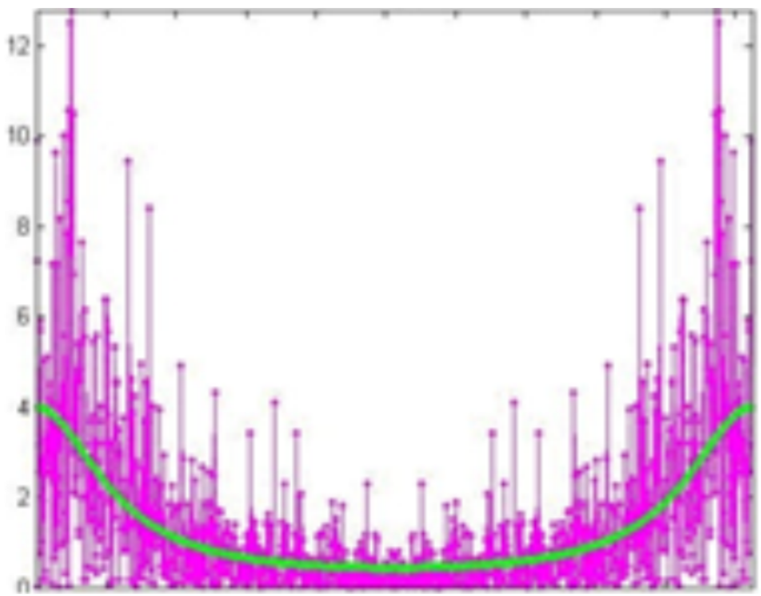
Norbert Wiener

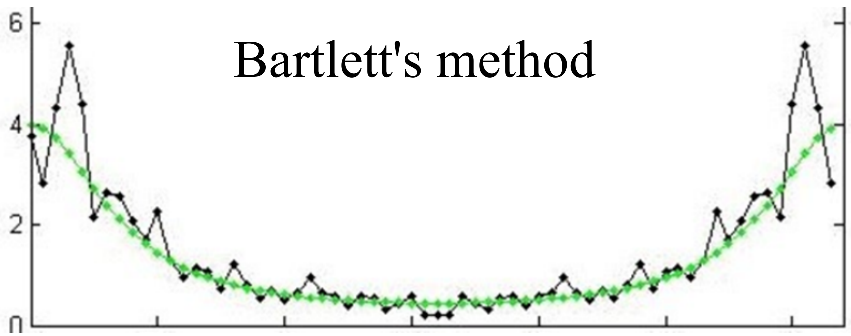


X(*n*)

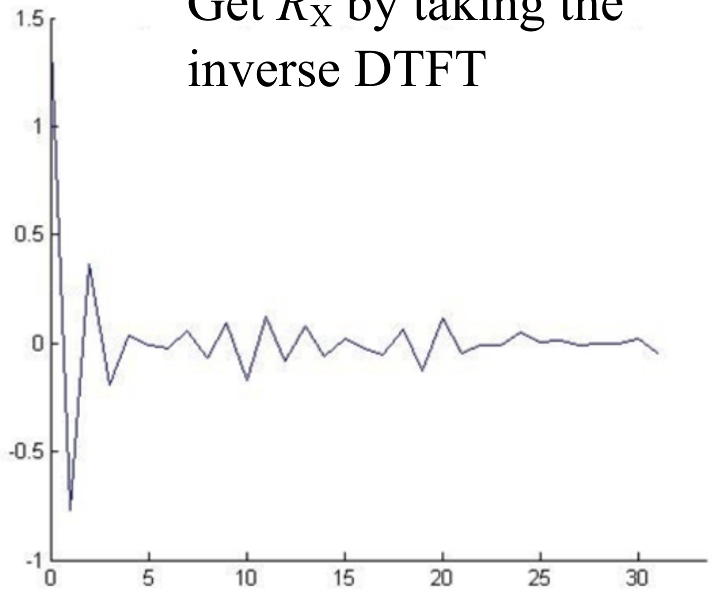


4096 *#^$#* data points!!!!!


12
10
8
6
4
2
0

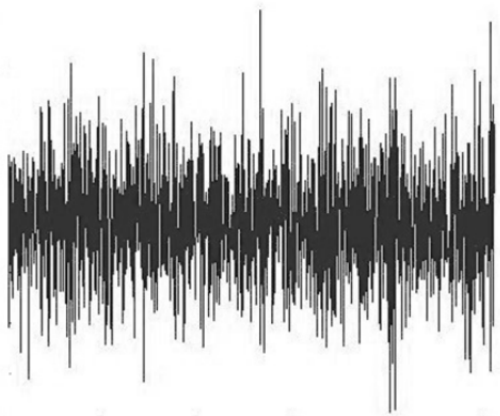

Bartlett's method
6
4
2
0

Get R_X by taking the inverse DTFT



Exact Autocorrelation estimates (RX(0) ≡ 4/3, RX(1) ≡ -2/3, RX(2) ≡
1/3, RX(3) ≡ -1/6)

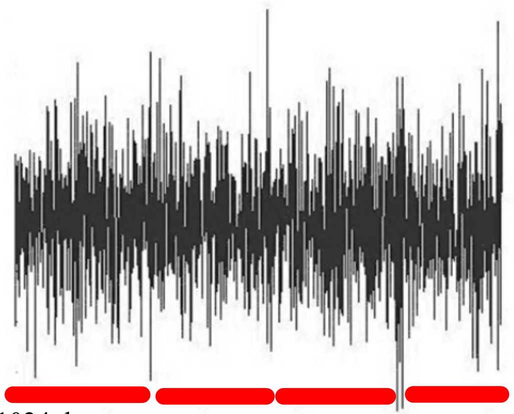
How did Bartlett do it?



4096 data

4096 data, yields estimates for $R_X(0)$, $R_X(1)$, ..., $R_X(4095)$

How did Bartlett do it?



4096

= 1024 x 4

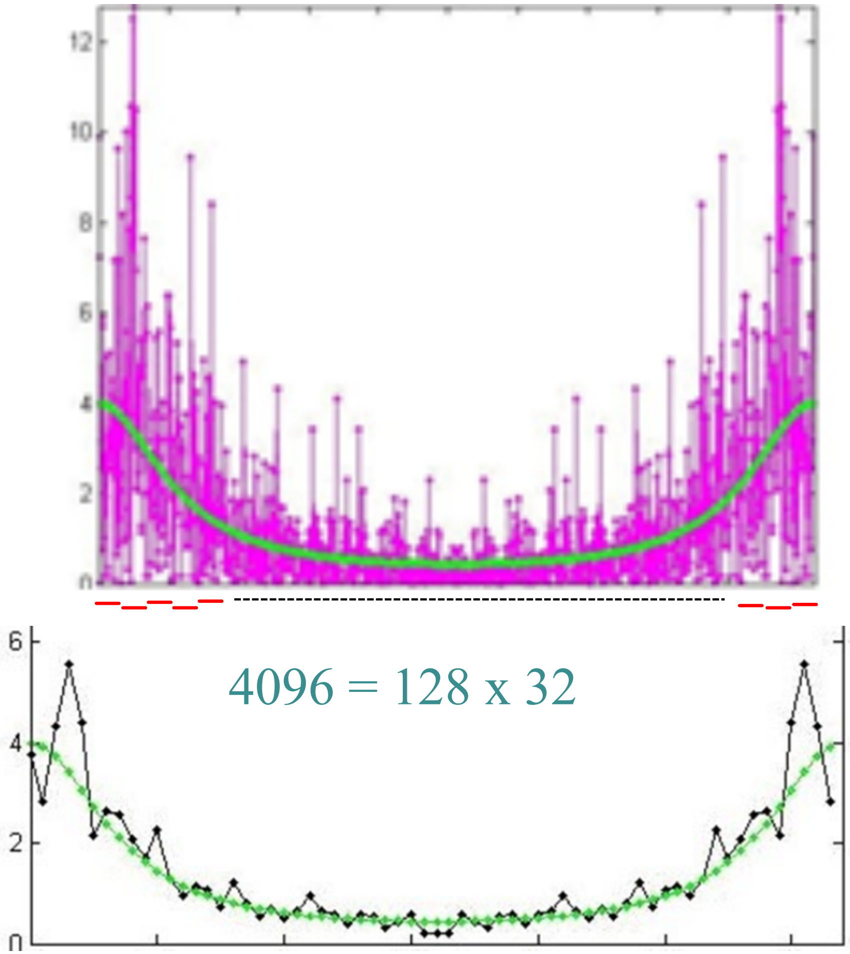
1024 data => $R_X(0), R_X(1), \ldots, R_X(1023)$

1024 data => $R_X(0), R_X(1), \ldots, R_X(1023)$

1024 data => $R_X(0), R_X(1), \ldots, R_X(1023)$

1024 data => $R_X(0), R_X(1), \ldots, R_X(1023)$

Now average them


12
10
8
6
4
2
0
6
4
2
0
4096 = 128 x 32

Lecture 12

Feb. 27, 2017

Wiener Khintchine Theorem (?)

The Power Spectral Density is the DTFT of the autocorrelation

$$\lim_{N\to\infty} \frac{|\tilde{X}_N(f)|^2}{2N} = \sum_{m=-\infty}^{\infty} R_X(m)\, e^{-j2\pi mf}$$

$$= \text{DTFT of } R_X \equiv S_X(f)$$

$$R_X(0) = \int_{-1/2}^{1/2} S_X(f)\, df$$

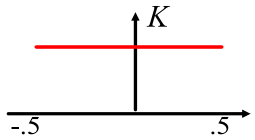
Wiener Khintchine Theorem (?)

The Power Spectral Density is the DTFT of the autocorrelation

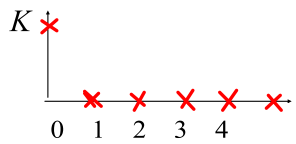
$$\lim_{N\to\infty}\frac{|\tilde{X}_N(f)|^2}{2N}=\sum_{m=-\infty}^{\infty}R_X(m)e^{-j2\pi mf}=\text{DTFT of } R_X\equiv S_X(f)$$

$$R_X(n)=\int_{-1/2}^{1/2}S_X(f)e^{j2\pi nf}\,df \qquad R_X(0)=\int_{-1/2}^{1/2}S_X(f)\,df$$

What does a ***broad band*** spectrum random process look like? Suppose $S_X(f) = K$



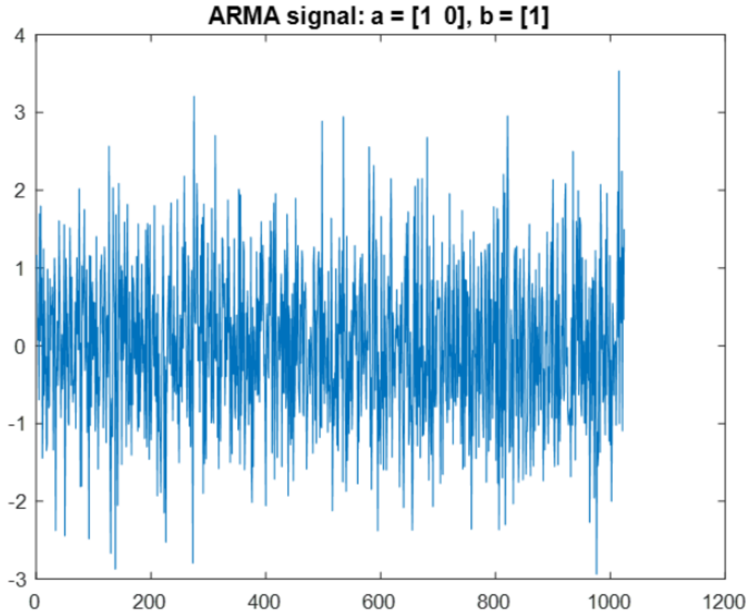
PSD S_X



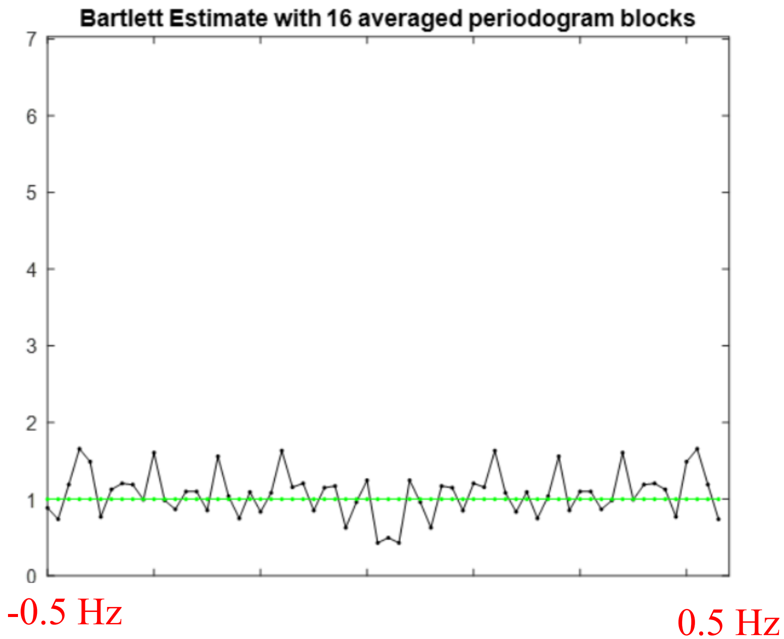
Autocorrelation R_X

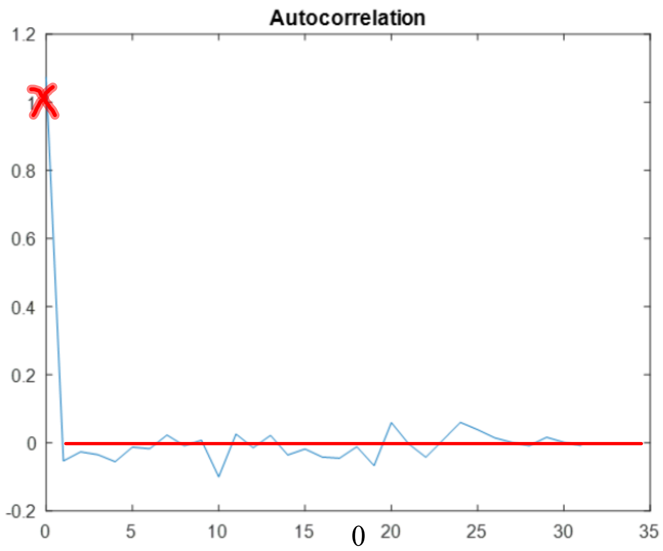
This is called ***White Noise***

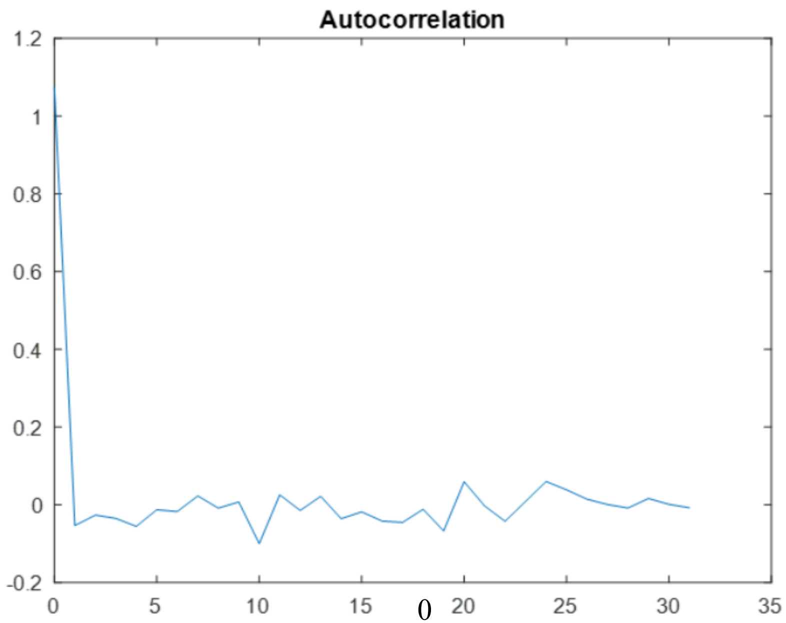
White Noise



PSD

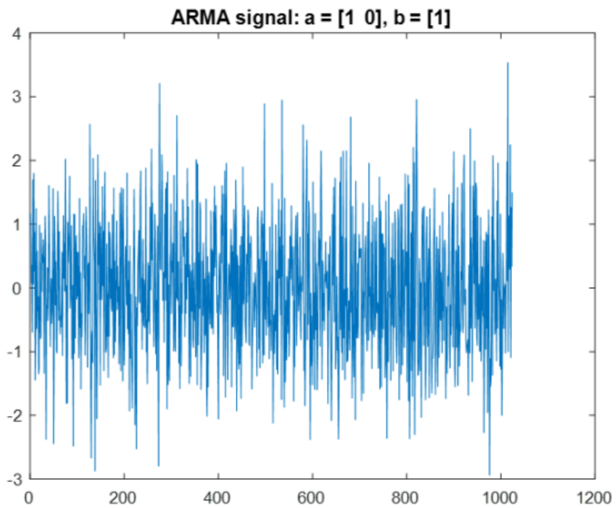

Bartlett Estimate with 16 averaged periodogram blocks
7
6
5
4
3
2
1
0
-0.5 Hz
0.5 Hz


Autocorrelation
1.2
0.8
0.6
0.4
0.2
0
-0.2
0
5
10
15
0
20
25
30
35



0

White Noise



Wiener Khintchine Theorem (?)

The Power Spectral Density is the DTFT of the autocorrelation

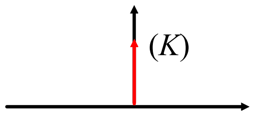
$$\lim_{N\to\infty}\frac{|\tilde{X}_N(f)|^2}{2N}=\sum_{m=-\infty}^{\infty}R_X(m)e^{-j2\pi mf}\qquad R_X(n)=\int_{-1/2}^{1/2}S_X(f)e^{j2\pi nf}\,df$$

$$=\text{DTFT of } R_X \equiv S_X(f)\qquad R_X(0)=\int_{-1/2}^{1/2}S_X(f)\,df$$

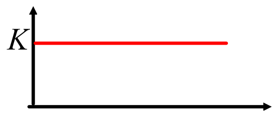
What does a ***narrow band*** spectrum random process look like? Suppose $S_X(f) = K\delta(f)$

$$R_X(n)=\int_{-1/2}^{1/2}S_X(f)e^{j2\pi nf}\,df$$

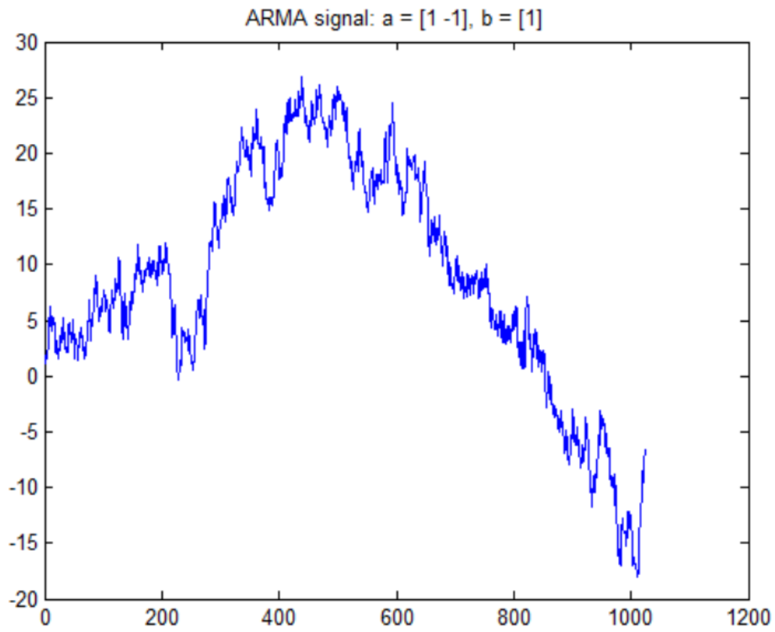
$$=K\int_{-1/2}^{1/2}[\delta(f)]e^{j2\pi nf}\,df=K$$

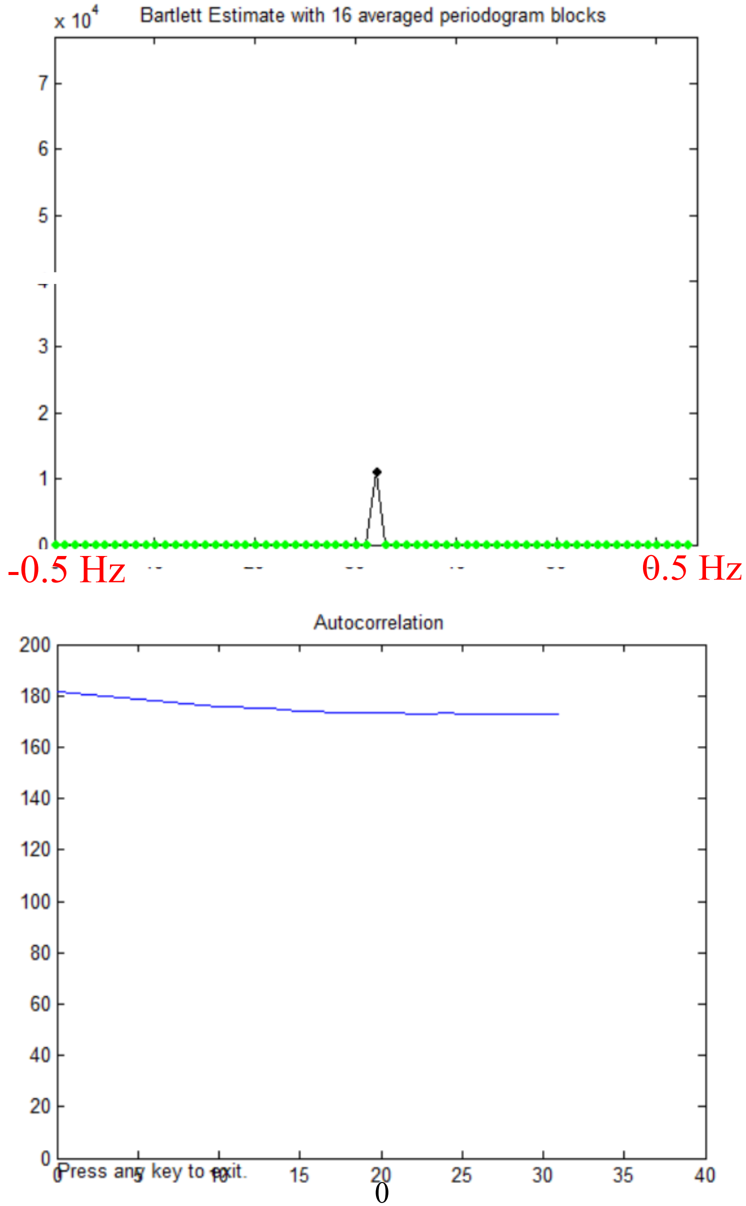


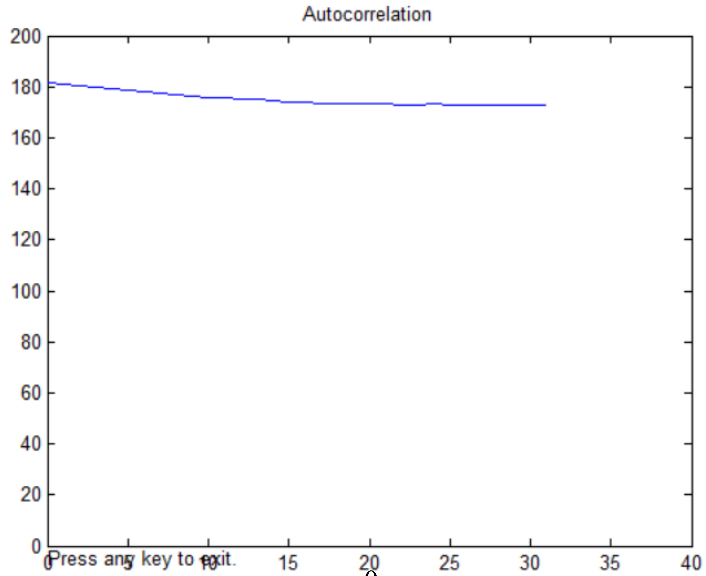
PSD S_X



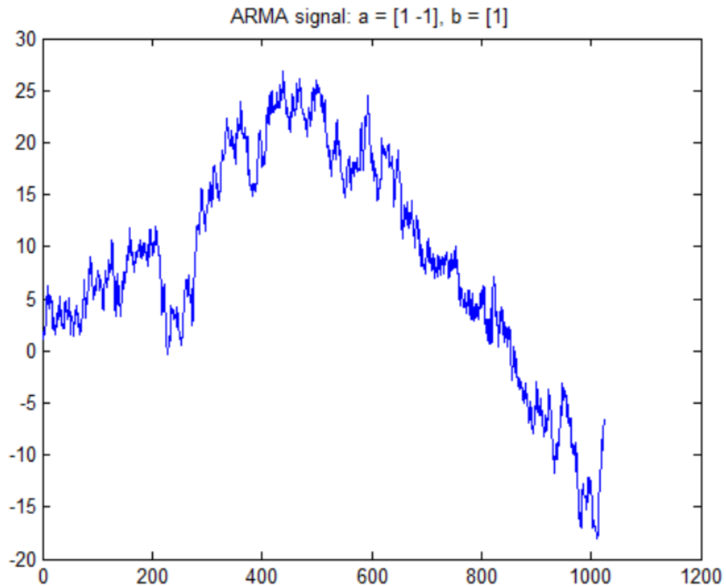
Autocorrelation R_X


ARMA signal: a = [1 -1], b = [1]
30
25
20
15
10
5
0
-5
-10
-15
-20
0
200
400
600
800
1000
1200


x 10^4
Bartlett Estimate with 16 averaged periodogram blocks
7
6
5
3
2
1
0
-0.5 Hz
0.5 Hz
Autocorrelation
200
180
160
140
120
100
80
60
40
20
0
5
10
15
20
25
30
35
40
Press any key to exit.



0



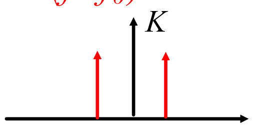
Wiener Khintchine Theorem (?)

The Power Spectral Density is the DTFT of the autocorrelation

$$\lim_{N\to\infty} \frac{|\tilde{X}_N(f)|^2}{2N} = \sum_{m=-\infty}^{\infty} R_X(m) e^{-j2\pi mf} \quad R_X(n) = \int_{-1/2}^{1/2} S_X(f) e^{j2\pi nf}\, df$$

$$= \text{DTFT of } R_X \equiv S_X(f) \qquad R_X(0) = \int_{-1/2}^{1/2} S_X(f)\, df$$

What does a *narrow band* spectrum random process look like? Suppose $S_X(f) = K\delta(f-f_0) + K\delta(f+f_0)$



PSD S_X

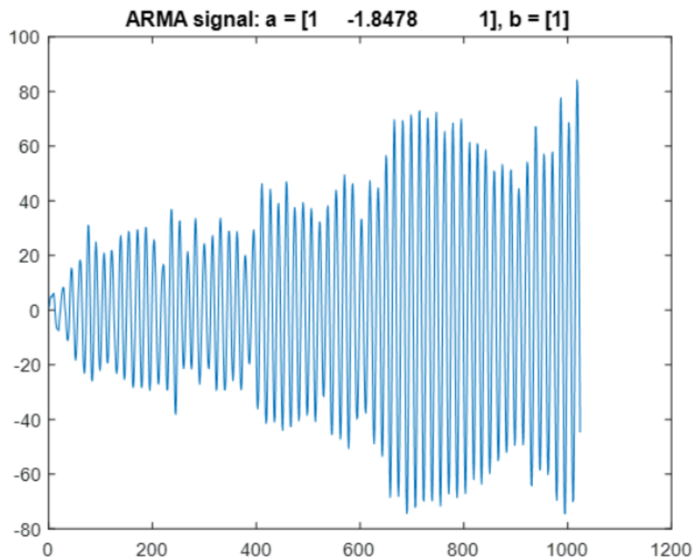
$$R_X(n) = \int_{-1/2}^{1/2} S_X(f) e^{j2\pi nf}\, df$$

$$= K \int_{-1/2}^{1/2} [\delta(f-f_o) + \delta(f+f_o)] e^{j2\pi nf}\, df$$

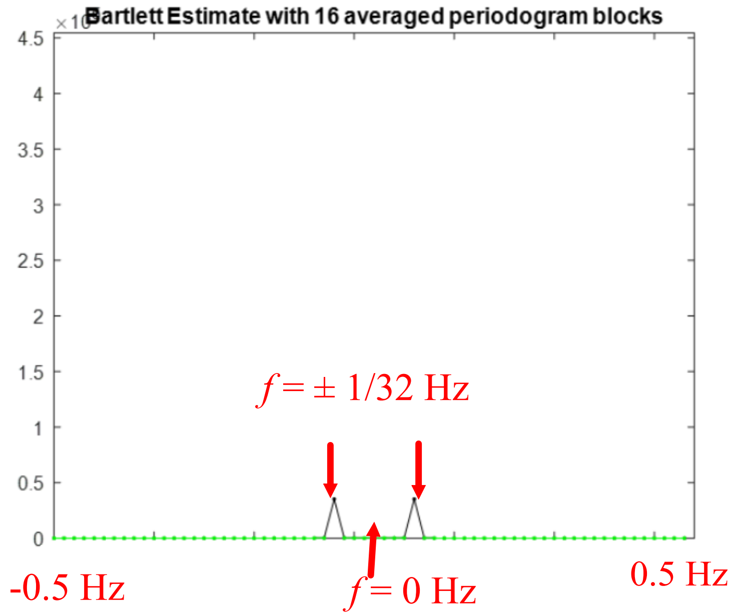
$$= K[e^{j2\pi nf_0} + e^{-j2\pi nf_0}] = 2K \cos 2\pi f_0 n$$

Autocorrelation R_X

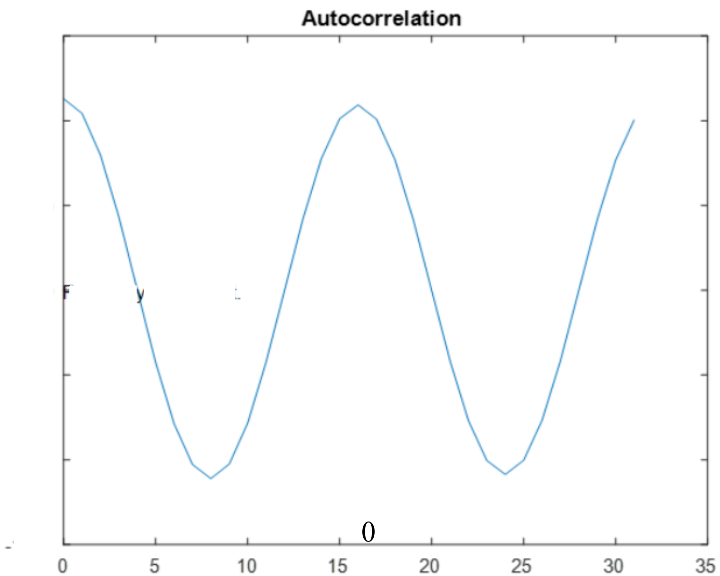
Narrowband spectrum



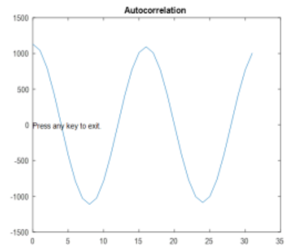
PSD



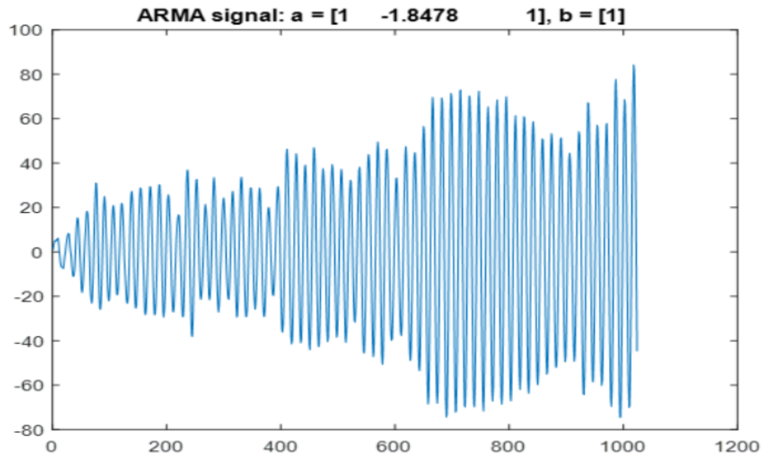
Bartlett Estimate with 16 averaged periodogram blocks
×10
4.5
4
3.5
3
2.5
2
1.5
1
0.5
0
f = ± 1/32 Hz
-0.5 Hz
f = 0 Hz
0.5 Hz



Autocorrelation
0
0
5
10
15
20
25
30
35



arrowband spectrum $a = \left[1 - 2 * \cos\left(\frac{2\pi}{16}\right)\right.$



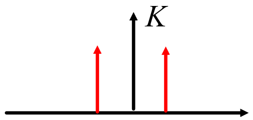
Wiener Khintchine Theorem (?)

The Power Spectral Density is the DTFT of the autocorrelation

$$\lim_{N\to\infty} \frac{|\tilde{X}_N(f)|^2}{2N} = \sum_{m=-\infty}^{\infty} R_X(m)\, e^{-j2\pi m f} \qquad R_X(n) = \int_{-1/2}^{1/2} S_X(f)\, e^{j2\pi n f}\, df$$

$$= \text{DTFT of } R_X \equiv S_X(f) \qquad R_X(0) = \int_{-1/2}^{1/2} S_X(f)\, df$$

What does a ***narrow band*** spectrum random process look like? Suppose $S_X(f) = K\delta(f-f_0) + K\delta(f+f_0)$



PSD S_X

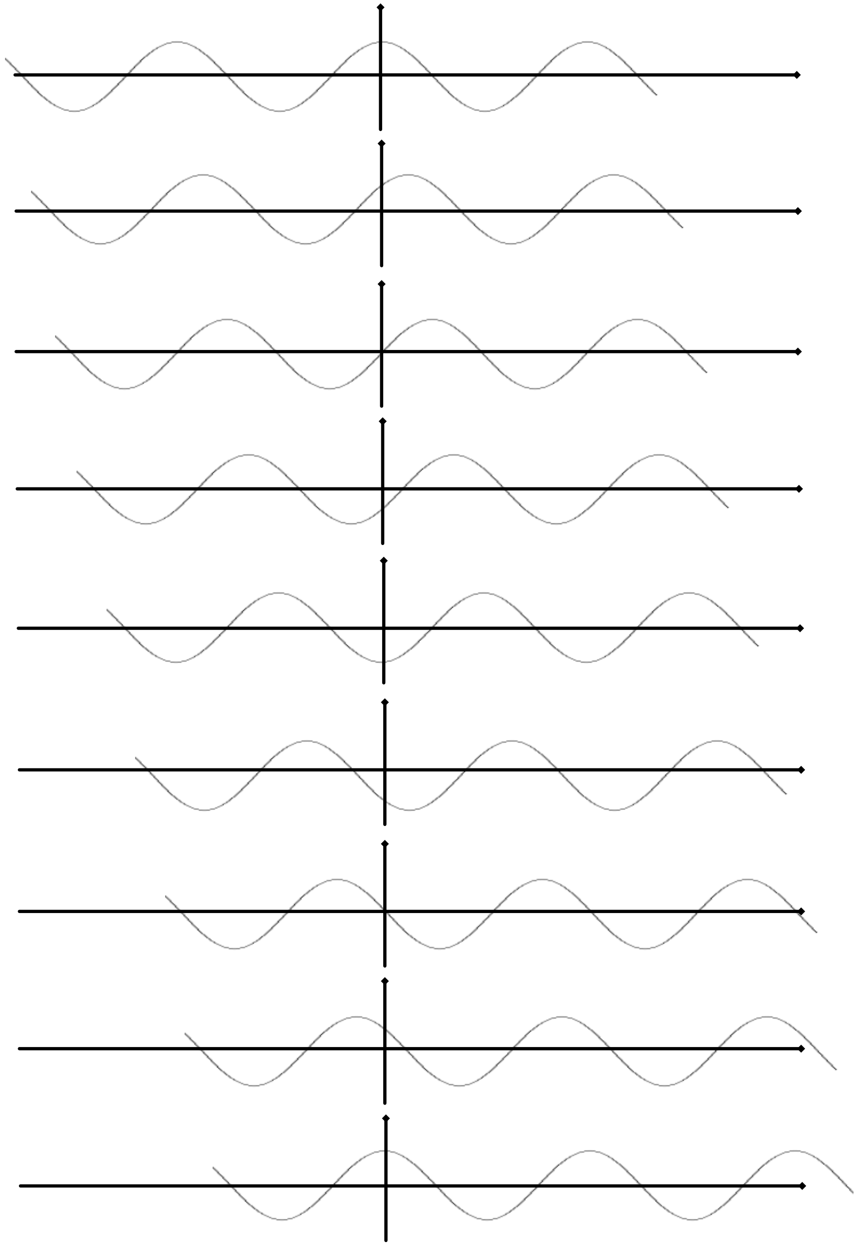
$$R_X(n) = \int_{-1/2}^{1/2} S_X(f)\, e^{j2\pi n f}\, df$$

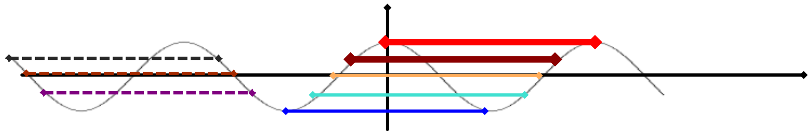
$$= K \int_{-1/2}^{1/2} [\delta(f - f_o) + \delta(f + f_o)] e^{j2\pi n f}\, df$$

$$= K[e^{j2\pi n f_0} + e^{-j2\pi n f_0}] = 2K \cos 2\pi f_0 n$$

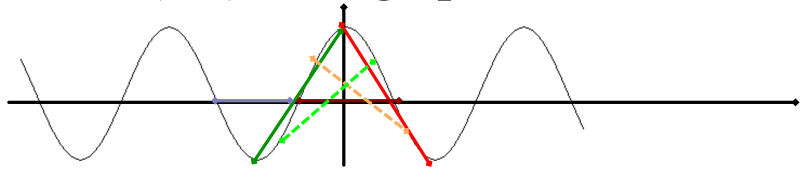
Autocorrelation R_X

Random-Phase Sine Wave

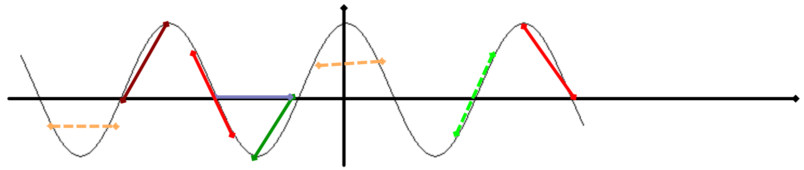




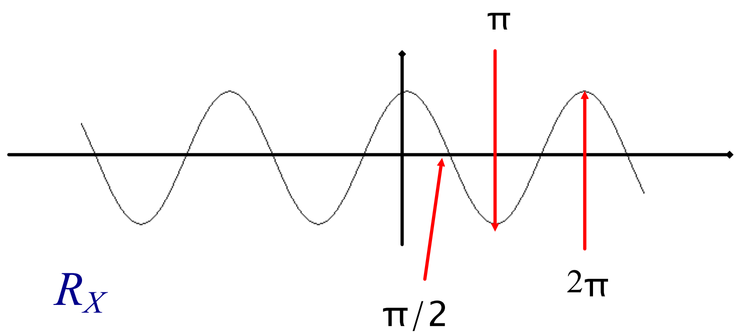
$R_X(2\pi)$ is large positive



$R_X(\pi)$ is large negative

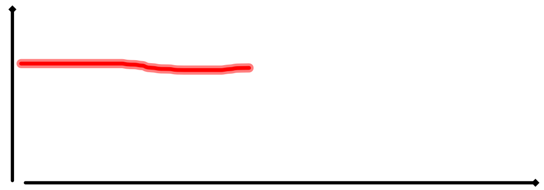


$R_X(\pi/2) = 0$

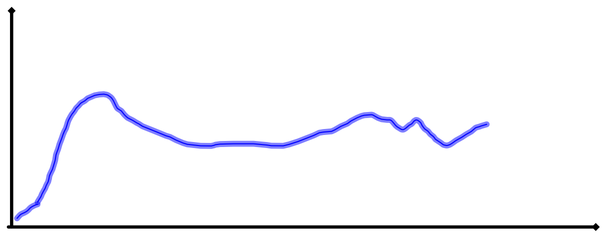


Spectral properties of linear time invariant systems

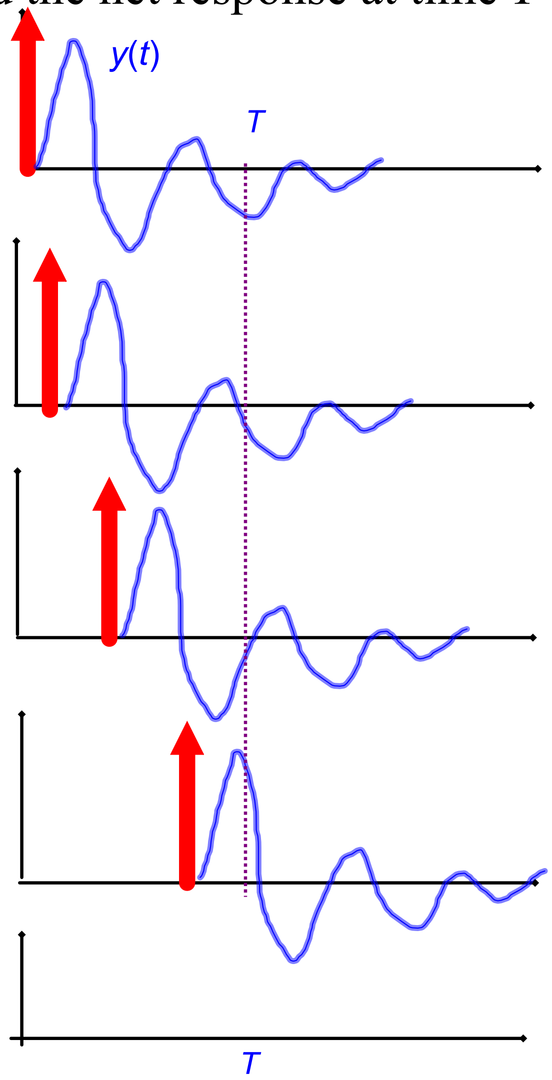
Force



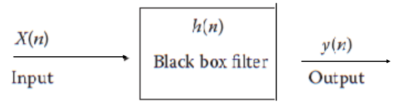
Response



and the net response at time T equals

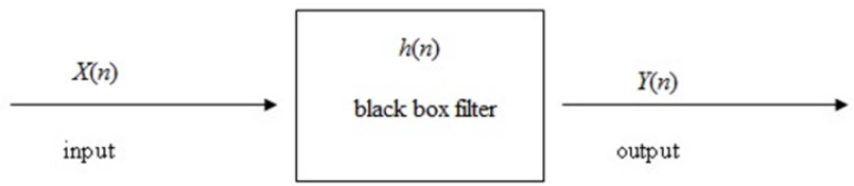


$F(0)y(T) + F(\tau)y(T-\tau) + F(2\tau)y(T-2\tau)$
$+ F(3\tau)y(T-3\tau) + \ldots$
(CONVOLUTION!)



				Input					Output
$\cdots$	0	1	0	$\cdots$	0	0	0	$\cdots$	$h(n)$
$\cdots$	0	0	0	$\cdots$	0	1	0	$\cdots$	$h(n-m)$
$\cdots$	0	0	0	$\cdots$	0	$X(m)$	0	$\cdots$	$X(m)h(n-m)$
$\cdots$	$X(-1)$	$X(0)$	$X(1)$	$\cdots$	$X(m-1)$	$X(m)$	$X(m+1)$	$\cdots$	$\sum_{m=-\infty}^{\infty} X(m)h(n-m)$

Figure 3.5 Linear time invariant system responses



$$Y(n) = \sum_{m=-\infty}^{\infty} X(m)h(n-m)$$

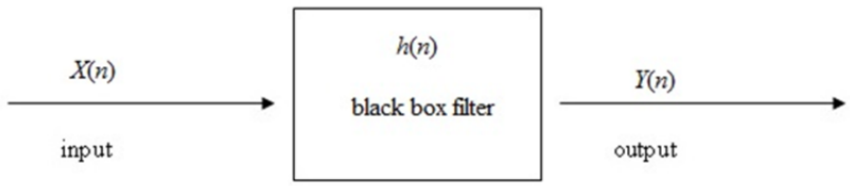
Impulse response

Transfer function

$$\hat{Y}(f) = \hat{H}(f)\hat{X}(f)$$

(Convolution Theorem)

Figure 3.5 Linear time invariant system responses



$$Y(n)=\sum_{m=-\infty}^{\infty} X(m)h(n-m)$$

Impulse response

Transfer function

$$\hat{Y}(f)=\hat{H}(f)\hat{X}(f)$$

$$X(n)=\int_{-1/2}^{1/2}\tilde{X}(f)e^{j2\pi nf}\,df$$

$$Y(n)=\int_{-1/2}^{1/2}\hat{Y}(f)e^{j2\pi nf}\,df=\int_{-1/2}^{1/2}\hat{H}(f)\hat{X}(f)e^{j2\pi nf}\,df$$

Some students had trouble following this discussion of transfer functions. The topic is revisited in lectures 18 and 20.

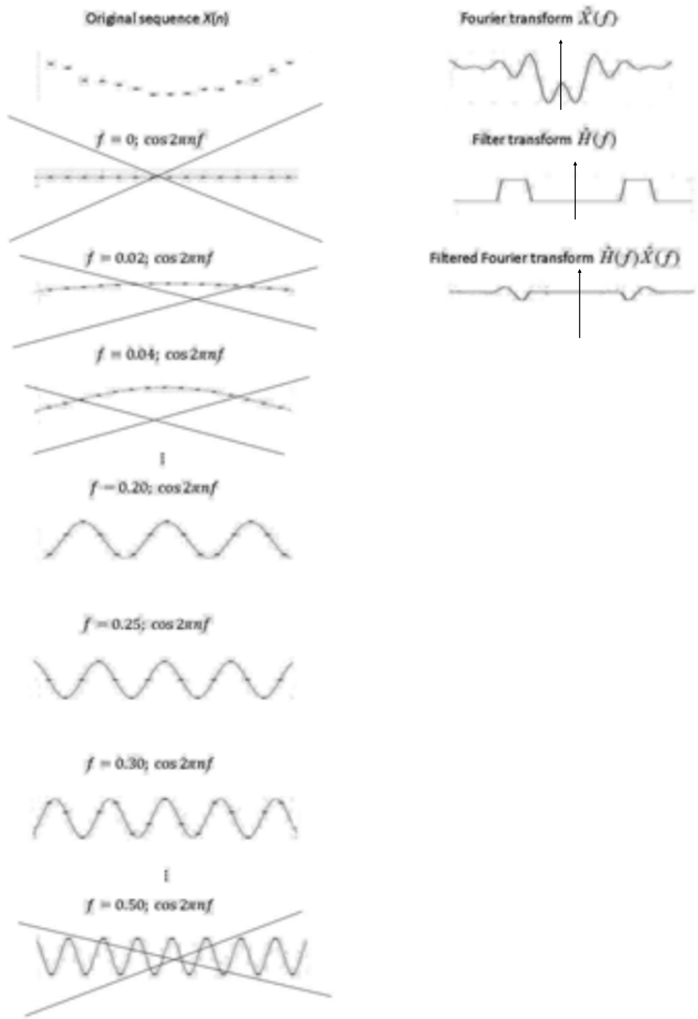


Figure 3.6 Narrow band pass filter

How do you find $H(f)$?   EASY

Let the input be a sinusoid:

$$X(n) = e^{j2\pi nf}$$

Since it's a linear time invariant system
the outwill be easy to find.
And this output will be $H(f)$ times $e^{j2\pi nf}$

because $\hat{Y}(f) = \hat{H}(f)\hat{X}(f)$

So erase the $e^{j2\pi nf}$ and read off $H(f)$

Let the input be a sinusoid: $X(n) = e^{j2\pi nf}$

Since it's a linear time invariant system the outwill be easy to find.
And this output will be $H(f)$ times $e^{j2\pi nf}$

because $\hat{Y}(f) = \hat{H}(f)\hat{X}(f)$

So erase the $e^{j2\pi nf}$ and read off $H(f)$

Example. Suppose the system is

$$Y(n) = 0.5Y(n-1) + 0.1Y(n-2) + 2X(n) + X(n-1)$$

Let the input be $X(n) = e^{j2\pi nf}$. Then

$$\hat{H}(f)e^{j2\pi nf} = 0.5\hat{H}(f)e^{j2\pi(n-1)f} + 0.1\hat{H}(f)e^{j2\pi(n-2)f} + 2e^{j2\pi nf} + e^{j2\pi(n-1)f}$$

$$\hat{H}(f) = \frac{2 + e^{-j2\pi f}}{1 - 0.5e^{-j2\pi f} - 0.1e^{-j4\pi f}}$$

What if the input is a random process with PSD $S_X(f)$ and autocorrelation $R_X(n)$? What's the PSD and autocorrelation of the output $R_Y(n)$?

$$Y(n) = \sum_{m=-\infty}^{\infty} X(m)h(n-m)$$

$$R_Y(n) = \mathrm{E}\{Y(n)Y(0)\}$$

$$Y(n)Y(0) = \sum_{m=-\infty}^{\infty} h(n-m)\,X(m)\,Y(0) = \sum_{m=-\infty}^{\infty} h(n-m)\,X(m) \sum_{p=-\infty}^{\infty} h(-p)\,X(p)$$

$$R_Y(n) = \sum_{m=-\infty}^{\infty} h(n-m) \sum_{p=-\infty}^{\infty} h(-p)\,R_X(m-p)\text{; thus}$$

$$R_Y(n) = \sum_{m=-\infty}^{\infty} h(n-m)\,Q(m) \;\; and \;\; Q(m) = \sum_{p=-\infty}^{\infty} h(-p)\,R_X(m-p).$$

(a "double convolution")

Now take Fourier transforms:

$$S_Y(y) = \hat{H}(f)\widehat{Q(f)} = \widehat{H}(f)\widehat{H}_{rev}(f)S_X(f) = |\hat{H}(f)|^2 S_X(f)$$

A little unnecessary nomenclature:

$$Y(m) = \sum_{r=-\infty}^{\infty} h(m-r)X(r) = \sum_{p=-\infty}^{\infty} h(p)X(m-p)$$

$$Y(m)X(n) = \sum_{p=-\infty}^{\infty} h(p)X(m-p)X(n)$$

"Cross Correlation"

$$R_{YX}(m,n) = \sum_{p=-\infty}^{\infty} h(p)R_X(m-n-p)$$

$$= \sum_{p=-\infty}^{\infty} h(p)R_X([m-n]-p) = R_{YX}(m-n)$$

"Cross Power Spectral Density"

$$S_{YX}(f) = FT\{R_{YX}(m-n)\} = \widehat{H}(f)S_X(f)$$

Example. Suppose the system is

$$Y(n) = 0.5Y(n-1) + 0.1Y(n-2) + 2X(n) + X(n-1)$$

Let the input be $X(n) = e^{j2\pi nf}$. Then

$$\hat{H}(f)e^{j2\pi nf} = 0.5\hat{H}(f)e^{j2\pi(n-1)f} + 0.1\hat{H}(f)e^{j2\pi(n-2)f}$$
$$+2e^{j2\pi nf} + e^{j2\pi(n-1)f}$$

$$\hat{H}(f) = \frac{2+e^{-j2\pi f}}{1-0.5e^{-j2\pi f} - 0.1e^{-j4\pi f}}$$

Suppose $X(n)$ is white noise with power = 5.

$$S_X(f) \equiv 5$$

$$S_Y(y) = \hat{H}(f)\widehat{Q(f)} = \hat{H}(f)\hat{H}_{rev}(f)S_X(f) = |\hat{H}(f)|^2 S_X(f)$$

$$S_Y(f) = \left|\frac{2+e^{-j2\pi f}}{1-0.5e^{-j2\pi f} - 0.1e^{-j4\pi f}}\right|^2 5$$

What's the autocorrelation of the output?

$$R_Y(n) = \int_{-0.5}^{0.5} \left|\frac{2+e^{-j2\pi f}}{1-0.5e^{-j2\pi f} - 0.1e^{-j4\pi f}}\right|^2 5 \; e^{j2\pi nf}\, df$$

What's the power in the output?

$$R_Y(0) = \int_{-0.5}^{0.5} \left|\frac{2+e^{-j2\pi f}}{1-0.5e^{-j2\pi f} - 0.1e^{-j4\pi f}}\right|^2 5 \; df$$

Summary: Spectral Properties of Discrete $X(n)$ and Continuous $X(t)$ Random Processes

Discrete: Frequency f is measured in cycles per sample: $-\frac{1}{2} < f < \frac{1}{2}$

Continuous: Frequency f is measured in cycles per second:

Discrete ergodicity:
$$E\{X_n\} = \lim_{N\to\infty} \frac{X_1 + X_2 + \ldots + X_N}{N}$$

$$R_n = \lim_{N\to\infty} \frac{\sum_1^N X_m X_{m-n}}{N}$$

Continuous:
$$E\{X(t)\} = \lim_{T\to\infty} \frac{\int_{-T/2}^{T/2} X(\tau)\,d\tau}{T}$$

$$R(t) = \lim_{T\to\infty} \frac{\int_{-T/2}^{T/2} X(\tau)X(\tau - t)\,d\tau}{T}$$

Fourier Transform

Discrete
$$\tilde{X}(f) = \sum_{n=-\infty}^{\infty} X(n)\, e^{-j2\pi nf} \qquad X(n) = \int_{-1/2}^{1/2} \tilde{X}(f)\, e^{j2\pi nf}\, df$$

Continuous
$$\tilde{X}(f) = \int_{-\infty}^{\infty} X(t)\, e^{-i2\pi tf}\, dt$$
$$X(f) = \int_{-\infty}^{\infty} \tilde{X}(f)\, e^{i2\pi tf}\, df$$

Parseval

Discrete
$$\sum_{n=-\infty}^{\infty} X(n)^2 = \int_{-1/2}^{1/2} |\tilde{X}(f)|^2\, df$$

Continuous
$$\int_{-\infty}^{\infty} |X(t)|^2\, dt = \int_{-\infty}^{\infty} |\tilde{X}(f)|^2\, df$$

Power Spectral Density

Discrete
$$S_X(f) = \sum_{k=-\infty}^{\infty} R_X(k) e^{-j2\pi kf}$$
$$R_X(n) = \int_{-1/2}^{1/2} S_X(f) e^{j2\pi nf}\, df$$

Continuous
$$S_X(f) = \int_{-\infty}^{\infty} R_X(t) e^{-j2\pi tf}\, dt$$
$$R_X(t) = \int_{-\infty}^{\infty} S_X(f) e^{j2\pi tf}\, df$$

White Noise

Discrete $S_W(f) = 1, \ -1/2 < f < 1/2$
$$R_W(n) = \delta_{n0} \quad (unit\ power)$$

Continuous
$$S_W(f) = 1, \ -\infty < f < \infty \quad (unit\ power\ density)$$
$$R_W(t) = \delta(t) \quad (infinite\ power)$$

Linear System Impulse Response

Discrete

$$Y(n) = \sum_{m=-\infty}^{\infty} h(n-m)\,X(m)$$

$$S_Y(f) = |\hat{H}(f)|^2\, S_X(f)$$

Continuous

$$Y(t) = \int_{-\infty}^{\infty} h(t-\tau)\,X(\tau)\,d\tau$$

$$S_Y(f) = |\hat{H}(f)|^2\, S_X(f)$$

Lecture 13

March 1, 2017

$$\frac{d^2x}{dt^2}+2\frac{dx}{dt}+2x=5\sin t,\ x(0)=1,\ \frac{dx(0)}{dt}=2$$

$$x(t)=x_{part}(t)+x_{\text{homog}}(t)$$

$$=\sin t-2\cos t+c_1 e^{-t}\cos t+c_2 e^{-t}\sin t$$

$$\frac{dx}{dt}=\lim_{\Delta t\to 0}\frac{x(t+\Delta t)-x(t)}{\Delta t}$$

$$\frac{dx}{dt}\approx\frac{x(t+\Delta t)-x(t)}{\Delta t}$$

$$\frac{d^2x}{dt^2}\approx\frac{x(t+2\Delta t)-2x(t+\Delta t)+x(t)}{\Delta t^2}$$

$$x(0)=1$$

$$x(\Delta t)=2\Delta t+1$$

$$x(2\Delta t)=\{2-2\Delta t\}x(\Delta t)+\{-1+2\Delta t-2\Delta t^2\}x(0)-5\,\Delta t^2\sin\ 0$$

$$x(3\Delta t)=\{2-2\Delta t\}x(2\Delta t)+\{-1+2\Delta t-2\Delta t^2\}x(\Delta t)-5\,\Delta t^2\sin\ \Delta t$$

$$\dots$$

$$x(n\Delta t)=\{2-2\Delta t\}x([n-1]\Delta t)+$$

$$\{-1+2\Delta t-2\Delta t^2\}x([n-2]\Delta t)-5\,\Delta t^2\sin(\ n-2)\Delta t$$

Autoregressive-Moving-Average ARMA(p,q)

$$x(n) = \underbrace{a(1)x(n-1) + a(2)x(n-2) + \cdots + a(p)x(n-p)}_{autoregressive\ of\ order\ p}$$

$$+ \underbrace{b(0)v(n) + b(1)v(n-1) + \cdots + b(q)v(n-q)}_{moving\ average\ of\ order\ q}$$

In most cases $v(n)$ is white noise:

$$E\{v(n)\} = 0\,,$$

$$E\{v(n)v(m)\} = \sigma^2\delta_{mn}$$

The *Yule-Walker* equations relate the coefficients $a(1)$, ..., $b(0)$, …
to the autocorrelation
$$R_X(k) = \mathrm{E}\{X(m)X(m+k)\}\ .$$

Given the coefficients, they can be solved for the autocorelations.

Given the autocorrelations, they can be solved for the coefficients.

To get the Yule-Walker equations,

1. write the ARMA(p,q) equation for $X(p+q)$. For ARMA(2,0):

$$X(2) = a(1)\,X(1) + a(2)\,X(0) + b(0)\,V(2)$$

2. Multiply this equation by $X(0)$, $X(1)$, up to $X(p+q)$; and by $V(p+q)$, $V(p+q-1)$, down to $V(q)$. Then take expected values.

$$E\{X(0)\,X(2)\} = a(1)\,E\{X(0)\,X(1)\} + a(2)\,E\{X(0)\,X(0)\} + b(0)\,E\{X(0)\,V(2)\}$$

$$R_X(2) = a(1)\,R_X(1) + a(2)\,R_X(0)$$

$$E\{X(1)\,X(2)\} = a(1)\,E\{X(1)\,X(1)\} + a(2)\,E\{X(1)\,X(0)\} + b(0)\,E\{X(1)\,V(2)\}$$

$$R_X(1) = a(1)\,R_X(0) + a(2)\,R_X(1)$$

$$E\{X(2)\,X(2)\} = a(1)\,E\{X(2)\,X(1)\} + a(2)\,E\{X(2)\,X(0)\} + b(0)\,E\{X(2)\,V(2)\}$$

$$X(2) = a(1)\,X(1) + a(2)\,X(0) + b(0)\,V(2)$$

2. Multiply this equation by $X(0)$, $X(1)$, up to $X(p+q)$; and by $V(p+q)$, $V(p+q-1)$, down to $V(q)$. Then take expected values.

$$E\{X(0)\,X(2)\} = a(1)\,E\{X(0)\,X(1)\} + a(2)\,E\{X(0)\,X(0)\} + b(0)\,E\{X(0)\,V(2)\}$$

$$R_X(2) = a(1)\,R_X(1) + a(2)\,R_X(0)$$

$$E\{X(1)\,X(2)\} = a(1)\,E\{X(1)\,X(1)\} + a(2)\,E\{X(1)\,X(0)\} + b(0)\,E\{X(1)\,V(2)\}$$

$$R_X(1) = a(1)\,R_X(0) + a(2)\,R_X(1)$$

$$E\{X(2)\,X(2)\} = a(1)\,E\{X(2)\,X(1)\} + a(2)\,E\{X(2)\,X(0)\} + b(0)\,E\{X(2)\,V(2)\}$$?

$$E\{V(2)\,X(2)\} = a(1)\,E\{V(2)\,X(1)\} + a(2)\,E\{V(2)\,X(0)\} + b(0)\,E\{V(2)\,V(2)\}$$

$$E\{V(2)X(2)\} = 0 + 0 + b(0)\sigma^2$$

$$R_X(0) = a(1)\,R_X(1) + a(2)\,R_X(2) + b(0)^2\,\sigma^2$$

$R_X(2) = a(1)\, R_X(1) + a(2)\, R_X(0)$.

$R_X(1) = a(1)\, R_X(0) + a(2)\, R_X(1)$.

$R_X(0) = a(1)\, R_X(1) + a(2)\, R_X(2) + b(0)^2\, \sigma^2$

$$\begin{bmatrix} -a(2) & -a(1) & 1 \\ -a(1) & [1-a(2)] & 0 \\ 1 & -a(1) & -a(2) \end{bmatrix} \begin{bmatrix} R_X(0) \\ R_X(1) \\ R_X(2) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ b(0)^2\sigma^2 \end{bmatrix}$$

$$\begin{bmatrix} R_X(1) & R_X(0) & 0 \\ R_X(0) & R_X(1) & 0 \\ R_X(1) & R_X(2) & \sigma^2 \end{bmatrix} \begin{bmatrix} a(1) \\ a(2) \\ b(0)^2 \end{bmatrix} = \begin{bmatrix} R_X(2) \\ R_X(1) \\ R_X(0) \end{bmatrix}$$

For higher n,

$X(2) = a(1)\, X(1) + a(2)\, X(0) + b(0)\, V(2)$

$R_X(n) = a(1)\, R_X(n\text{-}1) + a(2)\, R_X(n\text{-}2)$

To get the Yule-Walker equations,

1. write the ARMA(p,q) equation for $X(p+q)$. For ARMA(3,1):

$$X(4) = a(1)X(3) + a(2)X(2) + a(3)X(1) + b(0)V(4) + b(1)V(3)$$

2. Multiply this equation by $X(p+q)$, $X(p+q\text{-}1)$, down to $X(0)$; and by $V(p+q)$, $V(p+q\text{-}1)$, down to $V(q)$. Then take expected values.

$$X(4) = a(1)X(3) + a(2)X(2) + a(3)X(1)$$
$$+ b(0)V(4) + b(1)V(3)$$

$$R_X(4) = a(1)\ R_X(3) + a(2)\ R_X(2) + a(3)\ R_X(1)$$

$$R_X(3) = a(1)\ R_X(2) + a(2)\ R_X(1) + a(3)\ R_X(0)$$

$$R_X(2) = a(1)\ R_X(1) + a(2)\ R_X(0) + a(3)\ R_X(1)\ ,$$

$$R_X(1) = a(1)\ R_X(0) + a(2)\ R_X(1) + a(3)\ R_X(2)$$

$$+ b(1)\ E\{X(3)\ V(3)\}\ ,$$

$$R_X(0) = a(1)\ R_X(1) + a(2)\ R_X(2) + a(3)\ R_X(3)$$

$$+ b(0)\ E\{X(4)\ V(4)\} + b(1)\ E\{X(4)\ V(3)\}$$

$$E\{X(4)\ V(4)\} = b(0)\sigma^2$$

$$E\{X(j)V(j)\} = b(0)\sigma^2$$

$$E\{X(4)\ V(3)\} = a(1)\ E\{X(3)\ V(3)\} + (0) + b(1)\ \sigma^2$$

$$= a(1)\ b(0)\sigma^2 + b(1)\sigma^2$$

$$
\begin{bmatrix}
0 & -a(3) & -a(2) & -a(1) & 1 \\
-a(3) & -a(2) & -a(1) & 1 & 0 \\
-a(2) & [-a(3)-a(1)] & 1 & 0 & 0 \\
-a(1) & [1-a(2)] & -a(3) & 0 & 0 \\
1 & -a(1) & -a(2) & -a(3) & 0
\end{bmatrix}
\begin{bmatrix}
R_X(0) \\ R_X(1) \\ R_X(2) \\ R_X(3) \\ R_X(4)
\end{bmatrix}
$$

$$
= \begin{bmatrix}
0 \\ 0 \\ 0 \\ b(1)b(0)\sigma^2 \\ \{b(0)^2 + b(1)[a(1)b(0) + b(1)]\}\sigma^2
\end{bmatrix}
$$

For higher n,

$$
X(4) = a(1)X(3) + a(2)X(2) + a(3)X(1) + b(0)V(4) + b(1)V(3)
$$

$$
R_X(n) = a(1)\, R_X(n\text{-}1) + a(2)\, R_X(n\text{-}2) + a(3)\, R_X(n\text{-}3)
$$

$$\begin{bmatrix} 0 & -a(3) & -a(2) & -a(1) & 1 \\ -a(3) & -a(2) & -a(1) & 1 & 0 \\ -a(2) & [-a(3)-a(1)] & 1 & 0 & 0 \\ -a(1) & [1-a(2)] & -a(3) & 0 & 0 \\ 1 & -a(1) & -a(2) & -a(3) & 0 \end{bmatrix} \begin{bmatrix} R_X(0) \\ R_X(1) \\ R_X(2) \\ R_X(3) \\ R_X(4) \end{bmatrix}$$

$$= \begin{bmatrix} 0 \\ 0 \\ 0 \\ b(1)b(0)\sigma^2 \\ \{b(0)^2 + b(1)[a(1)b(0) + b(1)]\}\sigma^2 \end{bmatrix}$$

For higher *n*,

$$X(4) = a(1)X(3) + a(2)X(2) + a(3)X(1) + b(0)V(4) + b(1)V(3)$$

$$R_X(n) = a(1)\, R_X(n\text{-}1) + a(2)\, R_X(n\text{-}2) + a(3)\, R_X(n\text{-}3)$$

But the equations for the ARMA(*p*,*q*) coefficients are ugly when *q* > 0.

Here's a problem. If you type
> help filter

in MATLAB, you'll learn that MATLAB's notation for the ARMA equation is

```
a(1)*y(n) = - a(2)*y(n-1) - ... - a(na+1)*y(n-na)
            + b(1)*x(n) + b(2)*x(n-1) + ...
            + b(nb+1)*x(n-nb)
```

but ours is

$$x(n) = \underbrace{a(1)x(n-1) + a(2)x(n-2) + \cdots + a(p)x(n-p)}_{autoregressive\ of\ order\ p}$$

$$+ \underbrace{b(0)v(n) + b(1)v(n-1) + \cdots + b(q)v(n-q)}_{moving\ average\ of\ order\ q}$$

So MATLAB's a(1) is 1,
MATLAB's a(2) is our -a(1),
MATLAB's a(3) is our -a(2),
MATLAB's a(na+1) is -a(p),
MATLAB's b(1) is our b(0),
MATLAB's b(2) is our b(1), and
MATLAB's b(nb+1) is our b(q).

MATLAB's notation for the ARMA equation is

a(1)*y(n) = - a(2)*y(n-1) - ... - a(na+1)*y(n-na)
+ b(1)*x(n) + b(2)*x(n-1) + ...
+ b(nb+1)*x(n-nb)

but ours is

$$x(n) = \underbrace{a(1)x(n-1) + a(2)x(n-2) + \cdots + a(p)x(n-p)}_{autoregressive\ of\ order\ p}$$

$$+ \underbrace{b(0)v(n) + b(1)v(n-1) + \cdots + b(q)v(n-q)}_{moving\ average\ of\ order\ q}$$

So MATLAB's a(1) is 1,
MATLAB's a(2) is our -a(1),
MATLAB's a(3) is our -a(2),
MATLAB's a(na+1) is -a(p),
MATLAB's b(1) is our b(0),
MATLAB's b(2) is our b(1), and
MATLAB's b(nb+1) is our b(q).

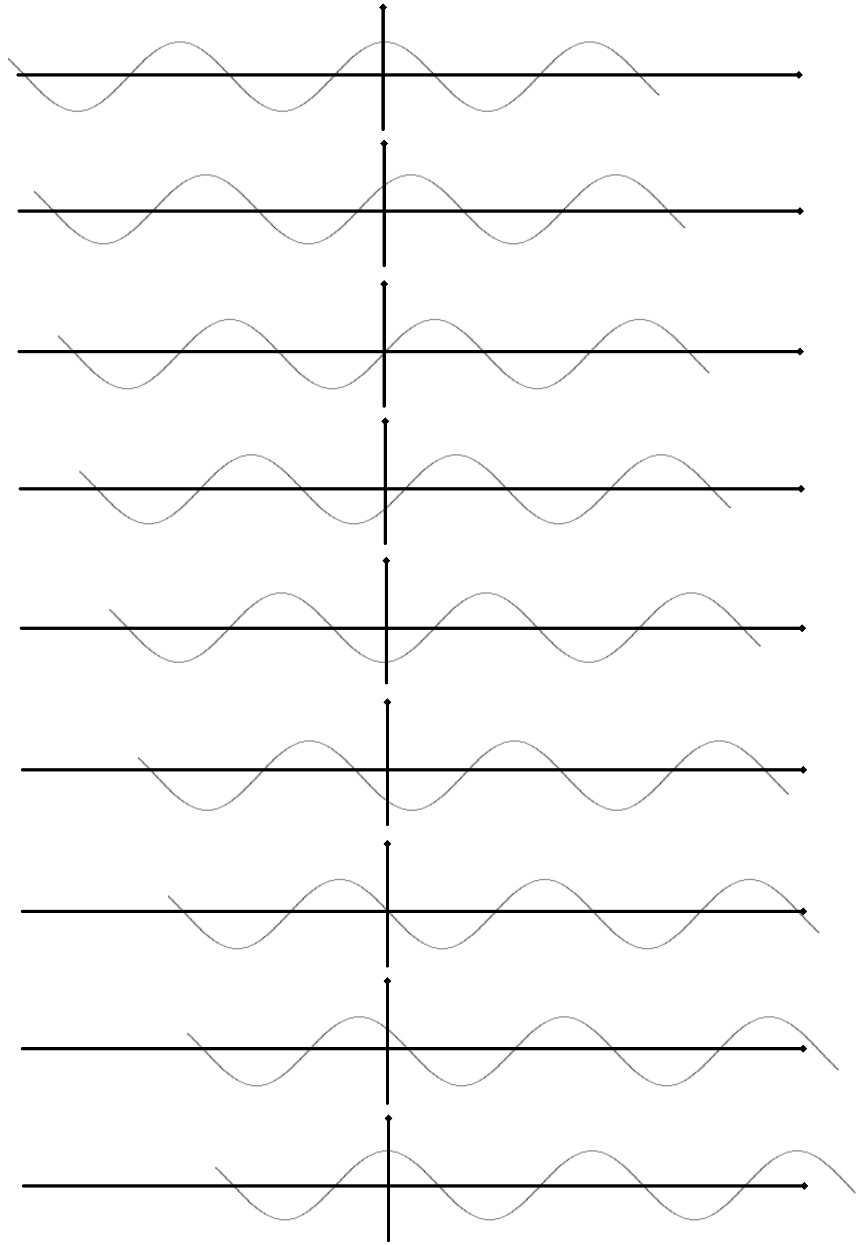
To use MATLAB's **Y = filter(B,A,X) ,**

in MATLAB define
X = [Y(1), Y(2), ...]
B = [b(0), b(1), b(2), ...]
A = [1, -a(1), -a(2), ...]

Lecture 14

March 6, 2017

Random-Phase Sine Wave



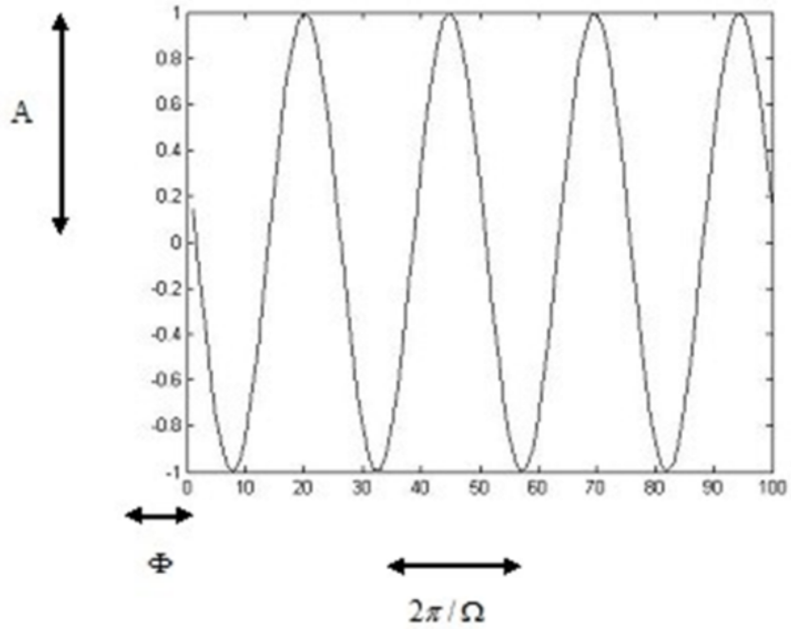


Figure 4.1. Random Sine Wave

$$X(t) = A\cos\left(\Omega t + \Phi\right)$$

$$X(t) = A\cos(\Omega t + \Phi)$$

$$\mu_X(t) = E\{X(t)\} =$$

$$\int_0^{\infty}\int_0^{\infty}\int_0^{2\pi} a\cos(\omega t + \phi)\, f_A(a) f_\Omega(\omega) f_\Phi(\phi)\, d\phi\, d\omega\, da$$

$$R_X(t_1, t_2) = E\{X(t_1)X(t_2)\} =$$

$$\int_0^{\infty}\int_0^{\infty}\int_0^{2\pi} a^2 \cos(\omega t_1 + \phi)\cos(\omega t_2 + \phi) f_A(a) f_\Omega(\omega) f_\Phi(\phi)\, d\phi\, d\omega\, da$$

$$= \int_0^{\infty}\int_0^{\infty}\int_0^{2\pi} a^2 \frac{\cos\omega[t_1 - t_2] + \cos(\omega[t_1 + t_2] + 2\phi)}{2} f_A(a) f_\Omega(\omega) f_\Phi(\phi)\, d\phi\, d\omega\, da$$

---

pdf $f_\Phi(\varphi)$ is uniform (equal to $1/2\pi$)

$$\mu_{\text{sine}}(t) = \mathbf{E}\{X(t)\} =$$

$$\frac{1}{2\pi}\int_0^{\infty}\int_0^{\infty} a\, f_A(a) f_\Omega(\omega) \int_0^{2\pi} \cos(\omega t + \phi)\, d\phi\, d\omega\, da = 0$$

$$X(t) = A\cos(\Omega t + \Phi)$$

pdf $f_\Phi(\varphi)$ is uniform (equal to $1/2\pi$)

$$\mu_{sine}(t) = \mathbf{E}\{X(t)\} = 0$$

$$R_X(t_1,t_2) = E\{X(t_1)X(t_2)\} =$$

$$\int_0^\infty\int_0^\infty\int_0^{2\pi} a^2 \cos(\omega t_1+\phi)\cos(\omega t_2+\phi) f_A(a) f_\Omega(\omega) f_\Phi(\phi)\, d\phi\, d\omega\, da$$

$$= \int_0^\infty\int_0^\infty\int_0^{2\pi} a^2 \frac{\cos\omega[t_1-t_2] + \cos(\omega[t_1+t_2]+2\phi)}{2} f_A(a) f_\Omega(\omega) f_\Phi(\phi)\, d\phi\, d\omega\, da$$

$$R_{sine}(t_1,t_2) =$$

$$\frac{1}{2\pi}\int_0^\infty a^2 f_A(a)\, da \int_0^\infty f_\Omega(\omega) \int_0^{2\pi} \frac{\cos\omega[t_1-t_2] + \cos(\omega[t_1+t_2]+2\phi)}{2}\, d\phi\, d\omega$$

$$= \frac{1}{2\pi}\int_0^\infty a^2 f_A(a)\, da \int_0^\infty f_\Omega(\omega) 2\pi \frac{\cos\omega[t_1-t_2] + (0)}{2}\, d\omega$$

$$= \frac{1}{2}\mathbf{E}\{A^2\}\int_0^\infty f_\Omega(\omega)\cos\omega[t_1-t_2]\, d\omega$$

$$X(t) = A\cos(\Omega t + \Phi)$$

pdf $f_\Phi(\varphi)$ is uniform (equal to $1/2\pi$)

$$\mu_{sine}(t) = \mathbf{E}\{X(t)\} = 0$$

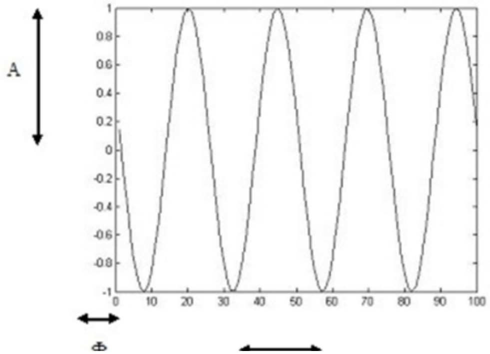
$R_X(t_1,t_2)$

$$= \frac{1}{2\pi}\int_0^\infty a^2 f_A(a)\,da \int_0^\infty f_\Omega(\omega) 2\pi \frac{\cos\omega[t_1 - t_2] + (0)}{2}\,d\omega$$

$$= \frac{1}{2}\mathbf{E}\{A^2\}\int_0^\infty f_\Omega(\omega)\cos\omega[t_1 - t_2]\,d\omega$$

Stationary!

Certainly not ergodic!



$$X(t) = A\cos(\Omega t + \Phi)$$

pdf $f_\Phi(\varphi)$ is uniform (equal to $1/2\pi$)

$$\mu_{\text{sine}}(t) = \mathbf{E}\{X(t)\} = 0$$

$R_X(t_1,t_2)$

$$= \frac{1}{2\pi}\int_0^\infty a^2 f_A(a)\,da \int_0^\infty f_\Omega(\omega) 2\pi \frac{\cos\omega[t_1 - t_2] + (0)}{2}\,d\omega$$

$$= \frac{1}{2}\mathbf{E}\{A^2\}\int_0^\infty f_\Omega(\omega)\cos\omega[t_1 - t_2]\,d\omega$$

$$S_{\text{sine}}(f) = \int_{-\infty}^{\infty} R_X(\tau) e^{-i2\pi f\tau}\,d\tau$$

$$= \frac{1}{2}E\{A^2\}\int_0^\infty f_\Omega(\omega)\int_{-\infty}^{\infty}\cos\omega\tau\, e^{-i2\pi f\tau}\,d\tau\,d\omega$$

$$= \frac{1}{2}E\{A^2\}\int_0^\infty f_\Omega(\omega)\frac{\delta(\frac{\omega}{2\pi} - f) + \delta(\frac{\omega}{2\pi} + f)}{2}\,d\omega$$

$$= \frac{\pi}{2}E\{A^2\} f_\Omega(|\,2\pi f\,|)$$

Bernoulli Process

Independent Coin Flips

Summary: Bernoulli Process

For $n = 1, 2, 3, \ldots,\quad X(n) = \begin{cases} a \text{ with probability } p \\ b \text{ with probability } 1-p \end{cases}$

$\mu_{\text{Bernoulli}} \equiv E\{X(n)\} = pa + (1\text{-}p)b$

$R_{\text{Bernoulli}}(n_1, n_2) \equiv E\{X(n_1)\, X(n_2)\}$

$$= \begin{cases} pa^2 + (1-p)b^2 & if\ n_1 = n_2 \\ \mu^2_{\text{Bernoulli}} & if\ n_1 \neq n_2 \end{cases}$$

$$R_{\text{Bernoulli}}(m) = \mu^2_{\text{Bernoulli}} + [pa^2 + (1-p)b^2 - \mu^2_{\text{Bernoulli}}]\delta_{m0}$$

$S_{\text{Bernoulli}}(f) = \mu^2_{\text{Bernoulli}}\ \delta(f) + pa^2 + (1-p)b^2 - \mu^2_{\text{Bernoulli}}$.

Summary: Bernoulli Process

$$\text{For } n = 1, 2, 3, \ldots, \quad X(n) = \begin{Bmatrix} a \text{ with probability } p \\ b \text{ with probability } 1-p \end{Bmatrix}$$

$$\mu_{\text{Bernoulli}} \equiv E\{X(n)\} = pa + (1\text{-}p)b$$

$$R_{\text{Bernoulli}}(n_1, n_2) \equiv E\{X(n_1)\, X(n_2)\}$$

$$= \begin{Bmatrix} pa^2 + (1-p)b^2 & \text{if } n_1 = n_2 \\ \mu^2_{\text{Bernoulli}} & \text{if } n_1 \neq n_2 \end{Bmatrix}$$

$$R_{\text{Bernoulli}}(m) = \mu^2_{\text{Bernoulli}} + [pa^2 + (1-p)b^2 - \mu^2_{\text{Bernoulli}}]\delta_{m0}$$

$$S_{\text{Bernoulli}}(f) = \mu^2_{\text{Bernoulli}}\ \delta(f) + pa^2 + (1-p)b^2 - \mu^2_{\text{Bernoulli}}\ .$$

Binomial Process

Accumulation of Bernoulli Trials

Summary: Bernoulli Process

For $n = 1, 2, 3, \ldots, \quad X(n) = \begin{cases} a \text{ with probability } p \\ b \text{ with probability } 1-p \end{cases}$

$\mu_{\text{Bernoulli}} \equiv E\{X(n)\} = pa + (1\text{-}p)b$

$R_{\text{Bernoulli}}(n_1, n_2) \equiv E\{X(n_1)\, X(n_2)\} = \begin{cases} pa^2 + (1-p)b^2 & if\, n_1 = n_2 \\ \mu^2_{\text{Bernoulli}} & if\, n_1 \neq n_2 \end{cases}$

$$R_{\text{Bernoulli}}(m) = \mu^2_{\text{Bernoulli}} + [pa^2 + (1-p)b^2 - \mu^2_{\text{Bernoulli}}]\delta_{m0}$$

$S_{\text{Bernoulli}}(f) = \mu^2_{\text{Bernoulli}}\,\delta(f) + pa^2 + (1-p)b^2 - \mu^2_{\text{Bernoulli}}.$

$$X_{binomial}(n) = \Sigma_{k=1}^{n} X_{Bernoulli}(k)$$

Summary: Binomial Process

$$\text{Prob}\{X{=}a\} = p,\ \text{Prob}\{X{=}b\} = 1\text{-}p$$

$$\mu_{\text{binomial}}(n) = n[pa + (1\text{-}p)b]$$

$$R_{\text{binomial}}(n_1, n_2) = n_1 n_2 \mu^2_{\text{Bernoulli}} + [pa^2 + (1-p)b^2 - \mu^2_{\text{Bernoulli}}]\min(n_1, n_2)$$

Binomial Process

$X_{binomial}(n) = \Sigma_{k=1}^{n} X_{Bernoulli}(k)$

$$\mu_{\text{Bernoulli}} \equiv E\{X(n)\} = pa + (1\text{-}p)b$$

$$\mu_{\text{binomial}}(n) = E\{X_{\text{binomial}}(n)\} =$$

$$E\{\sum_{r=1}^{n} X_{\text{Bernoulli}}(r)\} = n\mu_{\text{Bernoulli}} = n[pa + (1\text{-}p)b]$$

$$R_{\text{binomial}}(n_1, n_2) =$$
$$E\{(X_{\text{Bernoulli}}(1) + X_{\text{Bernoulli}}(2) + \ldots + X_{\text{Bernoulli}}(n_1))\times$$
$$(X_{\text{Bernoulli}}(1) + X_{\text{Bernoulli}}(2) + \ldots + X_{\text{Bernoulli}}(n_2))\}$$

$$R_{\text{Bernoulli}}(m) = \mu^2_{\text{Bernoulli}} + [pa^2 + (1-p)b^2 - \mu^2_{\text{Bernoulli}}]\delta_{m0}$$

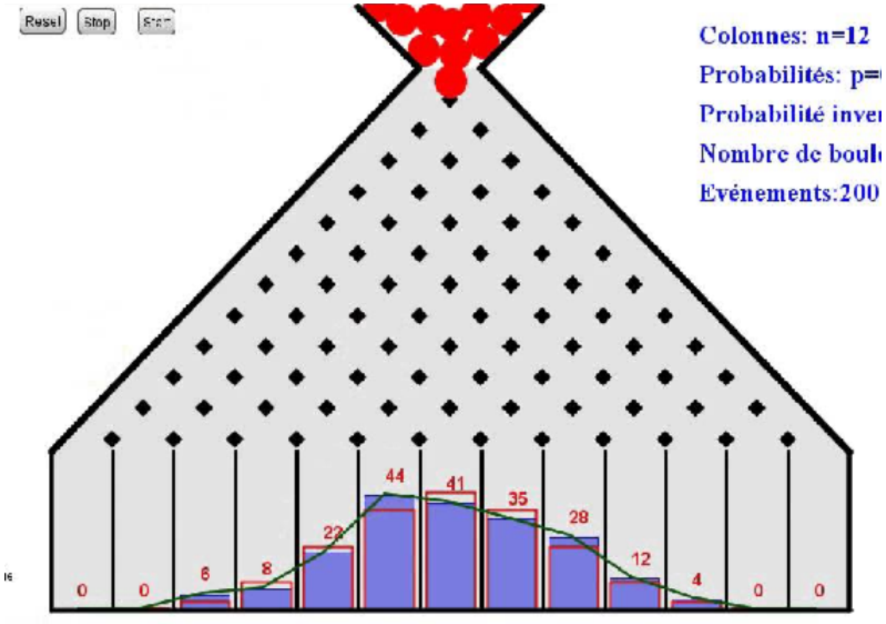
$$R_{\text{binomial}}(n_1, n_2)$$
$$= n_1[pa^2 + (1-p)b^2] + (n_1 n_2 - n_1)\mu^2_{\text{Bernoulli}}$$
$$= n_1 n_2 \mu^2_{\text{Bernoulli}}$$
$$+ [pa^2 + (1-p)b^2 - \mu^2_{\text{Bernoulli}}]\min(n_1, n_2)$$

Binomial Process

Probability of r a's in n trials =

$$\text{Prob}\{X_{binomial}(n) \equiv \sum_{j=1}^{n} X_{Bernoulli}(n) = ra + (n-r)b\}$$

$$= \binom{n}{r} p^r (1-p)^{n-r} \ .$$



Reset
Stop
Colonnes: n=12
Probabilités: p=
Probabilité inve
Nombre de boul
Evénements:200
0
0
6
8
22
44
41
35
28
12
4
0
0

Figure 4.2 Galton machine
© UCL Galton Collection
(University College London)

Approximation of Binomial by Normal

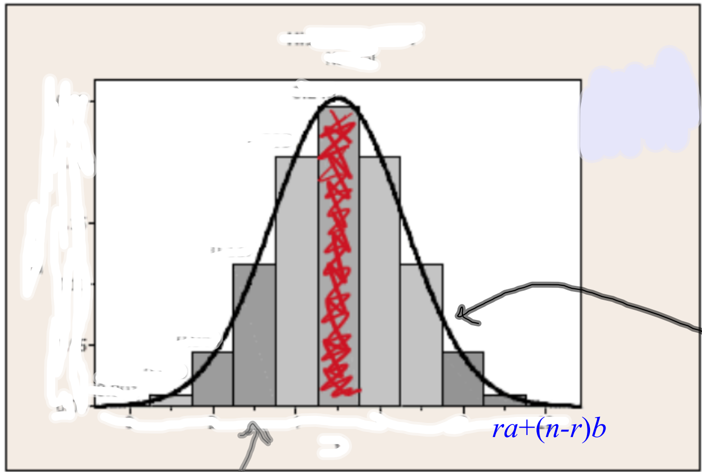
$$\text{Prob}\{X{=}a\} = p,\ \text{Prob}\{X{=}b\} = 1\text{-}p$$

$$\mu_{\text{binomial}}(n) = n[pa + (1\text{-}p)b]$$

$$R_{\text{binomial}}(n_1, n_2) = n_1 n_2 \mu^2_{\text{Bernoulli}} + [pa^2 + (1-p)b^2 - \mu^2_{\text{Bernoulli}}]\min(n_1, n_2)$$

$$\text{Prob}\{X_{binomial}(n) \equiv \textstyle\sum_{j=1}^{n} X_{Bernoulli}(n) = ra + (n-r)b\}$$

$$= \binom{n}{r} p^r (1-p)^{n-r}\ .$$



Binomial

$$N\{\mu_{binomial}(n), \sqrt{R_{binomial}(n,n)}\}$$

Poisson Process

Shot Noise

(Random Arrivals)

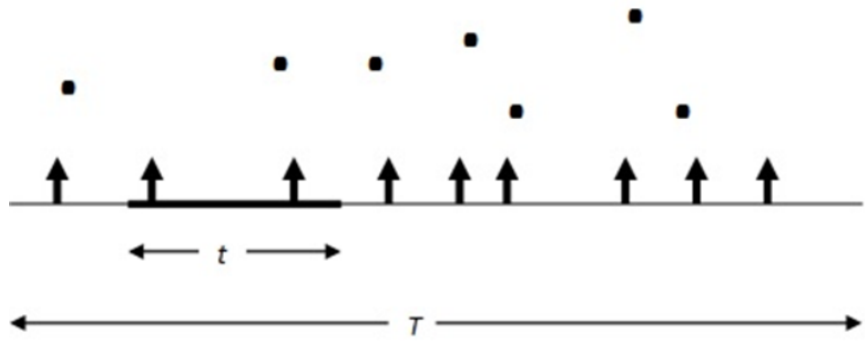
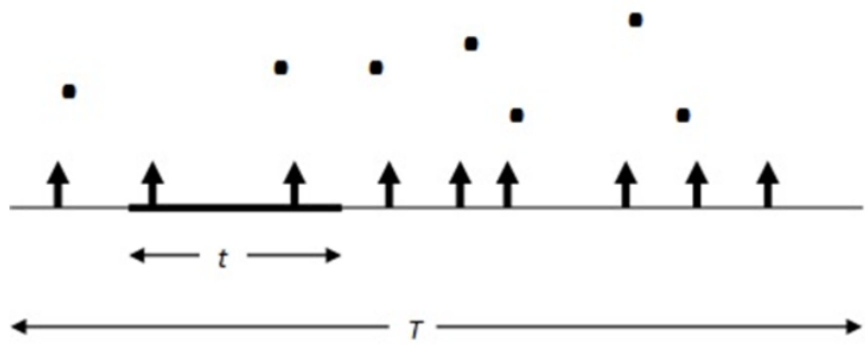


Figure 4.3 m=9 events in an interval T

One ampere equals 6.24150975e+18 electrons per second.



the probability of n of the m events

landing in the interval I_t is given by

$$\binom{m}{n}\left(\frac{t}{T}\right)^{n}\left(\frac{T-t}{T}\right)^{m-n}$$

let T and m go to infinity in this

model while $\frac{m}{T} \equiv \lambda$

$$\frac{m!}{n!(m-n)!}(\frac{t}{T})^n(\frac{T-t}{T})^{m-n}$$

$$= \frac{t^n}{n!} \quad \frac{m(m-1)(m-2)...(m-n+1)}{T^n} \quad (1-\frac{t}{T})^{-n} \quad (1-\frac{t}{T})^m$$

$$= \frac{t^n}{n!}(\frac{m}{T})^n \frac{m(m-1)(m-2)...(m-n+1)}{m \quad m \quad m \quad ... \quad m} (1-\frac{t}{T})^{-n} (1-\frac{t}{T})^{(\frac{T}{t})(\frac{m}{T})t}$$

$$\| \qquad \downarrow_{m\to\infty} \qquad \downarrow_{T\to\infty} \qquad \downarrow_{T\to\infty}$$

$$\frac{(\lambda t)^n}{n!} \qquad 1 \qquad 1 \qquad (\frac{1}{e})^{(\frac{m}{T})t}$$

$$P(n;t) = \frac{(\lambda t)^n}{n!} e^{-\lambda t}.$$

Lecture 15

March 8, 2017

Recall binomial process: probability of *r* successes in *n* tries.

$$\text{Prob}\{X=a\} = p, \ \text{Prob}\{X=b\} = 1\text{-}p$$

$$\mu_{\text{binomial}}(n) = n[pa + (1\text{-}p)b]$$

$$R_{\text{binomial}}(n_1, n_2) = n_1 n_2 \mu^2_{\text{Bernoulli}} + [pa^2 + (1-p)b^2 - \mu^2_{\text{Bernoulli}}]\min(n_1, n_2)$$

$$\text{Prob}\{X_{binomial}(n) \equiv \sum_{j=1}^{n} X_{Bernoulli}(n) = ra + (n-r)b\}$$

$$= \binom{n}{r} p^r (1-p)^{n-r} \, .$$

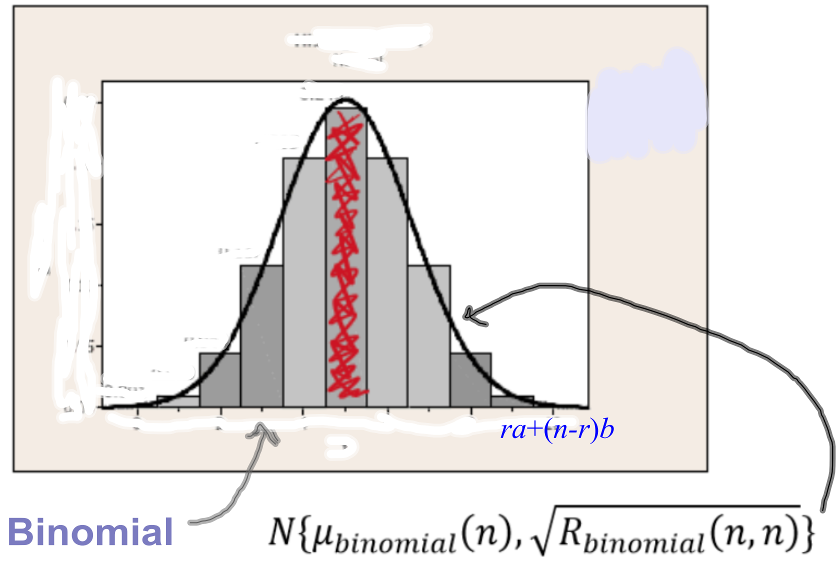
Approximation of Binomial by Normal

$$\text{Prob}\{X=a\} = p,\ \text{Prob}\{X=b\} = 1\text{-}p$$

$$\mu_{\text{binomial}}(n) = n[pa + (1\text{-}p)b]$$

$$R_{\text{binomial}}(n_1, n_2) = n_1 n_2 \mu^2_{\text{Bernoulli}} + [pa^2 + (1-p)b^2 - \mu^2_{\text{Bernoulli}}]\min(n_1, n_2)$$

$$\text{Prob}\{X_{binomial}(n) \equiv \sum_{j=1}^{n} X_{Bernoulli}(n) = ra + (n-r)b\} = \binom{n}{r} p^r (1-p)^{n-r}\ .$$



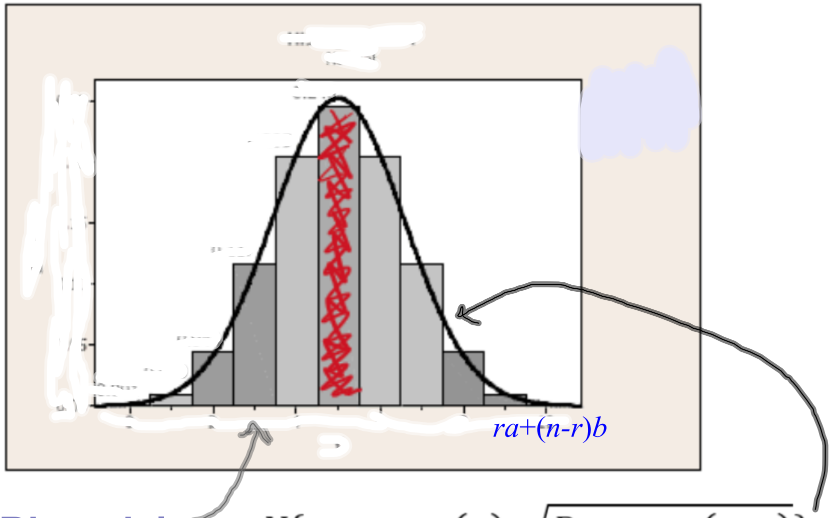
Approximation of Binomial by Normal

Prob{X=a} = p, Prob{X=b} = 1-p

$\mu_{\text{binomial}}(n) = n[pa + (1\text{-}p)b]$

$$R_{\text{binomial}}(n_1, n_2) = n_1 n_2 \mu^2_{\text{Bernoulli}} + [pa^2 + (1-p)b^2 - \mu^2_{\text{Bernoulli}}]\min(n_1, n_2)$$

$$\text{Prob}\{X_{binomial}(n) \equiv \sum_{j=1}^{n} X_{Bernoulli}(n) = ra + (n-r)b\}$$
$$= \binom{n}{r} p^r (1-p)^{n-r} .$$



$N\{\mu_{binomial}(n), \sqrt{R_{binomial}(n,n)}\}$

The binomial process is certainly not stationary. These curves migrate to the right and broaden as n incleases.

Recall binomial process: probability of *r* successes in *n* tries.

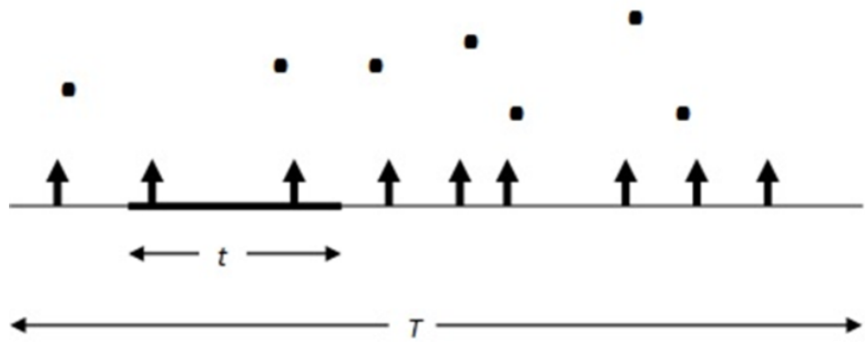
$$\text{Prob}\{X=a\} = p, \ \text{Prob}\{X=b\} = 1\text{-}p$$

$$\mu_{\text{binomial}}(n) = n[pa + (1\text{-}p)b]$$

$$R_{\text{binomial}}(n_1, n_2) = n_1 n_2 \mu^2_{\text{Bernoulli}} + [pa^2 + (1-p)b^2 - \mu^2_{\text{Bernoulli}}]\min(n_1, n_2)$$

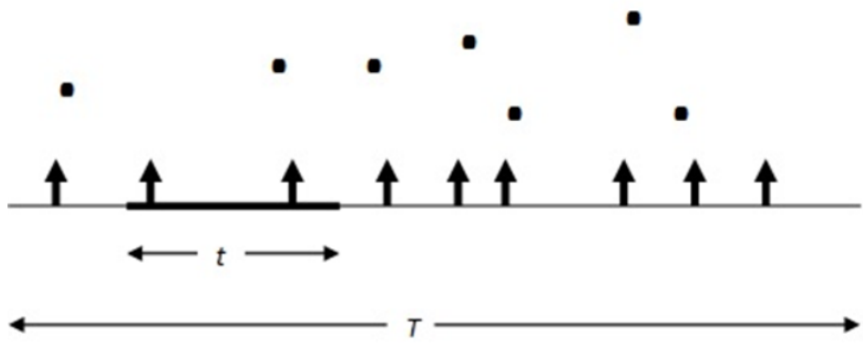
$$\text{Prob}\{X_{binomial}(n) \equiv \sum_{j=1}^{n} X_{Bernoulli}(n) = ra + (n-r)b\}$$

$$= \binom{n}{r} p^r (1-p)^{n-r} \ .$$

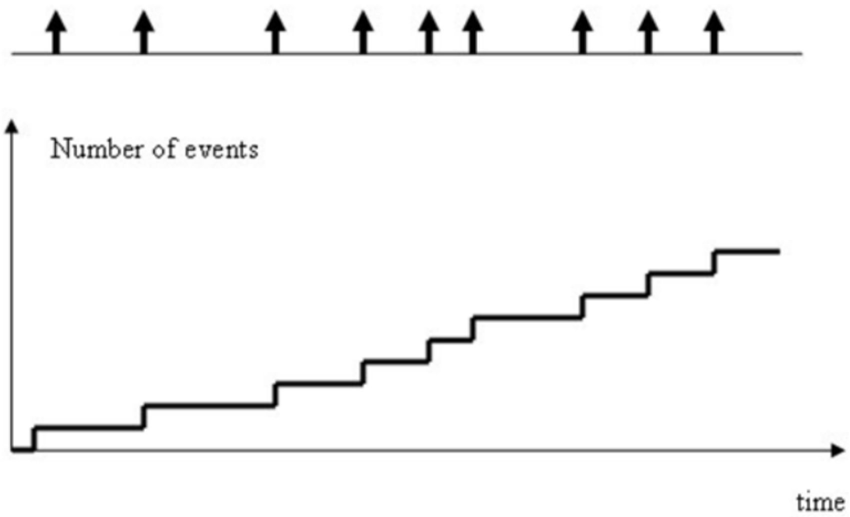


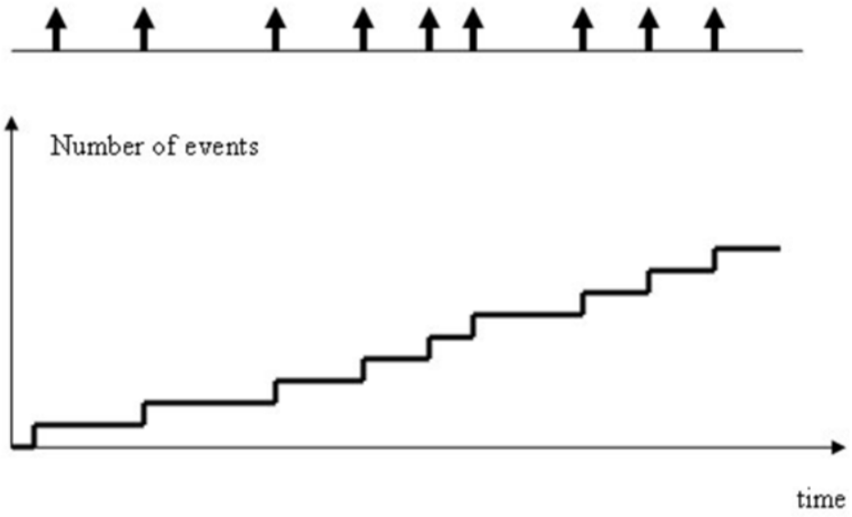
Random Arrivals:
probability of *r* arrivals in time *t* when the average number of arrivals per unit time is λ .

$$P(r;t) = \frac{(\lambda t)^r}{r!} e^{-\lambda t}$$



The *Poisson process* is the accumulation of random arrivals since $t = 0$.





electrons crossing a junction
customers arriving at MacDonald's
incoming calls at a information service
cars approaching a toll booth
airplanes coming in for service

QUEUEING THEORY:
How many servers should you hire to avoid long waiting times?

Random Arrivals:
probability of *r* arrivals in time *t* when the average number of arrivals per unit time is λ .

$$P(r;t) = \frac{(\lambda t)^r}{r!} e^{-\lambda t}$$

A little mathematical song and dance.
1. Is the total probability of 0, 1, 2, 3, ... arrivals in any interval equal to one?
2. Is the expected number of arrivals in time *t* equal to λt ?
3. Are the number of arrivals in distinct time intervals related, or independent?
4. What are the statistics for the waiting time, until the next arrival?
5. How about the next ***n*** arrivals?
6. What is the pdf, mean, and autocorrelation of the Poisson process (accumulated arrivals)?

$$P(r;t) = \frac{(\lambda t)^r}{r!} e^{-\lambda t}$$

1. Is the total probability of 0, 1, 2, 3, ... arrivals in any interval equal to one?

$$\sum_{n=0}^{\infty} \frac{(\lambda t)^n}{n!} e^{-\lambda t} = 1 \quad (\text{since } \sum_{n=0}^{\infty} \frac{(\lambda t)^n}{n!} = e^{\lambda t})$$

2. Is the expected number of arrivals in time *t* equal to λt ?

$$\sum_{n=0}^{\infty} n \frac{(\lambda t)^n}{n!} e^{-\lambda t} = \lambda t \sum_{n=0}^{\infty} \frac{(\lambda t)^{n-1}}{(n-1)!} e^{-\lambda t} = \lambda t(1)$$

$$(-1)! = \infty$$

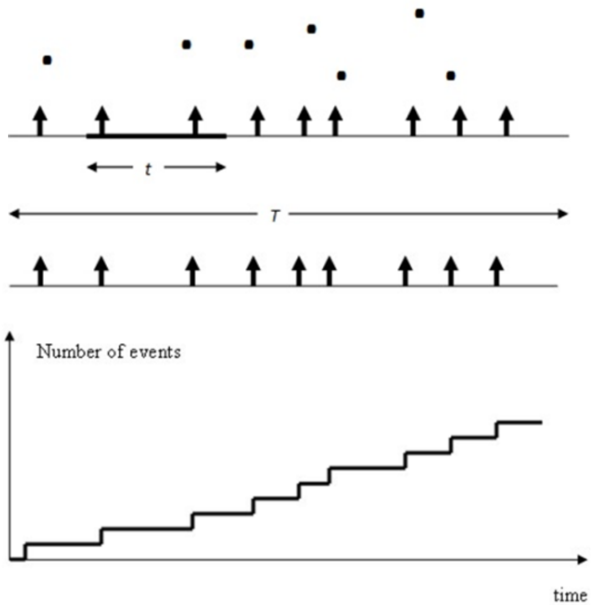
3. Are the number of arrivals in distinct time intervals related, or independent?

Yes. See page 131.

"Independent increments"

$$P(r;t) = \frac{(\lambda t)^r}{r!} e^{-\lambda t}$$

6. What is the pdf, mean, and autocorrelation of the Poisson process ?



$$\int_0^N f_{X_{Poisson}(t)}(x)\, dx =$$

probability that $x \leq N$

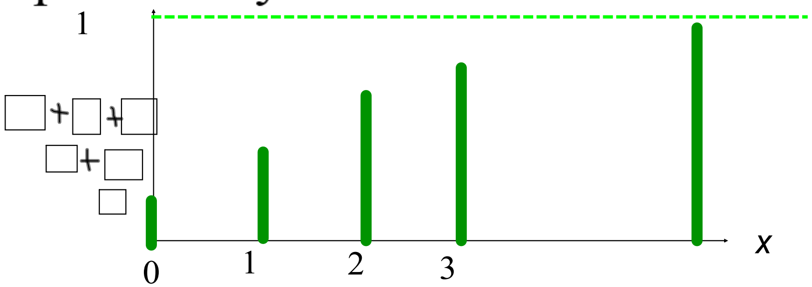
$$= \sum_{n=0}^{N} \frac{(\lambda t)^n}{n!} e^{-\lambda t}$$

As a function of *x* (with *t* fixed)

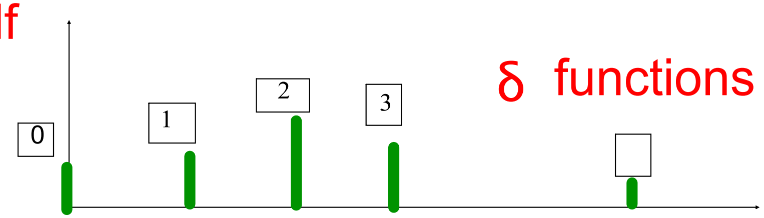
$$\int_0^N f_{X_{Poisson(t)}}(x)\,dx =$$

$$= \sum_{n=0}^{N} \frac{(\lambda t)^n}{n!} e^{-\lambda t}$$

probability that $x \leq N$



pdf



δ functions

$$f_{X_{Poisson(t)}}(x) = \sum_{n=0}^{\infty} \frac{(\lambda t)^n}{n!} e^{-\lambda t} \delta(x-n)$$

mean

$$f_{X_{Poisson}(t)}(x) = \sum_{n=0}^{\infty} \frac{(\lambda t)^n}{n!} e^{-\lambda t} \delta(x - n)$$

$$\mu_{Poisson}(t) = E\{X_{Poisson}(t)\} = \int_{-\infty}^{\infty} x\, f_{X_{Poisson}(t)}(x)\, dx$$

$$= \sum_{n=0}^{\infty} n \frac{(\lambda t)^n}{n!} e^{-\lambda t} = \lambda t \sum_{n=1}^{\infty} \frac{(\lambda t)^{n-1}}{(n-1)!} e^{-\lambda t}$$

$$= \lambda t \{ \sum_{r=0}^{\infty} \frac{(\lambda t)^r}{r!} e^{-\lambda t} \} = \lambda t \cdot (1)$$

autocorrelation

$$P(r;t) = \frac{(\lambda t)^r}{r!} e^{-\lambda t}$$

$$\sum_{n=0}^{\infty} \frac{(\lambda t)^n}{n!} e^{-\lambda t} = 1 \qquad \sum_{n=0}^{\infty} n \frac{(\lambda t)^n}{n!} e^{-\lambda t} = \lambda t$$

assume $t_1 < t_2$

$R_{\text{Poisson}}(t_1, t_2) =$

$$\sum_{n_1=0}^{\infty} \sum_{p=0}^{\infty} n_1 (n_1 + p) \frac{(\lambda t_1)^{n_1}}{n_1!} e^{-\lambda t_1} \frac{(\lambda [t_2 - t_1])^p}{p!} e^{-\lambda [t_2 - t_1]}$$

$$= \sum_{n_1=0}^{\infty} n_1^2 \frac{(\lambda t_1)^{n_1}}{n_1!} e^{-\lambda t_1} \sum_{p=0}^{\infty} \frac{(\lambda [t_2 - t_1])^p}{p!} e^{-\lambda [t_2 - t_1]}$$

$$+ \sum_{n_1=0}^{\infty} n_1 \frac{(\lambda t_1)^{n_1}}{n_1!} e^{-\lambda t_1} \sum_{p=0}^{\infty} p \frac{(\lambda [t_2 - t_1])^p}{p!} e^{-\lambda [t_2 - t_1]}$$

$$= \sum_{n_1=0}^{\infty} n_1^2 \frac{(\lambda t_1)^{n_1}}{n_1!} e^{-\lambda t_1} (1) + \sum_{n_1=0}^{\infty} n_1 \frac{(\lambda t_1)^{n_1}}{n_1!} e^{-\lambda t_1} (\lambda [t_2 - t_1])$$

$$= \sum_{n_1=0}^{\infty} n_1^2 \frac{(\lambda t_1)^{n_1}}{n_1!} e^{-\lambda t_1} + \lambda t_1 (\lambda [t_2 - t_1])$$

$$= \sum_{n_1=0}^{\infty} n_1 \frac{(\lambda t_1)^{n_1}}{(n_1 - 1)!} e^{-\lambda t_1} + \lambda t_1 (\lambda [t_2 - t_1])$$

autocorrelation   assume $t_1 < t_2$

$$= \sum_{n_1=0}^{\infty} n_1 \frac{(\lambda t_1)^{n_1}}{(n_1 - 1)!} e^{-\lambda t_1} + \lambda t_1 (\lambda [t_2 - t_1])$$

$$= \lambda t_1 \sum_{n_1=0}^{\infty} (n_1 - 1) \frac{(\lambda t_1)^{n_1 - 1}}{(n_1 - 1)!} e^{-\lambda t_1}$$

$$+ \lambda t_1 \sum_{n_1=0}^{\infty} (1) \frac{(\lambda t_1)^{n_1 - 1}}{(n_1 - 1)!} e^{-\lambda t_1} + \lambda t_1 (\lambda [t_2 - t_1])$$

$$= (\lambda t_1)(\lambda t_1) + \lambda t_1 (1) + \lambda t_1 (\lambda [t_2 - t_1])$$

$$= \lambda t_1 + \lambda^2 t_1 t_2$$

$$R_{\text{Poisson}}(t_1, t_2) = \lambda \min(t_1, t_2) + \lambda^2 t_1 t_2$$

Autocovariance

(Second moment identity)

$$C_{\text{Poisson}}(t_1, t_2) = R_{\text{Poisson}}(t_1, t_2) - E\{X(t_1)\}E\{(X(t_2)\}$$

$$= \lambda \min(t_1, t_2) + \lambda^2 t_1 t_2$$

$$-(\lambda t_1)(\lambda t_2) = \lambda \min(t_1, t_2)$$

$$\sigma^2 = \lambda t$$ (not stationary)

The mean and the variance of the Poisson probability are the same.

Autocovariance

(Second moment identity)

$$C_{\text{Poisson}}(t_1, t_2) = R_{\text{Poisson}}(t_1, t_2) - E\{X(t_1)\}E\{(X(t_2)\}$$

$$= \lambda \min(t_1, t_2) + \lambda^2 t_1 t_2$$

$$- (\lambda t_1)(\lambda t_2) = \lambda \min(t_1, t_2)$$

$$\sigma^2 = \lambda t$$

The mean and the variance of the Poisson probability are the same.

4. What are the statistics for the waiting time, until the next arrival?

$$P(r;t) = \frac{(\lambda t)^r}{r!} e^{-\lambda t}$$

1 arrival in *dt*, 0 arrivals in previous *t*

$$f(t)dt = \frac{\lambda dt\, e^{-\lambda dt}}{1!} \frac{(\lambda t)^0 e^{-\lambda t}}{0!}$$

$$= \lambda e^{-\lambda t} dt\, e^{-\lambda dt} = \lambda e^{-\lambda t} dt + O(dt^2)$$

$$f(t) = \lambda e^{-\lambda t}$$

The waiting time problem has an exponential pdf.

Mean = $1/\lambda$.

Standard deviation = $1/\lambda$.

5. How about the next ***n*** arrivals?

1 arrival in *dt*,

n-1 arrivals in previous *t*

$$f(t)\,dt = \frac{\lambda\,dt\,e^{-\lambda\,dt}}{1!}\,\frac{(\lambda t)^{n-1}\,e^{-\lambda t}}{(n-1)!}$$

$$f_{Erlang}(t) = \lambda\,\frac{(\lambda t)^{n-1}\,e^{-\lambda t}}{(n-1)!}$$

Lecture 16
March 20, 2017

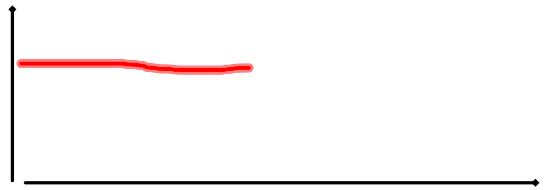
Consider two different ways to analyze a linear time invariant system:

1. Using the impulse response;

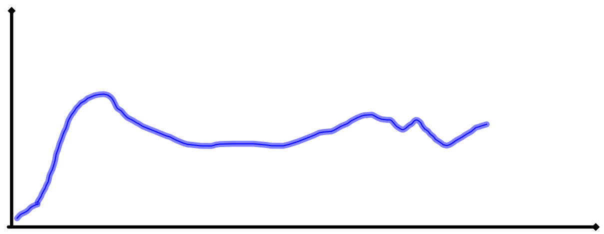
2. Using the frequency domain.

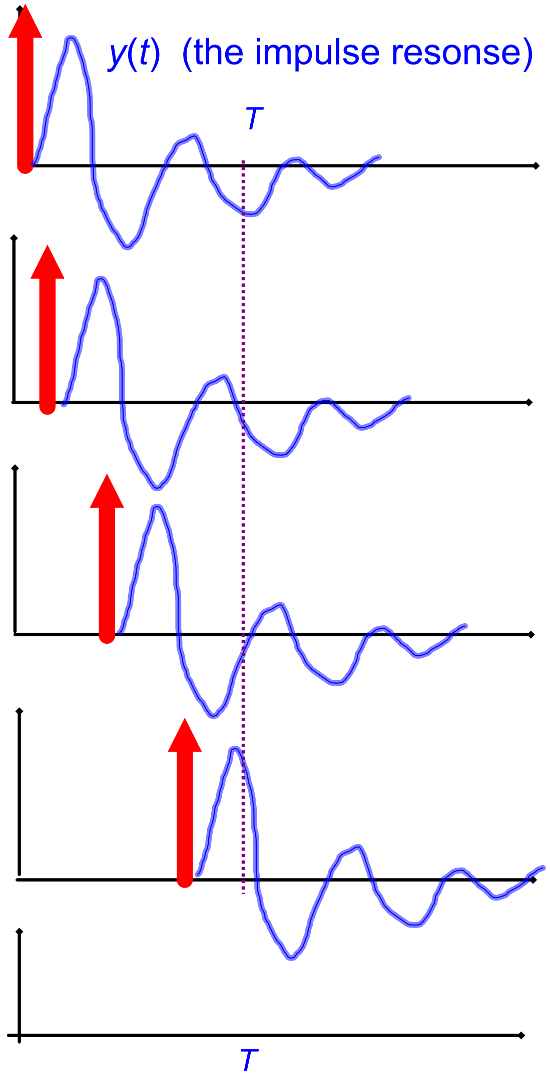
1. Use the impulse response.

Force



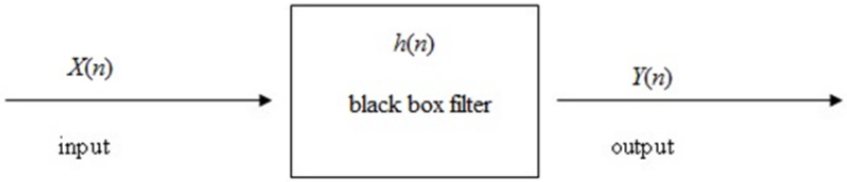
Response





$F(0)y(T) + F(\tau)y(T-\tau) + F(2\tau)y(T-2\tau)$
$+ F(3\tau)y(T-3\tau) + \ldots$
(CONVOLUTION!)

Figure 3.5 Linear time invariant system responses



$$Y(n) = \sum_{m=-\infty}^{\infty} X(m) h(n-m)$$

Impulse response

According to the Convolution Theorem

Transfer function

$$\hat{Y}(f) = \hat{H}(f)\hat{X}(f)$$

$$X(n) = \int_{-1/2}^{1/2} \tilde{X}(f) e^{j2\pi nf}\, df$$

$$Y(n) = \int_{-1/2}^{1/2} \hat{Y}(f) e^{j2\pi nf}\, df = \int_{-1/2}^{1/2} \hat{H}(f)\hat{X}(f) e^{j2\pi nf}\, df$$

Some students had trouble following this discussion of the transfer function. The topic is revisited in lectures 18 and 20.

Bottom Line:

If X(*f*) is the Fourier Transform of the input, and H(*f*) is the FT out the impulse response, then

H(*f*)X(*f*) is the FT of the output.

Frequency Domain Analysis

Express the input as a superposition of sinusoids:

Fourier Transform

Discrete

$$\tilde{X}(f) = \sum_{n=-\infty}^{\infty} X(n) e^{-j2\pi nf} \qquad X(n) = \int_{-1/2}^{1/2} \tilde{X}(f) e^{j2\pi nf}\, df$$

Continuous

$$\tilde{X}(f) = \int_{-\infty}^{\infty} X(t)\, e^{-i2\pi tf}\, dt$$

$$X(f) = \int_{-\infty}^{\infty} \tilde{X}(f)\, e^{i2\pi tf}\, df$$

Find the response to a single sinusoid $e^{j2\pi ft}$ (call it $Y_{sin}(t)$ (or $Y_{sin}(t;f)$).

It will be a multiple of $e^{j2\pi ft}$:

$$Y_{sin}(t;f) = \hat{Y}_{sin}(f) e^{j2\pi ft}$$

Then the output will be

$$Y(t) = \int_{-\infty}^{\infty} X(f)\, Y_{sin}(t;f) df$$

$$= \int_{-\infty}^{\infty} X(f)\, \hat{Y}_{sin}(f) e^{j2\pi ft} df$$

Compare.

$$Y(n) = \int_{-1/2}^{1/2} Y(f) e^{j2\pi nf} df = \int_{-1/2}^{1/2} H(f) X(f) e^{j2\pi nf} df$$

$$Y(t) = \int_{-\infty}^{\infty} X(f)\, Y_{sin}(t; f) df$$

$$= \int_{-\infty}^{\infty} X(f)\, \hat{Y}_{sin}(f) e^{j2\pi ft} df$$

So the transfer function is the response to the input $e^{j2\pi ft}$

(divided by $e^{j2\pi ft}$ **).**

Let the input be a sinusoid: $X(n) = e^{j2\pi nf}$

Since it's a linear time invariant system the outwill be easy to find.

And this output will be $H(f)$ times $e^{j2\pi nf}$

because $\hat{Y}(f) = \hat{H}(f)\hat{X}(f)$

So erase the $e^{j2\pi nf}$ and read off $H(f)$

Example. Suppose the system is

$$Y(n) = 0.5Y(n-1) + 0.1Y(n-2) + 2X(n) + X(n-1)$$

Let the input be $X(n) = e^{j2\pi nf}$. Then

$$\hat{H}(f)e^{j2\pi nf} = 0.5\hat{H}(f)e^{j2\pi(n-1)f} + 0.1\hat{H}(f)e^{j2\pi(n-2)f} + 2e^{j2\pi nf} + e^{j2\pi(n-1)f}$$

$$\hat{H}(f) = \frac{2 + e^{-j2\pi f}}{1 - 0.5e^{-j2\pi f} - 0.1e^{-j4\pi f}}$$

What if the input is a random process with PSD $S_X(f)$ and autocorrelation $R_X(n)$? What's the PSD and autocorrelation of the output $R_Y(n)$?

$$Y(n) = \sum_{m=-\infty}^{\infty} X(m)h(n-m)$$

$$R_Y(n) = \mathrm{E}\{Y(n)Y(0)\}$$

$$Y(n)Y(0) = \sum_{m=-\infty}^{\infty} h(n-m)\,X(m)\,Y(0) = \sum_{m=-\infty}^{\infty} h(n-m)\,X(m) \sum_{p=-\infty}^{\infty} h(-p)\,X(p)$$

$$R_Y(n) = \sum_{m=-\infty}^{\infty} h(n-m) \sum_{p=-\infty}^{\infty} h(-p)\,R_X(m-p)\text{; thus}$$

$$R_Y(n) = \sum_{m=-\infty}^{\infty} h(n-m)\,Q(m) \quad and \quad Q(m) = \sum_{p=-\infty}^{\infty} h(-p)\,R_X(m-p).$$

(a "double convolution")

Now take Fourier transforms:

$$S_Y(y) = \hat{H}(f)\widehat{Q(f)} = \hat{H}(f)\hat{H}_{rev}(f)S_X(f) = |\hat{H}(f)|^2\, S_X(f)$$

Example. Suppose the system is

$$Y(n) = 0.5Y(n-1) + 0.1Y(n-2) + 2X(n) + X(n-1)$$

Let the input be $X(n) = e^{j2\pi nf}$. Then

$$\hat{H}(f)e^{j2\pi nf} = 0.5\hat{H}(f)e^{j2\pi(n-1)f} + 0.1\hat{H}(f)e^{j2\pi(n-2)f} + 2e^{j2\pi nf} + e^{j2\pi(n-1)f}$$

$$\hat{H}(f) = \frac{2 + e^{-j2\pi f}}{1 - 0.5e^{-j2\pi f} - 0.1e^{-j4\pi f}}$$

Suppose $X(n)$ is white noise with power = 5.

$$S_X(f) \equiv 5$$

$$S_Y(y) = \hat{H}(f)\widehat{Q(f)} = \hat{H}(f)\hat{H}_{rev}(f)S_X(f) = |\hat{H}(f)|^2 S_X(f)$$

$$S_Y(f) = \left|\frac{2 + e^{-j2\pi f}}{1 - 0.5e^{-j2\pi f} - 0.1e^{-j4\pi f}}\right|^2 5$$

What's the autocorrelation of the output?

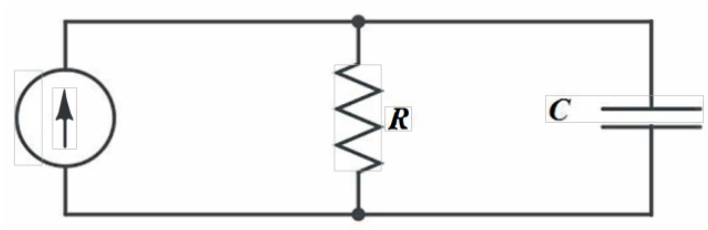
$$R_Y(n) = \int_{-0.5}^{0.5} \left|\frac{2 + e^{-j2\pi f}}{1 - 0.5e^{-j2\pi f} - 0.1e^{-j4\pi f}}\right|^2 5 \; e^{j2\pi nf}\, df$$

What's the power in the output?

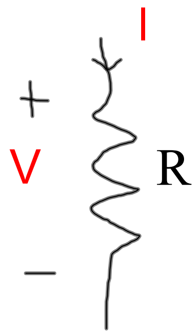
$$R_Y(0) = \int_{-0.5}^{0.5} \left|\frac{2 + e^{-j2\pi f}}{1 - 0.5e^{-j2\pi f} - 0.1e^{-j4\pi f}}\right|^2 5 \; df$$

7. a. What is the unit impulse voltage response of the RC circuit in Figure 3.11 to a current input?

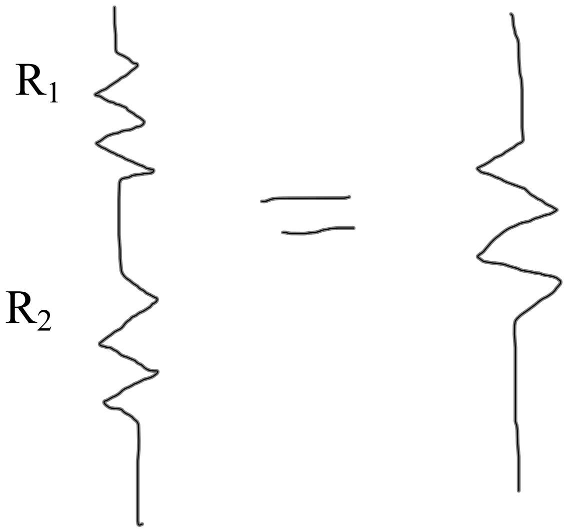
b. If the input current $X(t)$ has zero mean and autocorrelation (τ) and the output voltage $Y(t)$ has average power 10, express the integral for the output autocorrelation $R_Y(\tau)$.



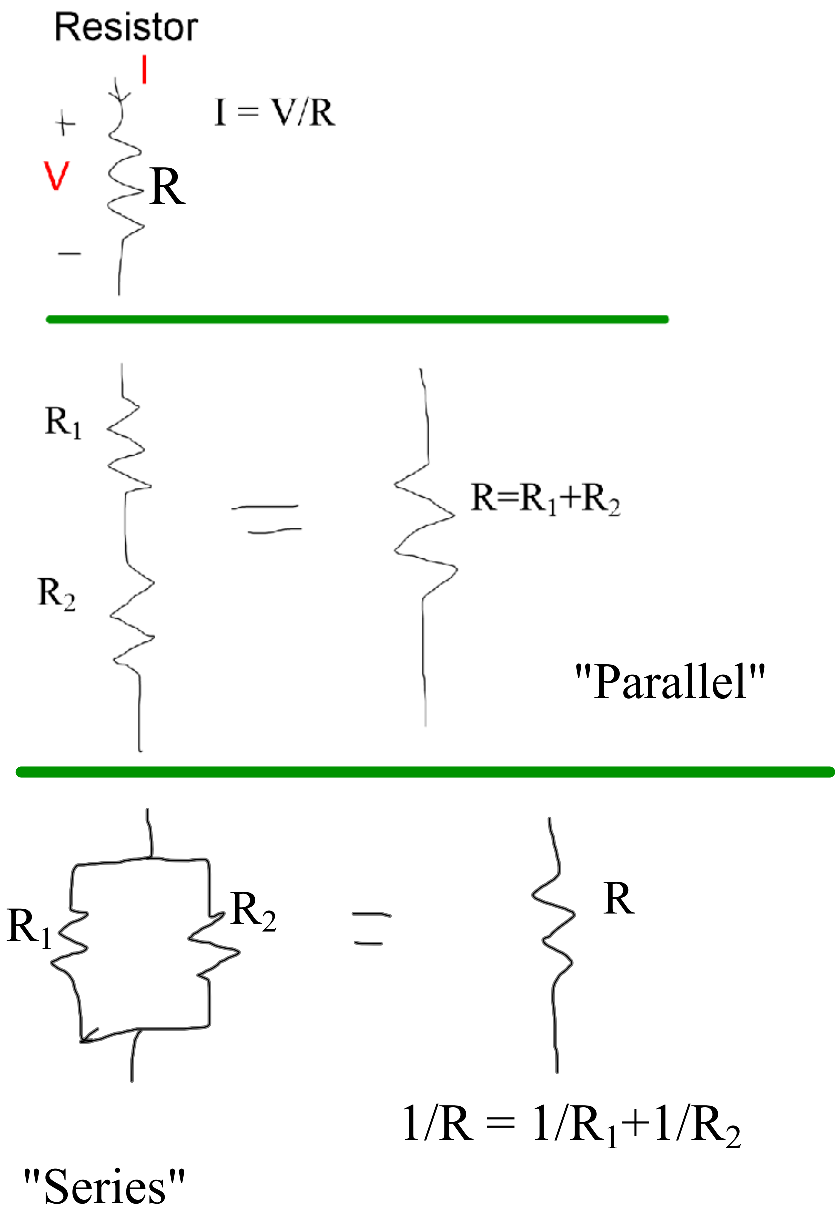
Resistor



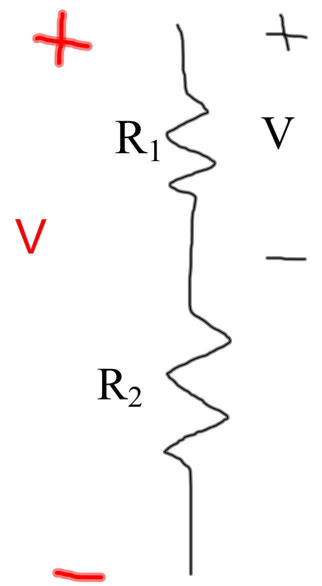
$$I = V/R$$



$$R=R_1+R_2$$

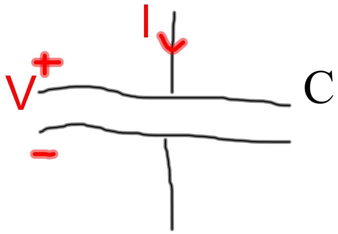

Resistor
I
I = V/R
+
V
R
−
R1
R2
R=R1+R2
"Parallel"
R1
R2
R
1/R = 1/R1+1/R2
"Series"

Voltage Divider



$V/\mathbf{V} = R_1/(R_1+R_2)$

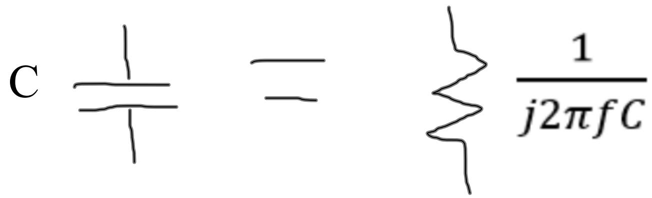
Capacitor in the frequency domain



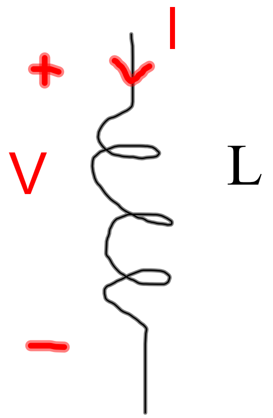
$$I = C\frac{dV}{dt}$$

$$V = e^{j2\pi ft} \qquad I = Cj2\pi f\; e^{j2\pi ft}$$

$$\frac{V}{I} = \frac{1}{j2\pi fC} = \text{"}R\text{"}$$



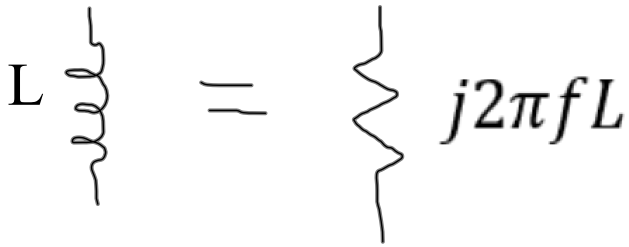
Inductor in the frequency domain



V = L dI/dt

$I = e^{j2\pi ft}$ $\quad$ $V = Lj2\pi f\ e^{j2\pi ft}$

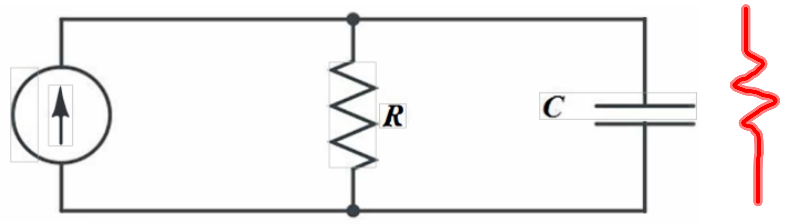
$$\frac{V}{I} = j2\pi fL = "R"$$



So the transfer function is the response to the input $e^{j2\pi ft}$

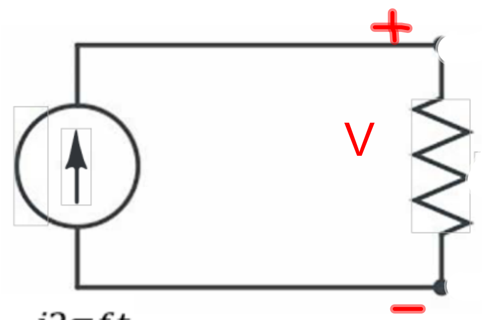
(divided by $e^{j2\pi ft}$ **).**

$$S_Y(y) = \hat{H}(f)\widehat{Q(f)} = \hat{H}(f)\hat{H}_{rev}(f)S_X(f) = |\hat{H}(f)|^2 S_X(f)$$



$e^{j2\pi ft}$

$$\frac{1}{j2\pi fC}$$

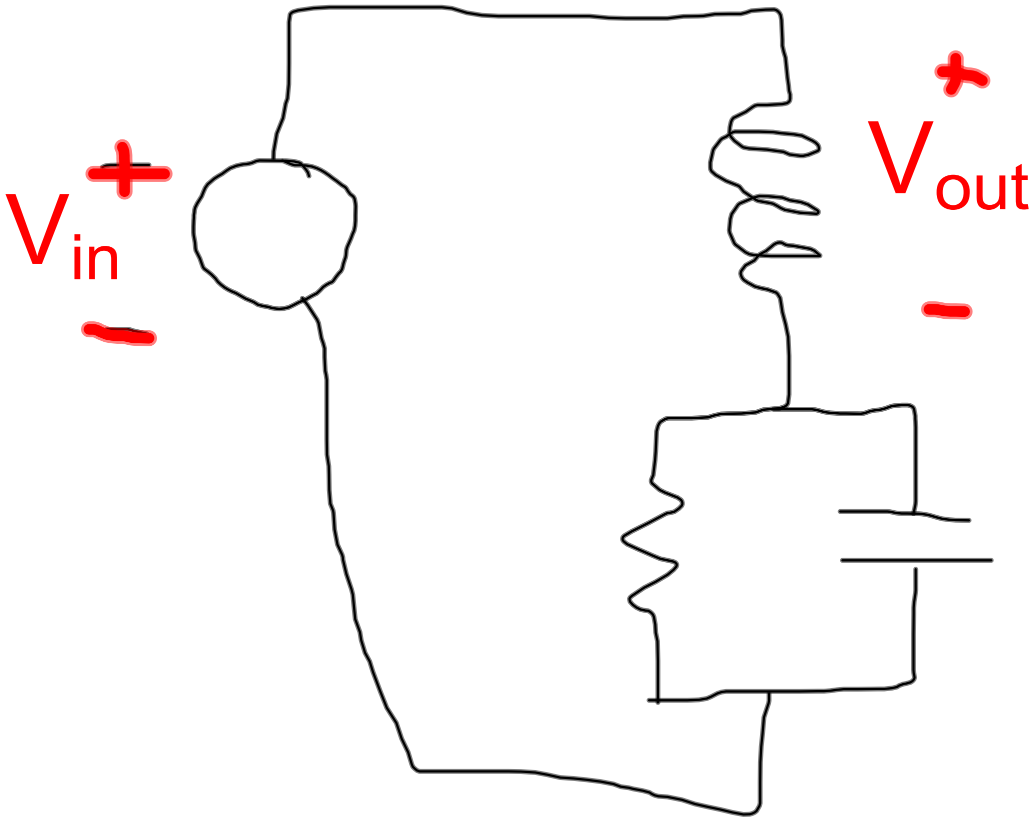


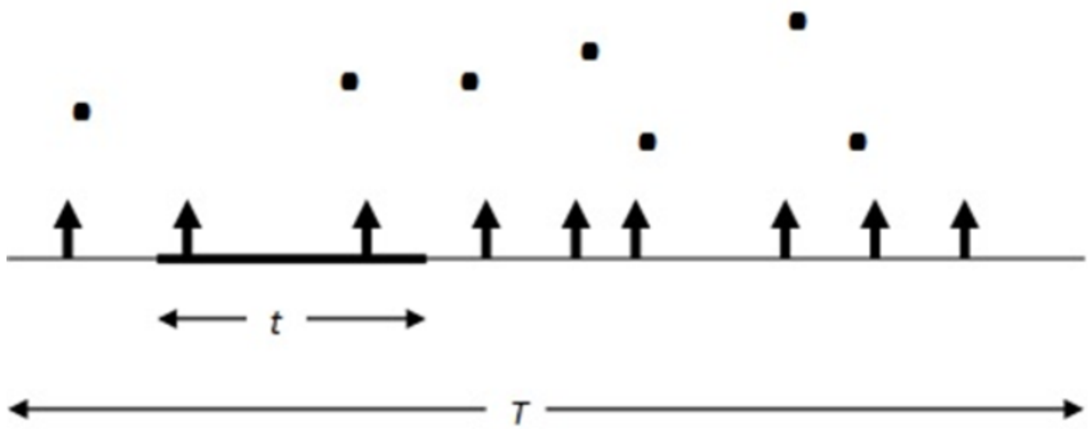
$$R_{AC} = \frac{1}{\frac{1}{R} + j2\pi fC}$$

$e^{j2\pi ft}$

$$V = IR_{AC} = e^{j2\pi ft} \frac{1}{\frac{1}{R} + j2\pi fC}$$

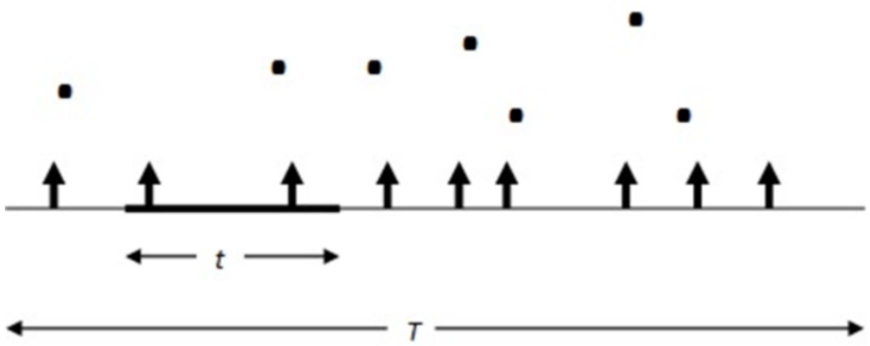
Next Assignment



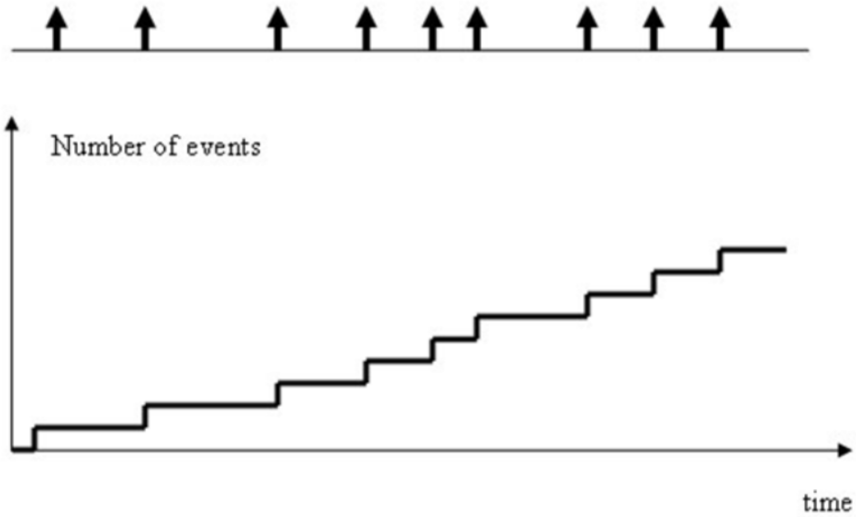


Random Arrivals:
probability of *r* arrivals in time *t* when the average number of arrivals per unit time is λ .

$$P(r;t) = \frac{(\lambda t)^r}{r!} e^{-\lambda t}$$



The *Poisson process* is the accumulation of random arrivals since $t = 0$.



Random Arrivals:

probability of *r* arrivals in time *t* when the average number of arrivals per unit time is λ .

$$P(r;t) = \frac{(\lambda t)^r}{r!} e^{-\lambda t}$$

A little mathematical song and dance.

1. Is the total probability of 0, 1, 2, 3, ... arrivals in any interval equal to one?
2. Is the expected number of arrivals in time *t* equal to λt ?
3. Are the number of arrivals in distinct time intervals related, or independent?
4. What are the statistics for the waiting time, until the next arrival?
5. How about the next ***n*** arrivals?
6. What is the pdf, mean, and autocorrelation of the Poisson process (accumulated arrivals)?

4. What are the statistics for the waiting time, until the next arrival?

$$f(t) = \lambda e^{-\lambda t}$$

The waiting time problem has an exponential pdf.

Mean = $1/\lambda$.

Standard deviation = $1/\lambda$.

5. How about the next ***n*** arrivals?

$$f_{Erlang}(t) = \lambda \frac{(\lambda t)^{n-1} e^{-\lambda t}}{(n-1)!}$$

6. What is the pdf, mean, and autocorrelation of the Poisson process (accumulated arrivals)?

$$f_{X_{Poisson}(t)}(x) = \sum_{n=0}^{\infty} \frac{(\lambda t)^n}{n!} e^{-\lambda t} \delta(x-n)$$

$$\mu_{Poisson}(t) = E\{X_{Poisson}(t)\} = \int_{-\infty}^{\infty} x\, f_{X_{Poisson}(t)}(x)\, dx$$

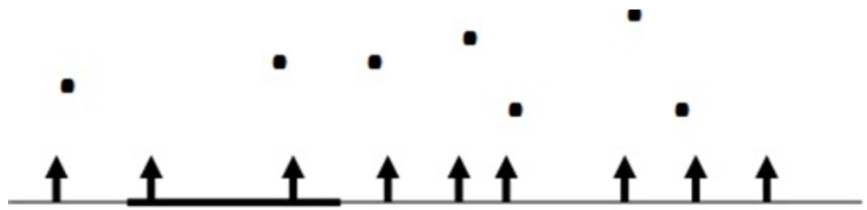
$$= \sum_{n=0}^{\infty} n \frac{(\lambda t)^n}{n!} e^{-\lambda t} = \lambda t \sum_{n=1}^{\infty} \frac{(\lambda t)^{n-1}}{(n-1)!} e^{-\lambda t}$$

$$= \lambda t \{\sum_{r=0}^{\infty} \frac{(\lambda t)^r}{r!} e^{-\lambda t}\} = \lambda t \cdot (1)$$

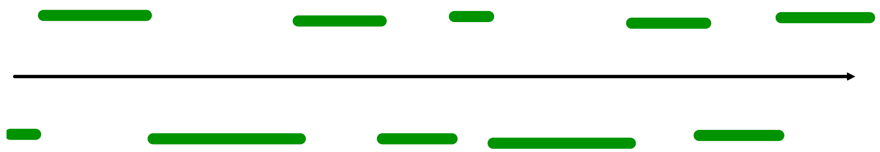
$$R_{\text{Poisson}}(t_1, t_2) = \lambda \min(t_1, t_2) + \lambda^2 t_1 t_2$$

$$C_{\text{Poisson}}(t_1, t_2) = \lambda \min(t_1, t_2)$$

$$P(r;t) = \frac{(\lambda t)^r}{r!} e^{-\lambda t}$$



Random Telegraph Signal: +a, -a



Mean = 0 .

Autocorrelation:

$X(t_1)X(t_2) = a^2$ (even # in $[t_1, t_2]$)

$= a^2$ (odd # in $[t_1, t_2]$)

$$R_{telegraph}(t_1,t_2) = E\{X(t_1)X(t_1+\tau)\}$$

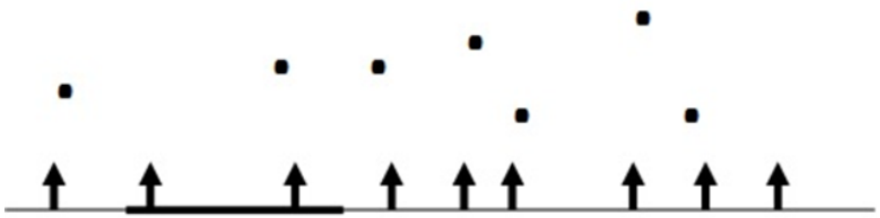
$$= +a^2 \sum_{n=0,2,\ldots} \frac{(\lambda\tau)^n}{n!} e^{-\lambda\tau} - a^2 \sum_{n=1,3,\ldots} \frac{(\lambda\tau)^n}{n!} e^{-\lambda\tau}$$

$$= a^2 \sum_{n=0}^{\infty} \frac{(-\lambda\tau)^n}{n!} e^{-\lambda\tau} = a^2 e^{-2\lambda\tau} \sum_{n=0}^{\infty} \frac{(-\lambda\tau)^n}{n!} e^{+\lambda\tau}$$

$$= a^2 e^{-2\lambda\tau}$$

$$S_{telegraph}(f) = a^2 \frac{\lambda}{(\pi f)^2 + \lambda^2}$$

$$P(r;t) = \frac{(\lambda t)^r}{r!} e^{-\lambda t}$$



SHOT Noise

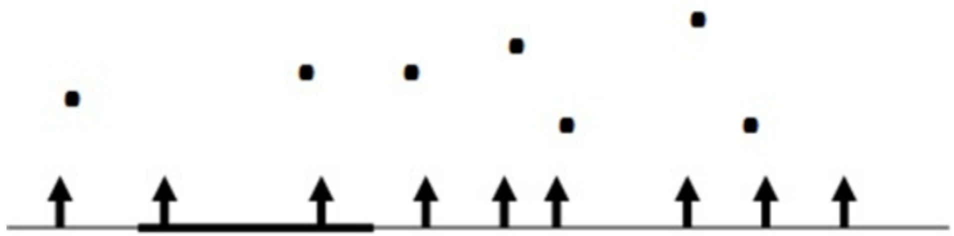
$$X_{Poisson}(t) = \int_{-\infty}^{t} X_{shot}(\tau) d\tau$$

$$X_{shot}(t) = \frac{d}{dt} X_{Poisson}(t)$$

$$\mu_{shot} = E\{X_{shot}(t)\} = E\{\frac{d}{dt} X_{Poisson}(t)\}$$

$$= \frac{d}{dt} E\{X_{Poisson}(t)\} = \frac{d}{dt}(\lambda t) = \lambda$$

2. Is the expected number of arrivals in time *t* equal to λt ? ✓



SHOT Noise $X_{shot}(t) = \frac{d}{dt} X_{Poisson}(t)$

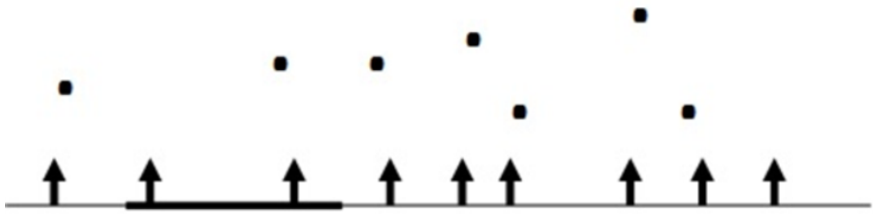
$$R_{X'}(t_1, t_2) \equiv E\{\frac{dX(t_1)}{dt_1} \frac{dX(t_2)}{dt_2}\}$$

$$\frac{dX(t_1)}{dt_1} \frac{dX(t_2)}{dt_2} = \frac{\partial}{\partial t_1}[\frac{\partial}{\partial t_2} X(t_1)X(t_2)]$$

$$E\{\frac{\partial}{\partial t_1}[\frac{\partial}{\partial t_2} X(t_1)X(t_2)]\} = \frac{\partial}{\partial t_1}[\frac{\partial}{\partial t_2} E\{X(t_1)X(t_2)\}]$$

6. What is the pdf, mean, and autocorrelation of the Poisson process (accumulated arrivals)?

$$R_{\text{Poisson}}(t_1, t_2) = \lambda \min(t_1, t_2) + \lambda^2 t_1 t_2$$



SHOT Noise $X_{shot}(t) = \frac{d}{dt} X_{Poisson}(t)$

$$R_{X'}(t_1,t_2) \equiv E\{\frac{dX(t_1)}{dt_1}\frac{dX(t_2)}{dt_2}\}$$

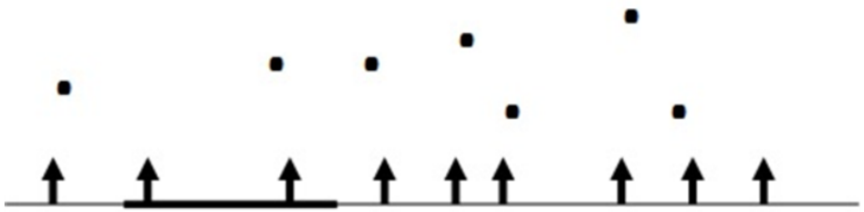
$$\frac{dX(t_1)}{dt_1}\frac{dX(t_2)}{dt_2} = \frac{\partial}{\partial t_1}[\frac{\partial}{\partial t_2}X(t_1)X(t_2)]$$

$$E\{\frac{\partial}{\partial t_1}[\frac{\partial}{\partial t_2}X(t_1)X(t_2)]\} = \frac{\partial}{\partial t_1}[\frac{\partial}{\partial t_2}E\{X(t_1)X(t_2)\}]$$

$$R_{\mathrm{Poisson}}(t_1, t_2) = \begin{matrix} \lambda t_1 + \lambda^2 t_1 t_2 & if\ t_2 \geq t_1, \\ \lambda t_2 + \lambda^2 t_1 t_2 & if\ t_2 \leq t_1 \end{matrix}$$

$$\frac{\partial R_{Poisson}(t_1,t_2)}{\partial t_2} = \begin{matrix} \lambda^2 t_1 & if\ t_2 \geq t_1 \\ \lambda + \lambda^2 t_1 & if\ t_2 \leq t_1 \end{matrix}$$

$$= \lambda u(t_1 - t_2) + \lambda^2 t_1$$



SHOT Noise $X_{shot}(t) = \frac{d}{dt} X_{Poisson}(t)$

$$R_{\text{Poisson}}(t_1, t_2) = \begin{cases} \lambda t_1 + \lambda^2 t_1 t_2 & if \; t_2 \geq t_1 , \\ \lambda t_2 + \lambda^2 t_1 t_2 & if \; t_2 \leq t_1 \end{cases}$$

$$\frac{\partial R_{Poisson}(t_1,t_2)}{\partial t_2} = \begin{cases} \lambda^2 t_1 & if \; t_2 \geq t_1 \\ \lambda + \lambda^2 t_1 & if \; t_2 \leq t_1 \end{cases}$$

$$= \lambda u(t_1 - t_2) + \lambda^2 t_1$$

$$R_{shot}(t_1,t_2) = \frac{\partial}{\partial t_1} \frac{\partial R_{Poisson}(t_1,t_2)}{\partial t_2}$$

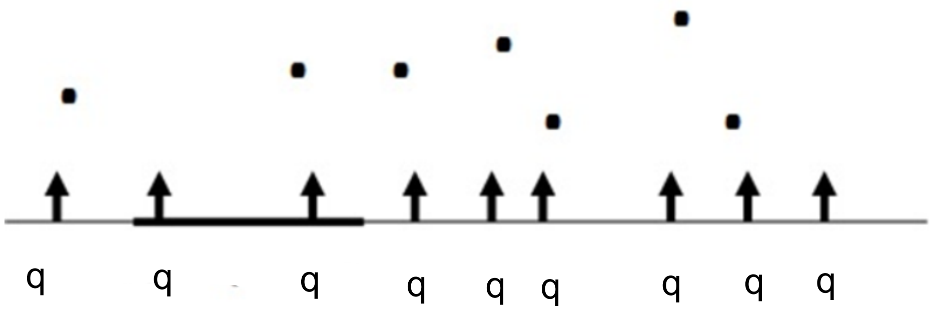
$$= \lambda \delta(t_1 - t_2) + \lambda^2 = R_{shot}(| t_1 - t_2 |)$$

$$R_{shot}(t_1,t_2) = \frac{\partial}{\partial t_1}\frac{\partial R_{Poisson}(t_1,t_2)}{\partial t_2}$$

$$= \lambda\delta(t_1 - t_2) + \lambda^2$$

$$S_{shot}(f) = \int_{-\infty}^{\infty} R_{shot}(\tau)e^{-j2\pi f\tau}\, d\tau = \lambda^2\delta(f) + \lambda$$

Current Noise



$$I(t) = q\, X_{shot}(t)$$

$$E\{I(t)\} = \qquad q\lambda = I_{DC}$$

$$R_I(t_1, t_2)$$

$$= q^2 \lambda \delta(\tau) + q^2 \lambda^2 = q I_{DC} \delta(\tau) + I_{DC}^2$$

$$S_I(f) = I_{DC}^2 \delta(f) + q I_{DC}$$

Fourier Transform

$$\widehat{F(f)} = \int_{-\infty}^{\infty} F(t)\, e^{-j2\pi f t} dt$$

Fourier Integral Representation

$$F(t) = \int_{-\infty}^{\infty} \widehat{F(f)}\, e^{j2\pi f t} df$$

$$\text{all } t \ in \ (-\infty, \infty)$$

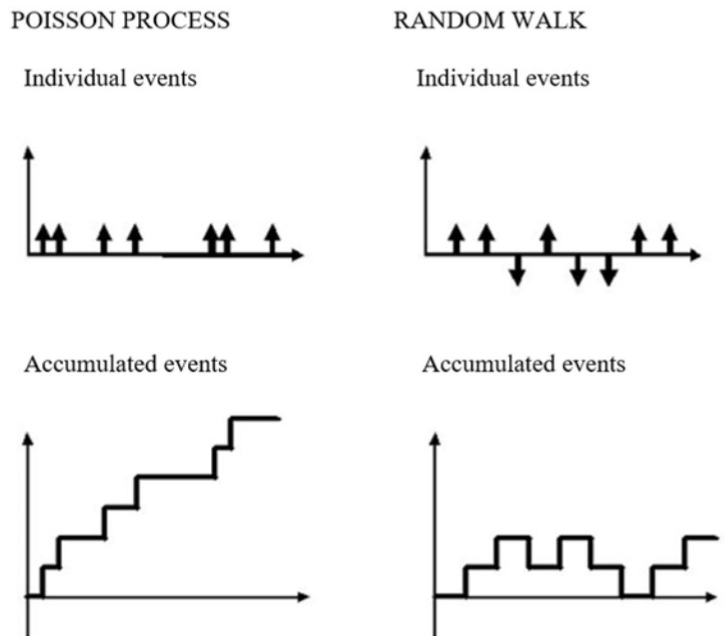
http://www.thefouriertransform.com/pairs/fourier.php

If $F(t)$ is an impulse $\delta(t)$, then

$$\widehat{F(f)} = \int_{-\infty}^{\infty} F(t)\, e^{-j2\pi f t} dt \ = 1$$

Lecture 17

March 22, 2017



Prob. that $X(n) = r$ is $\binom{n}{r} p^r (1-p)^{n-r}$

$$\mu_{\text{binomial}}(n) = E\{X(n)\} = np$$

$$R_{\text{binomial}}(n_1, n_2) = n_1 n_2 p^2 + (p - p^2)\min(n_1, n_2)$$

$$\mu_{\text{RandWalk}} = E\{X_{\text{RandWalk}}(n\tau)\} = nE\{x_1\} = n(2p-1)s$$

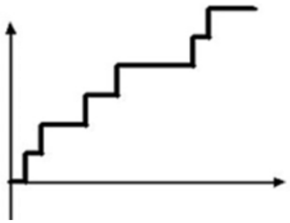
$$R_{\text{RandWalk}}(n_1\tau, n_2\tau) = s^2 \min(n_1, n_2)$$

(if $p = ½$)

POISSON PROCESS

Individual events

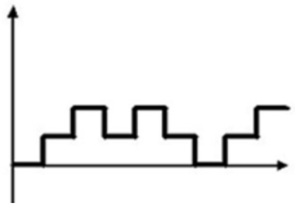
Accumulated events



RANDOM WALK

Individual events

Accumulated events



Prob. that $X(n) = r$ is $\binom{n}{r} p^r (1-p)^{n-r}$

ps + (1-p)(-s)

$\mu_{\text{binomial}}(n) = E\{X(n)\} = np$

$R_{\text{binomial}}(n_1, n_2) = n_1 n_2 p^2 + (p - p^2)\min(n_1, n_2)$

$\mu_{\text{RandWalk}} = E\{X_{\text{RandWalk}}(n\tau)\} = nE\{x_1\} = n(2p-1)s$

$$R_{\text{RandWalk}}(n_1\tau, n_2\tau) = s^2 \min(n_1, n_2)$$

(if $p = ½$)

$$\mu_{RandWalk} = E\{X_{RandWalk}(n\tau)\} = nE\{x_1\} = n(2p-1)s$$

$$R_{RandWalk}(n_1\tau, n_2\tau) = s^2 \min(n_1, n_2)$$

(if $p = ½$)

WIENER Process

step size s -> 0

time step τ -> 0

p = 1/2

$$E\{X_{Wiener}(t)\} = 0,$$

$$R_{Wiener}(t_1, t_2) = \frac{s^2}{\tau} \min(t_1, t_2)$$

$$R_{Wiener}(t,t) = \alpha t$$

WIENER Process

$$E\{X_{\text{Wiener}}(t)\} = 0,$$

$$R_{\text{Wiener}}(t_1, t_2) = \frac{s^2}{\tau} \min(t_1, t_2)$$

$$R_{\text{Wiener}}(t,t) = \alpha t$$

BROWNIAN Motion is the derivative of the Wiener Process.

$$E\{dW(t)/dt\} = 0$$

$$R_{dW/dt}(t_1, t_2) = \frac{\partial}{\partial t_1} \frac{\partial}{\partial t_2} R_{\text{Wiener}}(t_1, t_2)$$

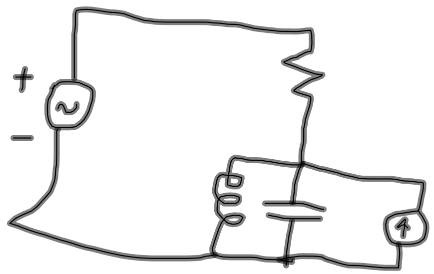
$$= \alpha\delta(t_1 - t_2) = \alpha\delta(\tau)$$

White Noise

Markov Processes is deferred to
Lecture 21, April 12, 2017

Lecture 18

March 27, 2017



Summary: Important Facts about Change of Variable

$$\text{pdf of } y = g(x):\quad f_Y(y) = \sum_{\substack{preimages\\ of\ y}} \frac{f_X(x_i)}{|g'(x_i)|}$$

$$E\{aX+b\} = a\mu_X + b\,,\quad \sigma_{aX+b} = |a|\,\sigma_X\,.$$

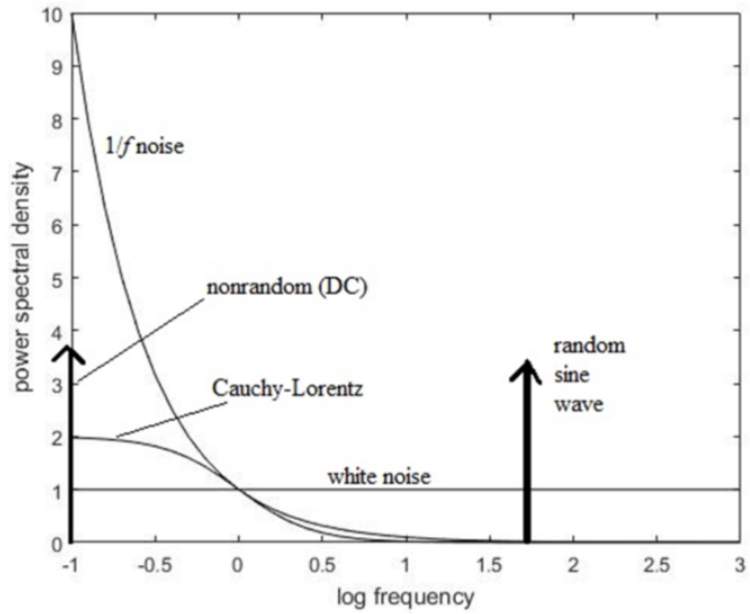
So the statistics of an ideal voltage source and a (zero-mean) noise voltage are:

mean = ideal;

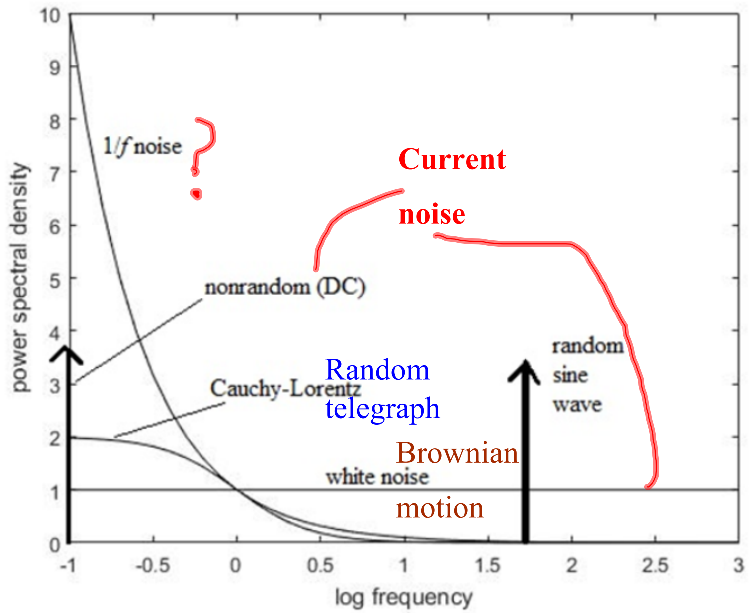
standard deviation = noise

(We will elaborate on this in the next lecture.)

This part of the lecture, on noise in electrical circuits, was incomprehensible to many students. The next lecture will go over this again in more detail, so don't panic yet!



power spectral density
log frequency
1/f noise
nonrandom (DC)
Cauchy-Lorentz
white noise
random
sine
wave



1/f noise
Current noise
nonrandom (DC)
Random telegraph
Cauchy-Lorentz
Brownian motion
white noise
random sine wave
power spectral density
log frequency

Chapter 5. Least Mean-Square Error Predictors

5.1 The Optimal Constant Predictor

5.2 The Optimal Constant-Multiple Predictor

5.3 Digression: Orthogonality

5.4 Multivariate LMSE Prediction: The Normal Equations

5.5 The Bias

5.6 Best Straight-Line Predictor

5.7 Prediction for a Random Process

5.8 Interpolation, Smoothing, Extrapolation, and Back-Prediction

5.9 The Wiener Filter

The Optimal Constant Predictor

Try to predict Y with
some constant: $\hat{Y} = C$

$E\{(Y\text{-}C)^2\}$ = [mean of $(Y$ - $C)]^2$ + [standard deviation of $(Y$ - $C)]^2$

$$(\bar{Y} - C)^2 + \sigma_Y^2 .$$

$$C = \bar{Y}$$

The *root*-mean-squared "RMS" error is its standard deviation σ_Y

Optimal Constant-Multiple Predictors

$$\hat{Y} = wX$$

$$MSE = E\{(\hat{Y} - Y)^2\} = E\{(wX - Y)^2\}$$

$$= E\{w^2 X^2 + Y^2 - 2wXY\} = w^2 \overline{X^2} + \overline{Y^2} - 2w\overline{XY}$$

$$\frac{\partial}{\partial w} E\{(\hat{Y} - Y)^2\} = 2w\overline{X^2} - 2\overline{XY} = 0$$,

or $$w = \frac{\overline{XY}}{\overline{X^2}} = \frac{R_{XY}}{\overline{X^2}}$$,

$$\hat{Y} = wX = \frac{R_{XY}}{\overline{X^2}} X$$.

Propose a formula for the predictor;

formulate the prediction *error* by subtracting the quantity to be predicted;

square and take expected value to formulate the mean-squared error;

minimize the mean squared error (MSE) by setting its derivatives,
with respect to the parameters, equal to zero.

Predict Y using *two* data

$$\hat{Y} = w_1 X_1 + w_2 X_2$$

$$E\{(w_1X_1 + w_2X_2 - Y)^2\}$$
$$= E\{w_1^2X_1^2 + w_2^2X_2^2 + Y^2 + 2w_1w_2X_1X_2 - 2w_1X_1Y - 2w_2X_2Y\}$$
$$= w_1^2\overline{X_1^2} + w_2^2\overline{X_2^2} + \overline{Y^2} + 2w_1w_2\overline{X_1X_2} - 2w_1\overline{X_1Y} - 2w_2\overline{X_2Y}$$

$$\frac{\partial}{\partial w_1}E\{(w_1X_1 + w_2X_2 - Y)^2\} = 2w_1\overline{X_1^2} + 2w_2\overline{X_1X_2} - 2\overline{X_1Y} = 0$$

$$\frac{\partial}{\partial w_2}E\{(w_1X_1 + w_2X_2 - Y)^2\} = 2w_2\overline{X_2^2} + 2w_1\overline{X_1X_2} - 2\overline{X_2Y} = 0$$

$$w_1 = \frac{\overline{X_1Y}\ \overline{X_2^2} - \overline{X_2Y}\ \overline{X_1X_2}}{\overline{X_1^2}\ \overline{X_2^2} - \overline{X_1X_2}^2},\quad w_2 = \frac{\overline{X_2Y}\ \overline{X_1^2} - \overline{X_1Y}\ \overline{X_1X_2}}{\overline{X_1^2}\ \overline{X_2^2} - \overline{X_1X_2}^2}$$

$$\hat{Y} = w_1 X_1 + w_2 X_2$$

$$E\{(w_1 X_1 + w_2 X_2 - Y)^2\}$$
$$= E\{w_1^2 X_1^2 + w_2^2 X_2^2 + Y^2 + 2w_1 w_2 X_1 X_2 - 2w_1 X_1 Y - 2w_2 X_2 Y\}$$
$$= w_1^2 \overline{X_1^2} + w_2^2 \overline{X_2^2} + \overline{Y^2} + 2w_1 w_2 \overline{X_1 X_2} - 2w_1 \overline{X_1 Y} - 2w_2 \overline{X_2 Y}$$

$$\frac{\partial}{\partial w_1} E\{(w_1 X_1 + w_2 X_2 - Y)^2\} = 2w_1 \overline{X_1^2} + 2w_2 \overline{X_1 X_2} - 2\overline{X_1 Y} = 0$$

$$\frac{\partial}{\partial w_2} E\{(w_1 X_1 + w_2 X_2 - Y)^2\} = 2w_2 \overline{X_2^2} + 2w_1 \overline{X_1 X_2} - 2\overline{X_2 Y} = 0$$

$$w_1 = \frac{\overline{X_1 Y}\ \overline{X_2^2} - \overline{X_2 Y}\ \overline{X_1 X_2}}{\overline{X_1^2}\ \overline{X_2^2} - \overline{X_1 X_2}^2},\quad w_2 = \frac{\overline{X_2 Y}\ \overline{X_1^2} - \overline{X_1 Y}\ \overline{X_1 X_2}}{\overline{X_1^2}\ \overline{X_2^2} - \overline{X_1 X_2}^2}$$

$$2w_1\overline{X_1^2} + 2w_2\overline{X_1X_2} - 2\overline{X_1Y} = 0$$

$$2w_2\overline{X_2^2} + 2w_1\overline{X_1X_2} - 2\overline{X_2Y} = 0$$

$$\hat{Y} = w_1X_1 + w_2X_2 + \cdots + w_nX_n$$

$$\begin{bmatrix} \overline{X_1^2} & \overline{X_1X_2} & \cdots & \overline{X_1X_n} \\ \overline{X_2X_1} & \overline{X_2^2} & \cdots & \overline{X_2X_n} \\ & & \vdots & \\ \overline{X_nX_1} & \overline{X_nX_2} & \cdots & \overline{X_n^2} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} = \begin{bmatrix} \overline{X_1Y} \\ \overline{X_2Y} \\ \vdots \\ \overline{X_nY} \end{bmatrix}$$

$$\hat{Y} = w_1 X_1 + w_2 X_2 + \cdots + w_n X_n$$

$$\begin{bmatrix} \overline{X_1^2} & \overline{X_1 X_2} & \cdots & \overline{X_1 X_n} \\ \overline{X_2 X_1} & \overline{X_2^2} & \cdots & \overline{X_2 X_n} \\ & & \vdots & \\ \overline{X_n X_1} & \overline{X_n X_2} & \cdots & \overline{X_n^2} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} = \begin{bmatrix} \overline{X_1 Y} \\ \overline{X_2 Y} \\ \vdots \\ \overline{X_n Y} \end{bmatrix}$$

NORMAL EQUATIONS

$$\hat{Y} = w_1 X_1 + w_2 X_2 + \cdots + w_n X_n$$

$$\begin{bmatrix} X_1 & X_2 & \cdots X_n \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} \approx Y$$

Premultiply by $\begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix}$ and take expected values

$$\hat{Y} = w_1 X_1 + w_2 X_2 + \cdots + w_n X_n$$

$$\begin{bmatrix} X_1 & X_2 & \cdots X_n \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} \approx \qquad Y$$

$$\begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix} \qquad\qquad \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix}$$

$$\begin{bmatrix} \overline{X_1^2} & \overline{X_1 X_2} & \cdots & \overline{X_1 X_n} \\ \overline{X_2 X_1} & \overline{X_2^2} & \cdots & \overline{X_2 X_n} \\ & & \vdots & \\ \overline{X_n X_1} & \overline{X_n X_2} & \cdots & \overline{X_n^2} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} = \begin{bmatrix} \overline{X_1 Y} \\ \overline{X_2 Y} \\ \vdots \\ \overline{X_n Y} \end{bmatrix}$$

UNBIASED Predictor

$$E\{\hat{Y}\} = E\{Y\}$$

If X_1 and X_2 are zero-mean,

$$E\{\hat{Y}\} = w_1 E\{X_1\} + w_2 E\{X_2\} = 0.$$

The cure: $\hat{Y} = w_1 X_1 + w_2 X_2 + w_3 X_3$

"random variable" X_3 is the *fixed* number "1".

Normal equations

$$\begin{bmatrix} \overline{X_1^2} & \overline{X_1 X_2} & \overline{X_1 X_3} \\ \overline{X_2 X_1} & \overline{X_2^2} & \overline{X_2 X_3} \\ \overline{X_3 X_1} & \overline{X_3 X_2} & \overline{X_3^2} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} = \begin{bmatrix} \overline{X_1 Y} \\ \overline{X_2 Y} \\ \overline{X_3 Y} \end{bmatrix}$$

$$\begin{bmatrix} \overline{X_1^2} & \overline{X_1 X_2} & \overline{X_1} \\ \overline{X_2 X_1} & \overline{X_2^2} & \overline{X_2} \\ \overline{X_1} & \overline{X_2} & 1 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} = \begin{bmatrix} \overline{X_1 Y} \\ \overline{X_2 Y} \\ \overline{Y} \end{bmatrix}$$

UNBIASED Predictor

$$E\{\hat{Y}\} = E\{Y\}$$

The cure: $\hat{Y} = w_1 X_1 + w_2 X_2 + w_3 X_3$

Normal equations

$$\begin{bmatrix} \overline{X_1^2} & \overline{X_1 X_2} & \overline{X_1} \\ \overline{X_2 X_1} & \overline{X_2^2} & \overline{X_2} \\ \overline{X_1} & \overline{X_2} & 1 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} = \begin{bmatrix} \overline{X_1 Y} \\ \overline{X_2 Y} \\ \overline{Y} \end{bmatrix}$$

$$\hat{Y} = w_1 X_1 + w_2 X_2 + C$$

$$C = E\{\hat{Y}\} - w_1 E\{X_1\} - w_2 E\{X_2\}$$

Should one always use bias-corrected predictors?

(See page 157)

Best Straight-Line Predictor

$$y = mx + b$$

$$\hat{Y} = wX + C$$

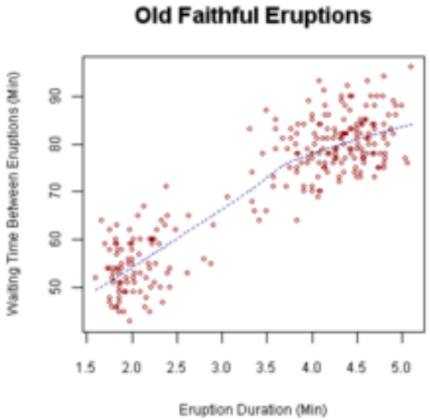
$$= w_1 X_1 + w_2 X_2$$

$$w_1 = w,\ X_1 = X,\ w_2 = C,\ and\ X_2 \equiv 1$$

Solution of normal equations

$$w_1 = \frac{\overline{XY} - \overline{X}\,\overline{Y}}{\overline{X^2} - \overline{X}^2} = \frac{\rho\sigma_X\sigma_Y}{\sigma_X^2},\ C = \frac{\overline{X^2}\,\overline{Y} - \overline{X}\,\overline{XY}}{\overline{X^2} - \overline{X}^2} = \overline{Y} - w_1\overline{X}$$

"scatter diagram" $\{x_1,y_1;\ x_2,y_2;\ x_3,y_3;\ldots;\ x_k,y_k\}$



$$y = mx + b$$

$$m = \frac{k\sum_{i=1}^{k} x_i y_i - \sum_{i=1}^{k} x_i \sum_{i=1}^{k} y_i}{k\sum_{i=1}^{k} x_i^2 - (\sum_{i=1}^{k} x_i)^2}$$

$$b = \frac{(\sum_{i=1}^{k} x_i^2)(\sum_{i=1}^{k} y_i) - (\sum_{i=1}^{k} x_i)(\sum_{i=1}^{k} x_i y_i)}{k\sum_{i=1}^{k} x_i^2 - (\sum_{i=1}^{k} x_i)^2}$$

$$\overline{X} \approx \frac{\sum_{i=1}^{k} x_i}{k}$$

$$\overline{X^2} \approx \frac{\sum_{i=1}^{k} x_i^2}{k}$$

$$\overline{XY} \approx \frac{\sum_{i=1}^{k} x_i y}{k}$$

$$[X_1\ X_2 \cdots X_n]\begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} \approx Y$$

$$\begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix} \qquad \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix}$$

$$\begin{bmatrix} \overline{X_1^2} & \overline{X_1X_2} & \cdots & \overline{X_1X_n} \\ \overline{X_2X_1} & \overline{X_2^2} & \cdots & \overline{X_2X_n} \\ & & \vdots & \\ \overline{X_nX_1} & \overline{X_nX_2} & \cdots & \overline{X_n^2} \end{bmatrix}\begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} = \begin{bmatrix} \overline{X_1Y} \\ \overline{X_2Y} \\ \vdots \\ \overline{X_nY} \end{bmatrix}$$

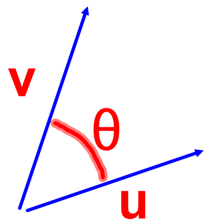
$$\begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix} \left[[X_1\ X_2 \cdots X_n]\begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} - Y \right] = 0$$

(error)

$\vec{v} \cdot \vec{u} = |\vec{v}| \cos\theta \; |\vec{u}|,$

$\overline{XY} = \sigma_X \, \rho \, \sigma_Y \,,$

(if X &Y are zero mean)



$\vec{v} \cdot \vec{v} = |\vec{v}|^2$ $\qquad$ $\vec{u} \cdot \vec{u} = |\vec{u}|^2$

$\overline{XX} = \sigma_X^2 \,,$ $\qquad$ $\overline{YY} = \sigma_Y^2$

X and Y are "orthogonal" if $E\{XY\} = 0$

(even if X &Y are *not* zero mean)

$$\begin{bmatrix} \overline{X_1^2} & \overline{X_1X_2} & \cdots & \overline{X_1X_n} \\ \overline{X_2X_1} & \overline{X_2^2} & \cdots & \overline{X_2X_n} \\ & \vdots & & \\ \overline{X_nX_1} & \overline{X_nX_2} & \cdots & \overline{X_n^2} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} = \begin{bmatrix} \overline{X_1Y} \\ \overline{X_2Y} \\ \vdots \\ \overline{X_nY} \end{bmatrix}$$

$$\begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix} \left\{ [X_1\ X_2 \cdots X_n] \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} - Y \right\} = 0$$

error

$$[X_1\ X_2 \cdots X_n] \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} - Y$$

"the error is orthogonal to each X_i"

Prediction for a Random Process

Recall

$$\hat{Y} = w_1 X_1 + w_2 X_2 + \cdots + w_n X_n$$

$$\begin{bmatrix} X_1 & X_2 & \cdots & X_n \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} \approx Y$$

$$\begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix} \qquad \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix}$$

$$\begin{bmatrix} \overline{X_1^2} & \overline{X_1 X_2} & \cdots & \overline{X_1 X_n} \\ \overline{X_2 X_1} & \overline{X_2^2} & \cdots & \overline{X_2 X_n} \\ & & \vdots & \\ \overline{X_n X_1} & \overline{X_n X_2} & \cdots & \overline{X_n^2} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} = \begin{bmatrix} \overline{X_1 Y} \\ \overline{X_2 Y} \\ \vdots \\ \overline{X_n Y} \end{bmatrix}$$

$$\hat{Y} = \begin{bmatrix} X(n) & X(n-1) & \dots & X(1) \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \dots \\ w_n \end{bmatrix}$$

$\hat{Y}$ is the predictor for $X(n+1)$. Multiply both sides by $X(n)$, $X(n-1)$, ..., $X(1)$ and take expectation.

$$\begin{bmatrix} R_X(n,n) & R_X(n,n-1) \dots & R_X(n,1) \\ R_X(n-1,n) & R_X(n-1,n-1) \dots & R_X(n-1,1) \\ & \dots & \\ R_X(1,n) & R_X(1,n-1) \dots & R_X(1,1) \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \dots \\ w_n \end{bmatrix} = \begin{bmatrix} R_X(n,n+1) \\ R_X(n-1,n+1) \\ \dots \\ R_X(1,n+1) \end{bmatrix}$$

Stationary?

$$\begin{bmatrix} R_X(0) & R_X(1) & R_X(2) & R_X(3) & R_X(4) \\ R_X(1) & R_X(0) & R_X(1) & R_X(2) & R_X(3) \\ R_X(2) & R_X(1) & R_X(0) & R_X(1) & R_X(2) \\ R_X(3) & R_X(2) & R_X(1) & R_X(0) & R_X(1) \\ R_X(4) & R_X(3) & R_X(2) & R_X(1) & R_X(0) \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \\ w_5 \end{bmatrix} = \begin{bmatrix} R_X(1) \\ R_X(2) \\ R_X(3) \\ R_X(4) \\ R_X(5) \end{bmatrix}$$

Interpolation, Smoothing,

Extrapolation, and Back-Prediction

$$\hat{Y} = X(n) \approx w_1 X(n-1) + w_2 X(n+1)$$

$$\begin{bmatrix} X_1 & X_2 & \cdots & X_n \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} \approx Y$$

$$\begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix} \qquad \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix}$$

$$\begin{bmatrix} R_X(n-1,n-1) & R_X(n-1,n+1) \\ R_X(n+1,n-1) & R_X(n+1,n+1) \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} R_X(n-1,n) \\ R_X(n+1,n) \end{bmatrix}$$

$$\begin{bmatrix} R_X(0) & R_X(2) \\ R_X(2) & R_X(0) \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} R_X(1) \\ R_X(1) \end{bmatrix}$$

Interpolation, Smoothing,

Extrapolation, and Back-Prediction

$$X(n) \approx w_1 X(n-1) + w_2 X(n) + w_3 X(n+1)$$

Y

$$\begin{bmatrix} X_1 & X_2 & \cdots & X_n \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} \approx Y$$

$$\begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix} \qquad \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix}$$

$$\begin{bmatrix} R_X(n-1,n-1) & R_X(n-1,n) & R_X(n-1,n+1) \\ R_X(n,n-1) & R_X(n,n) & R_X(n,n+1) \\ R_X(n+1,n-1) & R_X(n+1,n) & R_X(n+1,n+1) \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} = \begin{bmatrix} R_X(n-1,n) \\ R_X(n,n) \\ R_X(n+1,n) \end{bmatrix}$$

Interpolation, Smoothing,

Extrapolation, and Back-Prediction

$$X(n+m) \approx w_1 X(n) + w_2 X(n-1) + w_3 X(n-2)$$

$$\begin{bmatrix} X_1 & X_2 & \cdots & X_n \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} \approx Y$$

$$\begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix} \qquad \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix}$$

$$\begin{bmatrix} R_X(0) & R_X(1) & R_X(2) \\ R_X(1) & R_X(0) & R_X(1) \\ R_X(2) & R_X(1) & R_X(0) \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} = \begin{bmatrix} R_X(m) \\ R_X(m+1) \\ R_X(m+2) \end{bmatrix}$$

Interpolation, Smoothing,

Extrapolation, and Back-Prediction

$$X(n) \approx w_1 X(n+1) + w_2 X(n+2)$$

$$\begin{bmatrix} X_1 & X_2 & \cdots & X_n \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} \approx Y$$

$$\begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix} \qquad \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix}$$

$$\begin{bmatrix} R_X(0) & R_X(1) \\ R_X(1) & R_X(0) \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} R_X(1) \\ R_X(2) \end{bmatrix}$$

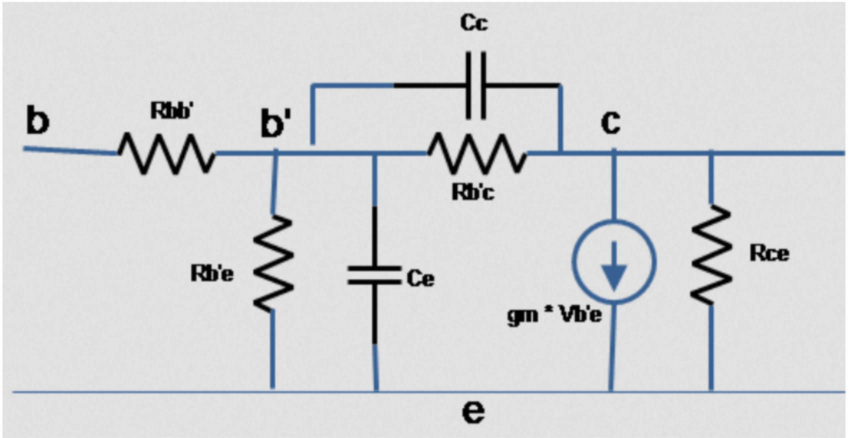
Lecture 18

March 29, 2017

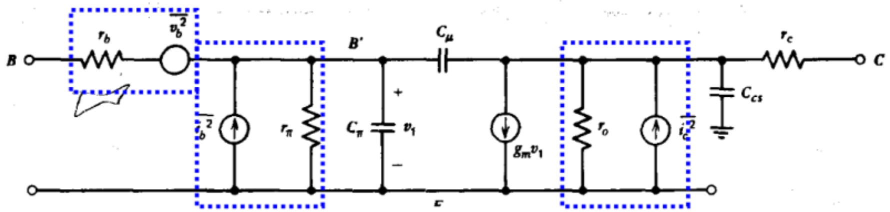
A little more discussion about
circuit noise analysis.

(Not in the book)

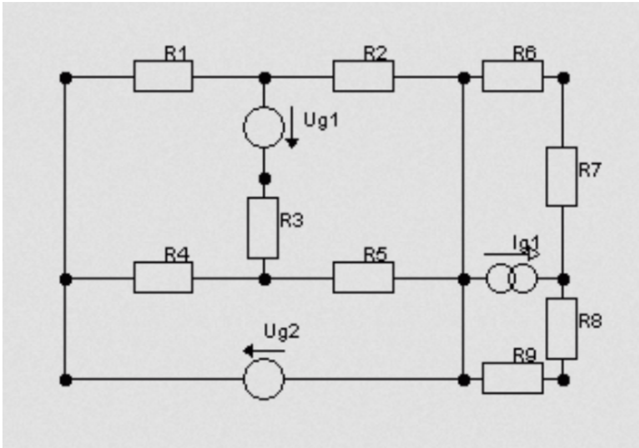
Bipolar junction transistor model



Add resistor and junction noise sources



Superposition



$$\begin{bmatrix} Z & Z & Z & Z & Z & Z \\ Z & Z & Z & Z & Z & Z \\ Z & Z & Z & Z & Z & Z \\ Z & Z & Z & Z & Z & Z \\ Z & Z & Z & Z & Z & Z \end{bmatrix} \begin{bmatrix} V \\ I \\ V \\ I \\ V \end{bmatrix} = \begin{bmatrix} V \\ V \\ I \\ I \\ V \end{bmatrix}$$

unknowns

sources

$$
\begin{bmatrix} V \\ I \\ V \\ I \\ V \end{bmatrix} = \begin{vmatrix} Z & Z & Z & Z & Z & Z \\ Z & Z & Z & Z & Z & Z \\ Z & Z & Z & Z & Z & Z \\ Z & Z & Z & Z & Z & Z \\ Z & Z & Z & Z & Z & Z \end{vmatrix}^{-1} \begin{vmatrix} V \\ V \\ I \\ I \\ V \end{vmatrix}
$$

$$
\begin{bmatrix} V \\ I \\ V \\ I \\ V \end{bmatrix} = \begin{vmatrix} Z & Z & Z & Z & Z & Z \\ Z & Z & Z & Z & Z & Z \\ Z & Z & Z & Z & Z & Z \\ Z & Z & Z & Z & Z & Z \\ Z & Z & Z & Z & Z & Z \end{vmatrix}^{-1} \begin{vmatrix} V \\ 0 \\ 0 \\ 0 \\ 0 \end{vmatrix}
$$

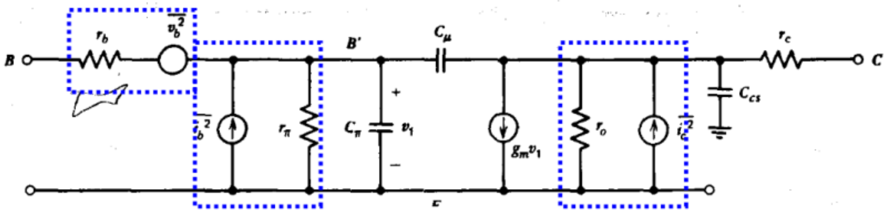
$+$

$$
\begin{bmatrix} V \\ I \\ V \\ I \\ V \end{bmatrix} = \begin{vmatrix} Z & Z & Z & Z & Z & Z \\ Z & Z & Z & Z & Z & Z \\ Z & Z & Z & Z & Z & Z \\ Z & Z & Z & Z & Z & Z \\ Z & Z & Z & Z & Z & Z \end{vmatrix}^{-1} \begin{vmatrix} 0 \\ V \\ 0 \\ 0 \\ 0 \end{vmatrix}
$$

$+ \dots$

General principle for multiple sources:
zero out all but one source, find the voltages and currents;
pick another source, zero out the others, find the voltages and currents;

- - -
add the voltages and currents.



In the frequency domain (AC analysis), the voltage (current) at any one location due to a source at another location equals the transfer function (between the two locations) time the source voltage (current).

$$V_{out}(f)e^{j2\pi ft} = \sum_{sources\ k} H_{to\ "out"}^{from\ "k"}(f)\ [V(f) \text{ or } I(f)]_{source\ "k"}\ e^{j2\pi ft}$$

$$= \sum_k H_{out}^k(f)\ V_k(f)\ e^{j2\pi ft}$$

Some students had trouble following the transfer function formulation. It is reviewed in more detail in lecture 20.

$$V_{out}(f)e^{j2\pi ft}$$

$$= \sum\nolimits_{ideal\ sources\ k} H_{to\ "out"}^{from\ "k"}(f)V(f)_{source\ k}^{ideal}\ e^{j2\pi ft}$$

$$+ \sum\nolimits_{noise\ sources\ k} H_{to\ "out"}^{from\ "k"}(f)V(f)_{source\ k}^{noise}\ e^{j2\pi ft}$$

and

$$V_{out}(t) = \int_{-\infty}^{\infty} V_{out}(f)e^{j2\pi ft}\,df$$

$$= \sum\nolimits_{ideal} \cdots + \sum\nolimits_{noise} \cdots$$

We don't know the noise voltage - it's random.

We know two things about the noise:
its mean (zero);
its power spectral density.

$$V_{out}(f)e^{j2\pi ft}$$

$$= \sum\nolimits_{ideal\ sources\ k} H_{to\ "out"}^{from\ "k"}(f)V(f)_{source\ k}^{ideal}\ e^{j2\pi ft}$$

$$+ \sum\nolimits_{noise\ sources\ k} H_{to\ "out"}^{from\ "k"}(f)V(f)_{source\ k}^{noise}\ e^{j2\pi ft}$$

$$V_{out}(t) = \int_{-\infty}^{\infty} V_{out}(f)e^{j2\pi ft}\, df$$

$$= \sum\nolimits_{ideal} \cdots + \sum\nolimits_{noise} \cdots$$

Take the expected value. Noise drops out. Perform regular circuit analysis to get the expected output.) (Use superposition of currents and voltages.)

$$V_{out}(f)e^{j2\pi ft}$$

$$= \sum\nolimits_{ideal\ sources\ k} H_{to\ "out"}^{from\ "k"}(f)V(f)_{source\ k}^{ideal}\ e^{j2\pi ft}$$

~~$+ \sum\nolimits_{noise\ sources\ k} H_{to\ "out"}^{from\ "k"}(f)V(f)_{source\ k}^{noise}\ e^{j2\pi ft}$~~

Now look at power.

$$V_{out}^2(t) = \{\int_{-\infty}^{\infty} V_{out}(f)e^{j2\pi ft}\,df\}^2$$

$$= \int_{f_1=-\infty}^{\infty}\int_{f_2=-\infty}^{\infty}\{\sum_{ideal} HV + \sum_{noise} HV\}\{\sum_{ideal} HV + \sum_{noise} HV\}\cdots$$

Take the expected value. Noise voltages are uncorrelated with each other and with ideal voltages.

$$\mathbf{E}\{V_{out}^2(t)\} = \int_{f_1=-\infty}^{\infty}\int_{f_2=-\infty}^{\infty}\{\sum_{ideal} HV\}\{\sum_{ideal} HV\}\cdots$$

$$+ \sum_{noise\ sources} \mathbf{E}\{HV_k\}\{HV_k\}\cdots$$

$$= |\mathbf{E}\{V_{out}(t)\}|^2$$

$$+ \sum\nolimits_{noise\ sources}[output\ noise\ power]$$

$$\mathbf{E}\{V_{out}^2(t)\}$$

$$= \left|\mathbf{E}\{V_{out}(t)\}\right|^2$$

$$+\sum\nolimits_{noise\ sources}[output\ noise\ power]$$

To compute output power with noise, you don't use superposition of voltages and currents; you use superposition of power.

Compute the ideal power.
Add the output power of each noise source.

And of course remember, for each noise source,

$$S_{out}(f) = |H(f)|^2 S_{in}(f)$$

The Wiener Filter

X_1, X_2, and Y

BUT $Z_1 = X_1 + e_1$, $Z_2 = X_2 + e_2$

and we seek

$$\hat{Y} = w_1 Z_1 + w_2 Z_2$$

$$\begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix} \begin{bmatrix} X_1 & X_2 & \cdots & X_n \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} \approx \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix} Y$$

$$\begin{bmatrix} \overline{Z_1^2} & \overline{Z_1 Z_2} \\ \overline{Z_1 Z_2} & \overline{Z_2^2} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} \overline{Z_1 Y} \\ \overline{Z_2 Y} \end{bmatrix}$$

$$Z_1 = X_1 + e_1 \ , \ Z_2 = X_2 + e_2$$

$$\begin{bmatrix} \overline{Z_1^2} & \overline{Z_1 Z_2} \\ \overline{Z_1 Z_2} & \overline{Z_2^2} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} \overline{Z_1 Y} \\ \overline{Z_2 Y} \end{bmatrix}$$

$$\overline{Z_1^2} = E\{(X_1 + e_1)^2\} = \overline{X_1^2} + 2\overline{e_1 X_1} + \overline{e_1^2}$$
$$= \overline{X_1^2} + 2\overline{e_1}\,\overline{X_1} + \overline{e_1^2} = \sigma_{X_1}^2 + 0 + \sigma_{e_1}^2 \ ,$$

$$\overline{Z_2^2} = \sigma_{X_2}^2 + \sigma_{e_2}^2$$

$$Z_1 = X_1 + e_1\ ,\ Z_2 = X_2 + e_2$$

$$\begin{bmatrix} \overline{Z_1^2} & \overline{Z_1 Z_2} \\ \overline{Z_1 Z_2} & \overline{Z_2^2} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} \overline{Z_1 Y} \\ \overline{Z_2 Y} \end{bmatrix}$$

$$\overline{Z_2^2} = \sigma_{X_2}^2 + \sigma_{e_2}^2$$

$$\overline{Z_1 Z_2} = E\{(X_1 + e_1)(X_2 + e_2)\}$$
$$= \overline{X_1 X_2} + \overline{e_1}\,\overline{X_2} + \overline{e_2}\,\overline{X_1} + \overline{e_1}\,\overline{e_2}$$
$$= \overline{X_1 X_2} + 0 + 0 + 0$$

$$\overline{Z_1 Y} = E\{(X_1 + e_1)(Y)\}$$
$$= \overline{X_1 Y} + \overline{e_1}\,\overline{Y} = \overline{X_1 Y} + 0$$

$$\overline{Z_2 Y} = \overline{X_2 Y}$$

$$Z_1 = X_1 + e_1 \,,\; Z_2 = X_2 + e_2$$

$$\begin{bmatrix} \overline{Z_1^2} & \overline{Z_1 Z_2} \\ \overline{Z_1 Z_2} & \overline{Z_2^2} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} \overline{Z_1 Y} \\ \overline{Z_2 Y} \end{bmatrix}$$

$$\begin{bmatrix} \sigma_{X_1}^2 + \sigma_{e_1}^2 & \overline{X_1 X_2} \\ \overline{X_1 X_2} & \sigma_{X_2}^2 + \sigma_{e_2}^2 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} \overline{X_1 Y} \\ \overline{X_2 Y} \end{bmatrix}$$

$$\left(\begin{bmatrix} \sigma_{X_1}^2 & \overline{X_1 X_2} & \ldots & \overline{X_1 X_p} \\ \overline{X_2 X_1} & \sigma_{X_2}^2 & \ldots & \overline{X_2 X_p} \\ & & \ldots & \\ \overline{X_p X_1} & \overline{X_p X_2} & \ldots & \sigma_{X_p}^2 \end{bmatrix} + \begin{bmatrix} \sigma_{e_1}^2 & 0 & \ldots & 0 \\ 0 & \sigma_{e_2}^2 & \ldots & 0 \\ & & \ldots & \\ 0 & 0 & \ldots & \sigma_{e_p}^2 \end{bmatrix} \right) \begin{bmatrix} w_1 \\ w_2 \\ \ldots \\ w_p \end{bmatrix} = \begin{bmatrix} \overline{X_1 Y} \\ \overline{X_2 Y} \\ \ldots \\ \overline{X_p Y} \end{bmatrix}$$

Wiener filter

infinite-length Wiener interpolator

$$\begin{bmatrix} \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \iddots \\ \cdots & R_X(0)+R_e(0) & R_X(-1)+R_e(-1) & R_X(-2)+R_e(-2) & R_X(-3)+R_e(-3) & R_X(-4)+R_e(-4) & \cdots \\ \cdots & R_X(1)+R_e(1) & R_X(0)+R_e(0) & R_X(-1)+R_e(-1) & R_X(-2)+R_e(-2) & R_X(-3)+R_e(-3) & \cdots \\ \cdots & R_X(2)+R_e(2) & R_X(1)+R_e(1) & R_X(0)+R_e(0) & R_X(-1)+R_e(-1) & R_X(-2)+R_e(-2) & \cdots \\ \cdots & R_X(3)+R_e(3) & R_X(2)+R_e(2) & R_X(1)+R_e(1) & R_X(0)+R_e(0) & R_X(-1)+R_e(-1) & \cdots \\ \cdots & R_X(4)+R_e(4) & R_X(3)+R_e(3) & R_X(2)+R_e(2) & R_X(1)+R_e(1) & R_X(0)+R_e(0) & \cdots \\ \iddots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} \times$$

$$\begin{bmatrix} \vdots \\ w(-2) \\ w(-1) \\ w(0) \\ w(1) \\ w(2) \\ \vdots \end{bmatrix} = \begin{bmatrix} \vdots \\ R_X(-2) \\ R_X(-1) \\ R_X(0) \\ R_X(1) \\ R_X(2) \\ \vdots \end{bmatrix}.$$

On the diagonal $R_X(0) + R_e(0)$

Off the diagonal $R_X(1) + R_e(1)$

$$R_X(n) = \sum_{p=-\infty}^{\infty} [R_X(n-p) + [R_e(n-p)]w(p)$$

$$[S_X(f) + S_e(f)]W(f) = S_X(f)$$

$$W(f) = \frac{S_X(f)}{S_X(f)+S_e(f)}$$

infinite-length Wiener interpolator

$$W(f) = \frac{S_X(f)}{S_X(f)+S_e(f)}$$

The optimal infinite-length linear interpolator

can be found by taking the inverse Fourier

transform of the ratio of the signal PSD

to the sum of the signal and noise PSDs.

Wiener deconvolution

(REVIEW)

The coefficients in the LMSE predictor for Y of the form

$$\hat{Y} = w_1 X_1 + w_2 X_2 + \cdots + w_n X_n$$

satisfy the *normal equations*

$$\begin{bmatrix} \overline{X_1^2} & \overline{X_1X_2} & \dots & \overline{X_1X_n} \\ \overline{X_2X_1} & \overline{X_2^2} & \dots & \overline{X_2X_n} \\ & & \dots & \\ \overline{X_nX_1} & \overline{X_nX_2} & \dots & \overline{X_n^2} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \dots \\ w_n \end{bmatrix} = \begin{bmatrix} \overline{X_1Y} \\ \overline{X_2Y} \\ \dots \\ \overline{X_nY} \end{bmatrix} .$$

When $X_1, X_2, \ldots, X_p$ are corrupted by iid uncorrelated measurement

the coefficients satisfy the Wiener equations

$$\left(\begin{bmatrix} \sigma_{X_1}^2 & \overline{X_1X_2} & \dots & \overline{X_1X_p} \\ \overline{X_2X_1} & \sigma_{X_2}^2 & \dots & \overline{X_2X_p} \\ & & \dots & \\ \overline{X_pX_1} & \overline{X_pX_2} & \dots & \sigma_{X_p}^2 \end{bmatrix} + \begin{bmatrix} \sigma_{e_1}^2 & 0 & \dots & 0 \\ 0 & \sigma_{e_2}^2 & \dots & 0 \\ & & \dots & \\ 0 & 0 & \dots & \sigma_{e_p}^2 \end{bmatrix} \right) \begin{bmatrix} w_1 \\ w_2 \\ \dots \\ w_p \end{bmatrix} = \begin{bmatrix} \overline{X_1Y} \\ \overline{X_2Y} \\ \dots \\ \overline{X_pY} \end{bmatrix} .$$

Chapter 6 The Kalman Filter
6.1 The Basic Kalman Filter
6.2 Kalman Filter with Transition: Model and Examples
6.3 The Scalar Kalman Filter with Noiseless Transition
6.4 The Scalar Kalman Filter with Noisy Transition
6.5 Iteration of the Scalar Kalman Filter
6.6 Matrix Formulation for the Kalman Filter

$$X_{old} - X_{true} = e_{old},$$

$$X_{new} - X_{true} = e_{new}$$

$$E\{e_{old}\, e_{new}\} = E\{e_{old}\}\, E\{e_{new}\} = 0 \cdot 0 = 0$$

σ_{old} and σ_{new} could be different

$$X_{Kalman} = K\, X_{new} + L\, X_{old}$$

$$X_{true} = E\{X_{Kalman}\} = E\{\, K\, X_{new} + L\, X_{old}\, \}$$

(unbiased)

$$= KE\{X_{new}\} + LE\{X_{old}\} = K\, X_{true} + L\, X_{true}$$

$$X_{Kalman} = K\, X_{new} + (1\text{-}K)\, X_{old}$$

$$X_{Kalman} = X_{old} + K\, \{X_{new} - X_{old}\}$$

(update)

Form the MSE

$$X_{Kalman} = X_{old} + K\,\{X_{new} - X_{old}\}$$

$$X_{true} = K\,X_{true} + (1\text{-}K)\,X_{true}$$

$$X_{Kalman} - X_{true}$$
$$= K\,[X_{new} - X_{true}] + (1\text{-}K)\,[X_{old}\text{-}X_{true}]$$
$$= K\,e_{new} + (1\text{-}K)\,e_{old}$$

$$\mathbf{E}\{[X_{Kalman} - X_{true}]^2\} = \mathbf{E}\{[Ke_{new} + (1\text{-}K)\,e_{old}]^2 =$$
$$K^2\,\mathbf{E}\{e_{new}{}^2\} + 2K(1\text{-}K)\,\mathbf{E}\{e_{new}\,e_{old}\} + (1\text{-}K)^2\,\mathbf{E}\{e_{old}{}^2\}$$
$$= K^2\,\sigma_{new}{}^2 \quad + \quad (0) \quad + (1\text{-}K)^2\,\sigma_{old}{}^2$$

$$\mathbf{E}\{[X_{Kalman} - X_{true}]^2\} =$$
$$K^2\,\sigma_{new}{}^2 \;+\; (1\text{-}K)^2\,\sigma_{old}{}^2$$

$$\frac{d}{dK}E\{[X_{Kalman} - X_{true}]^2\} =$$

$$2K\,\sigma_{new}{}^2 - 2(1\text{-}K)\,\sigma_{old}{}^2 = 0$$

$$K_{Kalman} = \frac{\sigma^2_{old}}{\sigma^2_{new} + \sigma^2_{old}}\ .$$

$$X_{Kalman} = K\,X_{new} + (1\text{-}K)\,X_{old}$$

$$\mathbf{E}\{[X_{Kalman} - X_{true}]^2\} =$$
$$K^2\, \sigma_{new}{}^2 \quad + \quad (1\text{-}K)^2\, \sigma_{old}{}^2$$

$$K_{Kalman} = \frac{\sigma_{old}^2}{\sigma_{new}^2 + \sigma_{old}^2}$$

$$\sigma_{Kalman}^2 = E\{[X_{Kalman} - X_{true}]^2\}$$

$$= \{\frac{\sigma_{old}^2}{\sigma_{new}^2 + \sigma_{old}^2}\}^2 \sigma_{new}^2 + \{1 - \frac{\sigma_{old}^2}{\sigma_{new}^2 + \sigma_{old}^2}\}^2 \sigma_{old}^2$$

$$= \frac{\sigma_{old}^4}{(\sigma_{new}^2 + \sigma_{old}^2)^2}\sigma_{new}^2 + \frac{\sigma_{new}^4}{(\sigma_{new}^2 + \sigma_{old}^2)^2}\sigma_{old}^2$$

$$= \frac{\sigma_{old}^2}{\sigma_{new}^2 + \sigma_{old}^2}\sigma_{new}^2$$

$$= K_{Kalman}\, \sigma_{new}^2$$

Quick Lemma

The minimum of

$$K^2Q + (1-K)^2R$$

occurs when

$$K = R/(R+Q)$$

The minimum value is

$$KQ\ .$$

The basic Kalman filter equations are

$$X_{Kalman} = X_{old} + K \{X_{new} - X_{old}\}$$

$$K_{Kalman} = \frac{\sigma_{old}^2}{\sigma_{new}^2 + \sigma_{old}^2}$$

$$\sigma_{Kalman}^2 = K_{Kalman}\, \sigma_{new}^2$$

Lecture 19

April 3, 2017

Kalman Filter with Transition: Model and Examples

Instead of

$$X_{old} - X_{true} = e_{old},$$

$$X_{new} - X_{true} = e_{new}$$

we interpret the X's as related:

$$X_{new}^{true} = AX_{old}^{true} + B$$

Assume the transition is noisy:

$$X_{new}^{true} = AX_{old}^{true} + B + e_{trans}$$

Assume these are *matrices*

Perhaps not all of the variables are measured:

$$S = DX_{new}^{true} + e_{meas} =$$

$$\begin{bmatrix} d_{11} & d_{12} & \dots & d_{1p} \\ d_{21} & d_{22} & \dots & d_{2p} \\ & & \dots & \\ d_{r1} & d_{r2} & \dots & d_{rp} \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ \dots \\ X_p \end{bmatrix}_{new,true} + \begin{bmatrix} e_1^{meas} \\ e_2^{meas} \\ \dots \\ e_r^{meas} \end{bmatrix}$$

EXAMPLES

1. (Last time) Perform two measurements

$$X_{new}^{true} = [1]\, X_{old}^{true}$$

$$X_{new}^{true} = AX_{old}^{true} + B + e_{trans}$$

$$S = [1]X_{new}^{true} + e_{meas}.$$

$$S = DX_{new}^{true} + e_{meas}$$

(EXAMPLES)

2. *constant velocity* tracking

$$X_{new}^{true} = X_{old}^{true} + [v\Delta t] + e_{trans}$$

$$X_{new}^{true} = AX_{old}^{true} + B + e_{trans}$$

$$S = [1]X_{new}^{true} + e_{meas}$$

(EXAMPLES)

3. *constant acceleration* tracking

$$X_{new}^{true} \equiv \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}_{new}^{true} = \begin{bmatrix} \text{(position)} \\ \text{(velocity)} \end{bmatrix}$$

$$= \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}_{old}^{true} + \begin{bmatrix} X_{2\,old}^{true} \Delta t + \frac{1}{2} a\,(\Delta t)^2 \\ a\,\Delta t \end{bmatrix}$$

$$+ \begin{bmatrix} e_1 \\ e_2 \end{bmatrix}_{trans}$$

$$= \underbrace{\begin{bmatrix} 1 & \Delta t \\ 0 & 1 \end{bmatrix}}_{A} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}_{old}^{true} + \underbrace{\begin{bmatrix} \frac{1}{2} a\,(\Delta t)^2 \\ a\,\Delta t \end{bmatrix}}_{B} + \underbrace{\begin{bmatrix} e_1 \\ e_2 \end{bmatrix}_{trans}}_{e_{trans}}$$

$$X_{new}^{true} = A X_{old}^{true} + B + e_{trans}$$

(EXAMPLES)

(3) *constant acceleration* tracking

$$X_{new}^{true} \equiv \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}_{new}^{true} = \begin{bmatrix} \text{(position)} \\ \text{(velocity)} \end{bmatrix}$$

and if you only measure position

$$S = \underbrace{[1\ 0]}_{D} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}_{new}^{true} + e_{meas}$$

$$S = DX_{new}^{true} + e_{meas}$$

(EXAMPLES)

4. If the transition is described by a first-order differential equation

$$\frac{dX}{dt} = a(t)X + g(t)$$

its solution can be written

$$X(t) \equiv e^{\int_{t_0}^{t} a(\tau)d\tau} [\int_{t_0}^{t} e^{-\int_{t_0}^{\tau} a(\theta)d\theta} g(\tau)d\tau + X(t_0)]$$

$$X_{new}^{true} = \underbrace{e^{\int_{t_0}^{t} a(\tau)d\tau}}_{A} X_{old}^{true} +$$

$$\underbrace{e^{\int_{t_0}^{t} a(\tau)d\tau} [\int_{t_0}^{t} e^{-\int_{t_0}^{\tau} a(\theta)d\theta} g(\tau)d\tau]}_{B}$$

(EXAMPLES)

(4. If the transition is described by a first-order differential equation)

$$X_{new}^{true} = \underbrace{e^{\int_{t_0}^{t} a(\tau)d\tau}}_{A} X_{old}^{true} +$$

$$\underbrace{e^{\int_{t_0}^{t} a(\tau)d\tau} [\int_{t_0}^{t} e^{-\int_{t_0}^{\tau} a(\theta)d\theta} g(\tau)d\tau]}_{B}$$

$$g(t) = g_{det}(t) + g_{noise}(t)$$

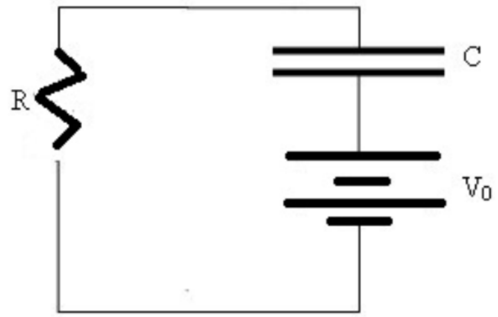
$$X_{new}^{true} = \underbrace{e^{\int_{t_0}^{t} a(\tau)d\tau}}_{A} X_{old}^{true} +$$

$$\underbrace{e^{\int_{t_0}^{t} a(\tau)d\tau} [\int_{t_0}^{t} e^{-\int_{t_0}^{\tau} a(\theta)d\theta} g_{det}(\tau)d\tau}_{B}$$

$$+ \underbrace{e^{\int_{t_0}^{t} a(\tau)d\tau} [\int_{t_0}^{t} e^{-\int_{t_0}^{\tau} a(\theta)d\theta} g_{noise}(\tau)d\tau}_{e_{trans}}$$

(EXAMPLES)

(4. If the transition is described by a first-order differential equation)



$$C\frac{dV}{dt} = -\frac{V+V_0}{R}$$

$$\text{or} \quad \frac{dV}{dt} = -\frac{V}{RC} - \frac{V_0}{RC}$$

$$V(t) = \underbrace{e^{-t/RC}}_{A} V(0) + \underbrace{V_0[e^{-t/RC} - 1]}_{B}$$

V_0 may be noisy, spawning e_{trans}

(EXAMPLES)

5. If the transition is described by a second-order differential equation

$$\frac{d^2X}{dt^2} = a(t)\frac{dX}{dt} + b(t)X + g(t)$$

general solution

$$X(t) = c_1 X_{\text{hom}}{}^{(1)}(t) + c_2 X_{\text{hom}}{}^{(2)}(t) + X_{\text{part}}(t)$$

assign initial conditions, rewrite

$$X(t) = \tilde{X}_1(t)X(0) + \tilde{X}_2(t)X'(0) + \tilde{X}_3(t)$$

differentiate

$$X'(t) = \tilde{X}_1'(t)X(0) + \tilde{X}_2(t)'X'(0) + \tilde{X}_3'(t)$$

You've got it!

$$\begin{bmatrix} X(t) \\ X'(t) \end{bmatrix} = \begin{bmatrix} \tilde{X}_1(t) & \tilde{X}_2(t) \\ \tilde{X}_1'(t) & \tilde{X}_2'(t) \end{bmatrix} \begin{bmatrix} X(0) \\ X'(0) \end{bmatrix} + \begin{bmatrix} \tilde{X}_3(t) \\ \tilde{X}_3'(t) \end{bmatrix}$$

$$X_{new}^{true} = AX_{old}^{true} + B + e_{trans}$$

(EXAMPLES)

6. If the transition is described by a linear system of differential equations

$$\frac{d}{dt}[X] = \mathbf{A}[X] + [g(t)]$$

the solution is

$$X(t) = \underbrace{\left[e^{\mathbf{A}t}\right]\left[e^{-\mathbf{A}t_0}\right]}_{A} [X(t_0)]$$

$$+ \underbrace{\left[e^{\mathbf{A}t}\right] \int_{t_0}^{t} \left[e^{-\mathbf{A}\tau}\right] g(\tau) d\tau}_{B}$$

(EXAMPLES)

7. ARMA

$$X(n) = a(1)\,X(n\text{-}1) + a(2)\,X(n\text{-}2)$$

$$+\,a(3)\,X(n\text{-}3) + b(0)\,V(n)$$

you're already there:

$$\underbrace{\begin{bmatrix} X(n) \\ X(n-1) \\ X(n-2) \end{bmatrix}}_{X_{new}^{true}} =$$

$$\underbrace{\begin{bmatrix} a(1) & a(2) & a(3) \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} X(n-1) \\ X(n-2) \\ X(n-3) \end{bmatrix}}_{X_{old}^{true}}$$

$$+ \underbrace{\begin{bmatrix} b(0)V(n) \\ 0 \\ 0 \end{bmatrix}}_{e_{trans}}$$

8. Spacecraft

$$X = \begin{bmatrix} x \text{ coordinate} \\ y \text{ coordinate} \\ z \text{ coordinate} \\ x \text{ velocity} \\ y \text{ velocity} \\ z \text{ velocity} \\ \text{angle from } x \text{ axis} \\ \text{angle from } y \text{ axis} \\ \text{angle from } z \text{ axis} \\ x\text{-angular velocity} \\ y\text{-angular velocity} \\ z\text{-angular velocity} \end{bmatrix}$$

$$X_{new}^{true} = AX_{old}^{true} + B + e_{trans}$$

$$S = DX_{new}^{true} + e_{meas}$$

The Scalar Kalman Filter with Noiseless Transition

$$X_{old} - X_{old}^{true} = e_{old} ,$$
$$X_{new} - X_{new}^{true} = e_{new}$$

$$X_{new} = AX_{old} + B$$

$$X_{Kalman} = K\, X_{new} + L\, [A\, X_{old} + B]$$

(apply the update
to the old measured
value)

$$X_{new} = AX_{old} + B$$

$$X_{Kalman} = K\,X_{new} + L\,[A\,X_{old} + B]$$

No bias:

$$E\{X_{Kalman}\} =$$
$$E\{\,K\,X_{new} + L\,[AX_{old} + B]\}$$
$$= K\,X_{true}^{new} + L\,[AX_{old}^{true} + B]$$
$$= (K+L)X_{true}^{new}$$

so $L = 1\text{-}K$ as before

$$X_{Kalman} = K\,X_{new} + (1\text{-}K)[A\,X_{old} + B]$$

$$X_{Kalman} = K\, X_{new} + (1\text{-}K)[A\, X_{old} + B]$$

to express the error,
from this subtract

$$X_{new}^{true} = K\, X_{new}^{true} + (1\text{-}K)\, X_{new}^{true}$$
$$= K\, X_{new}^{true} + (1\text{-}K)[AX_{old}^{true} + B]$$

$$X_{Kalman} - X_{new}^{true} =$$
$$K\,[X_{new} - X_{new}^{true}] + (1\text{-}K)A(X_{old} - X_{old}^{true})$$
$$= \underbrace{K\, e_{new} + (1\text{-}K)Ae_{old}}$$

$$X_{Kalman} - X_{new}^{true} =$$
$$K\,[X_{new} - X_{new}^{true}] + (1-K)A(X_{old} - X_{old}^{true})$$
$$= K\,e_{new} + (1-K)Ae_{old}$$

$$\text{MSE} = K^2\,e_{new}^2 + (1-K)^2A^2e_{old}^2$$
$$= K^2\,\sigma_{new}^2 + (1-K)^2\,A^2\sigma_{old}^2$$

The minimum of

$K^2Q + (1-K)^2R$

occurs when

$K = R/(R+Q)$

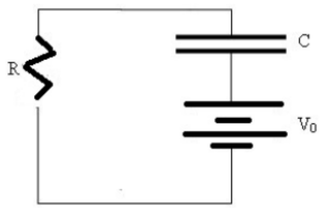
The minimum value is KQ .

$$K_{\text{Kalman}} = \frac{A^2\sigma_{old}^2}{\sigma_{new}^2 + A^2\sigma_{old}^2}$$

$$\sigma_{Kalman}^2 = \frac{A^2\sigma_{old}^2}{\sigma_{new}^2 + A^2\sigma_{old}^2}\,\sigma_{new}^2$$
$$= K_{Kalman}\sigma_{new}^2$$

The Scalar Kalman Filter with Noisy Transition

$$X_{new}^{true} = AX_{old}^{true} + B + e_{trans}$$



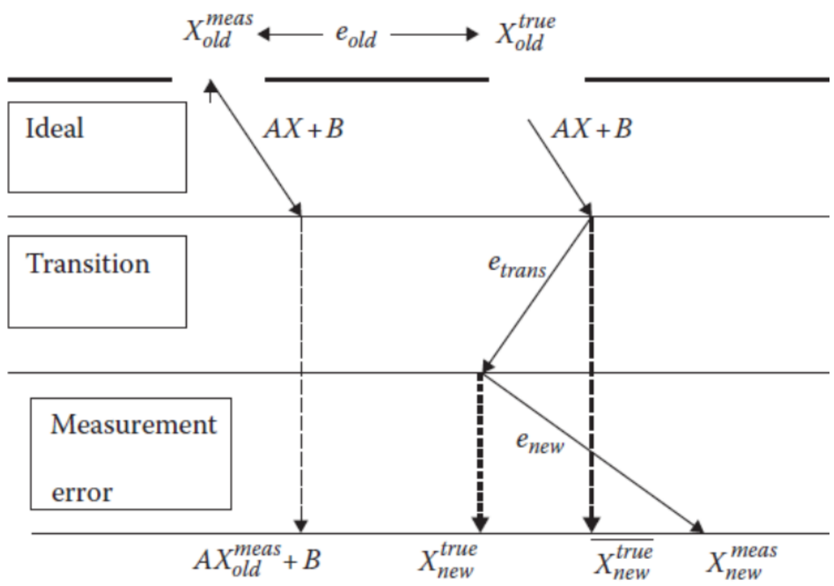
$$C\frac{dV}{dt} = -\frac{V+V_0}{R}$$

$$or \quad \frac{dV}{dt} = -\frac{V}{RC} - \frac{V_0}{RC}$$

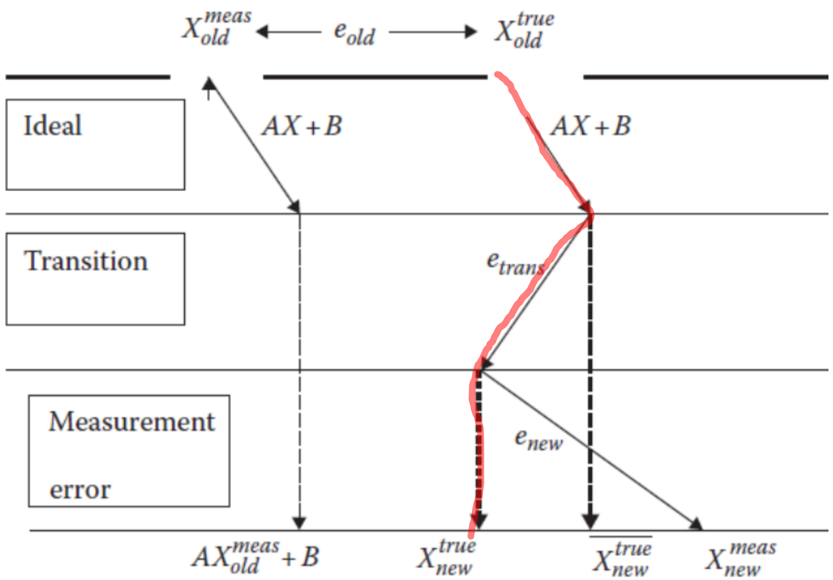
$$V(t) = \underbrace{e^{-t/RC}}_{A} V(0) + \underbrace{V_0[e^{-t/RC} - 1]}_{B}$$

$$V_0 = V_0^{true} + e_0$$

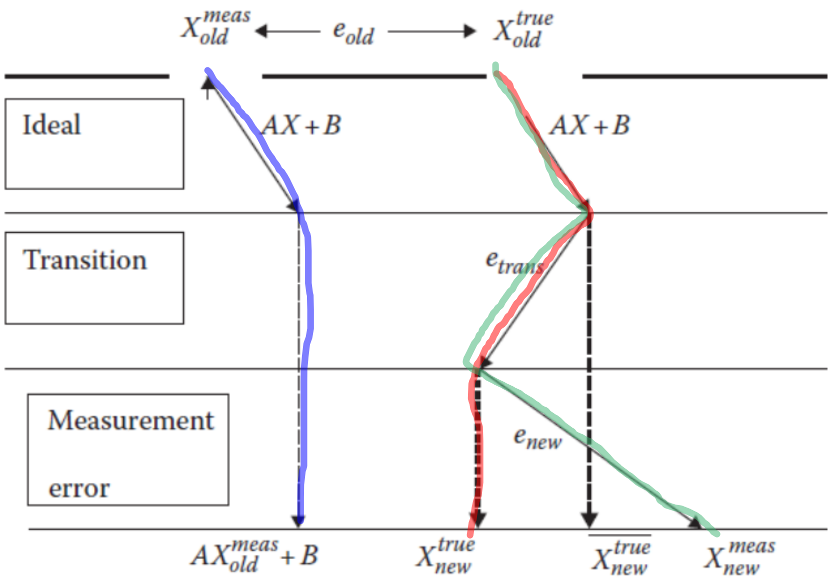
$$A = e^{-t/RC} \quad B = V_0[e^{-t/RC} - 1] \quad e_{trans} = e_0[e^{-t/RC} - 1]$$



$$X_{Kalman} = K\, X_{new}{}^{meas} + L\, [A\, X_{old}{}^{meas} + B]$$



$$X_{Kalman} = K\, X_{new}{}^{meas} + L\, [A\, X_{old}{}^{meas} + B]$$



$$X_{Kalman} = K\, X_{new}{}^{meas} + L\, [A\, X_{old}{}^{meas} + B]$$

Get the bias right

X_{new}^{true} is now a *random* variable

$E\{X_{Kalman}\} = X_{new}{}^{true}$ makes no sense

$$E\{X_{Kalman}\} = E\{X_{new}{}^{true}\}$$

$$= E\{AX_{old}^{true} + B + e_{trans}\}$$

$$= AX_{old}^{true} + B$$

BIAS: The usual song and dance

$$X_{Kalman} = K\,X_{new}{}^{meas} + L\,[A\,X_{old}{}^{meas} + B]$$

$$E\{X_{new}{}^{true}\} \quad KE\{X_{new}{}^{true}\} \quad LE\{X_{new}{}^{true}\}$$

$$L = 1\text{-}K$$

$$X_{Kalman} = K\,X_{new}{}^{meas} + (1\text{-}K)\,[A\,X_{old}{}^{meas} + B]$$

$$X_{Kalman} = K\,X_{new}^{meas} + (1\text{-}K)\,[A\,X_{old}^{meas} + B]$$

Get the formula for the error, square, take expected value, minimize: "MSE"

$$X_{new}^{true} = K\,X_{new}^{true} + (1\text{-}K)\,X_{new}^{true}$$

$$= KX_{new}^{true} + (1\text{-}K)\,[AX_{old}^{true} + B + e_{trans}]$$

$$e_{Kal} = X_{Kalman} - X_{new}^{true} =$$

$$K\,[X_{new}^{meas} - X_{new}^{true}] +$$

$$(1\text{-}K)\,[A(X_{old}^{meas} - X_{old}^{true}) - e_{trans}]$$

$$= K\,e_{new} + (1\text{-}K)\,Ae_{old} - (1\text{-}K)\,e_{trans}$$

$$e_{Kal} = K\, e_{new} + (1\text{-}K)\, A e_{old} - (1\text{-}K)\, e_{trans}$$

MSE =

$$K^2\, E\{e_{new}{}^2\} + (1\text{-}K)^2 A^2 E\{e_{old}{}^2\} + (1\text{-}K)^2 E\{e_{trans}{}^2\}$$

$$= K^2\, \sigma_{new}{}^2 + (1\text{-}K)^2\, [A^2\sigma_{old}{}^2 + \sigma_{trans}{}^2]$$

The minimum of

$K^2Q + (1\text{-}K)^2R$

occurs when

$K = R/(R+Q)$

The minimum value is KQ .

$$K_{\text{Kalman}} = \frac{A^2\sigma^2_{old} + \sigma^2_{trans}}{\sigma^2_{new} + A^2\sigma^2_{old} + \sigma^2_{trans}}$$

$$\sigma^2_{Kalman} = K_{Kalman}\, \sigma^2_{new}$$

Example $dy/dt = -y + g(t)$

$g(t)$ = zero mean white noise with power σ_g^2

$y_{meas}(0)$ and $y_{meas}(t)$

zero mean noise σ_0 and σ_1

$$y(t) = e^{-t}\, y(0) + e^{-t} \int_0^t e^{\tau}\, g(\tau)\, d\tau$$

$$= Ay(0) + B + e_{trans}$$

$$A = e^{-t}, B = 0, \quad e_{trans} = e^{-t} \int_0^t e^{\tau}\, g(\tau)\, d\tau$$

$$E\{e_{trans}\} = e^{-t} \int_0^t e^{\tau}\, E\{g(\tau)\}\, d\tau = 0$$

$$\sigma_{trans}^2 = E\left\{[e^{-t} \int_0^t e^{\tau}\, g(\tau)\, d\tau]^2\right\}$$

$$\sigma_{trans}^2 = E\left\{[e^{-t}\int_0^t e^{\tau}\, g(\tau)\, d\tau]^2\right\}$$

$$= E\{e^{-2t}\int_0^t e^{t_1}\, g(t_1)\, dt_1 \int_0^t e^{t_2}\, g(t_2)\, dt_2\}$$

$$= e^{-2t}\int_0^t \int_0^t \; e^{t_1+t_2} E\{g(t_1)g(t_2)\} dt_2\, dt_1$$

$$= e^{-2t}\int_0^t \int_0^t \; e^{t_1+t_2} \underline{R_g(t_1,t_2)}\, dt_2\, dt_1$$

(page 138)

$$R_g(t_1,t_2) = \sigma_g^2 \delta(t_1 - t_2)$$

$$\sigma_{trans}^2 = e^{-2t}\sigma_g^2 \int_0^t \int_0^t \; e^{t_1+t_2}\delta(t_1 - t_2) dt_2\, dt_1$$

$$= e^{-2t}\sigma_g^2 \int_0^t \; e^{t_1+t_1}\, dt_1 = \frac{\sigma_g^2}{2}[1 - e^{-2t}]$$

$$y(t) = Ay(0) + B + e_{trans}$$

$$\sigma_{trans}^2 = \frac{\sigma_g^2}{2}[1 - e^{-2t}]$$

$y_{meas}(0)$ and $y_{meas}(t)$. σ_0 and σ_1

$$A = e^{-t}, B = 0$$

$$X_{Kalman} = K X_{new}^{meas} + (1\text{-}K) [A X_{old}^{meas} + B]$$

$$K_{\text{Kalman}} = \frac{A^2\sigma_{old}^2 + \sigma_{trans}^2}{\sigma_{new}^2 + A^2\sigma_{old}^2 + \sigma_{trans}^2}$$

$$\sigma_{Kalman}^2 = K_{Kalman}\,\sigma_{new}^2$$

$$y(t) = Ay(0) + B + e_{trans}$$

$$\sigma_{trans}^2 = \frac{\sigma_g^2}{2}[1 - e^{-2t}]$$

$y_{meas}(0)$ and $y_{meas}(t)$. σ_0 and σ_1

$$A = e^{-t}, B = 0$$

$$X_{Kalman} = K\, X_{new}^{meas} + (1\text{-}K)\, [A\, X_{old}^{meas} + B]$$

$$K_{\text{Kalman}} = \frac{A^2\sigma_{old}^2 + \sigma_{trans}^2}{\sigma_{new}^2 + A^2\sigma_{old}^2 + \sigma_{trans}^2}$$

$$\sigma_{Kalman}^2 = K_{Kalman}\, \sigma_{new}^2$$

$$y(t) = Ay(0) + B + e_{trans}$$

$$\sigma_{trans}^2 = \frac{\sigma_g^2}{2}[1 - e^{-2t}]$$

$y_{meas}(0)$ and $y_{meas}(t)$. σ_0 and σ_1

$$A = e^{-t}, B = 0$$

$$X_{Kalman} = K\, X_{new}^{meas} + (1\text{-}K)\,[A\, X_{old}^{meas} + B]$$

$$K_{\mathrm{Kalman}} = \frac{A^2\sigma_{old}^2 + \sigma_{trans}^2}{\sigma_{new}^2 + A^2\sigma_{old}^2 + \sigma_{trans}^2}$$

$$\sigma_{Kalman}^2 = K_{Kalman}\,\sigma_{new}^2$$

What if we don't have the measurement at t?

$$X_{Kalman} = K\, X_{new}^{meas} + (1\text{-}K)\,[A\, X_{old}^{meas} + B]$$

$$K_{\mathrm{Kalman}} = \frac{A^2\sigma_{old}^2 + \sigma_{trans}^2}{\sigma_{new}^2 + A^2\sigma_{old}^2 + \sigma_{trans}^2}$$

$$\sigma_{Kalman}^2 = K_{Kalman}\,\sigma_{new}^2$$

What if we don't have the measurement at t?

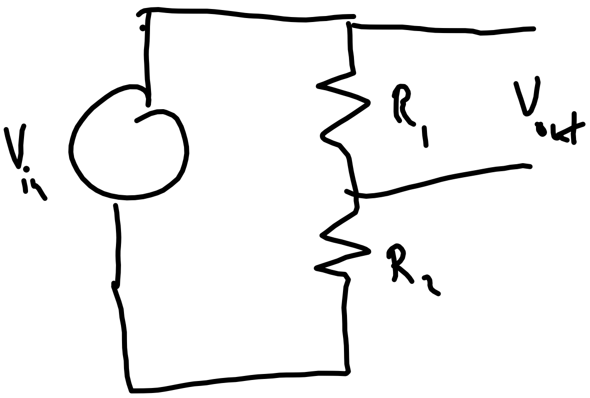
$$\sigma_{new} = \infty$$

$$K_{\mathrm{Kalman}} = \frac{A^2\sigma_{old}^2 + \sigma_{trans}^2}{\sigma_{new}^2 + A^2\sigma_{old}^2 + \sigma_{trans}^2} \rightarrow 0$$

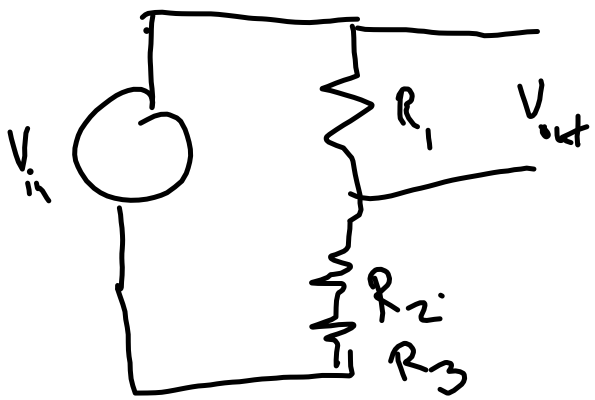
$$\sigma_{Kalman}^2 = \frac{A^2\sigma_{old}^2 + \sigma_{trans}^2}{\sigma_{new}^2 + A^2\sigma_{old}^2 + \sigma_{trans}^2}\,\sigma_{new}^2$$

$$\rightarrow A^2\sigma_{old}^2 + \sigma_{trans}^2$$

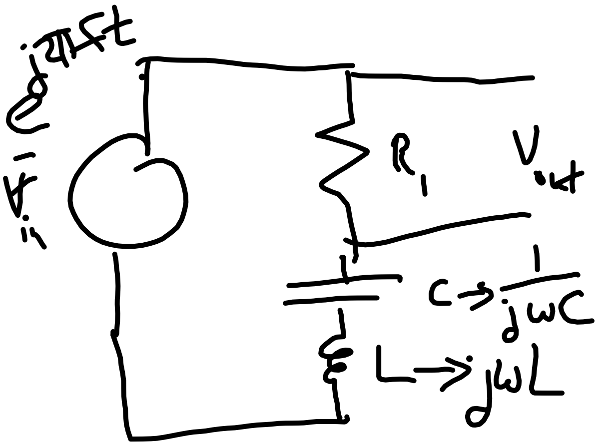
"Lecture 20"
This lecture was inserted off-the-cuff as a tutorial before lecture 20. This lecture is called
Lecture on 4/5/2017 (Wed)
There is some recording problem at 31:40. When you get there, simply "mouse" the timer ahead a few seconds and skip it.
The next lecture will be called "Lecture 21."



$$\frac{V_{out}}{V_{in}} = \frac{R_1}{R_1+R_2}$$



$$\frac{V_{out}}{V_{in}} = \frac{R_1}{R_1+R_2+R_3}$$



$$\omega = 2\pi f$$

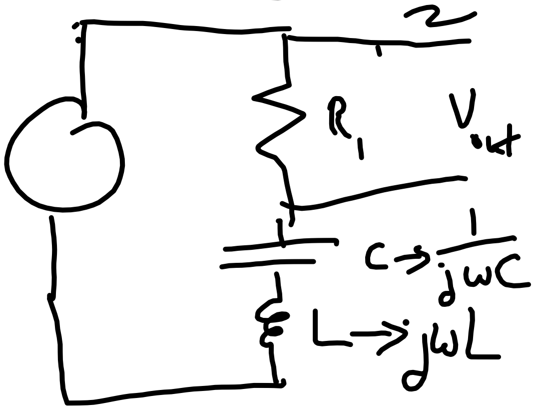
$$\frac{V_{out}}{V_{in}} = \frac{R_1}{R_1 + \frac{1}{j2\pi fC} + j2\pi fL}$$

$$V_{out} = \frac{R}{R + \frac{1}{j2\pi fC} + j2\pi fL} V_{in}$$

$$V_{in} \to e^{j2\pi ft}$$

$$V_{in} = \cos 2\pi f t = \frac{e^{j2\pi f t} + e^{-j2\pi f t}}{2}$$

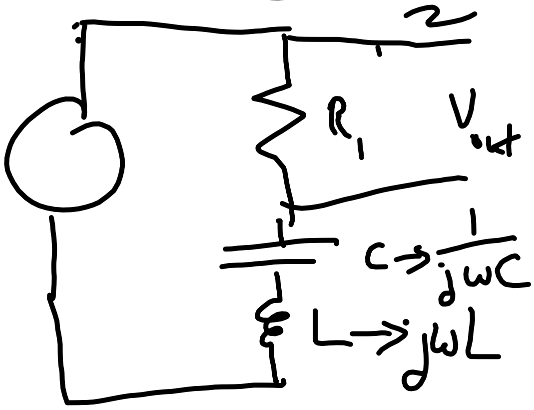
$$\omega = 2\pi f$$



$$\frac{V_{out}}{V_{in}} = \frac{R_1}{R_1 + \frac{1}{j2\pi f C} + j2\pi f L}$$

$$V_{out} = \frac{1}{2} \frac{R}{R + \frac{1}{j2\pi f C} + j2\pi f L} e^{j2\pi f t}$$

$$V_{in} = \cos 2\pi f t = \frac{e^{j2\pi f t} + e^{-j2\pi f t}}{2}$$



$$\omega = 2\pi f$$

$$\frac{V_{out}}{V_{in}} = \frac{R_1}{R_1 + \frac{1}{j2\pi f C} + j2\pi f L}$$

$$V_{out} = \frac{1}{2}\frac{R}{R + \frac{1}{j2\pi f C} + j2\pi f L} e^{j\pi f t}$$

$$+ \frac{1}{2}\frac{R}{R + \frac{1}{j(-1)2\pi f C} + j(-1)2\pi f L} e^{-j\pi f t}$$

$$V_{out} = \frac{1}{2}\,\frac{R}{R + \frac{1}{j2\pi fC} + j2\pi fL}\,e^{j2\pi ft} + \frac{1}{2}\,\frac{R}{R + \frac{1}{j2\pi(-f)C} + j2\pi(-f)L}\,e^{-j2\pi ft}$$

$$= \alpha + \overline{\alpha} = 2\,\mathrm{Re}\,\alpha$$

$$\mathrm{Re}\left[\frac{R}{R + \frac{1}{j2\pi(-f)C} + j2\pi(-f)L}\,e^{-j2\pi ft}\right]$$

$$V_{out} = \frac{1}{2}\frac{R}{R + \frac{1}{j2\pi fC} + j2\pi fL} e^{j2\pi ft} + \frac{1}{2}\frac{R}{R + \frac{1}{j2\pi fC} + j2\pi fL} e^{-j2\pi ft}$$

(-) (-)

$$\text{Power} = \frac{1}{2T}\int_{-T}^{T} V_{out}^2 \, dt = \frac{1}{2T}\int_{-T}^{T} (\alpha + \bar{\alpha})^2 dt$$

$$= \frac{1}{2T}\int_{-T}^{T} (\alpha^2 + \bar{\alpha}^2 + 2|\alpha|^2) dt$$

$\alpha = \square e^{j2\pi ft}$ $\alpha^2 = \square^2 e^{j4\pi ft}$

$$= \frac{1}{2T}\int_{-T}^{T}\left(\alpha^2 + \bar{\alpha}^2 + 2|\alpha|^2\right)dt$$

$\alpha = \square e^{j2\pi f t}$ $\alpha^2 = \square^2 e^{j4\pi f t}$

$= \frac{\square^2}{2T}\int_{-T}^{T} e^{j4\pi f t}\,dt$ is, at most, $\frac{\square^2}{2T} \to 0$

also the $\bar{\alpha}^2$ integral $\to 0$

$$= \frac{1}{2T}\int_{-T}^{T}\left(\alpha^2 + \bar{\alpha}^2 + 2|\alpha|^2\right)dt$$

$\alpha = \square e^{j\Pi f t}$ $\bar{\alpha}^2 = \square e^{2j\Pi f t}$

What about $\frac{1}{2T}\int_{-T}^{T} 2|\alpha|^2 dt$ $|\alpha|^2 = |\square|^2 \cdot 1$

$$= \frac{|\square|^2}{2T} 2\int_{-T}^{T} 1\, dt = 2|\square|^2$$

$= \frac{2}{4}|\text{TransFcn}|^2$

A tutorial lecture was inserted prior to this one, called
Lecture on 4/5/2017 (Wed).
THIS lecture was called
Lecture on 4/5/2017 (Wed) (Second Copy)

21

Lecture 20 April 5, 2017

Iteration of the Scalar Kalman Filter

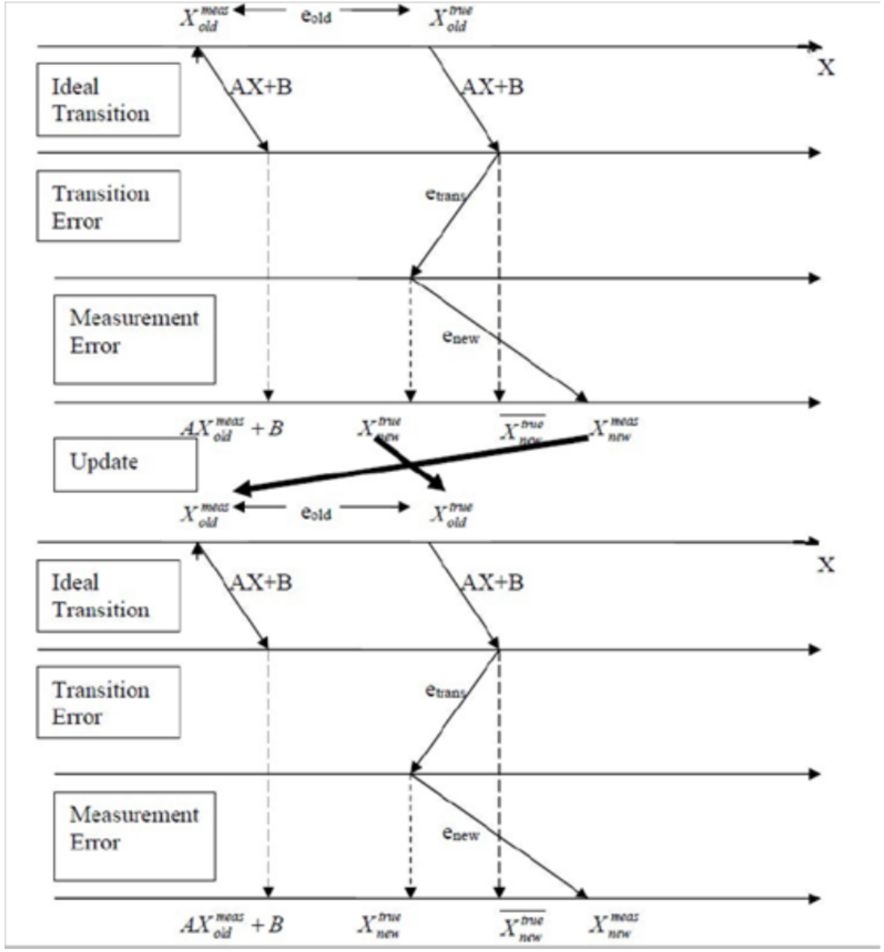


Figure 6.3 Consecutive Kalman Filtering

$$X_{Kalman} = K\, X_{new}{}^{meas} + L\, [A\, X_{old}{}^{meas} + B]$$

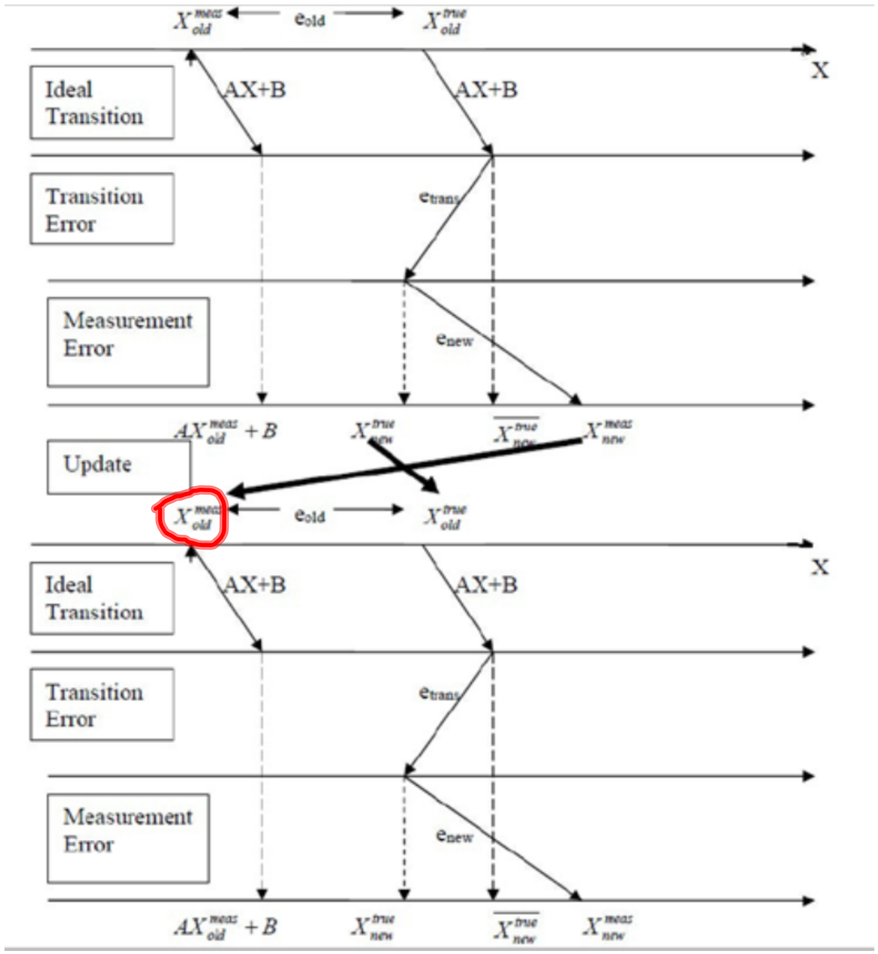


Figure 6.3 Consecutive Kalman Filtering

$$X_{Kalman} = K\, X_{new}{}^{meas} + L\, [A\, X_{old}{}^{meas} + B]$$

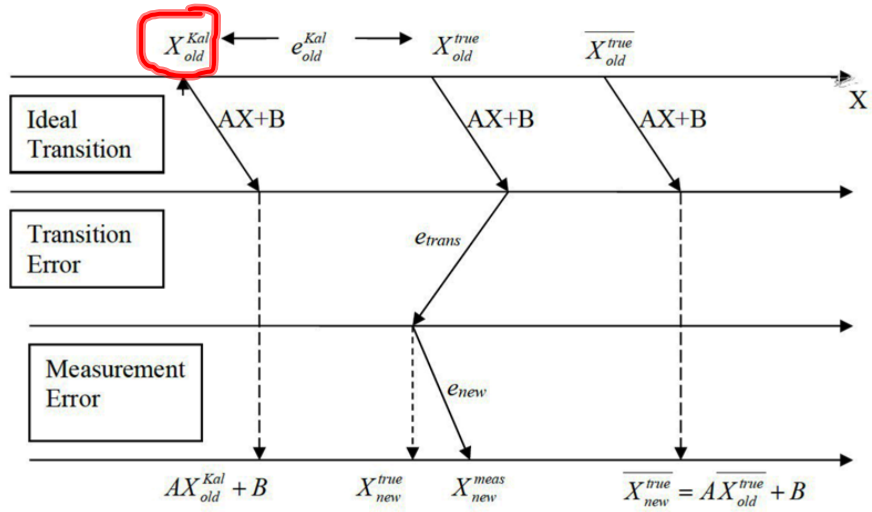


Figure 6.4 Consecutive Kalman Filtering

$$X_{Kalman} = K\,X_{new}^{meas} + (1\text{-}K)\,[A\,X_{old}^{Kal} + B]$$

BIAS

$$E\{X_{Kalman}\} = \quad E\{K\,X_{new}^{meas} +$$

$$L\,[A\,X_{old}^{Kal} + B]\}$$

$$= E\{K(X_{new}^{true} + e_{new})\} +$$

$$E\{L[A(X_{old}^{true} + e_{old}^{Kal}) + B]\}$$

$$= KE\{X_{new}^{true}\} + K\cdot 0 \;+$$

$$LE\{[A(X_{old}^{true}) + B + e_{trans} - e_{trans} + Ae_{old}^{Kal}]\}$$

$$= KE\{X_{new}^{true}\}$$

$$+\; LE\{X_{new}^{true}\} \qquad - L\cdot 0 \; + LA\cdot 0$$

$$= (K+L)\,E\{X_{new}^{true}\}$$

$$L = 1 - K$$

$$X_{Kalman} = K\,X_{new}^{meas} + (1\text{-}K)\,[A\,X_{old}^{Kal} + B]$$

$$e_{Kal} = X_{Kalman} - X_{new}{}^{true}$$
$$= K\, e_{new} + (1\text{-}K)\, Ae_{old}{}^{Kal} - (1\text{-}K)\, e_{trans}$$

The minimum of
$K^2Q + (1\text{-}K)^2R$
occurs when
$K = R/(R+Q)$
The minimum value is KQ .

$$K_{Kalman} = \frac{A^2 \sigma_{old}^{Kal^2} + \sigma_{trans}^2}{\sigma_{new}^2 + A^2 \sigma_{old}^{Kal^2} + \sigma_{trans}^2}$$

$$\sigma_{new}^{Kal^2} = K_{Kalman}\, \sigma_{new}^2$$

Matrix Formulation for the Kalman Filter

$$X = \begin{bmatrix} x \text{ coordinate} \\ y \text{ coordinate} \\ z \text{ coordinate} \\ x \text{ velocity} \\ y \text{ velocity} \\ z \text{ velocity} \\ \text{angle from } x \text{ axis} \\ \text{angle from } y \text{ axis} \\ \text{angle from } z \text{ axis} \\ x\text{-angular velocity} \\ y\text{-angular velocity} \\ z\text{-angular velocity} \end{bmatrix}$$

$$X_{new}^{true} = AX_{old}^{true} + B + e_{trans}$$

$$S = DX_{new}^{true} + e_{meas}$$

$$X_{new}^{true} = AX_{old}^{true} + B + e_{trans}$$

p by 1 *p* by *p* *p* by 1 *p* by 1 *p* by 1

$$S = DX_{new}^{true} + e_{meas}$$

r by 1 *r* by *p* *p* by 1 *r* by 1

$$X_{Kalman} = K\,S + L\,[A\,X_{old}{}^{est} + B]$$

(measured or Kalman)

$$X_{old}{}^{est} = X_{old}{}^{true} + e_{old}$$

p by 1 *p* by 1 *p* by 1

$$[e_{old}][e_{meas}]^T \equiv \begin{bmatrix} e_1^{old} \\ e_2^{old} \\ \cdots \\ e_p^{old} \end{bmatrix} \begin{bmatrix} e_1^{meas} & e_2^{meas} & \ldots & e_r^{meas} \end{bmatrix}$$

$$= \begin{bmatrix} e_1^{old} e_1^{meas} & e_1^{old} e_2^{meas} & \ldots & e_1^{old} e_r^{meas} \\ e_2^{old} e_1^{meas} & e_2^{old} e_2^{meas} & \ldots & e_2^{old} e_r^{meas} \\ & \cdots & & \\ e_p^{old} e_1^{meas} & e_p^{old} e_2^{meas} & \ldots & e_p^{old} e_r^{meas} \end{bmatrix}$$

All zero (uncorrelated, zero mean)

$$
\begin{aligned}
&E\{[e_{old}][e_{old}]^T\} \\
&= E\left\{\begin{bmatrix} e_1^{old}e_1^{old} & e_1^{old}e_2^{old} & \ldots & e_1^{old}e_p^{old} \\ e_2^{old}e_1^{old} & e_2^{old}e_2^{old} & \ldots & e_2^{old}e_p^{old} \\ & \ldots & & \\ e_p^{old}e_1^{old} & e_p^{old}e_2^{old} & \ldots & e_p^{old}e_p^{old} \end{bmatrix}\right\} \\
&= \begin{bmatrix} \sigma_1^{2,old} & \overline{e_1^{old}e_2^{old}} & \ldots & \overline{e_1^{old}e_p^{old}} \\ \overline{e_2^{old}e_1^{old}} & \sigma_2^{2,old} & \ldots & \overline{e_2^{old}e_p^{old}} \\ & \ldots & & \\ \overline{e_p^{old}e_1^{old}} & \overline{e_p^{old}e_2^{old}} & \ldots & \sigma_p^{2,old} \end{bmatrix} \\
&\equiv Q_{old}
\end{aligned}
$$

$$\text{MSE}(e_{old}) = \text{trace}(Q_{old})$$

$$X_{Kalman} = \quad K\,S + L\,[A\,X_{old}^{est} + B]$$

BIAS

$$E\{\,K\,(DX_{new}^{true} + e_{meas})\}\; +$$
$$E\{\,L\,[A\,(X_{old}^{true} + e_{old}) + B]\}$$
$$= \;K\,D\,E\{X_{new}^{true}\,\} + K\cdot 0$$
$$+ \;L\,E\{\,[A\,(X_{old}^{true})$$
$$+ B + e_{trans} - e_{trans} + A\,e_{old}]\}$$
$$= \;K\,D\,E\{X_{new}^{true}\,\}$$
$$+ \;L\,E\{\,X_{new}^{true}\,\} - L\cdot 0 + LA\cdot 0$$
$$= (KD{+}L)\;E\{X_{new}^{true}\}$$
$$KD + L = I$$

$$X_{Kalman} = [AX_{old}^{Kal} + B]$$
$$+ K[S - D(A\,X_{old}^{Kal} + B)]$$

$$X_{Kalman} = [AX_{old}^{Kal} + B] + K[S - D(A\, X_{old}^{Kal} + B)]$$

(p by 1)

$$e_{Kal} = X_{Kalman} - X_{new}^{true} =$$
$$Ae_{old} - e_{trans} + K[e_{meas} - D(Ae_{old} - e_{trans})]$$

$$\begin{bmatrix} \\ e_k \\ \\ \end{bmatrix} \begin{bmatrix} e_k^T & \end{bmatrix}$$

$$E\{[e_{Kal}][e_{Kal}]^T\} =$$

$$\begin{bmatrix} \sigma_1^{2,Kal} & \overline{e_1^{Kal} e_2^{Kal}} \dots & \overline{e_1^{Kal} e_p^{Kal}} \\ \overline{e_2^{Kal} e_1^{Kal}} & \sigma_2^{2,Kal} \dots & \overline{e_2^{Kal} e_p^{Kal}} \\ & \dots & \\ \overline{e_p^{Kal} e_1^{Kal}} & \overline{e_p^{Kal} e_2^{Kal}} \dots & \sigma_p^{2,Kal} \end{bmatrix}$$

$$e_{Kal} = X_{Kalman} - X_{new}{}^{true} =$$
$$Ae_{old} - e_{trans} + K[e_{meas} - D(Ae_{old} - e_{trans})]$$

$$Q_{Kal} = E\{e_{Kal} e_{Kal}^T\}$$

$$E\{(Ae_{old} - e_{trans})(e_{old}^T A^T - e_{trans}^T)\}$$
$$= AQ_{old} A^T + Q_{trans} + (0)$$

$$Q_{Kal} = AQ_{old} A^T + Q_{trans}$$
$$-(AQ_{old} A^T + Q_{trans}) D^T K^T$$
$$-KD(AQ_{old} A^T + Q_{trans})$$
$$+K[Q_{meas} + D(AQ_{old} A^T + Q_{trans}) D^T] K^T)$$

$$
\begin{aligned}
Q_{Kal} = {} & AQ_{old}A^T + Q_{trans} \\
& -(AQ_{old}A^T + Q_{trans})D^T K^T \\
& -KD(AQ_{old}A^T + Q_{trans}) \\
& +K[Q_{meas} + D(AQ_{old}A^T + Q_{trans})D^T]K^T)
\end{aligned}
$$

$$\frac{\partial tr\{KE\}}{\partial K_{ij}} = E_{ji} \qquad \frac{\partial tr\{KE\}}{\partial K} = E^T$$

$$\frac{\partial tr\{FK^T\}}{\partial K} = F$$

$$\frac{\partial tr\{KGK^T\}}{\partial K} = K(G+G^T)$$

$$
\begin{aligned}
& 2(AQ_{old}A^T + Q_{trans})D^T = \\
& 2K[Q_{meas} + D(AQ_{old}A^T + Q_{trans})D^T]
\end{aligned}
$$

$$K = (AQ_{old}A^T + Q_{trans})D^T$$
$$[Q_{meas} + D(AQ_{old}A^T + Q_{trans})D^T]^{-1}$$

$$Q_{Kal} = (AQ_{old}A^T + Q_{trans})(I - D^T K^T)$$
$$\equiv (I - KD)(AQ_{old}A^T + Q_{trans})$$

If $S = X_{new}{}^{true} + e_{meas}$

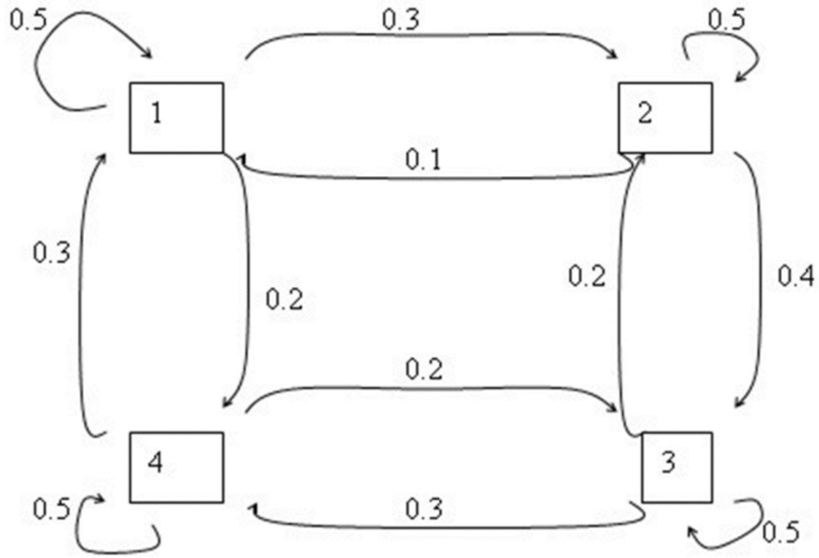
$$Q_{Kal} = KQ_{meas}$$

22

Lecture 21

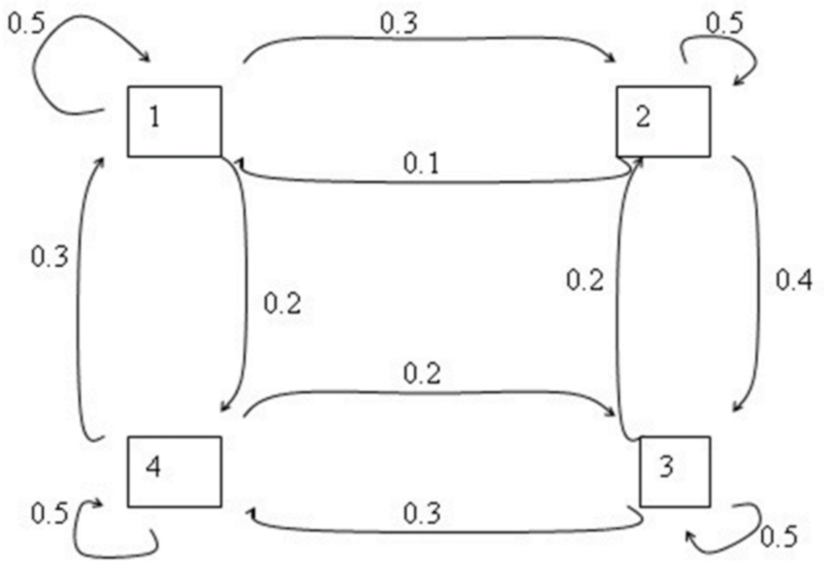
April 12, 2017

Markov Processes


0.5
0.3
0.5
1
2
0.1
0.3
0.2
0.2
0.4
0.2
4
3
0.5
0.3
0.5

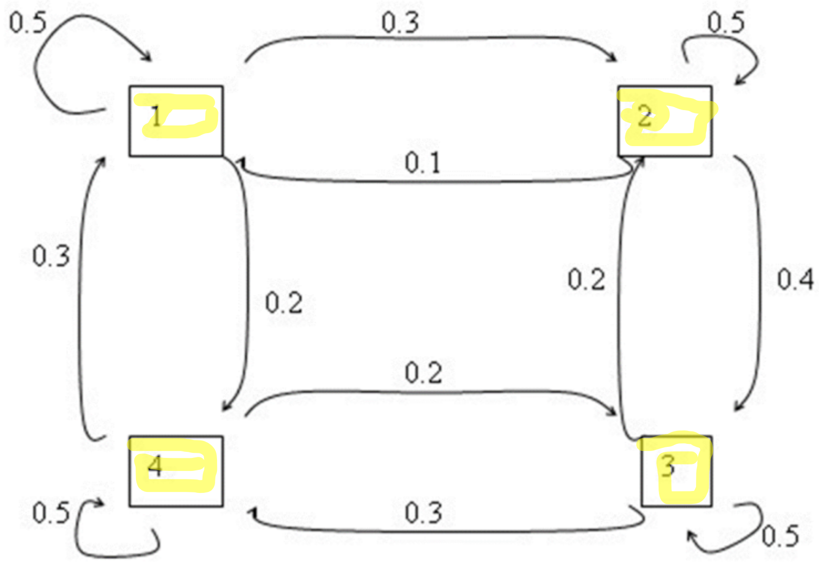
$$\begin{aligned}
&\text{Prob}\{X(8) = 1\} \\
&= \text{Prob}\{X(8) = 1 \ \& \ X(7) = 1\} \\
&\quad + \text{Prob}\{X(8) = 1 \ \& \ X(7) = 2\} \\
&\quad + \text{Prob}\{X(8) = 1 \ \& \ X(7) = 3\} \\
&\quad + \text{Prob}\{X(8) = 1 \ \& \ X(7) = 4\}
\end{aligned}$$

$$\begin{aligned}
&= \text{Prob}\{X(8) = 1 \mid X(7) = 1\}\text{Prob}\{X(7) = 1\} \\
&+ \text{Prob}\{X(8) = 1 \mid X(7) = 2\}\text{Prob}\{X(7) = 2\} \\
&+ \text{Prob}\{X(8) = 1 \mid X(7) = 3\}\text{Prob}\{X(7) = 3\} \\
&+ \text{Prob}\{X(8) = 1 \mid X(7) = 4\}\text{Prob}\{X(7) = 4\}
\end{aligned}$$



Markov Processes

$$f_{X(t)|X(t_1)\&X(t_2)\&..\&X(t_k)}(x|x_1,x_2,\ldots,x_k)$$
$$=f_{X(t)|X(t_k)}(x|x_k)$$



Transition probabilities

$\text{Prob}\{X(8) = 1\}$

$= \text{Prob}\{X(8) = 1 \ \& \ X(7) = 1\}$

$+ \text{Prob}\{X(8) = 1 \ \& \ X(7) = 2\}$

$+ \text{Prob}\{X(8) = 1 \ \& \ X(7) = 3\}$

$+ \text{Prob}\{X(8) = 1 \ \& \ X(7) = 4\}$

$= \text{Prob}\{X(8) = 1 \mid X(7) = 1\}\text{Prob}\{X(7) = 1\}$

$+ \text{Prob}\{X(8) = 1 \mid X(7) = 2\}\text{Prob}\{X(7) = 2\}$

$+ \text{Prob}\{X(8) = 1 \mid X(7) = 3\}\text{Prob}\{X(7) = 3\}$

$+ \text{Prob}\{X(8) = 1 \mid X(7) = 4\}\text{Prob}\{X(7) = 4\}$

$$\begin{bmatrix} \Pr ob\{X(8)=1\} \\ \Pr ob\{X(8)=2\} \\ \Pr ob\{X(8)=3\} \\ \Pr ob\{X(8)=4\} \end{bmatrix} \equiv P(8)$$

$$= \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1r} \\ p_{21} & p_{22} & \cdots & p_{2r} \\ & \cdots & & \\ p_{r1} & p_{r2} & \cdots & p_{rr} \end{bmatrix} \bullet P(7)$$

$$\begin{bmatrix} \text{Prob}\{X(7)=1\} \\ \text{Prob}\{X(7)=2\} \\ \text{Prob}\{X(7)=3\} \\ \text{Prob}\{X(7)=4\} \end{bmatrix}$$

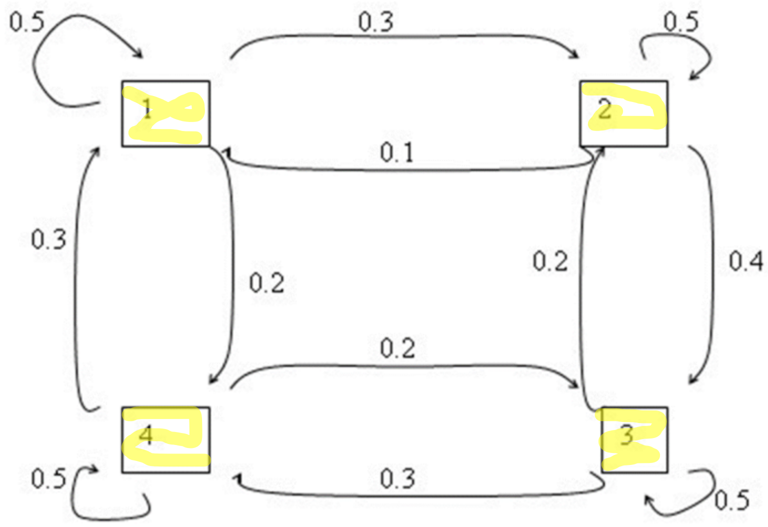
$\text{Prob}\{X(8) = 1\}$

$= \text{Prob}\{X(8) = 1 \;\&\; X(7) = 1\}$
$+ \text{Prob}\{X(8) = 1 \;\&\; X(7) = 2\}$
$+ \text{Prob}\{X(8) = 1 \;\&\; X(7) = 3\}$
$+ \text{Prob}\{X(8) = 1 \;\&\; X(7) = 4\}$

$= \text{Prob}\{X(8) = 1 \mid X(7) = 1\}\text{Prob}\{X(7) = 1\}$
$+ \text{Prob}\{X(8) = 1 \mid X(7) = 2\}\text{Prob}\{X(7) = 2\}$
$+ \text{Prob}\{X(8) = 1 \mid X(7) = 3\}\text{Prob}\{X(7) = 3\}$
$+ \text{Prob}\{X(8) = 1 \mid X(7) = 4\}\text{Prob}\{X(7) = 4\}$

$$\begin{bmatrix} \Pr ob\{X(8)=1\} \\ \Pr ob\{X(8)=2\} \\ \Pr ob\{X(8)=3\} \\ \Pr ob\{X(8)=4\} \end{bmatrix} \equiv P(8)$$

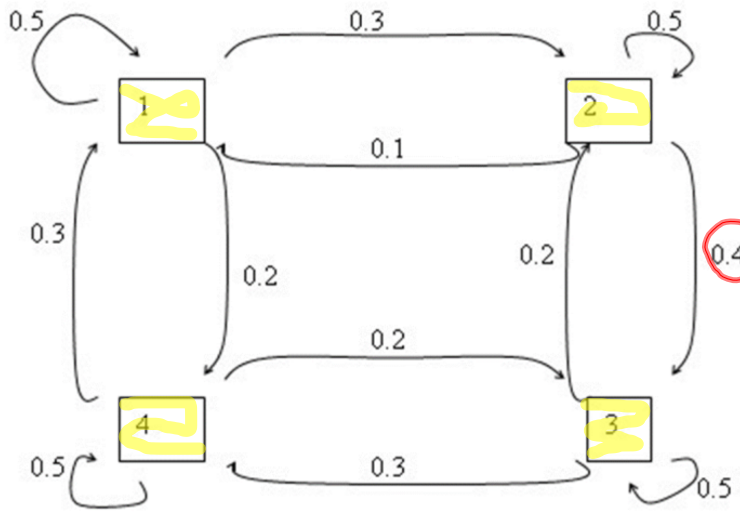
$$= \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1r} \\ p_{21} & p_{22} & \cdots & p_{2r} \\ & \cdots & & \\ p_{r1} & p_{r2} & \cdots & p_{rr} \end{bmatrix} \bullet P(7) \qquad \begin{bmatrix} \text{Prob}\{X(7)=1\} \\ \text{Prob}\{X(7)=2\} \\ \text{Prob}\{X(7)=3\} \\ \text{Prob}\{X(7)=4\} \end{bmatrix}$$



$$\mathbf{P}^{(1)}_{ij} = \text{Prob}\{X(n) = i \mid X(n\text{-}1) = j\}$$

$$\mathbf{P}^{(1)} = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \\ p_{41} & p_{42} & p_{43} & p_{44} \end{bmatrix}$$

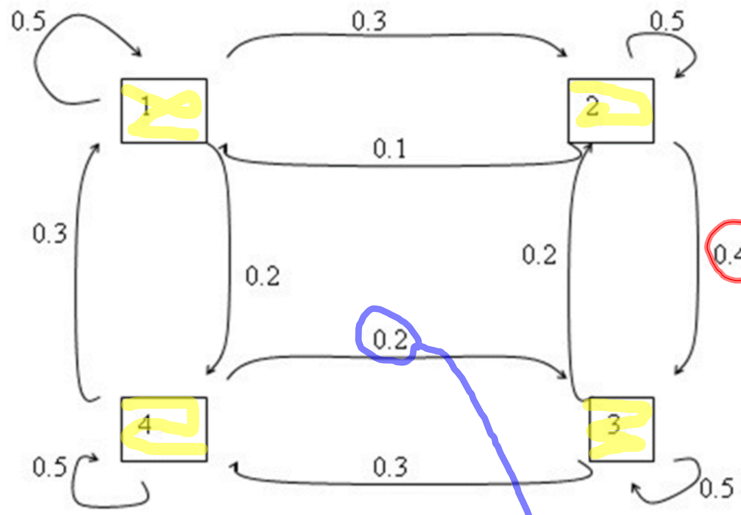
$$= \begin{bmatrix} .5 & .1 & 0 & .3 \\ .3 & .5 & .2 & 0 \\ 0 & .4 & .5 & .2 \\ .2 & 0 & .3 & .5 \end{bmatrix}$$



$$\mathbf{P}^{(1)}_{ij} = \text{Prob}\{X(n) = i \mid X(n\text{-}1) = j\}$$

$$\mathbf{P}^{(1)} = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \\ p_{41} & p_{42} & p_{43} & p_{44} \end{bmatrix}$$

$$= \begin{bmatrix} .5 & .1 & 0 & .3 \\ .3 & .5 & .2 & 0 \\ 0 & .4 & .5 & .2 \\ .2 & 0 & .3 & .5 \end{bmatrix}$$



$$\mathbf{P}^{(1)}{}_{ij} = \text{Prob}\{X(n) = i \mid X(n\text{-}1) = j\}$$

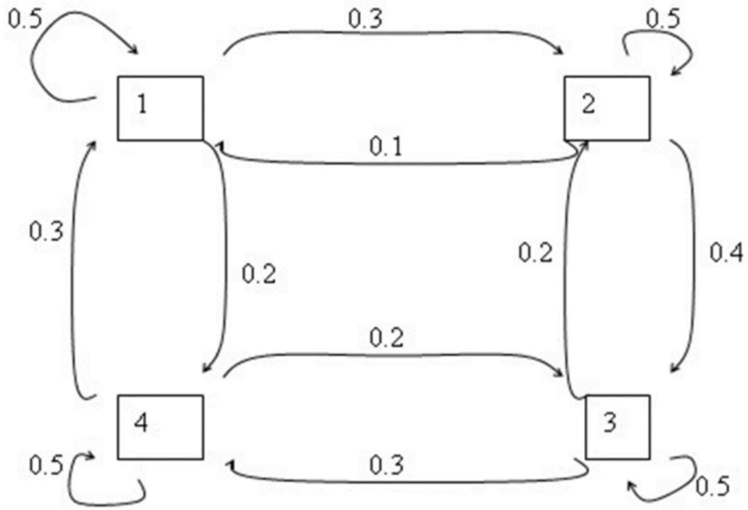
$$\mathbf{P}^{(1)} = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \\ p_{41} & p_{42} & p_{43} & p_{44} \end{bmatrix}$$

$$= \begin{bmatrix} .5 & .1 & 0 & .3 \\ .3 & .5 & .2 & 0 \\ 0 & .4 & .5 & .2 \\ .2 & 0 & .3 & .5 \end{bmatrix}$$

$$P(n) = \begin{bmatrix} \text{Prob}\{X(n) = 1\} \\ \text{Prob}\{X(n) = 2\} \\ \text{Prob}\{X(n) = 3\} \\ \text{Prob}\{X(n) = 4\} \end{bmatrix}$$

$$P(n+1) = \mathbf{P}^{(1)}\, P(n)$$

$$P(n+2) = \mathbf{P}^{(1)}\, P(n+1) = \{\mathbf{P}^{(1)}\}^2\, P(n) = \mathbf{P}^{(2)}\, P(n)$$



$$P^2 = \begin{bmatrix} .5 & .1 & 0 & .3 \\ .3 & .5 & .2 & 0 \\ 0 & .4 & .5 & .2 \\ .2 & 0 & .3 & .5 \end{bmatrix} \begin{bmatrix} .5 & .1 & 0 & .3 \\ .3 & .5 & .2 & 0 \\ 0 & .4 & .5 & .2 \\ .2 & 0 & .3 & .5 \end{bmatrix}$$

$$= \begin{bmatrix} .34 & .10 & .11 & .30 \\ .30 & .36 & .20 & .13 \\ .16 & .40 & .39 & .20 \\ .20 & .14 & .30 & .37 \end{bmatrix}$$

for $p > n > m$

$$\{\mathbf{P}^{(1)}\}^{p\text{-}m} = \{\mathbf{P}^{(1)}\}^{p\text{-}n}\{\mathbf{P}^{(1)}\}^{n\text{-}m}$$

$$\text{Prob}\{X(p) = i \mid X(m) = j\} =$$

$$\sum_k \text{Prob}\{X(p) = i \mid X(n) = k\}\ \text{Prob}\{X(n) = k \mid X(m) = j\}$$

$$f_{X(t_2)|X(t_1)}(x_2 \mid x_1) =$$

$$\int_{-\infty}^{\infty} f_{X(t_2)|X(t)}(x_2 \mid x)\, f_{X(t)|X(t_1)}(x \mid x_1)\, dx$$

$$(t_1 < t < t_2)$$

Chapman-Kolmogorov equation

$$P^1 = \begin{bmatrix} 0.5 & 0.3 & 0 & 0.2 \\ 0.2 & 0.5 & 0 & 0 \\ 0 & 0.2 & 1 & 0.3 \\ 0.3 & 0 & 0 & 0.5 \end{bmatrix}$$

$$\begin{bmatrix} 0.5 & 0.3 & 0 & 0.2 \\ 0.2 & 0.5 & 0 & 0 \\ 0 & 0.2 & 1 & 0.3 \\ 0.3 & 0 & 0 & 0.5 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

"Absorbing State"

$$\begin{bmatrix} .5 & .1 & 0 & .3 \\ .3 & .5 & .2 & 0 \\ 0 & .4 & .5 & .2 \\ .2 & 0 & .3 & .5 \end{bmatrix} \begin{bmatrix} 0.2037\ldots \\ 0.2407\ldots \\ 0.2963\ldots \\ 0.2593\ldots \end{bmatrix} = \begin{bmatrix} 0.2037\ldots \\ 0.2407\ldots \\ 0.2963\ldots \\ 0.2593\ldots \end{bmatrix}$$

"Equilibrium State"

(eigenvector)

2 questions

Does every finite-state stationary Markov process have equilibrium states?

Yes

$$\begin{bmatrix} .5 & .1 & 0 & .3 \\ .3 & .5 & .2 & 0 \\ 0 & .4 & .5 & .2 \\ .2 & 0 & .3 & .5 \end{bmatrix} \begin{bmatrix} 0.2037\ldots \\ 0.2407\ldots \\ 0.2963\ldots \\ 0.2593\ldots \end{bmatrix} = \begin{bmatrix} 0.2037\ldots \\ 0.2407\ldots \\ 0.2963\ldots \\ 0.2593\ldots \end{bmatrix}$$

"Equilibrium State"

$$\begin{bmatrix} .5 & .1 & 0 & .3 \\ .3 & .5 & .2 & 0 \\ 0 & .4 & .5 & .2 \\ .2 & 0 & .3 & .5 \end{bmatrix}^{50} =$$

$$\begin{bmatrix} 0.2037\ldots & 0.2037\ldots & 0.2037\ldots & 0.2037\ldots \\ 0.2407\ldots & 0.2407\ldots & 0.2407\ldots & 0.2407\ldots \\ 0.2963\ldots & 0.2963\ldots & 0.2963\ldots & 0.2963\ldots \\ 0.2593\ldots & 0.2593\ldots & 0.2593\ldots & 0.2593\ldots \end{bmatrix}$$

(You are eventually driven to an equilibrium state)

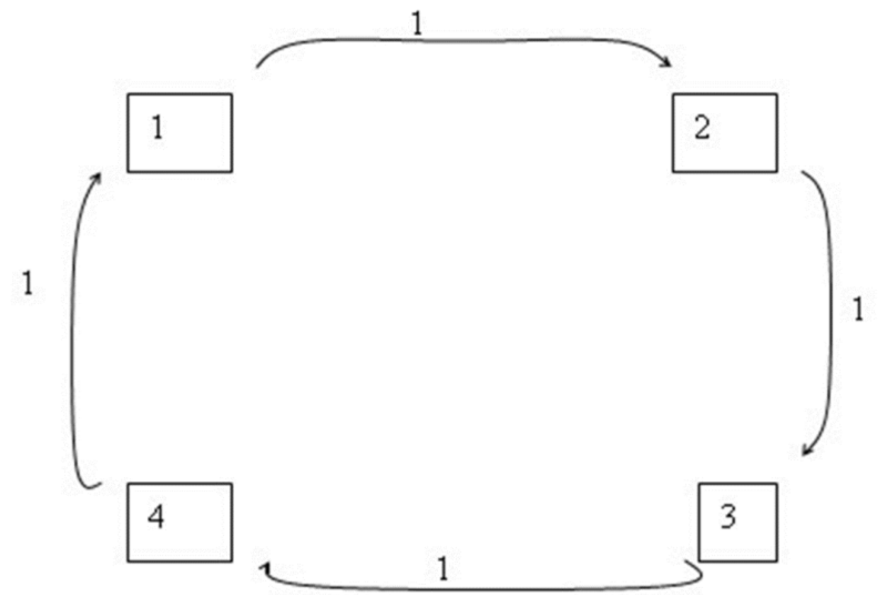
Does every initial configuration converge to an equilibrium state?

Yes, if P^1 contains no zeros.

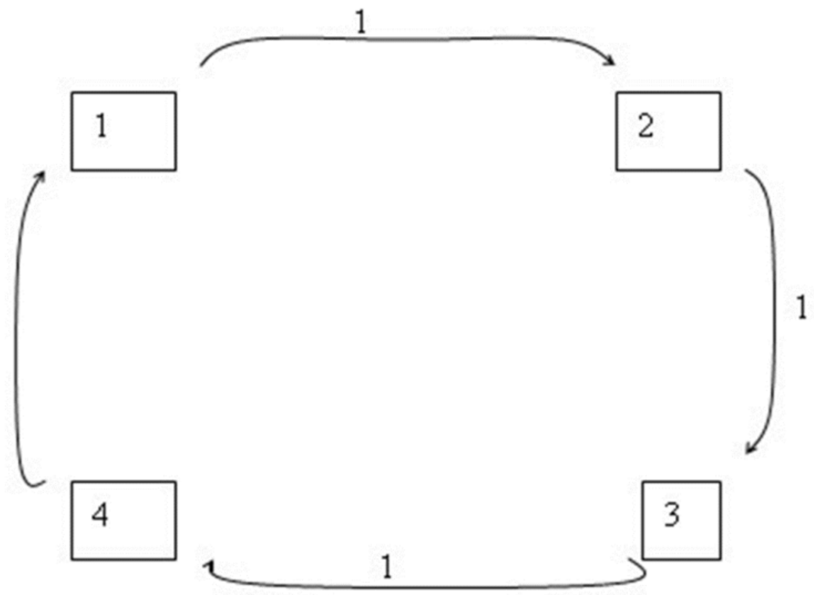
$$\begin{bmatrix} .5 & .1 & 0 & .3 \\ .3 & .5 & .2 & 0 \\ 0 & .4 & .5 & .2 \\ .2 & 0 & .3 & .5 \end{bmatrix}^{50} =$$

(although this P^1 does contain zeros)

$$\begin{bmatrix} 0.2037\ldots & 0.2037\ldots & 0.2037\ldots & 0.2037\ldots \\ 0.2407\ldots & 0.2407\ldots & 0.2407\ldots & 0.2407\ldots \\ 0.2963\ldots & 0.2963\ldots & 0.2963\ldots & 0.2963\ldots \\ 0.2593\ldots & 0.2593\ldots & 0.2593\ldots & 0.2593\ldots \end{bmatrix}$$



$$P^{(1)} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$



$$P^{(1)} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad [P^{(1)}]^2 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

$$[P^{(1)}]^3 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix} \quad [P^{(1)}]^4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$[P^{(1)}]^5 = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$