Lecture notes 2
Random Variables and Distributions

1. Random Variables:
A random variable is a well defined rule for assigning a numerical value to all possible outcome of an experiment.

A random variable is called discrete if it takes on either a finite or a countable number of possible values, and is called continuous if it takes on a continuum of possible values.

Example 1:
Experiment: Flipping a coin once
Outcomes: Head and Tail
Sample Space: Discrete and finite
A random variable can be defined as:

$$X = \begin{cases} 1, & if\ head\ occurs, \\ 0, & if\ tail\ occurs. \end{cases}$$

Example 2:
Experiment: Taking an exam.
Outcomes: Grade $A,\ B,\ C,\ D,\ F$
Sample Space: Discrete and finite
A random variable can be defined as:

$$Y = \begin{cases} 4, & if\ grade\ A, \\ 3, & if\ grade\ B, \\ 2, & if\ grade\ C, \\ 1, & if\ grade\ D, \\ 0, & if\ grade\ F. \end{cases}$$

Example 3:
Experiment: Investing \$1000 in a common stock.
Outcomes: Value of yield.
Sample Space: Continuous.
A random variable can be defined as: $X =$ value of yield, $0 < X < \infty$.

Since the value of a random variable is determined by the outcome of the experiment, we may assign probabilities to the possible value of the random variable.

Example 4.
For example 1, we may define

$$P(X=0) = \frac{1}{2}\ ;\ P(X=1) = \frac{1}{2}.$$

Of course $P(X=0) + P(X=1) = 1$.

The cumulative distribution function (cdf) (or simply the distribution function) $F(b)$ of the random variable X is defined for any real number b, $-\infty < b < \infty$, by

$$F_X(b) = F(b) = P(X \leq b).$$

Note:
1. $F(b)$ is a nondecreasing function of b.
2. $\lim_{b \to \infty} F(b) = F(\infty) = 1$.
3. $\lim_{b \to -\infty} F(b) = F(-\infty) = 0$.

2. Discrete Random Variables: For a discrete random variable X, we define the probability mass function (pmf) $f(a)$ of X by

$$f(a) = P(X = a).$$

(a) The Bernoulli Random Variable
Suppose that a trial, or an experiment, whose outcome can be classified as either a "success" or as "failure" is performed. If we let

$$X = \begin{cases} 1, & if\ success \\ 0, & if\ failure \end{cases}$$

and

$$f(1) = P(X = 1) = p\ ;\ f(0) = P(X = 0) = 1 - p$$

where $0 \leq p \leq 1$, then X is said to be a Bernoulli random variable with parameter p.

(b) The Binomial Random Variable
Suppose $X = X_1 + X_2 + \ldots + X_n$ where the X_i's are independent Bernoulli random variable with probability of success p, $i = 1, \ldots, n$. Then we call X to be a binomial random variable with parameters (n, p). We sometimes write $X \sim B(n, p)$. Note that

$$f(i) = C_i^n p^i (1 - p)^{n-i}.$$

(c) The Geometric Random Variable
Suppose that independent trials, each having a probability p of being a success, are performed until a success occurs. If we let X be the number of trials required until the first success, then X is said to be a geometric random variable with parameter p. Note that

$$f(n) = P(X = n) = (1 - p)^{n-1} p, \quad n = 1, 2, \ldots$$

(d) The Poisson Random Variable

A random variable X, taking on one of the values $0, 1, 2, \ldots$, is said to be a Poisson random variable with parameter λ, if for some $\lambda > 0$,

$$f(i) = P(X = i) = \frac{e^{-\lambda}\lambda^i}{i!}, \; i = 0, 1, \ldots$$

We sometimes write $X \sim P(\lambda)$.

The use of the Poisson distribution:

The Poisson distribution can be used to determine the probability of X occurrences per unit time if four basic assumptions are met:

i. Possible to divide time interval of interest into many small subintervals.
ii. Probability of an occurrence remains constant throughout the time interval.
iii. Probability of two occurrences in a subinterval is small enough to be ignored.
iv. The number of occurrences in any interval of time is independent of the number of occurrences in any other disjoint time interval.

Examples:

i. arrivals at a bank per hour.
ii. number of telephone calls arrived per minute.

An important property of the Poisson random variable is that it may be used to approximate a binomial random variable when the binomial parameter n and p where n is large and p is small. To see this, let $X \sim B(n, p)$ *and* $\lambda = np$, then

$$\begin{aligned} P(X = i) &= \frac{n!}{(n-i)!i!}\, p^i (1-p)^{n-i} \\ &= \frac{n!}{(n-i)!i!} (\frac{\lambda}{n})^i (1 - \frac{\lambda}{n})^{n-i} \\ &= \frac{n(n-1)\ldots(n-i+1)}{n^i} \frac{\lambda^i}{i!} \frac{(1-\frac{\lambda}{n})^n}{(1-\frac{\lambda}{n})^i}. \end{aligned}$$

Now, for n large and p small.

$$(1 - \frac{\lambda}{n})^n \approx e^{-\lambda}\ , \quad \frac{n(n-1)\ldots(n-i+1)}{n^i} \approx 1\ , \ (1 - \frac{\lambda}{n})^i \approx 1.$$

Hence, for n large and p small,

$$P(X = i) \approx \frac{e^{-\lambda}\lambda^i}{i!}.$$

3. Continuous Random Variables:
   We say that X is a continuous random variable if there exists a nonnegative function $f(x)$, defined for all real $x \in (-\infty, \infty)$, having the property that for any set B of real numbers
$$P(x \in B) = \int_B f(x)dx.$$
   The function $f(x)$ is called the probability density function (pdf) of the random variable X. Note that
$$P(a \leq X \leq b) = P(a < X \leq b) = P(a < X < b) = \int_a^b f(x)dx$$
   and
$$F(a) = P(-\infty < X \leq a) = \int_{-\infty}^a f(x)dx.$$
   (a) The Uniform Random Variable
   We say that X is a uniform random variable on the interval (α, β) if the pdf is given by
$$f(x) = \begin{cases} \frac{1}{\beta-\alpha}, & if\ \alpha < x < \beta \\ 0, & otherwise. \end{cases}$$
   (b) The Gamma Random Variable
   A random variable X has a gamma distribution if and only if its pdf is given by
$$f(x) = \begin{cases} \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta}, & if\ x > 0 \\ 0, & otherwise. \end{cases}$$
   where $\alpha > 0$ and $\beta > 0$.
   (c) The Exponential Random Variable
   A random variable X has an exponential distribution if and only if its pdf is given by
$$f(x) = \begin{cases} \frac{1}{\theta} e^{-x/\theta}, & if\ x > 0 \\ 0, & otherwise. \end{cases}$$
   where $\theta > 0$. In fact, it is a special case of the gamma distribution with $\alpha = 1$ and $\beta = \theta$.
   (d) The Normal Random Variable
   A random variable X has a normal distribution if and only if its pdf is given by
$$f(x; \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2}, \quad for\ -\infty < x < \infty$$
   where $\sigma > 0$. We write $X \sim N(\mu, \sigma^2)$. The normal distribution with $\mu = 0$ and $\sigma = 1$ is referred to as the standard normal distribution and is denoted as $N(0, 1)$.

4. Expectation of a Random Variable

(a) The discrete case
If X is a discrete random variable and $f(x)$ is the value of its probability distribution at x, the expected value (or expectation) of X is

$$E(X) = \sum_x x f(x).$$

(b) The continuous case
If X is a continuous random variable and $f(x)$ is the value of its probability density at x, the expected value (or expectation) of X is

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

(c) Expectation of a function of a random variable
Theorem. If X is a discrete random variable and $f(x)$ is the value of its probability distribution at x, the expected value of $g(X)$ is given by

$$E[g(X)] = \sum_x g(x) f(x).$$

Correspondingly, if X is a continuous random variable and $f(x)$ is the value of its probability density at x, the expected value $g(X)$ is

$$E[g(X)] = \int_{-\infty}^{\infty} g(x) f(x) dx.$$

Remarks:

i. If a and b are constants, then $E(aX + b) = aE(X) + b$.
ii. If $c_1, c_2, \ldots, c_n$ are constants, then $E[\sum_{i=1}^{n} c_i g_i(X)] = \sum_{i=1}^{n} c_i E[g_i(X)]$.

Definition: The variance of X is given by

$$Var(X) = E\{[X - E(X)]^2\} = E(X^2) - [E(X)]^2.$$

Example: Find $E[X]$ where X is the outcome when we roll a fair die.

Solution: Since $p(1) = p(2) = p(3) = p(4) = p(5) = p(6) = \frac{1}{6}$, we obtain

$$E[X] = 1(\frac{1}{6}) + 2(\frac{1}{6}) + 3(\frac{1}{6}) + 4(\frac{1}{6}) + 5(\frac{1}{6}) + 6(\frac{1}{6}) = \frac{7}{2}.$$

Example: Suppose X has the following probability distribution

$$f(0) = .2, \quad f(1) = .5, \quad f(2) = .3$$

Find $E(X^2)$.

Solution: $E(X^2) = 0^2 f(0) + 1^2 f(1) + 2^2 f(2) = 0^2(.2) + 1^2(.5) + 2^2(.3) = 1.7$.

Example: If $X \sim B(n,p)$, prove that

$$(1)\ E(X) = np, \quad (2)\ Var(X) = np(1-p).$$

Proof:
For (1):

$$\begin{aligned} E(X) &= \sum_{x=0}^{n} x \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} \\ &= \sum_{x=1}^{n} x \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} \\ &= np \sum_{x=1}^{n} \frac{(n-1)!}{(x-1)!(n-x)!} p^{x-1} (1-p)^{n-x}. \end{aligned}$$

Put $y = x - 1$. When $x = 1$, $y = 0$ and when $x = n$, $y = n - 1$. Therefore

$$\begin{aligned} E(X) &= np \sum_{y=0}^{n-1} \frac{(n-1)!}{y!(n-1-y)!} p^y (1-p)^{n-1-y} \\ &= np. \end{aligned}$$

For (2):
Note that $Var(X) = E(X^2) - [E(X)]^2 = E[X(X-1)] + E(X) - [E(X)]^2$ and

$$\begin{aligned} E[X(X-1)] &= \sum_{x=0}^{n} x(x-1) \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} \\ &= \sum_{x=2}^{n} x(x-1) \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} \\ &= n(n-1)p^2 \sum_{x=2}^{n} x \frac{(n-2)!}{(x-2)!(n-x)!} p^{x-2} (1-p)^{n-x}. \end{aligned}$$

Put $y = x - 2$. When $x = 2$, $y = 0$ and when $x = n$, $y = n - 2$. Therefore

$$\begin{aligned} E[X(X-1)] &= n(n-1)p^2 \sum_{y=0}^{n-2} \frac{(n-2)!}{y!(n-2-y)!} p^y (1-p)^{n-2-y} \\ &= n(n-1)p^2. \end{aligned}$$

Therefore

$$Var(X) = n(n-1)p^2 + np - (np)^2 = np - np^2 = np(1-p).$$

Example: If $X \sim Poisson(\lambda)$, prove that

$$(1)\ E(X) = \lambda, \quad (2)\ Var(X) = \lambda.$$

Proof:
For (1):

$$\begin{aligned} E(X) &= \sum_{x=0}^{\infty} x \frac{e^{-\lambda}\lambda^x}{x!} \\ &= \sum_{x=1}^{\infty} x \frac{e^{-\lambda}\lambda^x}{x!} \\ &= \lambda \sum_{x=1}^{\infty} \frac{e^{-\lambda}\lambda^{x-1}}{(x-1)!}. \end{aligned}$$

Put $y = x - 1$. When $x = 1$, $y = 0$ and when $x = \infty$, $y = \infty$. Therefore

$$\begin{aligned} E(X) &= \lambda \sum_{y=0}^{\infty} \frac{e^{-\lambda}\lambda^y}{y!} \\ &= \lambda. \end{aligned}$$

For (2):
Note that $Var(X) = E(X^2) - [E(X)]^2 = E[X(X-1)] + E(X) - [E(X)]^2$ and

$$\begin{aligned} E[X(X-1)] &= \sum_{x=0}^{\infty} x(x-1) \frac{e^{-\lambda}\lambda^x}{x!} \\ &= \sum_{x=2}^{\infty} x(x-1) \frac{e^{-\lambda}\lambda^x}{x!} \\ &= \lambda^2 \sum_{x=2}^{\infty} \frac{e^{-\lambda}\lambda^{x-2}}{(x-2)!}. \end{aligned}$$

Put $y = x - 2$. When $x = 2$, $y = 0$ and when $x = \infty$, $y = \infty$. Therefore

$$\begin{aligned} E[X(X-1)] &= \lambda^2 \sum_{y=0}^{\infty} \frac{e^{-\lambda}\lambda^y}{y!} \\ &= \lambda^2. \end{aligned}$$

Therefore

$$Var(X) = \lambda^2 + \lambda - \lambda^2 = \lambda.$$

(d) Jointly distributed random variables

For any two random variables X and Y, the joint cumulative probability distribution function of X and Y is defined by

$$F(x, y) = P(X \leq x, Y \leq y), \quad -\infty < x, \ y < \infty.$$

Moreover, if both X and Y are discrete, we define the joint probability mass function (joint p.m.f.) of X and Y by

$$f(x, y) = P(X = x, Y = y).$$

And we say that X and Y are jointly continuous if there exists a function $f(x, y)$, defined for all real x and y, having the property that for every sets A and B of real numbers

$$P(X \in A, Y \in B) = \int_B \int_A f(x, y) dx dy.$$

The function $f(x, y)$ is called the joint probability density function (or simply joint pdf) of X and Y.

If X and Y are discrete random variables and $f(x, y)$ is the value of their joint p.m.f. at (x, y), the function given by

$$g(x) = \sum_y f(x, y)$$

for each x of X is called the **marginal p.m.f.** of X. Correspondingly, the function given by

$$h(y) = \sum_x f(x, y)$$

for each y of Y is called the **marginal p.m.f.** of Y.

If X and Y are continuous random variables and $f(x, y)$ is the value of their joint p.d.f. at (x, y), the function given by

$$g(x) = \int_{-\infty}^{\infty} f(x, y) dy$$

for $-\infty < x < \infty$ is called the **marginal p.d.f.** of X. Correspondingly, the function given by

$$h(y) = \int_{-\infty}^{\infty} f(x, y) dx$$

for $-\infty < y < \infty$ is called the **marginal p.d.f.** of Y.

Theorem. If X and Y are discrete random variables and $f(x,y)$ is the value of their joint probability distribution at (x,y), the expected value of $g(X,Y)$ is

$$E[g(X,Y)] = \sum_x \sum_y g(x,y)f(x,y).$$

Correspondingly, if X and Y are continuous random variables and $f(x,y)$ is the value of their joint probability density at (x,y), the expected value of $g(X,Y)$ is

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

Theorem. If $c_1, c_2, \ldots$ and c_n are constants, then

$$E\ [\sum_{i=1}^{n} c_i g_i(X_1,\ldots,X_n)] = \sum_{i=1}^{n} c_i E[g_i(X_1,\ldots,X_n)].$$

Example: At a party N men throw their hats into the center of a room. The hats are mixed up and each man randomly selects one. Find the expected number of men that select their own hats.

Solution: Let X be the number of men that select their own hats and let

$$X_i = \begin{cases} 1, & if\ the\ i^{th}\ man\ selects\ his\ own\ hat \\ 0, & otherwise \end{cases}$$

Then $X = X_1 + X_2 + \ldots + X_N$.
Note that $P(X_i = 1) = 1/N$ and $P(X_i = 0) = (N-1)/N$. By definition

$$E(X_i) = 1[P(X_i = 1)] + 0[P(X_i = 0)] = \frac{1}{N}.$$

Therefore

$$E(X) = E(X_1) + E(X_2) + \ldots + E(X_N) = N(\frac{1}{N}) = 1.$$

(e) Independent random variables

The random variables X and Y are said to be independent if, for all a, b,

$$P(X \leq a, Y \leq b) = P(X \leq a)P(Y \leq b).$$

In term of the joint cdf F of X and Y , we say X and Y are independent if

$$F(a,b) = F_X(a)F_Y(b), \quad for\ all\ a, b,$$

where F_X and F_Y are the marginal cdf of X and Y respectively.

Moreover, when X and Y are discrete, X and Y are independent if

$$f(x, y) = f_X(x) f_Y(y),$$

where f, f_X and f_Y are the joint pmf of X and Y , the marginal pmf of X and marginal pmf of Y respectively.

And when X and Y are continuous, X and Y are independent if

$$f(x, y) = f_X(x) f_Y(y),$$

where f, f_X and f_Y are the joint pdf of X and Y , the marginal pdf of X and marginal pdf of Y respectively.

Proposition: If X and Y are independent, then for any functions h and g,

$$E[g(X)h(Y)] = E[g(X)][h(Y)].$$

The covariance of X and Y , denoted by $Cov(X, Y)$, is defined by

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

Example: Let X be a random variable with the following probability mass function and let $Y = X^2$:

x	-2	-1	1	2
Probability $f(x)$	.25	.25	.25	.25

i. Find the cumulative distribution function of X.
ii. Find the p.m.f. of Y.
iii. Find the joint p.m.f. of X and Y.
iv. Find $Cov(X, Y)$.
v. Is X and Y are independent? Why?

Solution:

i.

$$F_X(x) = P(X \leq x) = \begin{cases} 0 & if\ x < -2 \\ 0.25 & if\ -2 \leq x < -1 \\ 0.5 & if\ -1 \leq x < 1 \\ 0.75 & if\ 1 \leq x < 2 \\ 1 & if\ x \geq 2 \end{cases}$$

ii. The p.m.f. of Y is given by

$$f_Y(y) = P(Y = y) = \begin{cases} 0.5 & if\ y = 1 \\ 0.5 & if\ y = 4 \end{cases}$$

iii. The joint p.m.f. of X and Y is given by

$$f(x,y) = P(X = x, Y = y) = \begin{cases} 0.25 & if\ x = -2, y = 4 \\ 0.25 & if\ x = -1, y = 1 \\ 0.25 & if\ x = 1, y = 1 \\ 0.25 & if\ x = 2, y = 4 \end{cases}$$

iv. Note that
$E(X) = (-2)(0.25) + (-1)(0.25) + 1(0.25) + 2(0.25) = 0$,
$E(Y) = 1(0.5) + 4(0.5) = 2.5$ and
$E(XY) = (-2)(4)(0.25) + (-1)(1)(0.25) + (1)(1)(0.25) + (2)(4)(0.25) = 0$.
Therefore $Cov(X,Y) = E(XY) - E(X)E(Y) = 0$.

v. Note that $P(X = 1, Y = 4) = 0$ but $P(X = 1) = 0.25,\ P(Y = 4) = 0.5$.
Therefore $P(X = 1, Y = 4) \neq P(X = 1)P(Y = 4)$.
Hence X and Y are not independent.

Remarks:

i. $Var(X) = E(X^2) - [E(X)]^2$.
ii. $Var(aX + b) = a^2 Var(X)$.
iii. X and Y are independent implies $Cov(X,Y) = 0$.
iv. $Var(X + Y) = Var(X) + Var(Y) + 2Cov(X,Y)$.
v. If X and Y are independent, then $Var(X + Y) = Var(X) + Var(Y)$.
vi. If X and Y are independent having pdf $f(x)$ and $g(x)$ respectively, the pdf of the random variable $X + Y$ is given by

$$f_{X+Y}(a) = \int_{-\infty}^{\infty} f(a-y)g(y)dy.$$

Proof: Note that the joint cdf of $X + Y$ is given by

$$\begin{aligned} F_{X+Y}(a) &= P(X + Y \leq a) \\ &= \int\int_{x+y\leq a} f(x)g(y)dxdy \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{a-y} f(x)g(y)dxdy \\ &= \int_{-\infty}^{\infty}[\int_{-\infty}^{a-y} f(x)dx]g(y)dy \\ &= \int_{-\infty}^{\infty} F_X(a-y)g(y)dy \end{aligned}$$

where F_X is the cdf of X.

Therefore, the joint pdf of $X+Y$ is given by

$$\begin{aligned} f_{X+Y}(a) &= \frac{d}{da}F_{X+Y}(a) \\ &= \frac{d}{da}\int_{-\infty}^{\infty} F_X(a-y)g(y)dy \\ &= \int_{-\infty}^{\infty} \frac{d}{da}F_X(a-y)g(y)dy \\ &= \int_{-\infty}^{\infty} f(a-y)g(y)dy \end{aligned}$$

(f) The moment-generating function of a random variable X, where it exists, is given by

$$M_X(t) = E(e^{tX}).$$

Theorem.

$$\frac{d^r M_X(t)}{dt^r}|_{t=0} = E(X^r)$$

Remark: The moment generating function uniquely determines the distributions.

Example: If $X \sim B(n,p)$, prove that $M_X(t) = [1+p(e^t-1)]^n$.
Proof:

$$\begin{aligned} M_X(t) &= E(e^{tX}) \\ &= \sum_{x=0}^{n} e^{tX}\frac{n!}{x!(n-x)!}p^x(1-p)^{n-x} \\ &= \sum_{x=0}^{n} \frac{n!}{x!(n-x)!}(e^t p)^x(1-p)^{n-x} \\ &= [(e^t p)+(1-p)]^n \\ &= [1+p(e^t-1)]^n \end{aligned}$$

Example: If $X \sim Poisson(\lambda)$, prove that $M_X(t) = e^{\lambda(e^t-1)}$.
Proof:

$$\begin{aligned} M_X(t) &= E(e^{tX}) \\ &= \sum_{x=0}^{n} e^{tX}\frac{e^{-\lambda}\lambda^x}{x!} \\ &= e^{-\lambda}\sum_{x=0}^{n} \frac{(\lambda e^t)^x}{x!} \\ &= e^{-\lambda}e^{\lambda e^t} \\ &= e^{\lambda(e^t-1)} \end{aligned}$$

(g) Limit Theorems

i. (Markov's inequality)
If X is a random variable which takes only nonnegative values, then for any value $a > 0$,

$$P(X \geq a) \leq \frac{E(X)}{a}.$$

Proof: (For continuous case)

Let $f(x)$ be the pdf of X.

$$\begin{aligned} E(X) &= \int_0^\infty x f(x) dx \\ &= \int_0^a x f(x) dx + \int_a^\infty x f(x) dx \\ &\geq \int_a^\infty x f(x) dx \\ &\geq \int_a^\infty a f(x) dx \\ &= a \int_a^\infty f(x) dx \\ &= a P(X \geq a) \end{aligned}$$

ii. (Chebyshev's Theorem)
If μ and σ are the mean and the standard deviation of a random variable X, then for any positive constant k

$$P(|X - \mu| \geq k) \leq \frac{\sigma^2}{k^2}.$$

Proof: Note that $(X-\mu)^2$ is a nonnegative random variable. Apply Markov's inequality (with $a = k^2$), we have

$$P[(X-\mu)^2 \geq k^2] \leq \frac{E(X-\mu)^2}{k^2}.$$

Since $(X-\mu)^2 \geq k^2$ is equivalent to $|X-\mu| \geq k$. Thus we have

$$P(|X - \mu| \geq k) \leq \frac{E(X-\mu)^2}{k^2} = \frac{\sigma^2}{k^2}.$$

Definition: Converge in probability
X_n converge to X in probability, if for any $\epsilon > 0$,

$$\begin{aligned} \lim_{n\to\infty} P(|X_n - X| &\geq \epsilon) = 0 \qquad \textit{or} \\ \lim_{n\to\infty} P(|X_n - X| &< \epsilon) = 1 \end{aligned}$$

Denote by $X_n \overset{Pr.}{\to} X$ (weakly converge).

Definition: Converge in distribution
X_n converge in distribution to X with distribution function $F(x)$, if

$$\lim_{n\to\infty} F_n(x) = F(x)$$

for all points x at which $F(x)$ is continuous, where $F_n(x)$ is the distribution of X_n. Denote by $X_n \xrightarrow{d} X$ and say X_n has a limiting distribution with distribution of $F(x)$.

iii. **Theorem : Weak Law of Large Number(WLLN)**
Assume that $X_1, \ldots, X_n$ is a sequence of independent and identically distributed (iid) r.vs. and $E(X_n) = \mu < \infty$, $Var(X_n) = \sigma^2 < \infty$ exist, then

$$\bar{X_n} = \frac{sum_{i=1}^{n} X_i}{n} \xrightarrow{Pr.} \mu$$

iv. **Theorem : Central Limit Theorem**
Let $X_1, \ldots, X_n$ be a r.s. from a distribution that has mean $\mu < \infty$ and variance $\sigma^2 < \infty$ exist, then the r.v.

$$Y_n = \frac{(\sum_{i=1}^{n} X_i - n\mu)}{\sqrt{n}\sigma} = \frac{\sqrt{n}(\bar{X_n} - \mu)}{\sigma}$$

has a limiting distribution $N(0,1)$, i.e. $Y_n \xrightarrow{d.} Z \sim N(0,1)$.

5. Stochastic Processes

A stochastic process $\{X(t),\ t \in T\}$ is a collection of random variables. The index t is often interpreted as time and $X(t)$ as the state of the process at time t. The set T is called the index set of the process. When T is countable, the stochastic process is said to be a discrete-time process. When T is an interval of the real line, the stochastic process is said to be a continuous-time process.

Some basic rules of Differentiation

1. $\frac{d}{dx}C = 0$, C is a constant

2. $\frac{d}{dx}x^n = nx^{n-1}$

3. $\frac{d}{dx}(u+v) = \frac{du}{dx} + \frac{dv}{dx}$

4. $\frac{d}{dx}uv = u\frac{dv}{dx} + v\frac{du}{dx}$

5. $\frac{d}{dx}(\frac{u}{v}) = \frac{v\frac{du}{dx} - u\frac{dv}{dx}}{v^2}$

6. $\frac{d}{dx}e^x = e^x$

7. $\frac{d}{dx}ln\ x = \frac{1}{x}$

8. Chain rule:
$$\frac{dy}{dx} = \frac{dy}{dw}\frac{dw}{dx}$$

Some basic rules of Integration

1. $\int_a^b kf(x)dx = k\int_a^b f(x)dx$

2. $\int_a^b [f(x)+g(x)]dx = \int_a^b f(x)dx + \int_a^b g(x)dx$

3. $\int_a^b x^n dx = \frac{b^{n+1}-a^{n+1}}{n+1}$

4. $\int_a^b e^x dx = e^b - e^a$

5. $\int_a^b \frac{dx}{x} = ln|b| - ln|a|$

6. Integration by parts:
$$\int_a^b f(x)dg(x) = [f(b)g(b) - f(a)g(a)] - \int_a^b g(x)df(x)$$

More examples:

Example 1

Letting X denote the random variable that is defined as the sum of two fair dice; then

$$
\begin{aligned}
P\{X=2\} &= P\{(1,1)\} = \frac{1}{36}, \\
P\{X=3\} &= P\{(1,2),(2,1)\} = \frac{2}{36}, \\
P\{X=4\} &= P\{(1,3),(2,2),(3,1)\} = \frac{3}{36}, \\
P\{X=5\} &= P\{(1,4),(2,3),(3,2),(4,1)\} = \frac{4}{36}, \\
P\{X=6\} &= P\{(1,5),(2,4),(3,3),(4,2),(5,1)\} = \frac{5}{36}, \\
P\{X=7\} &= P\{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\} = \frac{6}{36}, \\
P\{X=8\} &= P\{(2,6),(3,5),(4,4),(5,3),(6,2)\} = \frac{5}{36}, \\
P\{X=9\} &= P\{(3,6),(4,5),(5,4),(6,3)\} = \frac{4}{36}, \\
P\{X=10\} &= P\{(4,6),(5,5),(6,4)\} = \frac{3}{36}, \\
P\{X=11\} &= P\{(5,6),(6,5)\} = \frac{2}{36}, \\
P\{X=12\} &= P\{(6,6)\} = \frac{1}{36}
\end{aligned}
$$

p.m.f. of X is given by

$$f(x) = P(X = x), \ \ x = 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12$$

Example 2

Four fair coins are flipped. If the outcomes are assumed independent, what is the probability that two heads and two tails are obtained?

Solution

Letting X equal the number of heads ("successes") that appear, then X is a binomial random variable with parameters $(n = 4, p = \frac{1}{2})$. Hence,

$$P(X = 2) = \left(\begin{array}{c} 4 \\ 2 \end{array} \right) (\frac{1}{2})^2 (\frac{1}{2})^2 = \frac{3}{8}$$

Example 3

(Expectation of an Exponential Random Variable): Let X be exponential distributed with p.d.f

$$f(x) = \lambda e^{-\lambda x},\ x > 0$$

Calculate E(X).

Solution

$$E(X) = \int_0^\infty xf(x)dx = \int_0^\infty x\lambda e^{-\lambda x}dx = -\int_0^\infty xe^{-\lambda x}d(-\lambda x) = -\int_0^\infty xde^{-\lambda x}$$

Integrating by parts yields

$$\begin{aligned} E(X) &= -xe^{-\lambda x}\Big|_0^\infty + \int_0^\infty e^{-\lambda x}dx \\ &= 0 - \frac{1}{\lambda}\int_0^\infty e^{-\lambda x}d(-\lambda x) \\ &= -\frac{e^{-\lambda x}}{\lambda}\Big|_0^\infty \\ &= \frac{1}{\lambda}. \end{aligned}$$

Example 4

Suppose there are 25 different types of coupons and suppose that each time one obtains a coupon, it is equally likely to be any one of the 25 types. Compute the expected number of different types that are contained in a set of 10 coupons.

Solution

Let X denote the number of different types in the set of 10 coupons. We compute $E(X)$ by using the representation

$$X = X_1 + \ldots + X_{25}$$

where

$$X_i = \begin{cases} 1, & \text{if at least one type } i \text{ coupon is in the set of 10} \\ 0, & \text{otherwise} \end{cases}$$

Now,

$$\begin{aligned} E(X_i) &= P(X_i = 1) \\ &= P\{\text{at least one type } i \text{ coupon is in the set of 10}\} \\ &= 1 - P\{\text{no type } i \text{ coupons are in the set of 10}\} \\ &= 1 - (\frac{24}{25})^{10} \end{aligned}$$

when the last equality follows since each of the 10 coupons will (independently) not be a type i with probability $\frac{24}{25}$. Hence,

$$E(X) = E(X_1) + \ldots + E(X_{25}) = 25[1 - (\frac{24}{25})^{10}].$$

Example 5

Suppose we know that the number of items produced in a factory during a week is a random variable with mean 500.
(a) What can be said about the probability that this week's production will be at least 1000?
(b) If the variance of a week's production is known to be equal 100, then what can be said about the probability that this week's production will be between 400 and 600?

Solution

Let X be the number of items that will be produced in a week.

1. By Markov's inequality,

$$P(X \geq 1000) \leq \frac{E(X)}{1000} = \frac{500}{1000} = \frac{1}{2}$$

2. By Chebyshev's inequality,

$$P(|X - 500| \geq 100) \leq \frac{\sigma^2}{(100)^2} = \frac{1}{100}$$

   Hence,

$$P(|X - 500| \leq 100) \geq 1 - \frac{1}{100} = \frac{99}{100}$$

   and so the probability that this week's production will be between 400 and 600, is at least 0.99.