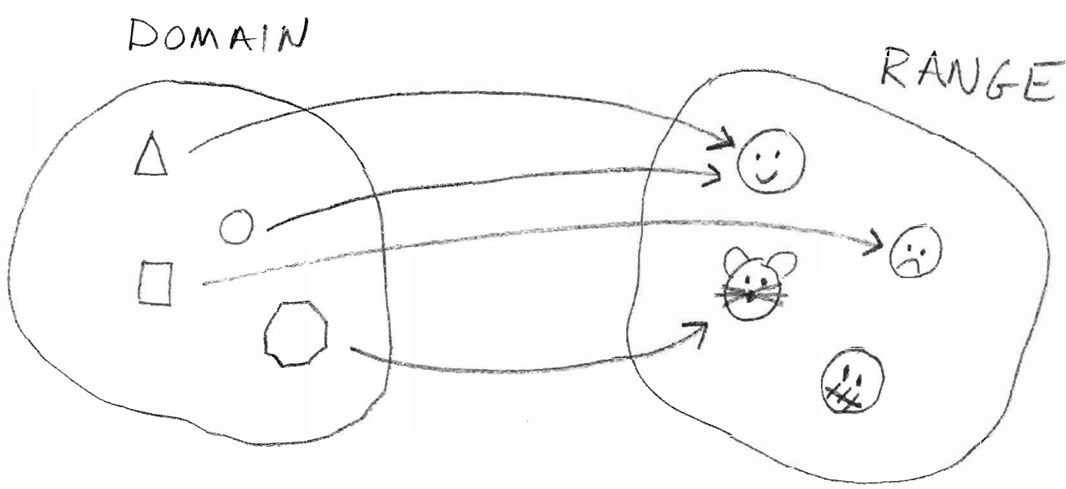
Probability & Random Variables

- Recall the formal definition of a function:
  - Associated with the function f are two sets, the <u>domain</u>, and the <u>range</u>.
  - The function is a <u>rule</u> that matches each member of the domain to one and only one member of the range.

<u>EX:</u>



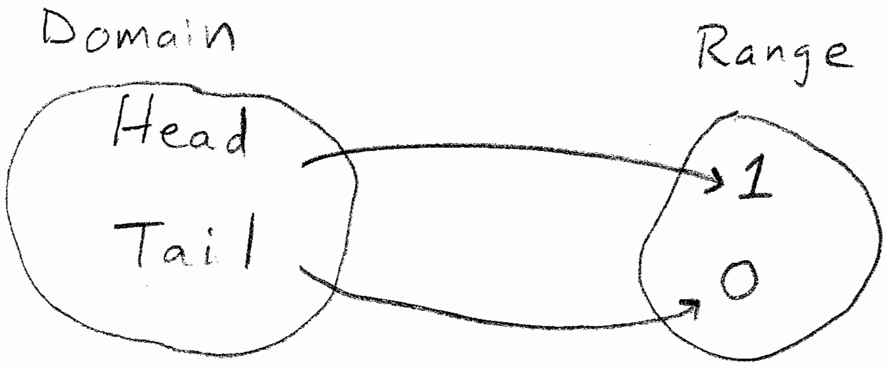
<u>EX:</u> Domain $= \mathbb{R}$

Range $= \mathbb{R}$

$f(x) = x^2$.

- A <u>random variable</u>, or <u>RV</u> is a function.
  - The domain is the set of possible outcomes of a chance, or random, experiment.
  - The range is a set of real numbers.

EX: A toss of a fair coin:



EX: The temperature at some specified place in Norman at a specified time on a specified day:

$$\text{Domain} = \{T : -200°F < T < 200°F\}$$

$$\text{Range} = \{x \in \mathbb{R} : -200 < x < 200\}$$

$$x = f(T) = T.$$

- The domain of an RV is called a <u>sample space</u>.
- The elements of a sample space must be disjoint; for any trial of the experiment, there will be one and only one outcome.
- We usually use capital letters like X, Y to denote random variables (RV's).
- We usually use lower case letters like x, y to denote an instance of the RV (a value from the range).

- If the number of elements in the domain (sample space) of an RV X is countable, then we call X a discrete RV.

- If the number of elements in the sample space is uncountable, then we call the RV X a continuous RV.

- The "set of events" is the set of all possible subsets of the sample space.

EX: Temperature in Norman.

Events:

$$T < 100° F$$

$$0 \leq T \leq 52° F$$

$$T = 71° F$$

$$-200° F \leq T \leq 200° F$$

$\emptyset$ (Null set -- No temperature)

EX: Draw a card at random from a fair deck.

Events:

Card = Queen of Spades

card = any Club

Card = any red

Card = not a face card

card = any card

- A <u>probability</u> <u>measure</u> is a function with
  - domain = the set of events
  - range = $[0,1]$.

<u>EX</u>: Toss a fair coin.

$P(\text{Head}) = P(1) = \frac{1}{2}$

$P(\text{Tail}) = P(0) = \frac{1}{2}$

$P(\text{Head or Tail}) = P(1 \cup 0) = 1$

$P\{(\text{Not Head}) \text{ AND } (\text{Not Tail})\} = P(\phi) = 0.$

<u>EX</u>: Draw a card from a fair deck.

$P(\text{Queen of Spades}) = \frac{1}{52}$

$P(\text{Any Club}) = \frac{1}{4}$

$P(\text{Any red}) = \frac{1}{2}$

$P(\text{Not a face card}) = \frac{40}{52}$

$P(\text{Any Card}) = 1$

<u>EX</u>: Temperature in Norman.

$P(-50°F \leq T \leq 150°F) = .999$

$P(T = \text{exactly } \pi°F) = 0.$

- A probability measure P must satisfy the three axioms:

  1) If A is an event, then $0 \le P(A) \le 1$.

  2) If S = the set of all possible outcomes, then

  $$P(S) = 1$$

  3) If $\{A_i\}$ is a set of <u>mutually</u> <u>exclusive</u> (disjoint) events, then

  $$P\left(\bigcup_i A_i\right) = \sum_i P(A_i).$$

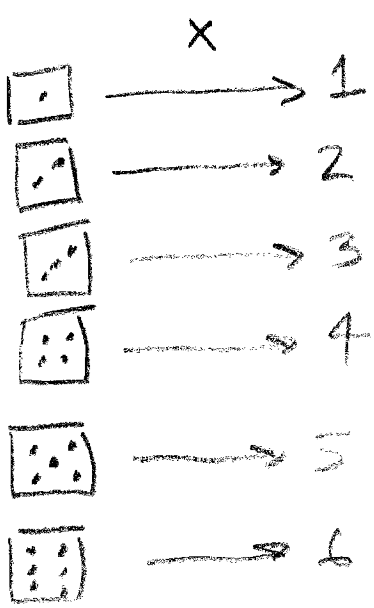
- The three axioms imply many things, including:

  1) $P(\text{NOT } A) = 1 - P(A)$

  2) $P(\phi) = 0$

  3) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

<u>EX</u>: Roll a fair die:



$$P(1) = P(2) = P(3) = P(4) = P(5) = P(6) = \frac{1}{6}$$

Probability Measure

"RV" X

- Let Event A = odd number of dots
$= 1 \cup 3 \cup 5$.

- Let Event B = number of dots < 4
$= 1 \cup 2 \cup 3$.

$$P(A) = P(B) = \frac{1}{2}$$

$$P(A \cap B) = P(1 \cup 3) = \frac{1}{6} + \frac{1}{6} - 0 = \frac{1}{3}$$

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

$$= \frac{1}{2} + \frac{1}{2} - \frac{1}{3} = \frac{2}{3}$$

DEF: The events A and B are called independent if $P(A \cap B) = P(A)P(B)$.

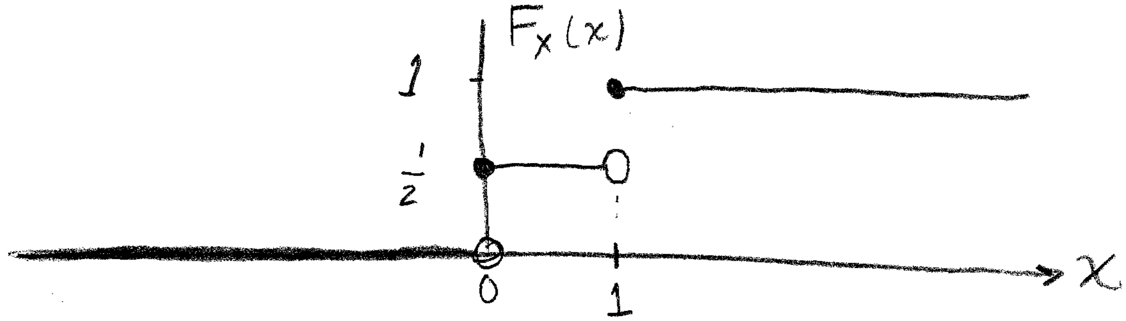
Probability Distribution Function

- For an RV X, the "cummulative distribution function", or "cdf", or simply "distribution", is defined by

$$F_X(x) = P(X \le x)$$

EX: Toss of a fair coin; Head = 1, Tail = 0.

$$F_X(x) = \begin{cases} 0, & x < 0 \\ \frac{1}{2}, & 0 \le x < 1 \\ 1, & x \geq 1 \end{cases}$$



Properties of cdf:

1. $\lim_{x \to -\infty} F_X(x) = 0$
2. $\lim_{x \to \infty} F_X(x) = 1$
3. $F_X(x)$ is nondecreasing in x.
4. $P(x_1 \le X \le x_2) = F_X(x_2) - F_X(x_1)$

Probability Density Function

- The "probability density function", or "pdf" is given by

$$f_X(x) = \frac{d}{dx} F_X(x)$$

pdf Properties

1. $F_X(x) = \int_{-\infty}^{x} f_X(\theta)\, d\theta$

2. $P(x_1 < X \leq x_2) = F_X(x_2) - F_X(x_1) = \int_{x_1^+}^{x_2} f_X(\theta)\, d\theta$

3. $\int_{-\infty}^{\infty} f_X(\theta)\, d\theta = 1$

4. $f_X(x) \geq 0$.

NOTES:

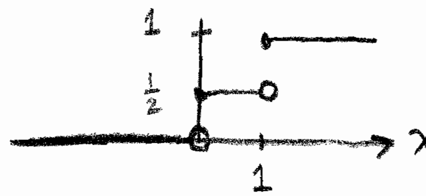
1. $f_X(x)$ can be interpreted as $P(X = x)$.
2. For a continuous RV, the probability of any specific outcome is generally zero; e.g., $P(\text{Temp} = \pi^\circ F) = 0$.

$\rightarrow$ You only get nonzero probability by considering events like

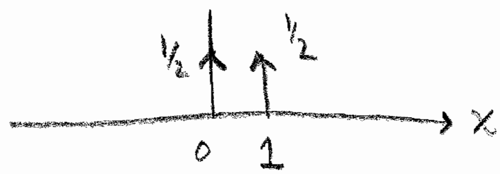
$$P(-100^\circ F \leq \text{Temp} \leq 150^\circ F) = 0.999$$

3. For a discrete RV like the fair coin toss, the density includes Dirac deltas and there are specific outcomes with nonzero probability:

$$\frac{d}{dx}\left\{F_X(x)\right\} = \frac{d}{dx}\left\{\quad\right\}$$



$$=$$



$$= \frac{1}{2}\delta(x) + \frac{1}{2}\delta(x-1).$$

Expectation

- If $g(X)$ is a function of the RV X,
  - → That is, if $g(\cdot)$ maps the number output by the RV X to another number,

  then the expected value of $g(X)$ is

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

Mean

- The mean, or expected value of the RV X is the expected value of the function $g(x) = x$:

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

- The mean is sometimes denoted $\bar{X}$, $E[X]$, EX, m_X, or μ.

EX: Fair coin toss:

$$E[X] = \int_{-\infty}^{\infty} x f_X(x) dx = \int_{-\infty}^{\infty} x \left[\tfrac{1}{2}\delta(x) + \tfrac{1}{2}\delta(x-1)\right] dx$$

$$= 0 \cdot \tfrac{1}{2} + 1 \cdot \tfrac{1}{2} = \tfrac{1}{2}.$$

EX: X is an RV with pdf $f_X(x)$, g_1 and g_2 are functions, α and β are constants:

$$E\left[\alpha g_1(x) + \beta g_2(x)\right] = \int_{-\infty}^{\infty} \left[\alpha g_1(x) + \beta g_2(x)\right] f_X(x) dx$$

$$= \alpha \int_{-\infty}^{\infty} g_1(x) f_X(x) dx + \beta \int_{-\infty}^{\infty} g_2(x) f_X(x)\, dx$$

$$= \alpha E\left[g_1(x)\right] + \beta E\left[g_2(x)\right].$$

Moments

- The "k^{th} moment" of an RV X is the expected value of the function $g(x) = x^k$:

$$E[X^k] = \int_{-\infty}^{\infty} x^k f_X(x)\, dx$$

- The zeroth moment is always equal to 1.
- The first moment is the mean.

Central Moments

- The "k^{th}" central moment of an RV X is the expected value of the function $g(x) = (x - \bar{X})^k$:

$$E\left[(X - \bar{X})^k\right] = \int_{-\infty}^{\infty} (x - \bar{X})^k f_X(x)\, dx.$$

- The zeroth central moment is always equal to 1.
- The first central moment is always zero.
- The second central moment is called the "Variance", usually denoted σ_X^2.

$$\sigma_X^2 = E[(X-\bar{X})^2]$$

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

$$= \int_{-\infty}^{\infty} (x^2 - 2x\bar{X} + \bar{X}^2) f_X(x)\, dx$$

$$= \int_{-\infty}^{\infty} x^2 f_X(x)\, dx - 2\bar{X}\int_{-\infty}^{\infty} x f_X(x)\, dx + \bar{X}^2 \int_{-\infty}^{\infty} f_X(x)\, dx$$

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

$$= E[X^2] - \left(E[X]\right)^2.$$

- The square root of the variance is called the "<u>standard deviation</u>", usually denoted σ_X:

$$\sigma_X = \sqrt{\sigma_X^2} = \sqrt{E[(X-\bar{X})^2]}$$

Gaussian RV

- The pdf for a "Gaussian" or "Normal" RV is

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]$$

  - The mean is $E[X] = \mu$.
  - The variance is $E[(X-\mu)^2] = \sigma^2$
  - The standard deviation is σ.

- The cdf cannot be written in closed form using elementary functions.
  - It is necessary to use tables or numerical integration

Two Random Variables

- Let X and Y be two RVs.
- The "joint distribution" of X and Y is

$$F_{X,Y}(x,y) = P\big([X \le x] \cap [Y \le y]\big).$$

- The "joint density" of X and Y is

$$f_{X,Y}(x,y) = \frac{\partial^2}{\partial x \partial y} F_{X,Y}(x,y)$$

- The joint cdf can be recovered from the joint pdf by integrating:

$$F_{X,Y}(x,y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{X,Y}(\alpha,\beta)\, d\beta\, d\alpha$$

- The joint pdf integrates to 1:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(\alpha,\beta)\, d\beta\, d\alpha = 1.$$

- The pdf of X <u>alone</u> is $f_X(x)$ and is called the "marginal density of X" in the bivariate case (the case where we have two RVs).

- Likewise, $f_Y(y)$ is called the marginal density of Y.

- The marginal densities can be recovered from the joint density by integrating out the unwanted variable:

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\, dy$$

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

- The marginal cdf's can then be obtained using integration as before, e.g.

$$F_X(x) = \int_{-\infty}^{x} f_X(\theta)\, d\theta .$$

- The RVs X and Y are called "independent" if

$$F_{X,Y}(x,y) = F_X(x)\,F_Y(y)$$

- This is the exact same thing as saying that the joint pdf is the product of the marginal pdfs:

$$\text{independent} \iff f_{X,Y}(x,y) = f_X(x)\,f_Y(y).$$

Functions of Two RVs

- Suppose $g: \mathbb{R}^2 \to \mathbb{R}$ is a function.
  - → For each trial of the experiment, the RVs X and Y map the outcome to a pair of numbers.
  - → $g(X,Y)$ maps this pair of numbers to another number.

- The expected value of $g(X,Y)$ is

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

- If X and Y are independent and $g(X,Y)$ is separable so that $g(X,Y) = g_1(X)\, g_2(Y)$, then

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

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g_1(x)\, g_2(y)\, f_X(x)\, f_Y(y)\, dx\, dy$$

$$= \left[\int_{-\infty}^{\infty} g_1(x) f_X(x)\, dx\right]\left[\int_{-\infty}^{\infty} g_2(y) f_Y(y)\, dy\right]$$

$$= E[g_1(X)]\, E[g_2(Y)].$$

EX: Suppose $g(X,Y) = \alpha X + \beta Y$, where α and β are constants. Let X and Y be independent. Then

$$E[g(X,Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} [\alpha x + \beta y] f_{X,Y}(x,y)\,dx\,dy$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} [\alpha x + \beta y] f_X(x) f_Y(y)\,dx\,dy$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \alpha x f_X(x) f_Y(y)\,dx\,dy + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \beta y f_X(x) f_Y(y)\,dx\,dy$$

$$= \alpha \int_{-\infty}^{\infty} x f_X(x) \underbrace{\int_{-\infty}^{\infty} f_Y(y)\,dy}_{1}\,dx + \beta \int_{-\infty}^{\infty} y f_Y(y) \underbrace{\int_{-\infty}^{\infty} f_X(x)\,dx}_{1}\,dy$$

$$= \alpha \int_{-\infty}^{\infty} x f_X(x)\,dx + \beta \int_{-\infty}^{\infty} y f_Y(y)\,dy$$

$$= \alpha E[X] + \beta E[Y].$$

Joint Moments

- The joint moment of order k, l of the RVs, X and Y is

$$E[X^k Y^l] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^k y^l f_{X,Y}(x,y)\,dx\,dy$$

Joint Central Moments

- The joint central moment of order k, l of the RVs, X and Y is

$$E\left[(X-E[X])^k (Y-E[Y])^l\right]$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x-E[X])^k (y-E[Y])^l f_{X,Y}(x,y)\,dx\,dy$$

- When $k = l = 1$, this is called the "<u>COVARIANCE</u>" of X and Y:

$$\mathrm{cov}(X,Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x-E[X])(y-E[Y])\, f_{X,Y}(x,y)\,dx\,dy$$

<u>FACT:</u> $|\mathrm{Cov}(X,Y)| \leq \sigma_X \sigma_Y$

- The correlation coefficient between X and Y is

$$\rho_{X,Y} = \frac{\mathrm{Cov}(X,Y)}{\sigma_X \sigma_Y}$$

$\Rightarrow$ so $|\rho_{X,Y}| \leq 1$.

- If $\rho_{X,Y} = 1$, then $X = \alpha Y$ where α is a positive constant.
- If $\rho_{X,Y} = -1$, Then $X = \alpha Y$ where α is a negative constant (anticorrelation).
- If $\rho_{X,Y} = 0$, then X and Y are uncorrelated.

$\Rightarrow$ $\rho_{X,Y} = 1$ or $\rho_{X,Y} = -1$: perfect correlation; each RV carries all the information about the other

$\Rightarrow$ $\rho_{X,Y} = 0$: no correlation; each RV carries no information about the other

Fact: $\text{Cov}(X,Y) = E[XY] - E[X]E[Y]$.

- If X and Y are independent, then

$$E[XY] = E[X]E[Y]$$

$\Rightarrow$ So independence implies $\text{cov}(X,Y) = 0$.

$\Rightarrow$ This implies $\rho_{X,Y} = 0$.

$\Rightarrow$ So independence <u>implies</u> uncorrelated.

☆: The converse is <u>NOT</u> true.

- The idea of two RVs easily extends to three or more.

trivariate distribution:

$$F_{X,Y,Z}(x,y,z) = P\left([X \le x] \cap [Y \le y] \cap [Z \le z]\right)$$

trivariate density:

$$f_{X,Y,Z}(x,y,z) = \frac{\partial^3}{\partial x \partial y \partial z} F_{X,Y,Z}(x,y,z)$$

Covariance Matrix

- Let $\vec{X}$ and $\vec{Y}$ be two N-dimensional vectors of RVs:

$$\vec{X} = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_N \end{bmatrix}, \quad \vec{Y} = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_N \end{bmatrix}$$

- Then the product $\vec{X}\vec{Y}^T$ is an $N \times N$ matrix:

$$\begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_N \end{bmatrix} \begin{bmatrix} Y_1 & Y_2 & \cdots & Y_N \end{bmatrix} = \begin{bmatrix} X_1Y_1 & X_1Y_2 & \cdots & X_1Y_N \\ X_2Y_1 & X_2Y_2 & \cdots & X_2Y_N \\ \vdots & & \ddots & \\ X_NY_1 & X_NY_2 & \cdots & X_NY_N \end{bmatrix}$$

- The <u>COVARIANCE MATRIX</u> of $\vec{X}$ and $\vec{Y}$ is

$$\mathrm{Cov}(\vec{X},\vec{Y}) = \begin{bmatrix} \mathrm{Cov}(X_1,Y_1) & \mathrm{Cov}(X_1,Y_2) & \cdots & \mathrm{Cov}(X_1,Y_N) \\ \mathrm{Cov}(X_2,Y_1) & \mathrm{Cov}(X_2,Y_2) & \cdots & \mathrm{Cov}(X_2,Y_N) \\ \vdots & & & \\ \mathrm{Cov}(X_N,Y_1) & \cdots & \cdots & \mathrm{Cov}(X_N,Y_N) \end{bmatrix}$$

$$= E\left[(\vec{X} - E[\vec{X}])(\vec{Y} - E[\vec{Y}])^T\right]$$

- $Cov(\vec{X}, \vec{X})$ is called the covariance matrix of the random vector $\vec{X}$, and is usually written $Cov(\vec{X})$.

FACT:

$$Cov(\vec{X}) = \begin{bmatrix} \sigma_{X_1}^2 & \rho_{X_1,X_2}\sigma_{X_1}\sigma_{X_2} & \cdots & \rho_{X_1,X_N}\sigma_{X_1}\sigma_{X_N} \\ \rho_{X_1,X_2}\sigma_{X_1}\sigma_{X_2} & \sigma_{X_2}^2 & \cdots & \rho_{X_2,X_N}\sigma_{X_2}\sigma_{X_N} \\ \vdots & & \ddots & \\ \rho_{X_1,X_N}\sigma_{X_1}\sigma_{X_N} & \cdots & \cdots & \sigma_{X_N}^2 \end{bmatrix}$$

$\Rightarrow$ $Cov(\vec{X})$ is symmetric

$\Rightarrow$ $Cov(\vec{X})$ is positive semidefinite: all eigenvalues are non-negative.

Joint Gaussian Variables

- Let $\vec{X} = \begin{bmatrix} X \\ Y \end{bmatrix}$, where X and Y are RV's.

- Write $m_{\vec{X}}$ for $E[\vec{X}] = \begin{bmatrix} E[X] \\ E[Y] \end{bmatrix}$

- Let $C_{\vec{X}}$ be the covariance matrix of $\vec{X}$:

$$C_{\vec{X}} = Cov(\vec{X}) = \begin{bmatrix} \sigma_X^2 & \rho\sigma_X\sigma_Y \\ \rho\sigma_X\sigma_Y & \sigma_Y^2 \end{bmatrix},$$

where $\rho = Corr(X, Y)$.

- Then $\det C_{\vec{X}} = \sigma_X \sigma_Y (1-\rho^2)$.

- Also,

$$C_{\vec{X}}^{-1} = \begin{bmatrix} \dfrac{1}{(1-\rho^2)\sigma_X^2} & \dfrac{-\rho}{(1-\rho^2)\sigma_X\sigma_Y} \\ \dfrac{-\rho}{(1-\rho^2)\sigma_X\sigma_Y} & \dfrac{1}{(1-\rho^2)\sigma_Y^2} \end{bmatrix}$$

- For an N-dimensional random vector

$$\vec{X} = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_N \end{bmatrix}$$

of jointly Gaussian variables $X_1, X_2, \ldots, X_N$, the multivariate joint density is

$$f_{\vec{X}}(\vec{x}) = \frac{1}{(2\pi)^{\frac{N}{2}} (\det C_{\vec{X}})^{1/2}} \exp\left\{ -\frac{1}{2} (\vec{x} - m_{\vec{X}})^T C_{\vec{X}}^{-1} (\vec{x} - m_{\vec{X}}) \right\}$$

- In the 2-D case, $\vec{X} = \begin{bmatrix} X \\ Y \end{bmatrix}$, this becomes

$$f_{\vec{X}}(\vec{x}) = f_{X,Y}(x,y) = \frac{1}{2\pi \sigma_X \sigma_Y \sqrt{1-\rho^2}} \exp\left\{ \frac{-1}{2(1-\rho^2)} \times \left[\frac{(x-E[X])^2}{\sigma_X^2} - \frac{2\rho (x-E[X])(y-E[Y])}{\sigma_X \sigma_Y} + \frac{(y-E[Y])^2}{\sigma_Y^2} \right] \right\}$$

- if we suppose further that X and Y are <u>uncorrelated</u>, then $\rho = 0$ and

$$C_{\vec{X}} = \text{diag}(\sigma_1^2 \ \sigma_2^2).$$

- In this case the pdf becomes

$$f_{X,Y}(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y} \exp\left\{-\frac{1}{2}\left[\frac{(x-E[X])^2}{\sigma_X^2} + \frac{(y-E[Y])^2}{\sigma_Y^2}\right]\right\}$$

$$= \frac{1}{\sqrt{2\pi\sigma_X^2}} e^{-(x-E[X])^2/2\sigma_X^2}$$

$$\times \frac{1}{\sqrt{2\pi\sigma_Y^2}} e^{-(y-E[Y])^2/2\sigma_Y^2}$$

$$= f_X(x) f_Y(y).$$

$\Rightarrow$ Thus, for <u>joint Gaussian</u> RV's,

uncorrelated $\Longleftrightarrow$ independent.

FACT: if X and Y are two independent RUs with marginal densities $f_X(x)$ and $f_Y(y)$, and we define a new RU $Z = X + Y$, then the density of Z is given by

$$f_Z(z) = f_X(z) * f_Y(z) \quad \{\text{convolution}\}$$

$$= \int_{-\infty}^{\infty} f_X(\theta) f_Y(z-\theta)\, d\theta$$

$$= \int_{-\infty}^{\infty} f_X(z-\theta) f_Y(\theta)\, d\theta.$$

Transformation of on RV

- Let X be an RV with pdf $f_X(x)$.
- Let $g: \mathbb{R} \to \mathbb{R}$ be a function that is "one-to-one".

$\Rightarrow$ This means $g^{-1}(\cdot)$ is also a function.

- Define the new RV $Y = g(X)$.
- Then the pdf of Y is given by

$$f_Y(y) = |g^{-1}(y)| \, f_X(g^{-1}(y))$$

<u>EX</u>: $Y = g(X) = 5X + 3$.

$$X = g^{-1}(Y) = \frac{Y-3}{5}$$

$$f_Y(y) = \left| \frac{y-3}{5} \right| f_X\left(\frac{y-3}{5}\right).$$

Stochastic Processes

- To model deterministic signals, we used ordinary functions.

  <u>EX</u>: $$x(t) = e^{-2t} \cos\left(\tfrac{5}{7\pi} t\right) u(t)$$

- For a statistical signal, we have no precise knowledge of the values the signal takes at specific times.

- We model a statistical signal with a thing called a "<u>stochastic process</u>."

- A stochastic process is a collection of RV's, one for each "time" on the "t-axis"

  → For a discrete-time statistical signal $x(n)$, there is an RV for each $n = \ldots -1, 0, 1, 2, \ldots$.

  → For a continuous-time statistical signal $x(t)$, there is an RV for each $t \in \mathbb{R}$.

  ⇒ All of the RV's in the stochastic process have the <u>same</u> <u>domain</u> (underlying experiment space).

- For any one trial of the experiment, each of the RV's maps the outcome to a number.

- This results in a number for each <u>time</u>.

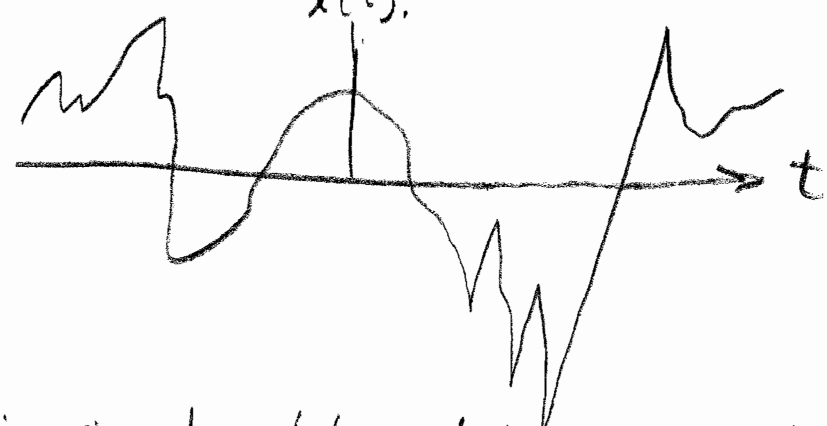
<u>EX</u>: $x(t)$ is a continuous-time stochastic process.

→ There is an RV at each time $t \in \mathbb{R}$.

→ The experiment is the rolling of a fair die.

→ When the die is rolled, each RV maps the outcome to a number.

→ This results in a number at every $t \in \mathbb{R}$, which gives us a continuous-time signal: $x(t)$.



⟹ This signal, obtained from a single experimental outcome, is called a "<u>sample function</u>", or a "<u>realization</u>" of the stochastic process $x(t)$.

- To completely describe a stochastic process $x(t)$, you must specify the joint density (or distribution) for <u>all</u> of the RV's that make up the process (i.e., for all times).

- often, this is not necessary and we describe the process by specifying certain joint moments and/or joint central moments.

$\Longrightarrow$ While this does not completely describe the stochastic process, it often gives enough information to solve many interesting problems.

<u>DEF</u>: Suppose $x(t)$ and $y(t)$ are two stochastic processes.

$\rightarrow$ Then, for each $t \in \mathbb{R}$, $x(t)$ is an RV and $y(t)$ is an RV.

- The processes $x(t)$ and $y(t)$ are <u>independent</u> if $x(\alpha)$ and $y(\beta)$ are independent RV's $\forall\, \alpha, \beta \in \mathbb{R}$.

Correlation & Covariance

- Suppose $x[n]$ and $y[n]$ are two discrete-time stochastic processes.

- The "Cross Correlation" function between $x[n]$ and $y[n]$ is defined by

$$R_{x,y}(k,\ell) = E\{x[k]\,y[\ell]\}$$

- The corresponding central moment is called the "Cross Covariance", given by

$$\begin{aligned}
C_{x,y}(k,\ell) &= \mathrm{Cov}(x[k], y[\ell]) \\
&= E\left\{\left(x[k] - E[x[k]]\right)\left(y[k] - E[y[k]]\right)\right\} \\
&= E\{x[k]\,y[\ell]\} - E\{x[k]\}\,E\{y[\ell]\}.
\end{aligned}$$

NOTE: $R_{x,y}(k,\ell) = \mathrm{Cov}(x[k], y[\ell]) + E\{x[k]\}\,E\{y[\ell]\}$

DEF: The "Autocorrelation" of the stochastic process $x[k]$ is the cross correlation of $x[k]$ with itself:

$$R_x(k,\ell) = R_{x,x}(k,\ell)$$

$$= E\{x[k]\,x[\ell]\}.$$

$\Rightarrow$ The concept of autocorrelation is super super important

- Now suppose that $x(t)$ and $y(t)$ are two continuous-time stochastic processes.
- Let α and β be two real variables.
- Then the "cross correlation" of $x(t)$ and $y(t)$ is given by

$$R_{x,y}(\alpha,\beta) = E\{x(\alpha)\,y(\beta)\}.$$

- The corresponding central moment is called "cross covariance", given by

$$C_{x,y}(\alpha,\beta) = \mathrm{Cov}(x(\alpha), y(\beta))$$
$$= E\{(x(\alpha)-E[x(\alpha)])(y(\beta)-E[y(\beta)])\}$$
$$= R_{x,y}(\alpha,\beta) - E\{x(\alpha)\}E\{y(\beta)\}.$$

<u>DEF</u>: the "<u>Autocorrelation</u>" of the process $x(t)$ is the cross correlation of $x(t)$ with itself:

$$R_x(\alpha,\beta) = R_{x,x}(\alpha,\beta)$$
$$= E\{x(\alpha)x(\beta)\}$$

$\Rightarrow$ This is an <u>important</u> concept. ★★★

- Recall: A stochastic process is a collection of RV's, one for each <u>time</u>.

→ If we choose a set of times $t_1, t_2, t_3, \ldots$ we can speak of the joint density of the RV's $x(t_1), x(t_2), x(t_3), \ldots$

Stationarity

- The stochastic process $x(t)$ is called "<u>strict sense stationary</u>" (SSS) if the joint density of any number of the involved RVs is <u>invariant</u> under time translation.

<u>EX:</u>

$$f_{x(t_1),x(t_2)}(x(t_1),x(t_2))$$

$$= f_{x(t_1+\Delta),x(t_2+\Delta)}(x(t_1+\Delta),x(t_2+\Delta))$$

- The stochastic process $X(t)$ is called "<u>wide sense stationary</u>" (WSS) if the autocorrelation and mean are invariant under time translation:

$$E\{X(t_0)\} = E\{X(t_0+\Delta)\}$$

$$R_X(t_0,t_1) = R_X(t_0+\Delta, t_1+\Delta).$$

$\Rightarrow$ In this case, $R_X(t_0,t_1)$ depends only on the "time difference" $|t_1-t_0|$.

- So we write

$$R_x(t_0, t_1) = R_x(t_0+\Delta, t_1+\Delta)$$
$$= R_x(\tau),$$

where $\tau = |t_1 - t_0|$.

$\Rightarrow$ For a WSS process, the first and second order moments do not change with time.

$\Rightarrow$ For a WSS process $x(t)$, the following properties hold;

1. $R_x(0) = E\{x^2(t)\} \geq 0$.

2. $R_x(-\tau) = E\{x(t-\tau)\,x(t)\}$
$$= E\{x(t)\,x(t+\tau)\}$$
$$= R_x(\tau)$$

$\rightarrow$ $R_x(\tau)$ is <u>even</u>.

3. $|R_x(\tau)| \leq R_x(0)$.

Joint Stationarity

- Suppose $x(t)$ and $y(t)$ are two WSS processes.
- The cross correlation between $x(t)$ and $y(t)$ is given by

$$R_{x,y}(\alpha,\beta) = E\{x(\alpha)y(\beta)\}.$$

- Suppose that the cross correlation depends only on $\tau = |\beta-\alpha|$, so that

$$R_{x,y}(\tau) = R_{x,y}(\alpha,\beta) = R_{x,y}(|\beta-\alpha|).$$

- Then the processes $x(t)$ and $y(t)$ are called "jointly wide sense stationary".

- The following properties then hold:

1. $R_{x,y}(0) = R_{y,x}(0)$
2. $R_{x,y}(\tau) = R_{y,x}(-\tau)$
3. $|R_{x,y}(\tau)| \leq [R_x(0)R_y(0)]^{1/2}$.

Power Spectral Density

- Suppose $X(t)$ is a WSS process.
- The "<u>Power Spectral Density</u>" of $X(t)$ is the Fourier Transform of the autocorrelation:

$$S_X(\omega) = \mathcal{F}[R_X(\tau)] = \int_{-\infty}^{\infty} R_X(\tau) e^{-j\omega\tau} d\tau .$$

- The power spectral density is also known as:
  - → PSD
  - → Power Spectrum
  - → Spectral Density

FACTS:

- $R_X(\tau)$ is real and even.
- $S_X(\omega)$ is real, even, and non-negative.

- The autocorrelation can be recovered from the PSD using the inverse Fourier Transform:

$$R_X(\tau) = \mathcal{F}^{-1}[S_X(\omega)] = \frac{1}{2\pi}\int_{-\infty}^{\infty} S_X(\omega) e^{j\omega\tau} d\omega$$

- Plugging in $\tau=0$, we obtain a usefull formula for the second moment of the WSS process $x(t)$:

$$R_x(0) = \frac{1}{2\pi}\int_{-\infty}^{\infty} S_x(\omega) e^{j\omega\tau} d\omega \Big|_{\tau=0}$$

$$= \frac{1}{2\pi}\int_{-\infty}^{\infty} S_x(\omega)\, d\omega$$

$$= E\{x^2(t)\}.$$

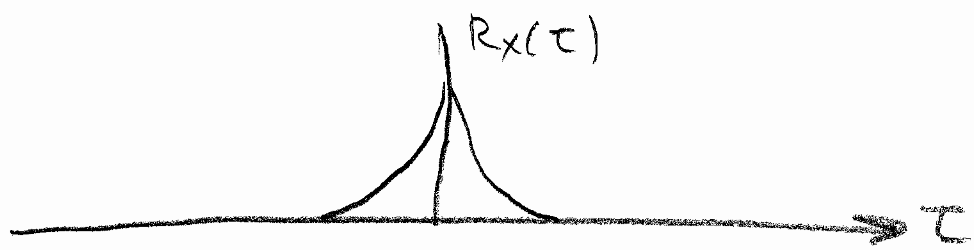
$\Rightarrow$ This is the <u>mean power</u> of the process $x(t)$.

<u>Interpretation</u>:

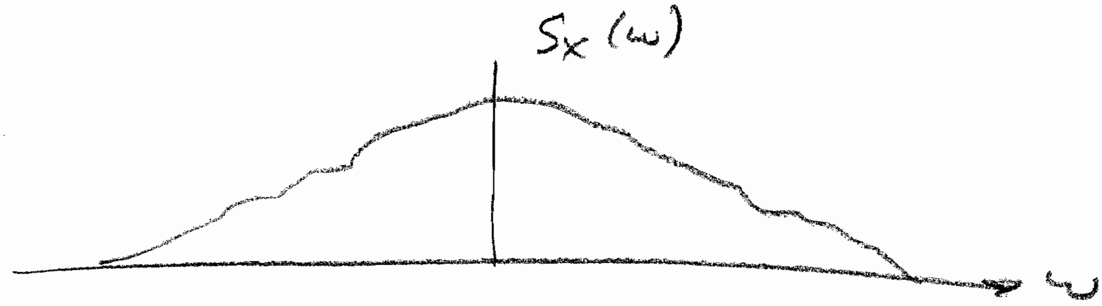
w.s.s.

- Suppose $x(t)$ varies <u>rapidly</u>, so that there is significant correlation between the RV's $x(t)$ and $x(t+\tau)$ only for $|\tau|$ small.

$\rightarrow$ Then $R_x(\tau)$ falls off rapidly:



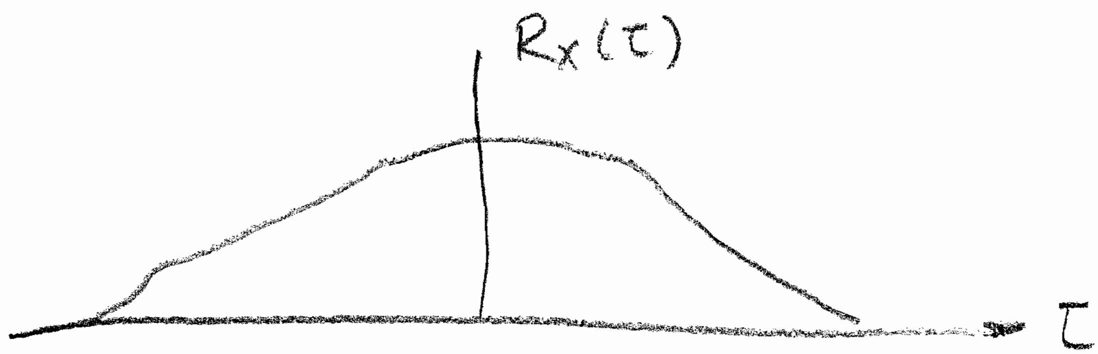
- By the reciprocal spreading principle, this <u>implies</u> that $S_x(\omega)$ falls off slowly;



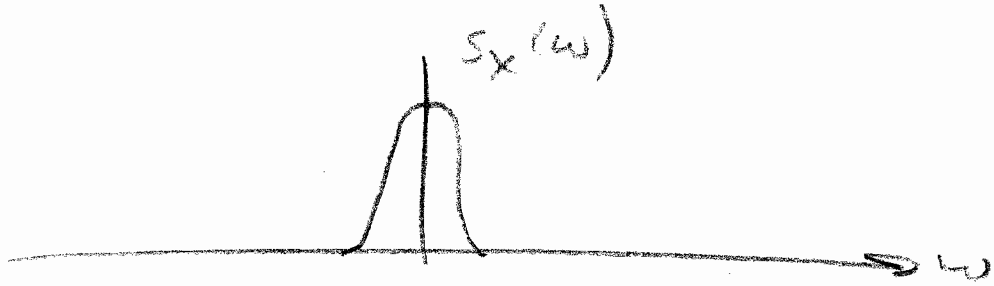
$\Rightarrow$ In other words: Rapid time variation implies significant high frequency content in the stochastic process $x(t)$.

- Likewise, suppose $x(t)$ is a W.S.S. process that varies <u>slowly</u> in time, so that there is significant correlation between $x(t)$ and $x(t+\tau)$ for $|\tau|$ large.

$\rightarrow$ Then $R_x(\tau)$ falls off slowly:



– The reciprocal spreading principle now suggests that $S_x(\omega)$ falls off rapidly:



$\Rightarrow$ In other words, slow time variation means that the process $x(t)$ is dominated by low frequencies.

NOTE: This seems a lot like our treatment of deterministic signals. Yet, here we have not specified the values of $X(t)$ at any times. In fact, we have only said something about the second-order statistics of $x(t)$.

– While we do know the autocorrelation of $x(t)$, we do not know the values of the signal at any specific times.

NOTE: Given the autocorrelation $R_X(\tau)$, we could also obtain a frequency representation using the Laplace transform:

$$S_X(s) = \mathcal{L}\{R_X(\tau)\} = \int_{-\infty}^{\infty} R_X(\tau) e^{-s\tau} d\tau$$

$\Rightarrow$ This is done often, and is also called the "power spectral density", or "PSD".

- For a discrete-time WSS process $X[n]$, the autocorrelation is

$$R_X(k) = E\{X[n] X[n+k]\}.$$

- As in the continuous-time case, the PSD is the Fourier transform of the autocorrelation:

$$S_X(e^{j\omega}) = \mathcal{F}[R_X(k)] = \sum_{k=-\infty}^{\infty} R_X(k) e^{-j\omega k}$$

$$R_X(k) = \mathcal{F}^{-1}[S_X(e^{j\omega})] = \frac{1}{2\pi}\int_{-\pi}^{\pi} S_X(e^{j\omega}) e^{j\omega k} d\omega.$$

- The discrete-time PSD can also be defined in terms of the Z-transform:

$$S_X(z) = \mathcal{Z}[R_X(k)] = \sum_{k=-\infty}^{\infty} R_X(k) z^{-k}$$

Cross Power Spectrum

- Let $x(t)$ and $y(t)$ be jointly W.S.S. processes with crosscorrelation functions $R_{xy}(\tau)$ and $R_{yx}(\tau)$,

$\Rightarrow$ Recall: $R_{y,x}(\tau) = R_{x,y}(-\tau)$

- The cross power spectra of $x(t)$ and $y(t)$ are given by

$$S_{x,y}(\omega) = \mathcal{F}\left[R_{x,y}(\tau)\right] = \int_{-\infty}^{\infty} R_{x,y}(\tau) e^{-j\omega\tau} d\tau$$

$$S_{y,x}(\omega) = \mathcal{F}\left[R_{y,x}(\tau)\right] = \int_{-\infty}^{\infty} R_{y,x}(\tau) e^{-j\omega\tau} d\tau$$

- The relationship between $S_{x,y}(\omega)$ and $S_{y,x}(\omega)$ is

$$S_{x,y}(\omega) = S_{y,x}^{*}(\omega).$$

- Analogous to the correlation coefficient between two RV's, we define the "coherence" of the WSS processes $x(t)$ and $y(t)$ by

$$\gamma_{x,y}^{2}(\omega) = \frac{|S_{x,y}(\omega)|^2}{S_x(\omega) S_y(\omega)}$$

$\Rightarrow$ This is like a "frequency domain" correlation coefficient.

- The magnitude of the coherence function is always less than or equal to one.

- In the maximum correlation case, we have

$$\gamma^2_{xx}(\omega) = \frac{|S_{xx}(\omega)|^2}{\underbrace{S_x(\omega)\, S_x(\omega)}_{\text{Real!!}}} = 1$$

Real!! so $S_x(\omega) S_x(\omega) = |S_x(\omega)|^2$

- The minimum coherence occurs when $x(t)$ and $y(t)$ have zero cross correlation. Then

$$\gamma^2_{x,y}(\omega) = \mathcal{F}[0] = 0.$$

- For two jointly WSS discrete-time processes $x[n]$ and $y[n]$, the cross power spectral density is given by

$$S_{x,y}(e^{j\omega}) = \mathcal{F}\left[R_{x,y}(k)\right] = \sum_{k=-\infty}^{\infty} R_{x,y}(k) e^{-j\omega k}$$

-or-

$$S_{x,y}(z) = \mathcal{Z}\left[R_{x,y}(k)\right] = \sum_{k=-\infty}^{\infty} R_{x,y}(k) z^{-k}$$

White Noise

- A continuous-time WSS process $x(t)$ is called a "white noise" if

$$R_x(\tau) = \alpha \delta(\tau), \quad \alpha \text{ constant.}$$

$\Rightarrow$ In this case, the PSD is given by

$$S_x(\omega) = \mathcal{F}[\alpha \delta(\tau)] = \alpha \quad (\text{constant}).$$

White Noise $\longleftrightarrow$ Contains equal amounts of all frequencies.

(Think of white light).

$\Rightarrow$ if $x(t)$ is a white noise, then

$$R_x(\tau) = \alpha \delta(\tau),$$

and the RV's @ different times are mutually uncorrelated.

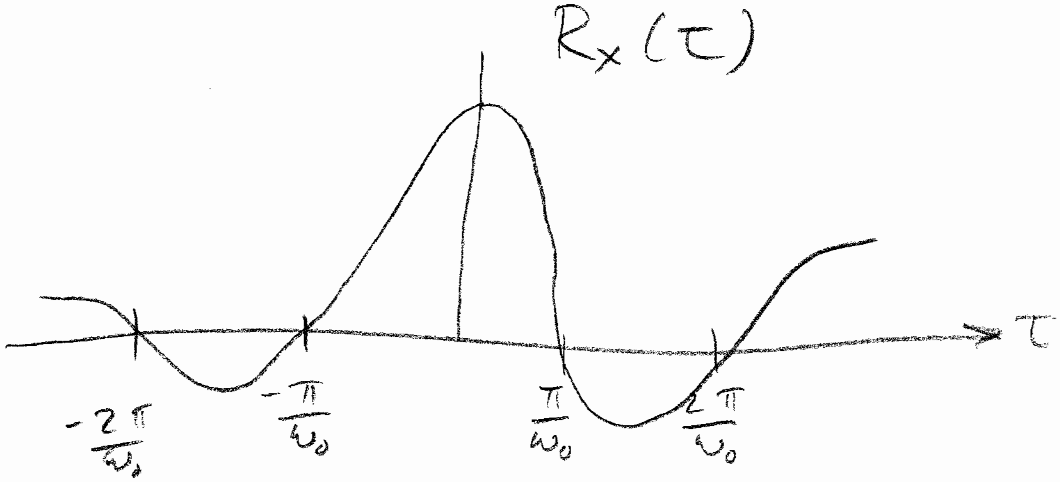
$\rightarrow$ For this reason, "white noise" is also called "uncorrelated noise".

- For a discrete-time white noise process,

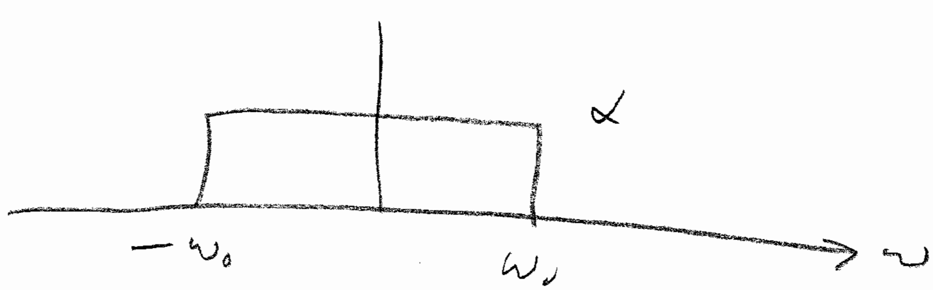
$$R_x(k) = \alpha\,\delta[k], \qquad \alpha \text{ constant.}$$

$$S_x(e^{j\omega}) = \alpha$$

<u>EX</u>: Suppose $R_x(\tau) = \alpha\omega_0 \dfrac{\sin \omega_0 \tau}{\pi \omega_0 \tau}$:



- Then the PSD is constant in the interval $-\omega_0 \le \omega \le \omega_0$, and zero outside this interval:



$\Rightarrow$ This is called "<u>Band Limited White Noise</u>"

IID Process

- If the RVs comprising the process $x(t)$ or $x[n]$ are all mutually independent and all have the same pdf, then the process $x(t)$ is called "Independent, Identically Distributed", or "IID".

⇒ IID implies zero mean.

⇒ IID implies that $x(t)$ is a white noise.

⇒ IID implies strict sense stationarity (SSS).

Gaussian Process

- If $x(t)$ is a stochastic process and the RVs at all times are jointly Gaussian, then $x(t)$ is called a "Gaussian Process".

Gauss-Markov Process

- The technical definition of a "Markov process" is outside the scope of this course.

- A Gauss-Markov process is one that is both Gaussian and Markovian.

- For a WSS Gauss-Markov process, the situation is quite elegant:

  ⇒ Exponential autocorrelation:

$$R_x(\tau) = \sigma^2 e^{-\beta|\tau|}, \quad \sigma, \beta \text{ constants.}$$

  ⇒ Gauss-Markov PSD:

$$S_x(\omega) = \frac{2\sigma^2\beta}{\omega^2 + \beta^2}$$

$$S_x(s) = \frac{2\sigma^2\beta}{-s^2 + \beta^2}$$

LTI Systems

- Suppose $x(t)$ is a stochastic process.
- Let H be an LTI system with impulse response $h(t)$ and transfer function $H(s)$.
- Let $y(t)$ be the system output (a stochastic process):

$$x(t) \rightarrow \boxed{H(s)} \rightarrow y(t).$$

- Suppose that H is BIBO stable.

$\Rightarrow$ If $x(t)$ is SSS, then so is $y(t)$.

$\Rightarrow$ If $x(t)$ is WSS, then so is $y(t)$.

$\Rightarrow$ If $x(t)$ is Gaussian, then so is $y(t)$.

- The situation is much like the deterministic case:

$$y(t) = x(t) * h(t)$$
$$Y(\omega) = X(\omega) H(\omega)$$
$$Y(s) = X(s) H(s).$$

– Henceforth, assume that the system input $x(t)$ is WSS

– First moment of output

$$E[y(t)] = E\left\{ \int_{-\infty}^{\infty} x(\theta) h(t-\theta)\, d\theta \right\}$$

$$= \int_{-\infty}^{\infty} E[x(\theta)]\, h(t-\theta)\, d\theta$$

$$= E[x(t)] \int_{-\infty}^{\infty} h(t-\theta)\, d\theta$$

$$= E[x(t)] \int_{-\infty}^{\infty} h(\theta)\, d\theta$$

$$= H(0)\, E[x(t)]$$

a number

a number

– Second moment of output:

$$E[y^2(t)] = E\left\{\int_{-\infty}^{\infty} x(t-\theta)h(\theta)d\theta \int_{-\infty}^{\infty} x(t-\lambda)h(\lambda)d\lambda\right\}$$

$$= E\left\{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} E\{x(t-\theta)x(t-\lambda)\}h(\theta)h(\lambda)\,d\theta\,d\lambda\right.$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} E\{x(t-\theta)\,x(t-\lambda)\}\,h(\theta)\,h(\lambda)\,d\theta\,d\lambda$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} R_x(t-\theta, t-\lambda)\,h(\theta)h(\lambda)\,d\theta\,d\lambda$$

(WSS)

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} R_x(\theta-\lambda)\,h(\theta)\,h(\lambda)\,d\theta\,d\lambda$$

$$= \int_{-\infty}^{\infty} h(\theta)\left[\int_{-\infty}^{\infty} R_x(\theta-\lambda)h(\lambda)\,d\lambda\right]d\theta$$

$$= \int_{-\infty}^{\infty} h(\theta)\left[R_x(\theta) * h(\theta)\right]d\theta$$

- Output Autocorrelation:

$$R_y(t_1,t_2) = E\{y(t_1)\,y(t_2)\}$$

$$= E\left\{\int_{-\infty}^{\infty} x(t_1-\theta)\,h(\theta)\,d\theta \int_{-\infty}^{\infty} x(t_2-\lambda)\,h(\lambda)\,d\lambda\right\}$$

$$= E\left\{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x(t_1-\theta)\,x(t_2-\lambda)\,h(\theta)\,h(\lambda)\,d\theta\,d\lambda\right\}$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} E\{x(t_1-\theta)\,x(t_2-\lambda)\}\,h(\theta)\,h(\lambda)\,d\theta\,d\lambda$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} R_x(t_1-\theta,\,t_2-\lambda)\,h(\theta)\,h(\lambda)\,d\theta\,d\lambda$$

WSS

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} R_x(t_2-t_1+\theta-\lambda)\,h(\theta)\,h(\lambda)\,d\theta\,d\lambda$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} R_x(\tau+\theta-\lambda)\,h(\theta)\,h(\lambda)\,d\theta\,d\lambda$$ ()

$$= \int_{-\infty}^{\infty} h(\theta)\int_{-\infty}^{\infty} R_x(\theta+\tau-\lambda)\,h(\lambda)\,d\lambda\,d\theta$$

$$= \int_{-\infty}^{\infty} h(\theta)\left[R_x(\theta+\tau) * h(\theta+\tau)\right] d\theta$$

$$R_y(\tau) = h(\tau) * \left[R_x(-\tau) * h(-\tau)\right].$$

$$= h(\tau) * R_x(\tau) * h(-\tau).$$

- Crosscorrelation of input and output:

$$R_{x,y}(t_1, t_2) = E\{x(t_1)\, y(t_2)\}$$

$$= E\left\{x(t_1) \int_{-\infty}^{\infty} x(t_2-\theta)\, h(\theta)\, d\theta\right\}$$

$$= \int_{-\infty}^{\infty} E\{x(t_1)x(t_2-\theta)\}\, h(\theta)\, d\theta$$

$$= \int_{-\infty}^{\infty} R_x(t_1, t_2-\theta)\, h(\theta)\, d\theta$$

WSS

$$= \int_{-\infty}^{\infty} R_x(t_2 - t_1 - \theta)\, h(\theta)\, d\theta$$

$$= \int_{-\infty}^{\infty} R_x(\tau-\theta)\, h(\theta)\, d\theta$$

$$R_{x,y}(\tau) = R_x(\tau) * h(\tau)$$

$$R_{Y,X}(t_1,t_2) = E\left[Y(t_1)\,X(t_2)\right]$$

$$= E\left[\int_{-\infty}^{\infty} X(t_1-\theta)\,h(\theta)\,d\theta\; X(t_2)\right]$$

$$= \int_{-\infty}^{\infty} E\left[X(t_1-\theta)\,X(t_2)\right] h(\theta)\,d\theta$$

$$= \int_{-\infty}^{\infty} R_X(t_1-\theta, t_2)\,h(\theta)\,d\theta$$

WSS

$$= \int_{-\infty}^{\infty} R_X(t_2-t_1+\theta)\,h(\theta)\,d\theta$$

$$= \int_{-\infty}^{\infty} R_X(\tau+\theta)\,h(\theta)\,d\theta \qquad (\lambda = -\theta)$$

$$= \int_{-\infty}^{\infty} R_X(\tau-\lambda)\,h(-\lambda)\,d\lambda$$

$$R_{Y,X}(\tau) = R_X(\tau) * h(-\tau)$$

- Output Power Spectrum:

→ Using eq. () on page SSP-52,

$$S_y(\omega) = \mathcal{F}[R_y(\tau)]$$

$$= \int_{-\infty}^{\infty} R_y(\tau) e^{-j\omega\tau}\, d\tau$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} R_x(\tau+\theta-\lambda) h(\theta) h(\lambda) e^{-j\omega\tau}\, d\theta\, d\lambda\, d\tau$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(\theta) h(\lambda) \left[\int_{-\infty}^{\infty} R_x(\tau+\theta-\lambda) e^{-j\omega\tau}\, d\tau\right] d\theta\, d\lambda$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(\theta) h(\lambda)\, S_x(\omega) e^{j\omega(\theta-\lambda)}\, d\theta\, d\lambda$$

$$= S_x(\omega) \int_{-\infty}^{\infty} h(\theta) e^{j\omega\theta}\, d\theta \int_{-\infty}^{\infty} h(\lambda) e^{-j\omega\lambda}\, d\lambda$$

$$= \underline{\underline{S_x(\omega) H(-\omega) H(\omega)}}$$

⇒ if $h(t)$ is real, then

$$S_y(\omega) = S_x(\omega)\, |H(\omega)|^2$$

⇒ Also,

$$S_y(s) = S_x(s) H(s) H(-s)$$

"Wiener-Khintchine relation"

- Input/output Cross Power

$$S_{x,y}(\omega) = \mathcal{F}[R_{x,y}(\tau)]$$

$$= \mathcal{F}[R_x(\tau) * h(\tau)]$$

$$= \underline{\underline{H(\omega)\, S_x(\omega)}}$$

$$S_{y,x}(\omega) = \mathcal{F}[R_{y,x}(\tau)]$$

$$= \mathcal{F}[R_x(\tau) * h(-\tau)]$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} R_x(\tau-\theta)\, h(-\theta)\, d\theta\, e^{-j\omega\tau}\, d\tau$$

$$= \int_{-\infty}^{\infty} h(-\theta) \int_{-\infty}^{\infty} R_x(\tau-\theta)\, e^{-j\omega\tau}\, d\tau\, d\theta \qquad \begin{aligned}\lambda &= \tau-\theta\\ \tau &= \lambda+\theta\end{aligned}$$

$$= \int_{-\infty}^{\infty} h(-\theta) \int_{-\infty}^{\infty} R_x(\lambda)\, e^{-j\omega(\lambda+\theta)}\, d\lambda\, d\theta$$

$$= \int_{-\infty}^{\infty} h(-\theta)\, e^{-j\omega\theta}\, d\theta \int_{-\infty}^{\infty} R_x(\lambda)\, e^{-j\omega\lambda}\, d\lambda \qquad \begin{aligned}\alpha &= -\theta\\ \theta &= -\alpha\end{aligned}$$

$$= \int_{-\infty}^{\infty} h(\alpha)\, e^{j\omega\alpha}\, d\alpha \int_{-\infty}^{\infty} R_x(\lambda)\, e^{-j\omega\lambda}\, d\lambda$$

$$= \underline{\underline{S_x(\omega)\, H(-\omega)}}$$

$\Rightarrow$ if $h(t)$ is real, $S_{yx}(\omega) = S_x(\omega)\, H^*(\omega)$.

- Second moment of output:

$$E[y^2(t)] = R_y(0)$$

$$= \mathcal{F}^{-1}\{S_y(\omega)\}\Big|_{\tau=0}$$

$$= \frac{1}{2\pi}\int_{-\infty}^{\infty} S_y(\omega)\, e^{j\omega\tau}\, d\omega \Big|_{\tau=0}$$

$$= \frac{1}{2\pi}\int_{-\infty}^{\infty} S_y(\omega)\, d\omega$$

$$= \frac{1}{2\pi}\int_{-\infty}^{\infty} S_x(\omega)\,|H(\omega)|^2\, d\omega$$

Spectral Factorization

- Let $x(t)$ be an arbitrary square-integrable signal with rational Laplace transform.

- define

$$x_+(t) = \begin{cases} x(t), & t \geq 0 \\ 0, & t < 0 \end{cases}$$

$$x_-(t) = \begin{cases} x(t), & t < 0 \\ 0, & t > 0 \end{cases}$$

- Clearly, $x(t) = x_+(t) + x_-(t)$.

- So, $X(s) = X_+(s) + X_-(s)$.

→ Since $x_+(t)$ is right-sided, all the poles of $X_+(s)$ are in the left half-plane.

→ Since $x_-(t)$ is ~~right~~ left-sided, all the poles of $X_-(s)$ are in the right half-plane.

- By performing "backwards partial fractions" we can write $X(s)$ as a product

$$X(s) = \mathcal{X}_+(s)\,\mathcal{X}_-(s)$$

Left half plane poles only

Right half-plane poles only.

⇒ This is called "spectral factorization"

EX $\quad x(t) = e^{-\alpha|t|}$

$$x_-(t) = \begin{cases} e^{\alpha t}, & t < 0 \\ 0, & t \geq 0 \end{cases}$$

$$x_+(t) = \begin{cases} e^{-\alpha t}, & t \geq 0 \\ 0, & t < 0 \end{cases}$$

$$X(s) = \frac{2\alpha}{\alpha^2 - s^2}$$

$$= \underbrace{\frac{1}{s+\alpha}}_{X_+(s)} + \underbrace{\frac{-1}{s-\alpha}}_{X_-(s)}$$

$$= \underbrace{\frac{\sqrt{2\alpha}}{s+\alpha}}_{\mathcal{X}_+(s)} \cdot \underbrace{\frac{\sqrt{2\alpha}}{s-\alpha}}_{\mathcal{X}_-(s)}$$

EX Let H be a first-order low-pass filter with frequency response

$$H(\omega) = \frac{1}{1+j\omega T}$$

- let the input be bandlimited white noise with PSD

$$S_X(\omega) = \begin{cases} A, & |\omega| \le \omega_c \\ 0, & |\omega| > \omega_c \end{cases}$$

- Then $S_Y(\omega) = S_X(\omega)\,|H(\omega)|^2$

$$= \begin{cases} \dfrac{A}{1+T^2\omega^2}, & |\omega| \le \omega_c \\ 0, & |\omega| > \omega_c \end{cases}$$

- The second moment of the output is

$$E[y^2(t)] = \frac{1}{2\pi}\int_{-\omega_c}^{\omega_c} \frac{A}{1+T^2\omega^2}\, d\omega$$

$$= \frac{A}{\pi T}\arctan(\omega_c T),$$

- now suppose the input $x(t)$ is white noise with PSD $S_x(\omega) = A$.

- then $S_y(\omega) = S_x(\omega)\,|H(\omega)|^2 = \dfrac{A}{1+T^2\omega^2}$

- The second moment of the output is then

$$E[y^2(t)] = \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{A}{1+T^2\omega^2}\,d\omega$$

$$= \lim_{B\to\infty} \frac{A}{\pi T}\arctan(B)$$

$$= \frac{A}{\pi T}$$

<u>NOTE</u> : This shows that, if the noise bandwidth is large compared to the passband of $H(\omega)$, there is little error in modelling the band-limited noise as white noise.

- In fact, this is the only real use for white noise, since it can't really exist.

Why?

→ By definition, if $x(t)$ is a white noise, then

$$R_x(\tau) = \delta(\tau).$$

→ So $R_x(0) = \mathrm{Var}[x(t)] \to \infty$.

⇒ So a <u>true</u> white noise would have to have infinite variance for all $t \in \mathbb{R}$.

Noise Shaping

- Often, we need to change the shape of the power spectrum of a statistical signal.

- Ex: Turn white noise into colored noise (or vice-versa).

→ Suppose $x(t)$ is a white noise with unity power spectrum:

$$S_x(\omega) = 1.$$

→ Suppose that the desired power spectrum is $S_y(\omega)$.

→ To design the required filter, observe that

$$S_y(\omega) = S_x(\omega)\,|H(\omega)|^2 = |H(\omega)|^2.$$

→ So $S_y(s) = H(s)\,H(-s)$,

which can be solved for $H(s)$. Then

$$x(t) \longrightarrow \boxed{H} \longrightarrow y(t)$$

where $x(t)$ is white and $y(t)$ will have the desired correlation structure.

EX: The desired power spectrum for $y(t)$ is

$$S_y(\omega) = \frac{\omega^2+1}{\omega^4+64}$$

$$S_y(s) = \frac{-s^2+1}{s^4+64}$$

$$= \underbrace{\frac{s+1}{s^2+4s+8}}_{H(s)} \cdot \underbrace{\frac{1-s}{s^2-4s+8}}_{H(-s)}.$$

$\Rightarrow$ So the required transfer function is

$$H(s) = \frac{s+1}{s^2+4s+8}$$

Discrete Time

- Results analagous to those we have obtained for continuous time can also be formulated for the discrete-time case by using similar calculations.

- Let H be a BIBO stable LTI system with impulse response $h[n]$ and frequency response $H(e^{j\omega})$.

- Let $x[n]$ be WSS.

$$x[n] \longrightarrow \boxed{H} \longrightarrow y[n].$$

- First moment of output:

$$E\left[y[n]\right] = \underbrace{E\left[x[n]\right]}_{\text{a number}} \underbrace{H(0)}_{\text{a number}}$$

- Second moment of output:

$$E[y^2[n]] = \sum_{i \in \mathbb{Z}} \sum_{j \in \mathbb{Z}} R_x(i-j)\, h[i]\, h[j]$$

- Output Autocorrelation:

$$R_y(m) = \sum_{i \in \mathbb{Z}} \sum_{j \in \mathbb{Z}} h[i]\, h[j]\, R_x(m+j-i)$$

- Input/output Crosscorrelation:

$$R_{xy}(m) = R_{yx}(-m) = h[m] * R_x(m)$$

- Output power spectrum:

$$S_y(z) = H\left(\tfrac{1}{z}\right) H(z)\, S_x(z)$$

"Discrete time Wiener-Kintchine relation".

- Input/output cross power:

$$S_{xy}(z) = H(z)\, S_x(z),$$

$$S_{yx}(z) = H\left(\tfrac{1}{z}\right) S_x(z).$$

Wiener Filter

- Let $x(t)$ be a WSS statistical signal.
- Let $n(t)$ be WSS noise.
- Suppose that $R_x(\tau)$, $R_n(\tau)$, and $R_{x,n}(\tau)$ are known.
- Design the "best" LTI filter to estimate $x(t+\alpha)$, given $z(t) = x(t) + n(t)$ ("additive noise").

$$z(t) = x(t) + n(t) \longrightarrow \boxed{G} \longrightarrow y(t) = \hat{x}(t+\alpha).$$

- "Best" means minimum mean squared error:

$$E\left\{ \left[x(t+\alpha) - \hat{x}(t+\alpha) \right]^2 \right\} \text{ minimized.}$$

- The solution to this problem is called the "Wiener filter.

- non-causal solution:

$$G(s) = \frac{S_{z,x}(s)e^{\alpha s}}{S_z(s)}$$

- Causal Solution:

$$G(s) = \frac{1}{S_z^+(s)} \mathcal{L}\left\{\mathcal{L}^{-1}\left[\frac{S_{z,x}(s)e^{\alpha s}}{S_z^-(s)}\right]u(t)\right\}$$

words:

1. Compute $\dfrac{S_{z,x}(s)e^{\alpha s}}{S_z^-(s)}$
2. Take the inverse Laplace transform and multiply the result by $u(t)$.
3. Take the Laplace transform of the result.
4. Divide by $S_z^+(s)$.