

MA203

Repeated trials

Repeated trials:

Given two experiments i.e., rolling a fair die and tossing a fair coin; find the probability that we get 2 on die and head on coin.

$$P(2 \text{ on die}) = 1/6$$

$$P(\text{head on coin}) = 1/2$$

Assuming statistically independent, $P(2 \text{ on die and head on coin}) = P(2 \text{ on die}) \times P(\text{head on coin}) = 1/12$

Other way,

$\{1,2,3,4,5,6\} \{H,T\}$

Considering both as a single experiment, sample space is:

$\{1H,2H,3H,4H,5H,6H,1T,2T,3T,4T,5T,6T\}$

$$P(2 \text{ on die and head on coin}) = 1/12$$

Cartesian product: given 2 sets S_1 and S_2 , product of the sets is a set S whose elements are all such pairs of elements of S_1 and S_2 .

$\{H,T\} \{H,T\} \{HH,HT,TH,TT\}$

Bernoulli Trials:

Out of a set of n distinct objects, if $k < n$ objects are taken out of n at a time, the total combinations are

$$\frac{n!}{(n-k)!k!} = \binom{n}{k} \quad (\text{orders are not considered})$$

Consider an experiment S and an event A with $P(A)=p$, $P(A^c)=q$ and $p+q=1$

Repeat the experiment n times and the resulting product space is $S_n = S \times S \times \dots \times S$ (n times)

Define $p_n(k) = P(A \text{ occurs } k \text{ times in any order})$

$P(A \text{ occurs } k \text{ times in a specific order}) = p^k q^{n-k}$

$$p_n(k) = \binom{n}{k} p^k q^{n-k} \quad \text{Fundamental theorem}$$

Considering event A as success and A^c as failure, $p_n(k)$ gives prob of k successes in n independent trials.

Eg. 1) A pair of dice is rolled n times. i) Find the probability that 7 will not show at all. ii) Find the probability of double six at least once.

Ans. Sample space of single roll of 2 dice consists of 36 elements.

i) $A = \{\text{seven}\} = \{(1, 6), (2, 5), (3, 4), (4, 3), (5, 2), (6, 1)\}$

$$P(A) = 6/36 = 1/6$$

$$P(A^c) = 5/6$$

$$P(7 \text{ will not show at all}) = \binom{n}{0} (1/6)^0 (5/6)^{n-0}$$

ii) X : double six at least once

$$P(X) = 1 - P(X^c) = 1 - (35/36)^n$$

Suppose we are interested in no. of throws required to ensure a 50% success of obtaining double six at least once.

$$\text{So, } 1 - (35/36)^n > 0.5$$

$$\text{or, } n = 24.605$$

Historical importance: one of the first problems solved by Pascal and correctly interpret gambler's choice.

For a fixed n , consider $p_n(k)$ as a function of k

As k increases, $p_n(k)$ increases reaching a maximum for $k=k_{\max}$

$$\frac{p_n(k-1)}{p_n(k)} = \frac{\binom{n}{k-1} p^{k-1} q^{n-k+1}}{\binom{n}{k} p^k q^{n-k}} = \frac{kq}{(n-k+1)p} < 1$$

if $k < (n+1)p$, $p_n(k-1) < p_n(k)$

If $k > (n+1)p$, then $p_n(k)$ is decreasing

Therefore, $p_n(k)$ is maximum for $k_{\max} = [(n+1)p]$

If $k_1 = (n+1)p$ is an integer, then $\frac{p_n(k-1)}{p_n(k)} = 1$

So, $p_n(k)$ has maximum values for $k=k_1$ and $k=k_1-1$

Most likely number of success out of n independent trials

Example 2. $n=10$, $p=1/3$, calculate most likely number of successes

Example 3. In New York state lottery, the player picks 6 numbers from a sequence of 1 through 51. At a lottery drawing, 6 balls are drawn at random from a box of 51 balls numbered 1 through 51. What is the prob that a player has k matches?

$$P(k \text{ matches}) = \frac{\binom{6}{k} \binom{51-6}{6-k}}{\binom{51}{6}}$$

For perfect match, k=6 and prob is 1/18009460

Random variable:

Finite, single-valued function which maps the function into sample space.

For every outcome a of sample space, we assign a number X(a) and this function is called RV.

Eg. S={HH,HT,TH,TT}

Function: number of heads

Outcomes of S (a)	HH	HT	TH	TT
RV X(a)	2	1	1	0

Consider, $S=\{1, 2, 3, 4, 5, 6\}$ of random experiment of rolling a fair die
 Let, function is $\{(\text{number of points on top less } 3)^2\}$

Outcomes of S (a)	1	2	3	4	5	6
RV X(a)	4	1	0	1	4	9

Two types: continuous and discrete

Since, defined over a sample space of a random experiment, each value is associated with a probability.

$\{X \leq x\}$ means a subset of S consisting of all outcomes a such that $X(a) \leq x$

$P\{X \leq x\} = F_X(x)$, distribution function or cumulative distribution function
 let, tossing of 2 coins

Outcomes of S (a)	HH	HT	TH	TT
RV X(a)	2	1	1	0

$$F(x) = P(X \leq x)$$

$$F(0) = P(X \leq 0) =$$

$$F(1) = P(X \leq 1) =$$

$$F(2) = P(X \leq 2) =$$

Properties:

1. $F(+\infty)=1, F(-\infty)=0$

$$F(+\infty)=P\{X\leq+\infty\}=P(S)=1$$

$$F(-\infty)=P\{X\leq-\infty\}=0$$

2. non-decreasing function of x i.e., if $x_1 < x_2$ then $F(x_1) \leq F(x_2)$

$\{X\leq x_1\}$ is a subset of $\{X\leq x_2\}$ if $x_1 < x_2$

So $P\{X\leq x_1\} \leq P\{X\leq x_2\}$ $F(x_1) \leq F(x_2)$

3. If $F(x_0)=0$ then $F(x)=0$ for every $x\leq x_0$

4. $P\{X>x\}=1-F(x)$

$\{X\leq x\} \cup \{X>x\}=S$ and they are mutually exclusive So $P\{X>x\}=1-F(x)$

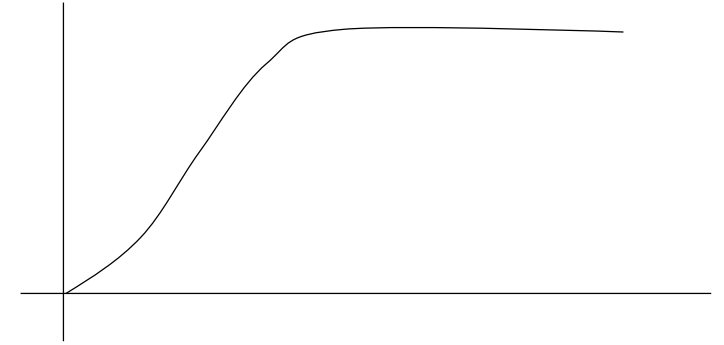
5. $P\{x_1<X\leq x_2\}=F(x_2)-F(x_1)$

$\{X\leq x_2\}=\{X\leq x_1\} \cup \{x_1<X\leq x_2\}$ [mutually exclusive]

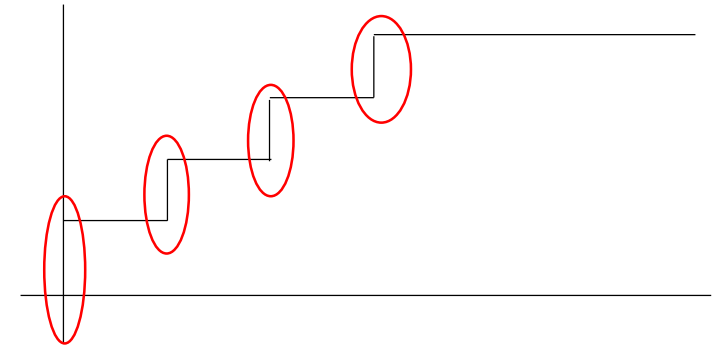
$$P\{X\leq x_2\}=P\{X\leq x_1\} + P\{x_1<X\leq x_2\}$$

$$P\{x_1<X\leq x_2\}=F(x_2)-F(x_1)$$

Continuous RV : if $F(x)$ is continuous



Discrete RV : if $F(x)$ is constant except for a finite number of jump discontinuities (piecewise constant/step)



Probability density function:

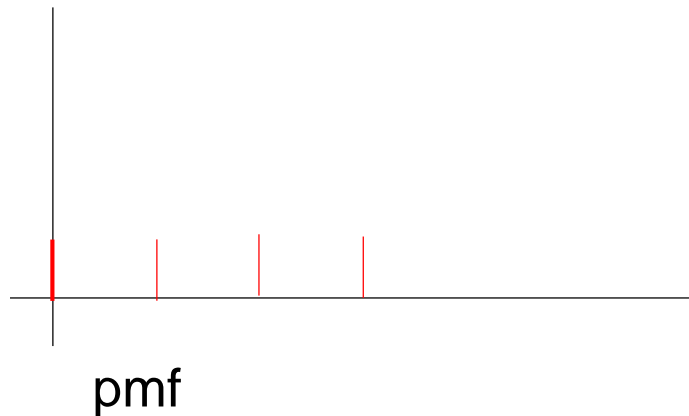
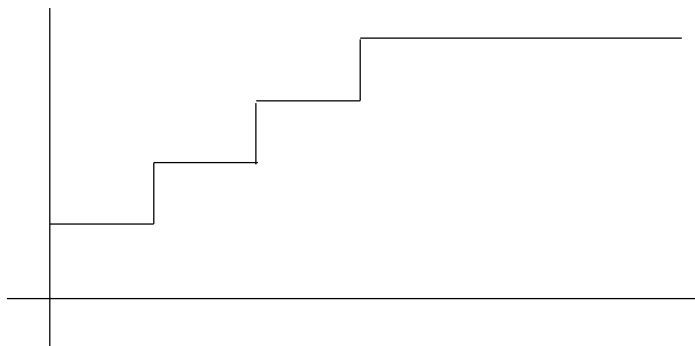
Derivative of distribution function $F(x)$

$$f_X(x) \triangleq \frac{dF_X(x)}{dx} = \lim_{\Delta x \rightarrow 0} \frac{F_X(x + \Delta x) - F_X(x)}{\Delta x} \geq 0$$

As $F(x)$ is monotonically increasing, so $f(x) \geq 0$ for all x

If X is continuous RV then $f(x)$ will be continuous

If X is discrete RV, then pdf has the general form $f_X(x) = \sum_i p_i \delta(x - x_i)$, x_i : jump discontinuities of $F(x)$
probability mass function (pmf)



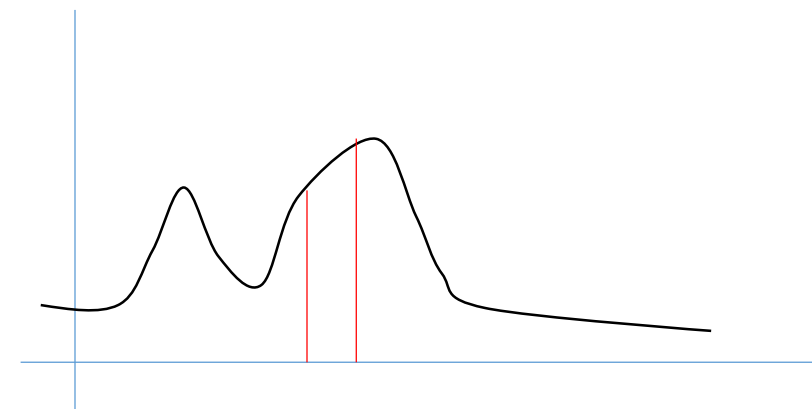
From definition,

$$F_X(x) = \int_{-\infty}^x f_X(u) du$$

Also, $\int_{-\infty}^{\infty} f_X(x) dx = 1$, so density function

$$P\{x_1 < X \leq x_2\} = F_X(x_2) - F_X(x_1) = \int_{x_1}^{x_2} f_X(x) dx$$

Area under the curve in the interval



Probability that a continuous RV takes any specified value is 0

Example 1. X is an RV with distribution function

$$F(x) = \begin{cases} 0, & x \leq 0 \\ x, & 0 < x \leq 1 \\ 1, & x > 1 \end{cases}$$

$$f(x) = F'(x) = \begin{cases} 0, & x < 0 \text{ or } x > 1 \\ 1, & 0 < x < 1 \end{cases}$$

$$P\{0.4 < X \leq 0.6\} = ?$$

2. X have the triangular pdf $f(x) = \begin{cases} x, & 0 < x \leq 1 \\ 2-x, & 1 \leq x \leq 2 \\ 0, & \text{otherwise} \end{cases}$

Check $f(x)$ is pdf

$$P\{0.3 < X \leq 1.5\} = ?$$

$$F(x) = 0, \quad x \leq 0$$

$$F(x) = \int_0^x t \, dt = \frac{x^2}{2}, \quad 0 < x \leq 1$$

$$f(x) = \int_0^1 t \, dt + \int_1^x (2-t) \, dt, \quad 1 \leq x \leq 2$$

$$= \frac{1}{2} + 2x - 2 - \frac{x^2}{2} + \frac{1}{2} = 2x - \frac{x^2}{2} - 1$$

$$F(x) = \int_0^1 t \, dt + \int_1^2 (2-t) \, dt, \quad x \geq 2$$

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

$$= 1$$

now,

$$P\{0.3 < X \leq 1.5\} = P(X \leq 1.5) - P(X \leq 0.3)$$

$$= \int_0^{1.5} f(x) \, dx$$

$$= \int_0^1 x \, dx + \int_1^{1.5} (2-x) \, dx$$

$$= \frac{1 - 0.09}{2} + 2(1.5 - 1) - \frac{2.25 - 1}{2}$$

$$= 0.8705 \approx 0.83$$

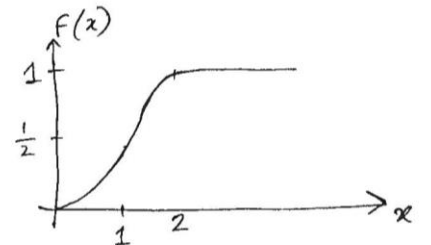
other way,

$$P\{0.3 < X \leq 1.5\} = F(1.5) - F(0.3)$$

$$= 2.15 - \frac{1.5^2}{2} - 1 - \frac{0.3^2}{2}$$

$$= 0.83$$

If we plot $F(x)$, it would be



For discrete RV, the collection of numbers $\{p_i\}$ satisfying $P(X=x_i)=p_i \geq 0$ for all i and $\sum_{i=1}^{\infty} p_i = 1$ is called pmf of RV X

DF is $F(x)=P(X \leq x)=\sum (p_i)$ for $x_i \leq x$

Example: A box contains good and defective items. If an item drawn is good, we assign the number 1 to the drawing and otherwise the number is 0. Let probability of drawing a good item at random is p .

$$\left. \begin{array}{l} P(X=1)=p \\ P(X=0)=1-p \end{array} \right\} \text{ pmf}$$

$$F(x)=P(X \leq x)= \left\{ \begin{array}{ll} 0 & , x < 0 \\ 1-p & , 0 \leq x < 1 \\ 1 & , x \geq 1 \end{array} \right\} \text{ DF}$$

Theoretical distribution: when a random experiment is theoretically assumed to serve as a model, the probability distribution of the RV associated with the random experiment is generally known as theoretical distribution.

Expectation – mean, variance, moments

Let, a discrete RV X assumes the values x_1, x_2, \dots, x_n with probabilities p_1, p_2, \dots, p_n respectively.

Expectation or expected value of X is

$$E(X) = \sum_{i=1}^n p_i x_i \quad \text{provided } \sum_{i=1}^n p_i |x_i| < \infty$$

$$\text{Similarly, } E(X^2) = \sum_{i=1}^n p_i x_i^2$$

Say, $g(X)$ is a function of RV X

$$E[g(X)] = \sum_{i=1}^n p_i g(x_i)$$

Expectation of a constant k is the constant k itself.

$$E(k) = \sum (k \cdot p_i) = k \quad [\text{since, } \sum (p_i) = 1]$$

Mean – of a RV X is $E(X) = \mu$

$$\text{Variance} - \sigma^2 = E(X - \mu)^2 = E(X^2 - 2X\mu + \mu^2) = E(X^2) - \mu^2$$

Standard deviation (σ) is the positive square root of variance.

Moments:

r-th moment about A is

$$m'_r = E(X - A)^r = \sum_{i=1}^n p_i (x_i - A)^r$$

r-th raw moment is

$$\mu'_r = E(X)^r = \sum_{i=1}^n p_i x_i^r$$

r-th central moment is

$$\mu_r = E(X - \mu)^r = \sum_{i=1}^n p_i (x_i - \mu)^r$$

Where $\mu = E(X)$

As per definition, $\mu'_0 = \mu_0 = 1$, $\mu'_1 = E(X) = \mu$, $\mu_1 = 0$, $\mu_2 = \sigma^2$

Central moments can be obtained from non-central moments as

$$\mu_2 = E(X - \mu)^2 = E(X)^2 - (E(X))^2 = \mu'_2 - (\mu'_1)^2$$

$$\mu_3 = \mu'_3 - 3\mu'_2\mu'_1 + 2(\mu'_1)^3$$

$$\mu_4 = \mu'_4 - 4\mu'_3\mu'_1 + 6\mu'_2(\mu'_1)^2 - 3(\mu'_1)^4$$

Expectation – mean, variance, moments

Let, a continuous RV X assumes the values between $-\infty$ to $+\infty$ with pdf $f(x)$

Expectation or expected value of X is

$$E(X) = \int_{-\infty}^{\infty} xf(x)dx \quad \text{provided } \int_{-\infty}^{\infty} |x|f(x)dx < \infty$$

Mean – of a RV X is $E(X) = \int_{-\infty}^{\infty} xf(x)dx = \mu$

Variance – $\sigma^2 = E(X - \mu)^2 = \int_{-\infty}^{\infty} (x - \mu)^2 f(x)dx = E(X)^2 - \mu^2$

Standard deviation (σ) is the positive square root of variance.

Moments:

r-th moment about A is

$$m'_r = E(X - A)^r = \int_{-\infty}^{\infty} (x - A)^r f(x)dx$$

r-th raw moment is

$$\mu'_r = E(X)^r = \int_{-\infty}^{\infty} (x)^r f(x)dx$$

r-th central moment is

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

Moment generating functions:

Consider X is an RV.

$M(s) = E(e^{sX})$ exists provided $\int_{-\infty}^{\infty} |e^{sx}| f(x) dx < \infty$

$M(s)$ is called moment generating function (MGF) of RV X

MGF uniquely determines the corresponding distribution function (DF) and if MGF exists, MGF is unique for an RV.

$$M(s) = \int_{-\infty}^{\infty} e^{sx} f(x) dx$$

Differentiating by n times,

$$M^n(s) = \int_{-\infty}^{\infty} x^n e^{sx} f(x) dx = E[X^n e^{sX}]$$

For $s=0$,

$$M^n(0) = E[X^n]$$

Which is n-th order raw moment.

Hence, MGF

$$\begin{aligned} M(s) &= \int_{-\infty}^{\infty} \left(1 + sx + \frac{s^2 x^2}{2!} + \dots \right) f(x) dx \\ M(s) &= 1 + s \int_{-\infty}^{\infty} x f(x) dx + \frac{s^2}{2!} \int_{-\infty}^{\infty} x^2 f(x) dx + \dots \\ M'(0) &= \int_{-\infty}^{\infty} x f(x) dx = E(x) \end{aligned}$$

$$M'(0) = E(X) = \mu$$

$$M''(0) = E(X^2) = \sigma^2 + (M'(0))^2$$

Similarly, third order moment : measure of skewness

Fourth order moment : measure of kurtosis

Example-4 Let, X have the PDF as

$$f(x) = \begin{cases} \frac{1}{2} e^{-x/2}, & x > 0 \\ 0, & \text{otherwise} \end{cases}$$

$$\begin{aligned} \therefore \text{MGF is } M(s) &= \int_0^{\infty} e^{sx} \frac{1}{2} e^{-x/2} dx \\ &= \int_0^{\infty} \frac{1}{2} e^{(s-\frac{1}{2})x} dx = \left[\frac{1}{2} \frac{e^{(s-\frac{1}{2})x}}{s-\frac{1}{2}} \right]_0^{\infty} \\ &= \frac{1}{1-2s}, \quad s < \frac{1}{2} \end{aligned}$$

$$\begin{aligned} M'(s) &= \frac{d}{ds} \left(\frac{1}{1-2s} \right) = \frac{-1(-2)}{(1-2s)^2} = \frac{2}{(1-2s)^2} \\ M''(s) &= \frac{d^2}{ds^2} \left(\frac{1}{1-2s} \right) = \frac{-2 \cdot 2(1-2s) \cdot (-2)}{(1-2s)^4} = \frac{8}{(1-2s)^3} \end{aligned}$$

$$\begin{aligned} \text{now, } M'(0) &= E(X) = 2 & \therefore \mu &= 2 \\ M''(0) &= E(X^2) = 8 & \sigma^2 &= 8 - 2^2 = 4 \end{aligned}$$

now, to check these results,

$$\begin{aligned} E(X) &= \int_0^{\infty} x \cdot \frac{1}{2} e^{-x/2} dx = \frac{1}{2} \int_0^{\infty} x e^{-x/2} dx \\ &= \frac{1}{2} \left[x \int e^{-x/2} dx - \int (1 \cdot \int e^{-x/2} dx) dx \right]_0^{\infty} \\ &= \frac{1}{2} \left[-2x e^{-x/2} - 4e^{-x/2} \right]_0^{\infty} \\ &= \frac{1}{2} \cdot 4 = 2 \end{aligned}$$

$$\begin{aligned}
 E(X^2) &= \int_0^{\infty} x^2 \frac{1}{2} e^{-x/2} dx = \frac{1}{2} \left[x^2 (-2e^{-x/2}) - \int 2x (-2e^{-x/2}) dx \right] \\
 &= \frac{1}{2} \left[-2x^2 e^{-x/2} + 4 \int x e^{-x/2} dx \right]_0^{\infty} \\
 &= \left[-x^2 e^{-x/2} \right]_0^{\infty} + 2 \cdot 4
 \end{aligned}$$

(from $E(X)$ solution)

$$\therefore \sigma^2 = 8 - 2^2 = 4 \quad (\text{hence verified})$$

... on x with

Suppose, X is a discrete RV taking values x_i with probability p_i

$$M(s) = \sum_i p_i e^{sx_i}$$

However, if X takes only integer values, then Z transform is preferable to define MGF

$$\Gamma(z) = E(z^X) = \sum_{n=-\infty}^{\infty} P(X = n)z^n = \sum_{n=-\infty}^{\infty} p_n z^n$$

Differentiating it k times,

$$\Gamma^{(k)}(z) = E\{X(X-1) \dots (X-k+1)z^{X-k}\}$$

With $z=1$,

$$\Gamma^{(k)}(1) = E\{X(X-1) \dots (X-k+1)\}$$

So,

$$\begin{aligned}\Gamma'(1) &= E(X) \\ \Gamma''(1) &= E(X(X-1)) = E(X^2 - X) = E(X^2) - E(X)\end{aligned}$$

Example 5. An RV X takes values 0 and 1 with $P(X=1)=p$ and $P(X=0)=q$. Find MGF.

ans. $\Gamma(z) = E(z^X) = P(X=1) \cdot z^1 + P(X=0) \cdot z^0$
 $= pz + q$

$$\therefore \Gamma'(1) = \frac{d}{dz} (pz + q) \Big|_{z=1} = p$$

$$\therefore E(X) = p$$

$$\Gamma''(1) = \frac{d^2}{dz^2} (pz + q) \Big|_{z=1} = 0$$

$$\therefore E(X^2) - E(X) = 0$$

$$\therefore E(X^2) = E(X) = p$$

$$\therefore \text{variance} = E(X^2) - (E(X))^2 = p - p^2 = p(1-p)$$

$$= pq$$

now to check,

$$E(X) = 1 \cdot p + 0 \cdot q = p$$

$$E(X^2) = 1^2 \cdot p + 0^2 \cdot q = p$$

$$\therefore \text{Variance} = E(X^2) - E^2(X) = p - p^2 = p(1-p) = pq$$

(hence verified)

Characteristic function:

$$\Phi_X(\omega) = \int_{-\infty}^{\infty} f(x)e^{j\omega x} dx$$

Since $f(x) \geq 0$,

$\Phi(\omega)$ is maximum at origin

$$|\Phi_X(\omega)| \leq \Phi_X(0) = 1$$

If $j\omega$ is substituted by s , resulting integral gives MGF

Second characteristic function of X : $\Psi_X(\omega) = \ln \Phi_X(\omega)$

If $Y = aX + b$, then

$$\Phi_Y(\omega) = \int_{-\infty}^{\infty} f(x)e^{j\omega(ax+b)} dx = e^{j\omega b} \int_{-\infty}^{\infty} f(x)e^{j\omega ax} dx = e^{j\omega b} \Phi_X(a\omega)$$

Characteristic function is the Fourier transform of $f(x)$.

So, $f(x)$ can be retrieved by using inverse transform

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

Characteristic function and MGF, both represent the distribution of RV

From MGF, using inverse transform $f(x)$ can be calculated and by integrating $f(x)$, DF is calculated uniquely.

Measure of variance of an RV near its mean μ is its variance σ^2 . The probability that X is outside an arbitrary interval $(\mu-\epsilon, \mu+\epsilon)$ is negligible if the ratio σ/ϵ is sufficiently small. This result is fundamental and known as Chebyshev inequality; i.e., for any $\epsilon > 0$

$$P\{|X - \mu| \geq \epsilon\} \leq \frac{\sigma^2}{\epsilon^2}$$

Proof:

$$P\{|X - \mu| \geq \epsilon\} = \int_{-\infty}^{-\mu-\epsilon} f(x)dx + \int_{\mu+\epsilon}^{\infty} f(x)dx = \int_{|x-\mu| \geq \epsilon} f(x)dx$$
$$\sigma^2 = \int_{-\infty}^{\infty} (x - \mu)^2 f(x)dx \geq \int_{|x-\mu| \geq \epsilon} (x - \mu)^2 f(x)dx \geq \epsilon^2 \int_{|x-\mu| \geq \epsilon} f(x)dx = \epsilon^2 P\{|X - \mu| \geq \epsilon\}$$

Note: 1. if variance is 0, the prob is 0 for any ϵ . Hence $X=\mu$ with probability 1.

2. For specific densities, the bound is too high. Eg. If X is Normal RV, then $P\{|X-\mu| \geq 3\sigma\} = 0.0027$ but this inequality gives $\leq 1/9$. the bound can be reduced by incorporating various assumptions of $f(x)$ (leads to Chernoff bound)

3. significance: it holds true for any pdf and therefore can be used when pdf is unknown.

Markov Inequality:

If $f(x)=0$ for $x<0$, then for any $\alpha>0$,

$$P\{X \geq \alpha\} \leq \mu/\alpha$$

Proof:

$$E(X) = \mu = \int_0^{\infty} xf(x)dx \geq \int_{\alpha}^{\infty} xf(x)dx \geq \alpha \int_{\alpha}^{\infty} f(x)dx = \alpha P\{X \geq \alpha\}$$
$$P\{X \geq \alpha\} \leq \frac{\mu}{\alpha}$$