King Fahd University of Petroleum & Minerals
Computer Engineering Dept

CSE 642 – Computer Systems Performance

Term 091

Dr. Ashraf S. Hasan Mahmoud

Rm 22-148-3

Ext. 1724

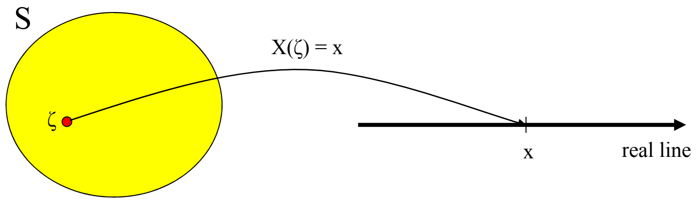
Email: ashraf@kfupm.edu.sa

Primer on Probability Theory

What is a Random Variable?

- **Random Experiment**
- **Sample Space**

- **Def: A random variable X is a function that assigns a number of X(ζ) to each outcome ζ in the sample space of S of the random experiment**



Set Functions

- Define **Ω** as the set of all possible outcomes
- Define **A** as set of events
- Define A as an event – subset of the set of all experiments outcomes
- Set operations:
  - Complementation A^c: is the event that event A does not occur
  - Intersection A ins B : is the event that event A and B occur
  - Union A un B: is the event that event A or B occur
  - Inclusion A in B: An event A occurring implying events B occurs

Set Functions

- Note:
  - Set of events **A** is closed under set operations
  - Φ – empty set
  - $A \cap B = \Phi$ ➔ are mutually exclusive or disjoint

Axioms of Probability

- Let P(A) denote probability of event A:
  1. For any event A belongs **A**, P(A) ≥ 0;
  2. For set of all possible outcomes **Ω**, P(**Ω**) = 1;
  3. If A and B are disjoint events, P(A un B) = P(A) + P(B)
  4. For countably infinite sets, A_1, A_2, … such that A_i ins $A_j = \Phi$ for $i \neq j$

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$$

Additional Properties

- For any event, $P(A) \leq 1$
- $P(A^C) = 1 - P(A)$
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
- $P(A) \leq P(B)$ for $A \subseteq B$

Conditional Probability

- Conditional probability is defined as

$$P(A/B) = \frac{P(A \cap B)}{P(B)}$$

- P(A/B) probability of event A conditioned on the occurrence of event B
- Note:
  - A and B are *independent* if $P(A \cap B) = P(A)P(B) \Rightarrow P(A/B) = P(A)$
  - Independent IS NOT EQUAL TO mutually exclusive

The Law of Total Probability

- A set of events A_i, i = 1, 2, …, n partitions the set of experimental outcomes if

$$\bigcup_{i=1}^{n} A_i = \Omega$$

and

$$A_i \cap A_j = \Phi$$

Then we can write any event B in terms of A_i, i = 1, 2, …, n as

$$B = \bigcup_{i=1}^{n} A_i \cap B$$

Furthermore,

$$P(B) = \sum_{i=1}^{n} P(A_i \cap B)$$

Bayes' Rule

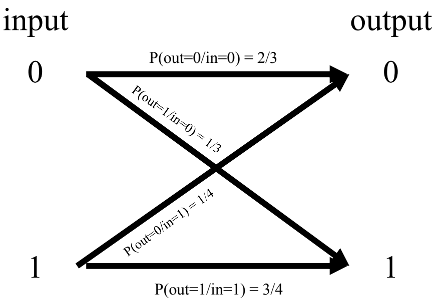
- Using the total law of probability and applying it to the definition of the conditional probability, yields

$$P(A_i / B) = \frac{P(A_i \cap B)}{\sum_{i=1}^{n} P(A_i \cap B)}$$

$$= \frac{P(A_i)P(B/A_i)}{\sum_{i=1}^{n} P(A_i)P(B/A_i)}$$

Example: Binary Symmetric Channel

- Given the binary symmetric channel depicted in figure, find P(input = j / output = i); i,j = 0,1. Given that P(input = 0) = 0.4, P(input = 1) = 0.6.



Solution:

- P(input=j/outpu=i) = P(input=j, output=i)/P(output=i)
- But the term P(input=j, output=i) is equal to P(output=i, input=j) = Prob(output=i/input=j)P(input=j)
- Therefore, the following joint probabilities are computed as:
  - P(input=0,output=0) = (2/3)(0.4) = 0.2667
  - P(input=0,output=1) = (1/3)(0.4) = 0.1333
  - P(input=1,output=0) = (1/4)(0.6) = 0.1500
  - P(input=1,output=1) = (3/4)(0.6) = 0.4500
- Also

P(out=0) = P(in=0)P(out=0/in=0) + P(in=1)P(out=0/in=1)
= (0.4)(2/3) + (0.6)(1/4) = 0.4167
P(out=1) = P(in=0)P(out=1/in=0) + P(in=1)P(out=1/in=1)
= (0.4)(1/3) + (0.6)(3/4) = 0.5833

- Hence,

P(in=0/out=0) = P(in=0,out=0)/P(out=0)
= (0.2667)/(0.4167) = 0.6400
P(in=0/out=1) = P(in=0,out=1)/P(out=1)
= (0.1333)/(0.5833) = 0.2286
Etc.

The Cumulative Distribution Function

- **The cumulative distribution function (cdf) of a random variable X is defined as the probability of the event {X ≤ x}:**

$$F_X(x) = \text{Prob}\{X \le x\} \quad \text{for } -\infty < x < \infty$$

i.e. it is equal to the probability the variable X takes on a value in the set (- ∞,x]

- **A convenient way to specify the probability of all semi-infinite intervals**

Properties of the CDF

- $0 \leq F_X(x) \leq 1$
- $\lim_{x \to \infty} F_X(x) = 1$
- $\lim_{x \to -\infty} F_X(x) = 0$
- $F_X(x)$ is a nondecreasing function → if $a < b$ → $F_X(a) \leq F_X(b)$
- $F_X(x)$ is continuous from the right → for $h > 0$,
  $$F_X(b) = \lim_{h \to 0} F_X(b+h) = F_X(b^+)$$
- $P[a < X \leq b] = F_X(b) - F_X(a)$
- $P[X = b] = F_X(b) - F_X(b^-)$

Example 1: Exponential Random Variable

- **Problem: The transmission time X of a message in a communication system obey the exponential probability law with parameter λ, that is**

  $$\text{Prob}\,[X > x] = e^{-\lambda x} \quad x > 0$$

Find the CDF of X. Find Prob $[T < X \leq 2T]$ where $T = 1/\lambda$

Example 1: Exponential Random Variable – cont'd

- **Answer:**

The CDF of X is

$$F_X(x) = \text{Prob}\{X \leq x\} = 1 - \text{Prob}\{X > x\}$$
$$= 1 - e^{-\lambda x} \quad x \geq 0$$
$$= 0 \quad x < 0$$

$$\text{Prob}\{T < X \leq 2T\} = F_X(2T) - F_X(T)$$
$$= 1-e^{-2} - (1-e^{-1})$$
$$= 0.233$$

Example 2: Use of Bayes Rule

- **Problem: The waiting time W of a customer in a queueing system is zero if he finds the system idle, and an exponentially distributed random length of time if he finds the system busy. The probabilities that he finds the system idle or busy are p and 1-p, respectively. Find the CDF of W**

Example 2: cont'd

- **Answer:**

The CDF of W is found as follows:

$F_X(x) = \text{Prob}\{W \leq x\}$

$= \text{Prob}\{W \leq x/\text{idle}\}p + \text{Prob}\{W \leq x/\text{busy}\}(1-p)$

Note $\text{Prob}\{W \leq x/\text{idle}\} = 1$ for any $x > 0$

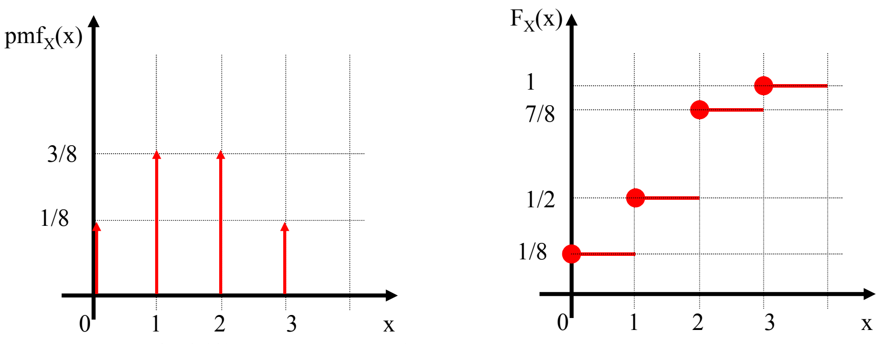
➔

$F_X(x) = 0 \qquad x < 0$

$= p+(1-p)(1- e^{-\lambda x}) \qquad x \geq 0$

Types of Random Variables

- **(1) Discrete Random Variables**
  - **CDF is right continuous, staircase function of x, with jumps at countable set x0, x1, x2, …**



Pmf: probability mass function

Types of Random Variables

- **(2) Continuous Random Variables**
    - **CDF is continuous for all values of x ➔ Prob { X = x} = 0 (recall the CDF properties)**
    - **Can be written as the integral of some non negative function**

$$F_X(x) = \int_{-\infty}^{\infty} f(t)dt$$

Or

$$f(t) = \frac{dF_X(x)}{dx}$$

f(t) is referred to as the probability density function or PDF

Types of Random Variables

- **(3) Random Variables of Mixed Types**

$$\mathbf{F_X(x) = p\ F_1(x) + (1\text{-}p)\ F_2(x)}$$

Probability Density Function

- **The PDF of X, if it exists, is define as the derivative of CDF $F_X(x)$:**

$$f_x(x) = \frac{dF_X(x)}{dx}$$

Properties of the PDF

- **$f_x(x) \geq 0$**
- $$P\{a \leq x \leq b\} = \int_a^b f_x(x)dx$$
- $$F_X(x) = \int_{-\infty}^{x} f_x(t)dt$$
- $$1 = \int_{-\infty}^{\infty} f_x(t)dt$$

A valid pdf can be formed from any nonnegative, piecewise continuous function g(x) that has a finite integral:

$$\int_{-\infty}^{\infty} g(x)dx = c < \infty$$

By letting $f_X(x) = g(x)/c$, we obtain a function that satisfies the normalization condition.
This is the scheme we use to generate pdfs from simulation results!

Conditional PDFs and CDFs

- If some event A concerning X is given, then conditional CDF of X given A is defined by

$$F_X(x/A) = \frac{P([X \le x] \cap A)}{P(A)} \qquad \text{if } P(A) > 0$$

The conditional pdf of X given A is then defined by

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

Expectation of a Random Variable

- Expectation of the random variable X can be computed by

$$E[X] = \sum_{\forall i} x_i P[X = x_i]$$

for discrete variables, or

$$E[X] = \int_{-\infty}^{\infty} t f_x(t) dt$$

for continuous variables.

nth Expectation of a Random Variable

- **The nth expectation of the random variable X can be computed by**

$$E[X^n] = \sum_{\forall i} x^n{}_i P[X = x_i]$$

for discrete variables, or

$$E[X^n] = \int_{-\infty}^{\infty} t^n f_x(t)\,dt$$

for continuous variables.

Central Moments a Random Variable

- **The nth central moment of a random variable is given by**

$$E\left[(X - E[X])^n\right]$$

Therefore, the variance of a r.v is given by

$$\begin{aligned}\sigma_X^2 \equiv Var[X] &= E\left[(X - E[X])^2\right] \\ &= E\left[X^2\right] - (E[X])^2\end{aligned}$$

The standard deviation is computed as

$$SD(X) = \sqrt{Var(X)} = \sigma_X$$

Expectation of a Function of the Random Variable

- Let g(x) be a function of the random variable x, the expectation of g(x) is given by

$$E[g(x)] = \sum_{\forall i} g(x_i) P[X = x_i]$$

for discrete variables, or

$$E[g(x)] = \int_{-\infty}^{\infty} g(t) f_x(t) dt$$

for continuous variables.

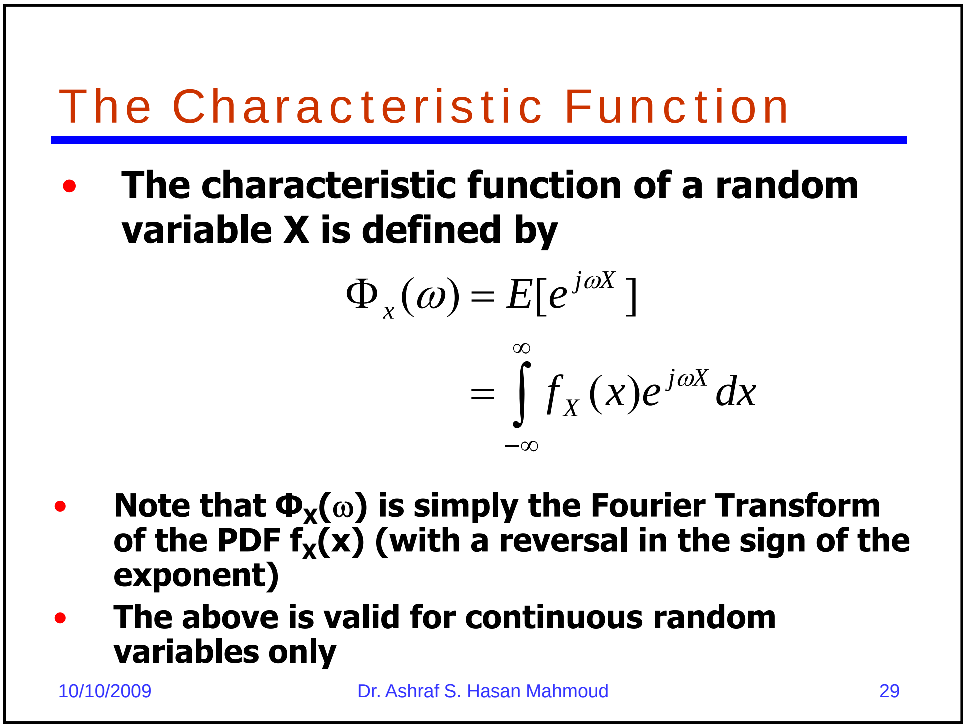
Example 3:

- Problem: For X nonnegative r.v. show that

for continuous X: $E[X] = \int_{0}^{\infty} \left(1 - F_x(t)\right) dt$, and

for discrete X: $E[X] = \sum_{k=0}^{\infty} P(X > k)$

Prove the above formulas

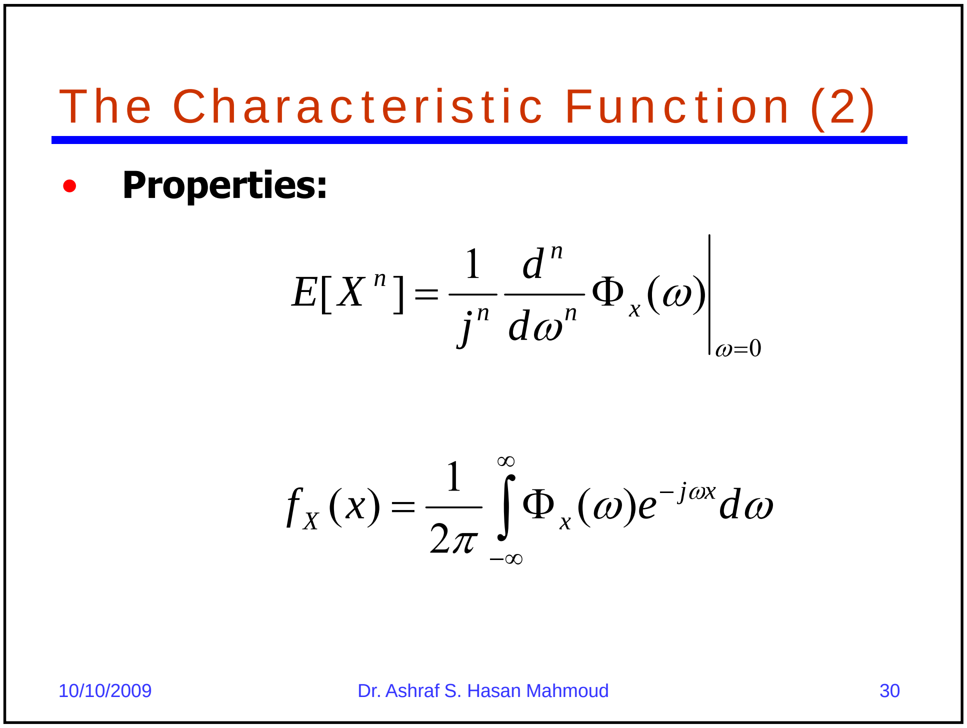


The Characteristic Function

- **The characteristic function of a random variable X is defined by**

$$\Phi_x(\omega) = E[e^{j\omega X}]$$
$$= \int_{-\infty}^{\infty} f_X(x) e^{j\omega X} dx$$

- **Note that $\Phi_X(\omega)$ is simply the Fourier Transform of the PDF $f_X(x)$ (with a reversal in the sign of the exponent)**
- **The above is valid for continuous random variables only**



The Characteristic Function (2)

- **Properties:**

$$E[X^n] = \frac{1}{j^n} \frac{d^n}{d\omega^n} \Phi_x(\omega) \Big|_{\omega=0}$$

$$f_X(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \Phi_x(\omega) e^{-j\omega x} d\omega$$

The Characteristic Function (3)

- **For discrete random variables,**

$$\begin{aligned}\Phi_x(\omega) &= E[e^{j\omega X}] \\ &= \sum_{\forall k} P(X = x_k) e^{j\omega x_k}\end{aligned}$$

- **For integer valued random variables,**

$$\Phi_x(\omega) = \sum_{k=-\infty}^{\infty} P(X = k) e^{j\omega k}$$

The Characteristic Function (4)

- **Properties**

$$P(X = k) = \frac{1}{2\pi} \int_0^{2\pi} \Phi_x(\omega) e^{-j\omega k} d\omega$$

for k=0, ±1, ±2, …

Probability Generating Function

- **A matter of convenience – compact representation**
- **The same as the z-transform**
- **If N is a non-negative integer-valued random variable, the probability generating function is defined as**

$$N(z) = E[z^N]$$

$$= \sum_{i=0}^{\infty} p(N=i) z^i$$

$$= P(N=0) + P(N=1)z + P(N=2)z^2 + ...$$

Probability Generating Function (2)

- **Properties:**
- **1** $$N(z)\big|_{z=1} = 1$$
- **2** $$P(N=i) = \frac{1}{i!}\frac{d^i}{dz^i} N(z)\bigg|_{z=0}$$
- **3** $$E[N] = \frac{dN(z)}{dz}\bigg|_{z=1}$$
- **4** $$Var[N] = N''(1) + N'(1) - \left[N'(1)\right]^2$$

Probability Generating Function (3)

- **For non-negative continuous random variables, let us define the Laplace transform of the PDF**

$$X^*(s) = \int_0^\infty f_X(x) e^{-sx} dx$$

$$= E[e^{-sx}]$$

Properties:

$$X(s)\big|_{s=0} = 1$$

$$X(s) = \Phi_X(js)$$

$$E[X^n] = (-1)^n \frac{d^n}{ds^n} X^*(s)\bigg|_{s=0}$$

Some Important Random Variables – Discrete Random Variables

- **Bernoulli**
- **Binomial**
- **Geometric**
- **Poisson**

Identities to remember:

$$\sum_{n=1}^{M} n = \frac{1}{2} M(M+1)$$

$$\sum_{n=1}^{M} n^2 = M(M+1)(2M+1)/6$$

$$\sum_{n=0}^{\infty} r^n = \frac{1}{1-r}; |r| < 1$$

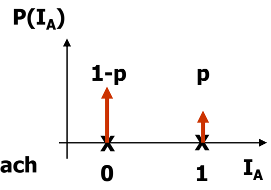
$$\sum_{n=0}^{\infty} n r^{n-1} = \frac{1}{(1-r)^2}; |r| < 1$$

$$\sum_{n=0}^{M} r^n = \frac{1-r^{M+1}}{1-r}; |r| < 1, M = 1,2,\ldots$$

$$\sum_{n=0}^{M} \binom{M}{n} r^n = (1+r)^M; |r| < 1$$

$$\sum_{n=0}^{M} n r^{n-1} = \frac{1+(Mr-M-1)r^M}{(1-r)^2}; |r| < 1$$

Bernoulli Random Variable



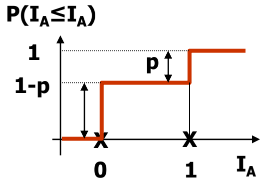
- **Let A be an event related to the outcomes of some random experiment. The indicator function for A is defined as**

$$I_A(\zeta) = 0 \quad \text{if } \zeta \text{ not in A}$$
$$= 1 \quad \text{if } \zeta \text{ is in A}$$

- **I_A is random variable since it assigns a number to each outcome in S**
- **It is discrete r.v. that takes on values from the set {0,1}**
- **PMF is given by**

$$p_I(0) = 1-p, \; p_I(1) = p$$

where P{A} = p



- **Describes the outcome of a Bernoulli trial**
- **E[X] = p, VAR[X] = p(1-p)**
- **$G_X(z) = (1-p+pz)$**

Binomial Random Variable

- **Suppose a random experiment is repeated n independent times; let X be the number of times a certain event A occurs in these n trials**

$$X = I1 + I2 + \ldots + In$$

i.e. X is the sum of Bernoulli trials (X's range = {0, 1, 2, …, n})

- **X has the following pmf**

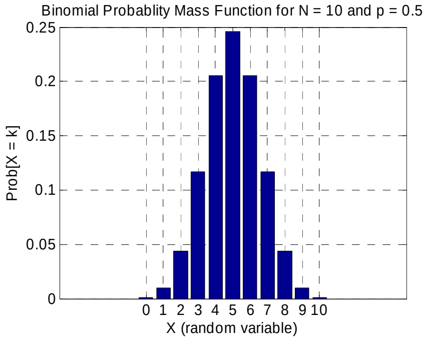
$$\Pr[X = k] = \binom{n}{k} p^k (1-p)^{n-k}$$

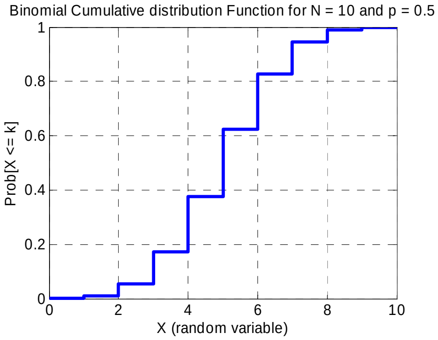
for k=0, 1, 2, …, n

- **E[X] = np, Var[X] = np(1-p)**
- **$G_X(z) = (1-p + pz)^n$**

Binomial Random Variable – cont'd

- **Example**


Binomial Probablity Mass Function for N = 10 and p = 0.5


Binomial Cumulative distribution Function for N = 10 and p = 0.5

Geometric Random Variable

- **Suppose a random experiment is repeated - We count the number of M of independent Bernoulli trials UNTIL the first occurrence of a success**
- **M is called geometric random variable**
  - **Range of M = 1, 2, 3, …**
- **X has the following pmf**

$$\Pr[X = k] = (1-p)^{k-1} p$$

for k=1, 2, 3, …

- **E[X] = 1/p,   Var[X] = (1-p)/p^2**
- **G$_X$(z) = pz/(1-(1-p)z))**

Geometric Random Variable - 2

- **Suppose a random experiment is repeated - We count the number of M of independent Bernoulli trials BEFORE the first occurrence of a success**
- **M is called geometric random variable**
  - **Range of M = 0, 1, 2, 3, …**
- **X has the following pmf**

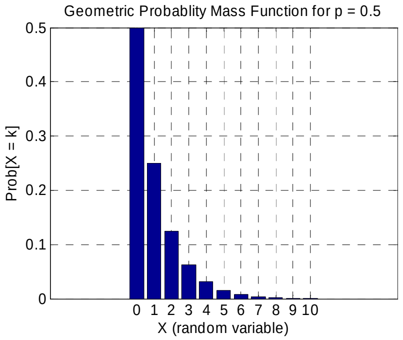
$$\Pr[X = k] = (1-p)^k p$$

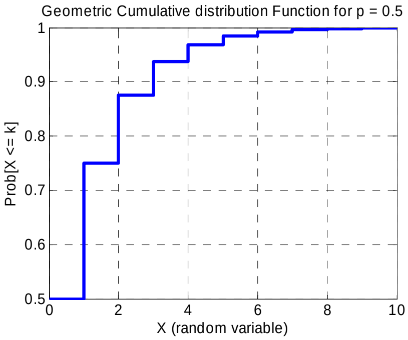
for k=0,1, 2, 3, …

- **$E[X] = (1-p)/p$, $Var[X] = (1-p)/p^2$**
- **$G_X(z) = pz/(1-(1-p)z))$**

Geometric Random Variable – cont'd

- **Example: p = 0.5; X is number of failures BEFORE a success (2^{nd} type)**
- **Note Matlab's version of geometric distribution is the 2^{nd} type**





Poisson Random Variable

- **In many applications we are interested in counting the number of occurrences of an event in a certain time period**
- **The pmf is given by**

$$\Pr[X = k] = \frac{\alpha^k}{k!} e^{-\alpha}$$

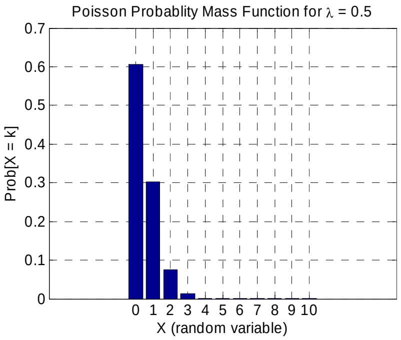
For k=0, 1, 2, … ; α is the average number of event occurrences in the specified interval

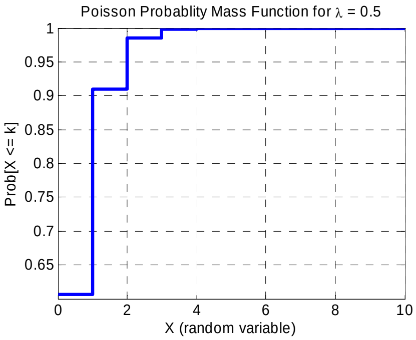
- **E[X] = α, Var[X] = α**
- **$G_X(z) = e^{\alpha(z-1)}$**

- **Remember: time between events is exponentially distributed!**
- **Poisson is the limiting case for Binomial as n→∞, p →0, such that np = α**

Poisson Random Variable – cont'd

- **Example:**





Example 4:

- **Calculate the probability generating function for the Poisson r.v.?**

- **Solution: Applying the definition**

$$N(Z) = \sum_{k=0}^{\infty} z^k \frac{\alpha^k}{k!} e^{-\alpha}$$

$$= e^{-\alpha} \sum_{k=0}^{\infty} \frac{(z\alpha)^k}{k!}$$

$$= e^{-\alpha} \times e^{\alpha z}$$

$$= e^{\alpha(z-1)}$$

Poisson Random Variable - 2

- **If the average rate of occurrence per time unit is λ, then the average number of occurrences in t seconds is equal to λt**

- **The probability of k occurrences in t seconds is given by**

$$P_k(t) = \frac{(\lambda t)^k}{k!} e^{-\lambda t} \qquad k = 0,1,2,\ldots$$

Compared to previous slide – we have replaced α by λt

Matlab Code to Plot Distributions

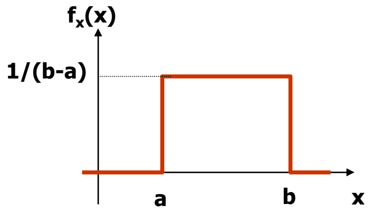
```
% plot distributions
% see "help stats"
clear all
FontSize = 14;
LineWidth = 3;
% Binomial
N = 10; X = [0:1:N]; P = 0.5;
ybp = binopdf(X, N, P); % get PMF
ybc = binocdf(X, N, P); % get CDF
figure(1); set(gca,'FontSize', FontSize);
bar(X, ybp);
title(['Binomial Probablity Mass Function for
    N = ' ...
    num2str(N) ' and p = ' num2str(P)]);
xlabel('X (random variable)');
ylabel('Prob[X = k]'); grid
figure(2); set(gca,'FontSize', FontSize);
stairs(X, ybc,'LineWidth', LineWidth);
title(['Binomial Cumulative distribution
    Function for N = ' ...
    num2str(N) ' and p = ' num2str(P)]);
xlabel('X (random variable)');
ylabel('Prob[X <= k]'); grid
% Geometric
P = 0.5; X = [0:10];
ygp = geopdf(X, P); % get pdf
ygc = geocdf(X, P); % get cdf
figure(3); set(gca,'FontSize', FontSize);
bar(X, ygp);
title(['Geometric Probablity Mass Function for
    p = ' num2str(P)]);
xlabel('X (random variable)');
ylabel('Prob[X = k]'); grid
figure(4); set(gca,'FontSize', FontSize);
stairs(X, ygc,'LineWidth', LineWidth);
title(['Geometric Cumulative distribution
    Function for p = ' num2str(P)]);
xlabel('X (random variable)');
ylabel('Prob[X <= k]'); grid
% Poisson
Lambda = 0.5; X     = [0:10];
ypp   = poisspdf(X, Lambda);
ypc   = poisscdf(X, Lambda);
figure(5); set(gca,'FontSize', FontSize);
bar(X, ypp);
title(['Poisson Probablity Mass Function for
    \lambda = ' num2str(Lambda)]);
xlabel('X (random variable)');
ylabel('Prob[X = k]'); grid
figure(6); set(gca,'FontSize', FontSize);
stairs(X, ypc,'LineWidth', LineWidth);
title(['Poisson Probablity Mass Function for
    \lambda = ' num2str(Lambda)]);
xlabel('X (random variable)');
ylabel('Prob[X <= k]'); grid
```

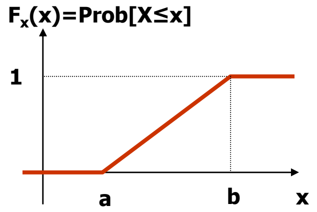
Some Important Random Variables – Continuous Random Variables

- **Uniform**
- **Exponential**
- **Gaussian (Normal)**
- **Rayleigh**
- **Gamma**
- **Pareto**

Uniform Random Variables

- **Realizations of the r.v. can take values from the interval [a, b]**
- **PDF $f_X(x) = 1/(b-a)$ $a \leq x \leq b$**
- **$E[X] = (a+b)/2$, $Var[X] = (b-a)^2/12$**
- **$\Phi_X(\omega) = [e^{j\omega b} - e^{j\omega a}]/(j\omega(b-a))$**





Example 5: Analog-to-Digital Conversion

Problem: compute the SNR for a uniform quantizer using 2^N representation values?

Exponential Random Variables

- **The exponential r.v. X with parameter λ has pdf**

$$f_X(x) = \begin{cases} 0 & x < 0 \\ \lambda e^{-\lambda x} & x \ge 0 \end{cases}$$

- **And CDF given by**

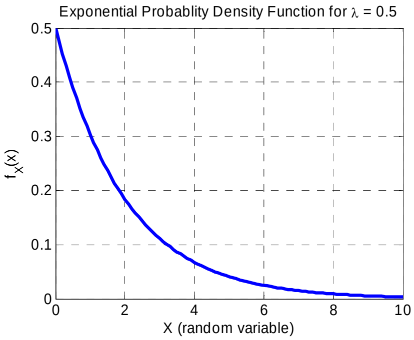
$$F_X(x) = \begin{cases} 0 & x < 0 \\ 1 - e^{-\lambda x} & x \ge 0 \end{cases}$$

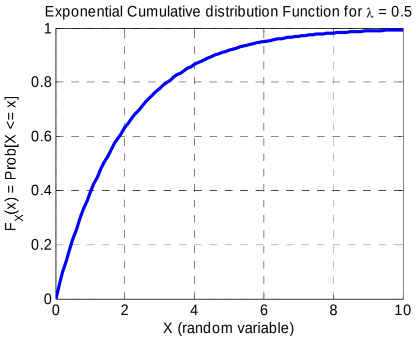
- **Range of X: [0, ∞)**
- **E[X] = 1/λ,   Var[X] = $1/\lambda^2$**
- **$\Phi_X(\omega) = \lambda/(\lambda - j\omega)$**

This means:
$\text{Prob}[X \le x] = 1 - e^{-\lambda x}$, or
$\text{Prob}[X > x] = e^{-\lambda x}$

Exponential Random Variables – cont'd

- **Example:**
  - **Note the mean is 1/λ = 2**





Exponential Random Variables – Memoryless Property

- **The exponential r.v. is the only continuous r.v. with the memoryless property!!**
- **Memoryless Property:**
  $P[X>t+h/ X>t] = P[X>h]$

i.e. the probability of having to wait h additional seconds given that one has already been waiting t second IS EXACTLY equal to the probability of waiting h seconds when one first begins to wait

<u>Proof</u>:

$$P[X>t+h/ X>t] = \frac{P[(X > t+h) \cap (X > t)]}{P[(X > t)]}$$

$$= \frac{P[(X > t+h)}{P[X > t]} = \frac{e^{-\lambda(t+h)}}{e^{-\lambda t}}$$

$$= e^{-\lambda h}$$

$$= P[X > h]$$

Gaussian (Normal) Random Variable

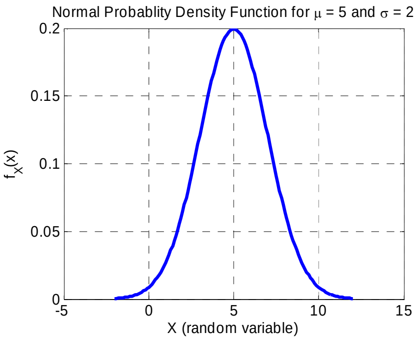
- **Rises in situations where a random variable X is the sum of a large number of "small" random variables – central limit theorem**
- **PDF** $$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/(2\sigma^2)}$$

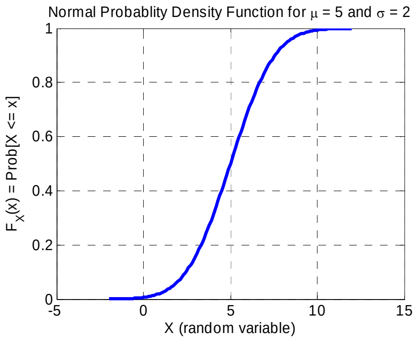
For $-\infty<x<\infty$; μ and $\sigma > 0$ are real numbers

- **$E[X] = \mu$, $Var[X] = \sigma^2$**
- $$\Phi_X(\omega) = e^{j\mu\omega-\sigma^2\omega^2/2}$$
- **Under wide range of conditions X can be used to approximate the sum of a large number of independent random variables**

Gaussian (Normal) Random Variable – cont’d

- **Example:**


Normal Probablity Density Function for μ = 5 and σ = 2


Normal Probablity Density Function for μ = 5 and σ = 2

Gaussian (Normal) Random Variable - 2

- **CDF given by**

$$F_X(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi\sigma}} e^{-(t-m)^2/(2\sigma^2)} dt$$

$$= 0.5 + 0.5 erf\left(\frac{x-m}{\sigma\sqrt{2}}\right)$$

where

$$erf(x) = \int_{0}^{x} e^{-t^2/2} dt$$

Note – the CDF can also be written in terms of the Q-function, where

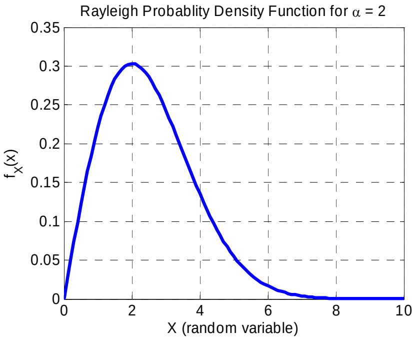
$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_{x}^{\infty} e^{-t^2/2} dt$$

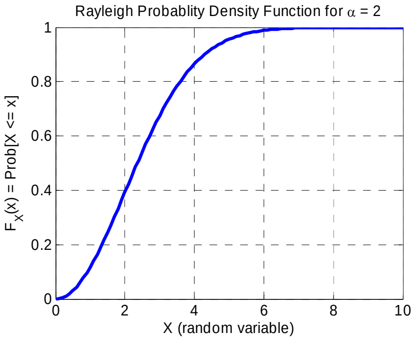
Rayleigh Random Variable

- **Arises in modeling of mobile channels**
- **Range: [0, ∞)**
- **PDF:** $$f_X(x) = \frac{x}{\alpha^2} e^{-x^2/(2\alpha^2)}$$
- **For $x \geq 0$, $\alpha > 0$**
- **$E[X] = \alpha\sqrt{(\pi/2)}$, $Var[X] = (2-\pi/2)\alpha^2$**

Rayleigh Random Variable – cont'd

- **Example:**
  - **Note that for Alpha = 2, the mean is $2\sqrt{(\pi/2)}$**





Matlab Code to Plot Distributions

```
% plot distributions
% see "help stats"
clear all
FontSize = 14;
LineWidth = 3;
% exponential
X = [0:.1:10]; Lambda = 0.5;
yep = exppdf(X, 1/Lambda); % get PDF
yec = expcdf(X, 1/Lambda); % get CDF
figure(1); set(gca,'FontSize', FontSize);
plot(X, yep, 'LineWidth', LineWidth);
title(['Exponential Probablity Density Function for \lambda = ' ...
     num2str(Lambda)]);
xlabel('X (random variable)');
ylabel('f_X(x)'); grid
figure(2); set(gca,'FontSize', FontSize);
plot(X, yec,'LineWidth', LineWidth);
title(['Exponential Cumulative Distribution Function for \lambda = ' ...
     num2str(Lambda)]);
xlabel('X (random variable)');
ylabel('F_X(x) = Prob[X <= x]'); grid
% normal
X = [-2:.1:12]; Mu = 5; Sigma = 2;
ynp = normpdf(X, Mu, Sigma); % get PDF
ync = normcdf(X, Mu, Sigma); % get CDF
figure(3); set(gca,'FontSize', FontSize);
plot(X, ynp, 'LineWidth', LineWidth);
title(['Normal Probablity Density Function for \mu = ' ...
     num2str(Mu) ' and \sigma = ' num2str(Sigma)]);
xlabel('X (random variable)');
ylabel('f_X(x)'); grid
figure(4); set(gca,'FontSize', FontSize);
plot(X, ync,'LineWidth', LineWidth);
title(['Normal Probablity Density Function for \mu = ' ...
     num2str(Mu) ' and \sigma = ' num2str(Sigma)]);
xlabel('X (random variable)');
ylabel('F_X(x) = Prob[X <= x]'); grid
% Rayleigh
X = [0:.1:10]; Alpha = 2;
yrp = raylpdf(X, Alpha); % get PDF
yrc = raylcdf(X, Alpha); % get CDF
figure(5); set(gca,'FontSize', FontSize);
plot(X, yrp, 'LineWidth', LineWidth);
title(['Rayleigh Probablity Density Function for \alpha = ' ...
     num2str(Alpha)]);
xlabel('X (random variable)');
ylabel('f_X(x)'); grid
figure(6); set(gca,'FontSize', FontSize);
plot(X, yrc,'LineWidth', LineWidth);
title(['Rayleigh Probablity Density Function for \alpha = ' ...
     num2str(Alpha)]);
xlabel('X (random variable)');
ylabel('F_X(x) = Prob[X <= x]'); grid
```

Gamma Random Variable

- **Versatile distribution ~ appears in modeling of lifetime of devices and systems**
- **Has two parameters: $\alpha > 0$ and $\lambda > 0$**
- **PDF:** $$f_X(x) = \frac{\lambda(\lambda x)^{\alpha-1}e^{-\lambda x}}{\Gamma(\alpha)}$$
- **For $0 < x < \infty$**
- **The quantity $\Gamma(z)$ is the gamma function and is specified by** $$\Gamma(z) = \int_0^\infty x^{z-1}e^{-x}dx$$
- **The gamma function has the following properties:**
  - **$\Gamma(1/2) = \sqrt{\pi}$**
  - **$\Gamma(z+1) = z\Gamma(z)$ for $z>0$**
  - **$\Gamma(m+1) = m!$ For m nonnegative integer**
- **$E[X] = \alpha/\lambda$, $Var[X] = \alpha/\lambda^2$**
- **$\Phi_X(\omega) = 1/(1-j\omega/\lambda)^{\alpha}$**

If $\alpha = 1$ ➔ gamma r.v. becomes exponential

Pareto Random Variable

- **Originally used by economists to model income and other soci-economic quantities.**
- **For α (shape parameter) > 0, β (scale parameter) > 0, the PDF is given by**

$$f_X(x) = \frac{\alpha\beta^{\alpha}}{x^{\alpha+1}} \qquad \beta \le x$$

- **The CDF is given by**

$$F_X(x) = 1 - \left(\frac{\beta}{x}\right)^{\alpha} \qquad \beta \le x$$

Pareto Random Variable - 2

- **nth moment (if it exists) is given by**

$$E\left[x^n\right] = \frac{\alpha\beta^n}{\alpha - n} \qquad n < \alpha$$

- **Expected value:** $E[x] = \dfrac{\alpha\beta}{\alpha-1} \qquad 1 < \alpha$
- **Variance:** $Var[x] = \dfrac{\alpha\beta^2}{(\alpha-1)^2(\alpha-2)} \qquad 2 < \alpha$

Example 6: Packet Size Modeling

- **Pareto distribution is used to model the packet size, P, in bytes for internet traffic as follows:** $P = \min(x, S_{\max})$

where x is a Pareto random variable with the following PDF

$$f_X(x) = \begin{cases} \dfrac{\alpha\beta^{\alpha}}{x^{\alpha+1}} & \beta \le x < S_{\max} \\ \theta & x = S_{\max} \end{cases}$$

θ is given by $\theta = 1 - F_X(S_{\max})$

Example 7: Packet Size Modeling

- **Calculate the expected value for packet size using the model proposed in Example 3?**

- **Models proposed to test ETSI/UMTS networks use the following parameters: α = 1.1, β = 81.5 Bytes, Smax = 66,666 Byte (this results in a mean packet size of 480 Bytes)**

Computer Methods for Generating Random Variables

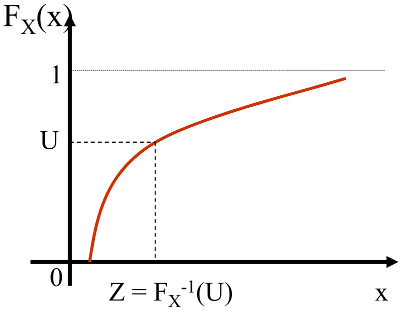
(1) The transformation method

Procedure:

a. **Obtain $F_X(x)$**

b. **Generate U ~ uniform between 0 and 1**

c. **Find $Z = F_X^{-1}(U)$ – Z follows the distribution specified by $f_X(x)$**



Example 8 – Generating Exponential r.v.

Problem: Generating exponential random variables with parameter λ
Answer:
To generate an exponentially distributed r.v. X with parameter λ (i.e. its mean is 1/λ), we need to find $F_X(x)$ and invert it.

$F_X(x) = 1 - e^{-\lambda x}$ (see example 1)

Therefore, $F_X^{-1}(x)$ is equal to

$$X = -(1/\lambda)\ln(1-U)$$

where ln(t) is the natural logarithm of t while U is a uniform r.v. between 0 and 1. Note that the above expression can be simplified to be

$$X = -(1/\lambda)\ln(U)$$

This is because 1-U is also a uniform random r.v. between 0 and 1

Example 9 – Generating Bounded Pareto Distribution

Problem: Generate a random variable conforming the bounded Pareto distribution specified in Example 4.

Answer: ?

Example 10 – Generating Gaussian Random Variable

Problem: Generate a Gaussian random variable of mean *m* and standard deviation equal to δ.

Answer: ?

Computer Methods for Generating Random Variables

(2) Rejection Method
(3) Composition Method
(4) Convolution Techniques
(5) Characterization Method

See references for details

Transformation method is sufficient for simulations required in this course

Joint Distributions of Random Variables

- **Def: The joint probability distribution of two r.v.s X and Y is given by**

$$F_{XY}(x,y) = P(X \leq x, Y \leq y)$$

where x and y are real numbers.

- **This refers to the JOINT occurrence of $\{X \leq x\}$ AND $\{Y \leq y\}$**
- **Can be generalized to any number of variables**

Joint Distributions of Random Variables - Properties

- $F_{XY}(-\infty, -\infty) = 0$
- $F_{XY}(\infty, \infty) = 1$
- $F_{XY}(x_1, y) \leq F_{XY}(x_2, y)$ for $x_1 \leq x_2$
- $F_{XY}(x, y_1) \leq F_{XY}(x, y_2)$ for $y_1 \leq y_2$
- The marginal distributions are given by
  - $F_X(x) = F_{XY}(x, \infty)$
  - $F_Y(y) = F_{XY}(\infty, y)$

Joint Distributions of Random Variables – Properties - 2

- **Density function:** $$f_{XY}(x,y) = \frac{\partial^2 F_{XY}(x,y)}{\partial x \partial y}$$

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

- **Marginal densities:** $$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x,y)\,dy$$

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

Joint Distributions of Random Variables – Independence

- **Two random variables are independent if the joint distribution functions are products of the marginal distributions:**

$$F_{XY}(x,y) = F_X(x)F_Y(y)$$

or

$$f_{XY}(x,y) = f_X(x)\,f_Y(y)$$

Joint Distributions of Random Variables – Discrete Nonnegative Variables

- **Def:**

$$F_{XY}(x,y) = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} P(X = x_i, Y = y_i)U(x - x_i)U(y - y_i)$$

where

P(X=xi, Y=yi) is the joint probability for the r.v.s X and Y

U(x) is 1 for x ≥ 0 and 0 otherwise

Example 11:

- **<u>Problem</u>: The number of bytes N in a message has a geometric distribution with parameter p. The message is broken into packets of maximum length M bytes. Let Q be the number of full packets in a message and let R be the number of bytes left over. Find the join pmf and the marginal pmfs of Q and R.**

Example 11: cont'd

- **<u>Solution</u>:**

N ~ geometric ➔ $P(N=k) = (1-p)p^k$

Message of N bytes ➔ Q full M-bytes packets +
R remaining bytes

Therefore: $Q \in \{0, 1, 2, \ldots\}$, $R \in \{0, 1, 2, \ldots, M-1\}$

The join pmf is given by:

$P(Q=q, R=r) = P(N = qM+r) = (1-p)p^{(qM+r)}$

Example 11: cont'd

- **Solution:**

The marginal pmfs:

$$P(Q=q)=\sum_{r=0}^{M-1}P(Q=q,R=r)$$
$$=\sum_{r=0}^{M-1}(1-p)p^{(qM+r)}$$
$$=\left(1-p^{M}\right)\left(p^{M}\right)^{q}\qquad q=0,1,2,\ldots$$

and

$$P(R=r)=\sum_{q=0}^{\infty}P(Q=q,R=r)$$
$$=\sum_{q=0}^{\infty}(1-p)p^{(qM+r)}$$
$$=\frac{(1-p)}{1-p^{M}}p^{r}\qquad r=0,1,\ldots,M-1$$

Verify the marginal pmfs add to ONE!!
P(R = r) is a truncated geometric r.v.

Independent Discrete R.V.s

- **For Discrete random variables:**

P(M=i, N=j) = P(M=i) P(N=j)

Example 12:

- **Problem: Are the Q and R random variables of Example 11 independent? Why?**

Conditional Distributions

- **Def: for continuous X and Y**

$$F_{Y/X}(y/x) = P(Y \le y / X \le x) = \frac{F_{XY}(x,y)}{F_X(x)}$$

Or

$$f_{Y/X}(y/x) = \frac{f_{XY}(x,y)}{f_X(x)}$$

- **For discrete M and N**

$$P(M = i / N = j) = \frac{P(M = i, N = j)}{P(N = j)}$$

Conditional Distributions - 2

- **For mixed types:**

$$F_X(x) = \sum_{i=0}^{\infty} P(N = j, X \le x)$$

$$= \sum_{j=0}^{\infty} P(N = j) P(X \le x / N = j)$$

or

$$P(N = j) = \int_{-\infty}^{\infty} P(N = j / X = x) f_X(x) dx$$

Conditional Distributions - 3

- **For mixed types:**

$$F_{X_1, X_2, \ldots, X_N}(x_1, x_2, \ldots, x_N) =$$

$$= F_{X_1}(x_1) \times F_{X_2/X_1}(x_2 / x_1) \times \ldots \times F_{X_N/X_1, \ldots, X_{N-1}}(x_N / x_1, \ldots, x_{N-1})$$

or

$$f_{X_1, X_2, \ldots, X_N}(x_1, x_2, \ldots, x_N) =$$

$$= f_{X_1}(x_1) \times f_{X_2/X_1}(x_2 / x_1) \times \ldots \times f_{X_N/X_1, \ldots, X_{N-1}}(x_N / x_1, \ldots, x_{N-1})$$

Example 14:

- **Problem: The number of customers that arrive at a service station during a time t is a Poisson random variable with parameter βt. The time required to service each customer is exponentially distributed with parameter α. Find the pmf for the number of customers N that arrive during the service time T of a specific customer. Assume the customer arrivals are independent of the customer service time.**

Example 14: cont'd

- **Solution:**

The PDF for T is given by $f_T(t) = \alpha e^{-\alpha t} \qquad t \ge 0$

Let N = number of arrivals during time t

➔ the arrivals conditional pmf is given by

$$P(N = j / T = t) = \frac{(\beta t)^j e^{-\beta t}}{j!} \qquad j = 0,1,\ldots \qquad t \ge 0$$

To find the arrivals pmf during service time T, we use:

$$P(N = j) = \int_{-\infty}^{\infty} P(N = j / T = t) f_T(t) dt$$

this reduces to:

$$= \int_{0}^{\infty} \frac{(\alpha\beta)^j}{j!} e^{-\alpha t} e^{-\beta t} dt$$

Note that:

$$\Gamma(j+1) = \int_0^\infty t^j e^{-t} dt = j!$$

$$P(N = j) = \left(\frac{\alpha}{\alpha+\beta}\right)\left(\frac{\beta}{\alpha+\beta}\right)^j \qquad j = 0,1,\ldots$$

Thus N is geometrically distributed with probability of success equal to α /(β+ α)

Joint Moments

- **For continuous X and Y:**

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

- **For discrete X and Y**

$$E[g(X,Y)] = \sum_{\forall i}\sum_{\forall j} g(x_i, y_j) P\left(X = x_i, Y = y_j\right)$$

Autocorrelation and Autocovariance Function

- **Autocorrelation:**
    - **For continuous X and Y:**

$$E[XY] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy f_{XY}(x,y)\,dxdy$$

-
    - **For discrete X and Y**

$$E[XY] = \sum_{\forall i}\sum_{\forall j} x_i y_j P\left(X = x_i, Y = y_j\right)$$

- **Autocovariance:**

$$Cov[X,Y] = E[(X - E[X])(Y - E[Y])]$$

Autocorrelation and Autocovariance Function - 2

- **X and Y are uncorrelated if**

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

or equivalently ➔ $Cov[X, Y] = 0$

- **Independent variables are uncorrelated, the reverse DOES NOT HOLD – Gaussian r.v.s are the exception**

Example 14: Joint Gaussian Variables

- **<u>Problem:</u> show that if X and Y are two uncorrelated Gaussian r.v.s, then they are independent**

Example 14: Joint Gaussian Variables – cont’d

- **Solution:**

The joint distribution for X and Y is given by

$$f_{XY}(x,y)=$$

$$\frac{\exp\left[\left(\frac{-1}{2\left(1-\rho_{XY}^{2}\right)}\right)\left\{\left(\frac{x-\mu_x}{\sigma_x}\right)^{2}-2\rho_{XY}\left(\frac{x-\mu_x}{\sigma_x}\right)\left(\frac{y-\mu_y}{\sigma_y}\right)+\left(\frac{y-\mu_y}{\sigma_y}\right)^{2}\right\}\right]}{2\pi\sigma_x\sigma_y\sqrt{1-\rho_{XY}^{2}}}$$

where:

μ_X and μ_X are equal to E[X] and E[Y], respectively

σ_X and σ_Y are the respective standard deviations

ρ_{XY} = Cov(X,Y)/($\sigma_X\sigma_Y$)

Example 14: Joint Gaussian Variables – cont’d

- **Solution:**

If X and Y are uncorrelated ➔ Cov(X,Y) = 0 or ρ_{XY} = 0. Rewriting the joint distribution yields

$$f_{XY}(x,y)=\frac{\exp\left\{-\left(x-\mu_x\right)^{2}/\left(2\sigma_x^{2}\right)-\left(y-\mu_y\right)^{2}/\left(2\sigma_y^{2}\right)\right\}}{2\pi\sigma_x\sigma_y}$$

but the last expression is equal to $f_X(x)f_Y(y)$

Therefore, X and Y are independent

Functions of a Random Variable

- **Problem setting:**
  - **Let X be a r.v.,**
  - **Let g(x) be a real-valued function**
  - **Y = g(X)**
  - **What is the probability distribution for Y?**

- **General Approach:**

Prob[Y in C] = Prob[g(X) in C]
= Prob[X in B]

These events are equivalent

Example 15: The MAX Function

- **Let $g(x) = (x)^+$**

$$g(x) = \begin{cases} 0 & \text{if } x < 0 \\ x & \text{if } x \geq 0 \end{cases}$$

Note g(x) can be written in other forms:
g(x) = max(x, 0)

e.g:

1. # of customers arriving in batch sizes greater than M ➔ $Y = (X - M)^+$

2. voltage output of a half-wave rectifier

Example 16: The MAX Function

- **<u>Problem:</u> Let X be an nonnegative integer-valued r.v. Let $P(X=i) = p_i$, for i=0,1, …**

  Y is defined as Y = max(X-M, 0) where M is +ve integer

  Find pmf for the r.v. Y

- **<u>Solution:</u>**

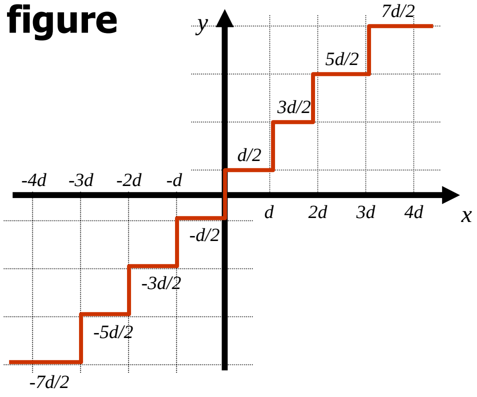
  $Y = \max(X-M,0) = (X-M)^+$ has the range {0, 1, …}

  $$P(Y = 0) = \text{Prob}[X \le M]$$
  $$= \Sigma\, p_i \qquad i=0,1, \ldots, M$$

  $$P(Y = k) = p_{k+M} \qquad k = 1,2, \ldots$$

Example 17: Quantization

- **Let Y = q(X) be the uniform quantization function defined in figure**



- **Note Y can be written as**

Y = floor(X) + 0.5d

e.g. PCM voice

Example 18: Quantization

- **Problem: Let X be a sample voltage of a speech waveform and suppose that X is uniform on the interval [-4d, 4d]. Let Y = q(X), where the quantizer input-output characteristic is as shown in previous example. Find the pmf for Y.**

Example 18: Quantization – cont'd

- **Solution:**

  Y = q(x), -4d ≤ x ≤ 4d, Y ∈ {±7d/2, ±5d/2, ±3d/2, ±d/2}

  The PDF for X is given by:

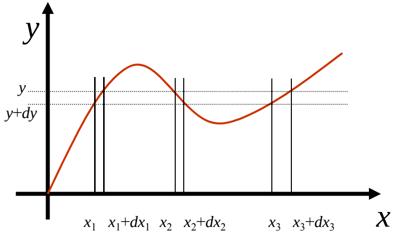
$$f_X(x) = 1/(8d) \qquad -4d \le x \le 4d$$

  Therefore, the PDF for Y is computed as:

$$P(Y = k) = \int_{k-0.5d}^{k+0.5d} f_X(x)dx$$
$$= 1/8 \qquad k \in \{\pm 7/2d, \pm 5/2d, \pm 3/2d, \pm 1/2d\}$$

General Rule

- **Problem: Let Y = g(x), if the PDF for X is given by $f_X(x)$, find the PDF for the r.v. Y.**



- **Solution:**

Prob[y < Y < y +dy] = $f_Y(y)$ |dy|

The event {y < Y < y +dy} is equivalent to the event {x_1 < X < x_1 +dx_1} ∪ {x_2 < X < x_2 +dx_2} ∪ {x_3 < X < x_3 +dx_3}

➔ $f_Y(y)\ |dy| = f_X(x_1)|dx_1|+f_X(x_2)|dx_2|+f_X(x_3)|dx_3|$

In general: $$f_Y(y) = \sum_k f_X(x)\left|\frac{dx}{dy}\right|_{x=x_k}$$

Linear Transformations – Y = aX + b

- **This is a special case of "Functions of Random Variables"**

- **The PDF of Y can be shown to be**

$$f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y-b}{a}\right)$$

- **One can also show that**

$$E[Y] = aE[X] + b$$

and

$$Var[Y] = a^2 Var[X]$$

Example 19: $Y = X^2$

- **<u>Problem:</u> Let $Y = X^2$, where X is a continuous r.v. Find the PDF of Y.**
- **<u>Solution:</u>**

$y = x^2$ ➔ has two solutions: $x_{0,1} = \pm\sqrt{y}$
$|dy/dx| = 2x = 2\sqrt{y}$
therefore $f_Y(y)$ is given by:

$$f_Y(y) = \frac{f_X(\sqrt{y})}{2\sqrt{y}} + \frac{f_X(-\sqrt{y})}{2\sqrt{y}}$$

Functions of Multiple Random Variables

- **Let $X_1, X_2, \ldots, X_n$ be random variables and let Z be defined as**

$$Z = g(X_1, X_2, \ldots, X_n)$$

The CDF of Z is found as follows:

$$\{Z \le z\} \equiv R_z = \{X = (x_1, x_2, \ldots, x_n): g(X) \le z\}$$

Therefore,

$$F_Z(z) = P(X \text{ in } R_z)$$

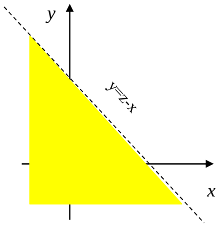
or

$$F_Z(z) = \int \ldots \int f_{X_1, X_2, \ldots, X_N}(x_1, x_2, \ldots, x_N)\, dx_1 dx_2 \ldots dx_N$$

where the integrals are carried over R_z

Example 20: Z = X + Y

- **<u>Problem:</u> Let Z = X + Y, find $F_Z(z)$ and $f_Z(z)$ in terms of $f_{XY}(x,y)$**



- **<u>Solution:</u>**

P(Z ≤ z) = P(X+Y ≤ z)

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

The PDF for Z is given by $$f_Z(z) = \int_{-\infty}^{\infty} f_{XY}(x', z-x')dx'$$

Note that is X and Y are independent, then $f_Z(z)$ can be written as: $$f_Z(z) = \int_{-\infty}^{\infty} f_X(x')f_Y(z-x')dx' = f_X(x) * f_Y(y)$$

The later relation is know as the convolution integral of the marginal PDFs for X and Y

One can also show that $$\Phi_Z(\omega) = \Phi_X(\omega)\Phi_Y(\omega)$$

where Φ(ω) is the characteristic function for the respective r.v.

Sum of Random Variables

- **Let $X_1, X_2, \ldots, X_n$ be random variables and let Y be defined as**

$$Y = X_1 + X_2 + \ldots + X_n$$

- **It is easy to show that**

$$E[Y] = E[X_1] + E[X_2] + \ldots + E[X_n]$$

Exercise: Prove these relations

This results holds whether X_is are independent or not

- **Furthermore,**

$$Var[Y] = \sum_{i=1}^{n} Var[X_i] + \sum_{i=1}^{n}\sum_{j=1, i\neq j}^{n} Cov(X_i, X_j)$$

For uncorrelated X_is, the relation reduces to

$$Var[Y] = \sum_{i=1}^{n} Var[X_i]$$

Sum of Random Variables – cont'd

- **Generalizing the results of Example 20, the PDF of the random variable Y is given by**

$$f_Y(y) = f_{X_1}(x_1) * f_{X_2}(x_2) * \ldots f_{X_N}(x_N)$$

or

$$\Phi_Y(\omega) = \Phi_{X_1}(\omega)\Phi_{X_2}(\omega)\ldots\Phi_{X_N}(\omega)$$

- **Note the above relation is valid for the probability generating function N(Z) and the Laplace transform X(s) as well.**

Sum of Two Nonnegative Integer-Valued Random Variables

- **Let $N = K_1 + K_2$, where K_1 and K_2 are nonnegative integer-valued random variables. The distribution for N is given by**

$$P(N = n) = P(i + j = n) \qquad \forall i, j = 0,1,\ldots$$
$$= \sum_{i=0}^{n} P(K_1 = i, K_2 = n - i) \qquad n = 0,1,\ldots$$

if the variables K_1 and K_2 are independent, then the distribution can be written as

$$P(N = n) = \sum_{i=0}^{n} P(K_1 = i)P(K_2 = n - i) \qquad n = 0,1,\ldots$$

which is the discrete form of the convolution integral introduced in Example 20

Example 21: Sum of Two Independent Poisson R.V.s

- **<u>Problem:</u> Define Y = $K_1 + K_2$, where K_1 and K_2 are two independent Poisson random variables with mean $\lambda_1 t$ and $\lambda_2 t$. Find the distribution of Y**

Example 21: Sum of Two Independent Poisson R.V.s – cont'd

- **<u>Solution 1</u>:**

The pmfs of K_1 and K_2 are given by

$$P(K_i = j) = \frac{(\lambda_i t)^j}{j!} e^{-\lambda_i t} \qquad i = 1,2;\ j = 0,1,\ldots$$

Using the convolution relation, the pmf for N is computed as

$$P(N = n) = \sum_{i=0}^{n} \frac{(\lambda_1 t)^i}{i!} e^{-\lambda_1 t} \frac{(\lambda_2 t)^{n-i}}{(n-i)!} e^{-\lambda_2 t}$$

$$= e^{-(\lambda_1+\lambda_2)t} \sum_{i=0}^{n} \frac{(\lambda_1 t)^i (\lambda_2 t)^{n-i}}{i!(n-i)!}$$

$$= \frac{[(\lambda_1 + \lambda_2)t]^n}{n!} e^{-(\lambda_1+\lambda_2)t}$$

Note:

$$\sum_{i=0}^{n} \binom{n}{i} x^i y^{n-i} = (x+y)^n$$

Examining the last expression, one can conclude that N itself follows the Poisson distribution with mean $(\lambda_1 t + \lambda_2 t)$

Example 21: Sum of Two Independent Poisson R.V.s – cont'd

- **Solution 2:**

  The pmfs of K_1 and K_2 are given by

$$P(K_i = j) = \frac{(\lambda_i t)^j}{j!} e^{-\lambda_i t} \qquad i = 1,2;\ j = 0,1,\ldots$$

Or equivalently, the respective probability generating functions are given by

$$N_{K_i}(z) = e^{-\lambda_i (1-z)} \qquad i = 1,2; |z| \le 1$$

Using the convolution relation, the probability generating function for the sum N is given by

$$\begin{aligned} N_N(z) &= N_{K_1}(z) N_{K_2}(z) \\ &= e^{-\lambda_1 (1-z)} e^{-\lambda_2 (1-z)} \\ &= e^{-(\lambda_1+\lambda_2)(1-z)} \end{aligned}$$

Examining the last expression, one can conclude that N itself follows the Poisson distribution with mean $(\lambda_1 + \lambda_2)t$

Example 22: Sum of Two Exponential Random Variables

- **Problem:** **Let $Y = X_1 + X_2$ where X_1 and X_2 are identical independent (iid) exponential r.v.s with parameter μ. Find the distribution of Y.**

- **Solution:**

  The exponential PDF is given by $f_{X_i}(x) = \mu e^{-\mu x} \qquad i = 1,2; x \ge 0$

  Using the convolution integral, the PDF for Y is computed as

$$\begin{aligned} f_Y(y) &= \int_0^\infty \mu e^{-\mu x} \times \mu e^{-\mu(y-x)} dx \\ &= y\mu^2 e^{-\mu y} \qquad y \ge 0 \end{aligned}$$

One can show in general that the distribution of the sum of k iid exponential r.v.s is given by

$$f_Y(y) = \frac{\mu(\mu y)^{k-1} e^{-\mu y}}{(k-1)!} \qquad y \ge 0; k = 1,2,\ldots$$

The above is referred to as k-stage Erlang distribution

It is useful to note

$$X_Y(s) = \left[\frac{\mu}{\mu + s}\right]^k$$

Exercise: prove that the $E[Y] = k/\mu$ and $Var[Y] = k/\mu^2$

Example 23: Sum of Random Number of Exponential Random Variables

- **Problem:** Let Y be the sum of k iid exponential r.v.s as in previous example. The number of random number k is itself a geometric r.v. with parameter p. Find the distribution of Y.
- **Solution:**

Using the results of previous example, the Laplace transform of the r.v. Y conditioned on the fact k is equal to n is given by

$$X_Y(s/k=n)=\left[\frac{\mu}{\mu+s}\right]^n$$

Therefore, the average Laplace transform is given by

$$\begin{aligned}X_Y(s)&=\sum_{n=1}^{\infty}X_Y(s/k=n)P(k=n)\\&=\sum_{n=1}^{\infty}(1-p)^{n-1}p\left[\frac{\mu}{\mu+s}\right]^n\\&=\frac{p\mu}{p\mu+s}\end{aligned}$$

Examining the last formula, one can conclude that the sum is itself an exponentially distributed r.v. with mean 1/(pμ)

Inequalities and Bounds

- **Markov Inequality**
- **Chebyshev Inequality**
- **Chernoff Bound**

Markov Inequality

- **Let X be a nonnegative random number and h(X) is a nondecreasing function of X, then the expectation h(X) can be written as**

$$E[h(X)] = \int_{-\infty}^{\infty} h(x) f_X(x)dx \geq \int_{t}^{\infty} h(x) f_X(x)dx \geq h(t)\int_{t}^{\infty} f_X(x)dx \geq h(t)P(X \geq t)$$

Therefore,

$$P(X \geq t) \leq \frac{E[h(X)]}{h(t)} \qquad t \geq 0$$

Two popular example of h(X), are h(X) = x and h(X) = e^{ax}

For h(X) = x, we can write

$$P(X \geq t) \leq \frac{E[X]}{t} \qquad t \geq 0$$

The above is referred to as the simple Markov inequality

Example 24: Markov Inequality

- **Problem:** **Find the simple Markov inequality and compare with the exact survivor function for an Erlang 4 distribution with μ = 2.**

- **Solution:**

E[X] is given by k/ μ = 2 (see example 23)

therefore, the simple Markov inequality is given by

$$P(X \geq t) \leq \frac{E[X]}{t} = \frac{2}{t} \qquad t \geq 0$$

The exact survivor function is evaluated as

$$P(X \geq t) = \int_{t}^{\infty} \frac{2(2x)^{k-1}}{(k-1)!} e^{-2x} dx$$

Example: Markov Inequality – cont'd

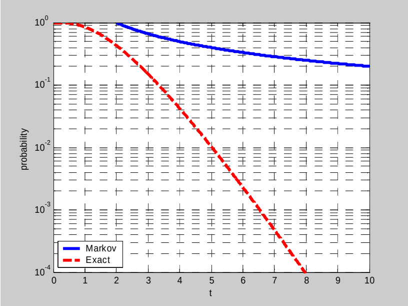
- **<u>Solution:</u>**

The former integral can be evaluated either numerically or using tables of integrals, using the latter,

$$P(X \ge t) = e^{-2t}\sum_{j=0}^{k-1}\frac{(2t)^{k-1-j}}{(k-1-j)!} \qquad k = 4$$

- **The simple Markov inequality and the exact survivor function are plotted in the graph. It is clear the computed bound it quite loose!!**

- **Actually for 0<t <2 → the bound value is greater than 1!**



Chebyshev Inequality

- **Using Markov's inequality one can write**

$$P(Y \ge \varepsilon^2) \le \frac{E[Y]}{\varepsilon^2}$$

Let Y = (X – E[X])² ➔ E[Y] = Var[X]

therefore,

$$P\left((X - E[X])^2 \ge \varepsilon^2\right) \le \frac{Var[X]}{\varepsilon^2}$$

but P(|X-E[X]|² ≥ ε²) is equal to P(|X-E[X]| ≥ ε); Hence the inequality can be rewritten as

$$P\left(|X - E[X]| \ge \varepsilon\right) \le \frac{Var[X]}{\varepsilon^2}$$

Example 25: Chebyshev Inequality

- **<u>Problem:</u> Find Chebyshev inequality for the Erlang 4 distribution of previous example.**

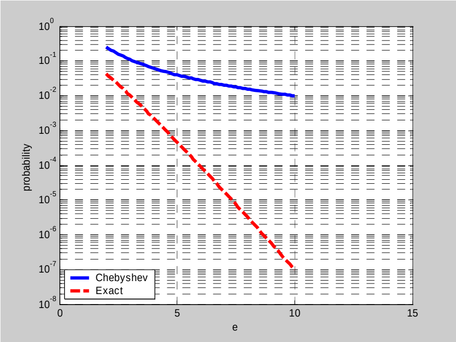
- **<u>Solution:</u>**

  The mean of the Erlang 4 distribution, E[X], is equal to k/μ=2, while variance is equal to k/μ^2= 1.

  Therefore, the inequality is then

$$P(|X-2| \geq \varepsilon) \leq \frac{1}{\varepsilon^2}$$

Example: Chebyshev Inequality – cont'd



- **<u>Solution:</u>**

  The exact solution of P(|X-2|≥ ε) is given by

$$\begin{aligned} P(|X-2| \geq \varepsilon) &= P(X \geq 2+\varepsilon) \qquad x \geq 2 \\ &= \int_{2+\varepsilon}^{\infty} f_X(x)dx \qquad x \geq 2 \\ &= e^{-2(2+\varepsilon)} \sum_{j=0}^{k} \frac{[2(2+\varepsilon)]^{k-1-j}}{(k-1-j)!} \qquad k=4 \end{aligned}$$

- **The graph shows a comparison between Chebyshev inequality and the exact solution. The bound is loose!**

Chernoff Bound

- **Using Markov's bound, and letting h(t) = e$^{\alpha t}$; where α ≥ 0, one can write**

$$P(X \geq d) \leq e^{-\alpha d} E\left[e^{\alpha X}\right] = e^{-\alpha d} X(-\alpha) \quad \alpha \geq 0$$

where X(- α) is the Laplace transform of the variable X evaluated at s = - α.

- **Note that for discrete r.v.s, Let z = e$^{\alpha}$ in the previous expression, this results in the following bound**

$$P(X \geq j) \leq z^{-j} N(-\ln \alpha) \quad \alpha \geq 0$$

where N(z) is the probability generating function for the r.v. X.

Example: Chernoff Bound

- **<u>Problem:</u> Find Chernoff bound for the Erlang 4 distribution of previous example.**
- **<u>Solution:</u>**

Substituting directly into the Chernoff bound formula,

$$P(X \geq t) \leq e^{-\alpha t}\left(\frac{\mu}{-\alpha + \mu}\right)^{k}$$

To determine the value of α that minimizes the RHS we differentiate with respect to α and solve for α

➔ α= μ – k/t

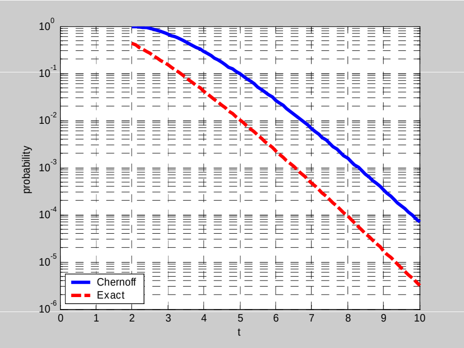
Example: Chernoff Bound – cont'd

- **Solution:**

Therefore,

$$P(X \geq t) \leq e^{-t\mu+k}\left(\frac{\mu t}{k}\right)^{k}$$

It is interesting to note that the bound decays at the same rate as that of the exact solution!



Weak Law of Large Numbers

- **Let $X_1, X_2, \ldots, X_n$ be a sequence of iid random variables with finite mean $E[X] = \mu$.**

Define $M_n = 1/n\,(X_1 + X_2 + \ldots + X_n)$

It can be shown that for $\varepsilon > 0$

$$\lim_{n \to \infty} P(|M_n - \mu| < \varepsilon) = 1$$

Prove this law

Note:

- **M_n is referred to as the sample mean**
- **The weak law states that for a large enough fixed value of n, the sample mean using n samples will be close to the true mean with high probability**
- **The above is true even if the variance is not finite!!**

References

- **Alberto Leon-Garcia, Probability and Random Processes for Electrical Engineering, Chapter 3, Addison Wesley, 1989**

- **J. Hayes and T.V.J. Ganesh Babu, Modeling and Analysis of Telecommunications Networks, Chapter 2, Wiley, 2004**