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 on die)=1/6
P(head on coin)=1/2
Assuming statistically independent, P(2 on die and head on coin)= P(2 on die)xP(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 on die and head on coin)=1/12

Cartesian product: given 2 sets S1 and S2, product of the sets is a set S whose elements are all such pairs of elements of S1 and S2.

{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

$\frac{n!}{(n-k)!k!} = \binom{n}{k}$ (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 $S_n=S \times S \times \ldots \times S$ (n times)

Define $p_n(k)$=P(A occurs k times in any order)

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

$p_n(k)=\binom{n}{k}p^k q^{n-k}$ 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={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 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.

So, $1-(35/36)^n > 0.5$

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=kmax

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

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 kmax=[(n+1)p]
If k1=(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=k1 and k=k1-1

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 matches) $= \dfrac{\binom{6}{k}\binom{51-6}{6-k}}{\binom{51}{6}}$

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 {(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<=x} means a subset of S consisting of all outcomes a such that X(a)<=x

P{X<=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<=x)

F(0)=P(X<=0)=
F(1)=P(X<=1)=
F(2)=P(X<=2)=

Properties:

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

$F(+\infty)=P\{X\le +\infty\}=P(S)=1$ $\qquad$ $F(-\infty)=P\{X\le -\infty\}=0$

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

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

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

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

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

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

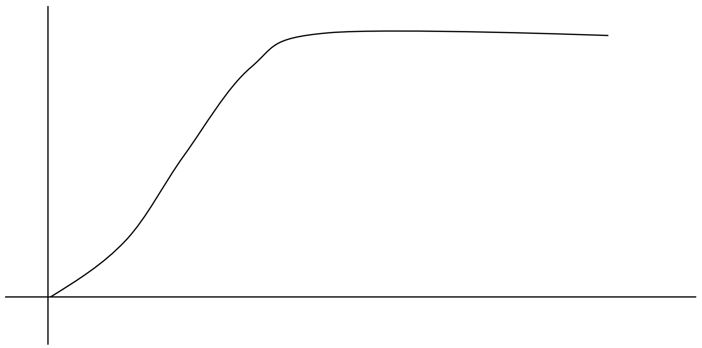
5. $P\{x_1<X\le x_2\}=F(x_2)-F(x_1)$

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

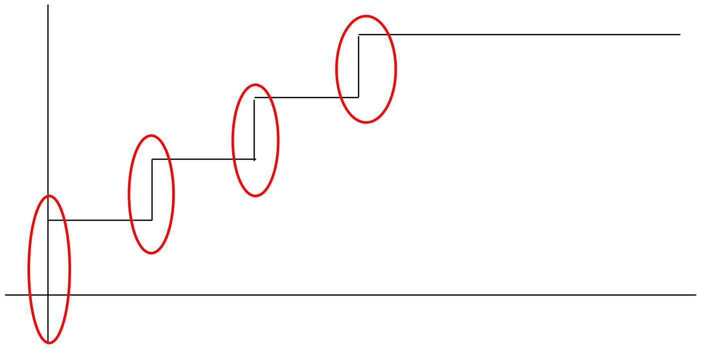
$P\{X\le x_2\}=P\{X\le x_1\} + P\{x_1<X\le x_2\}$

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

Continuous RV : if F(x) is continuous



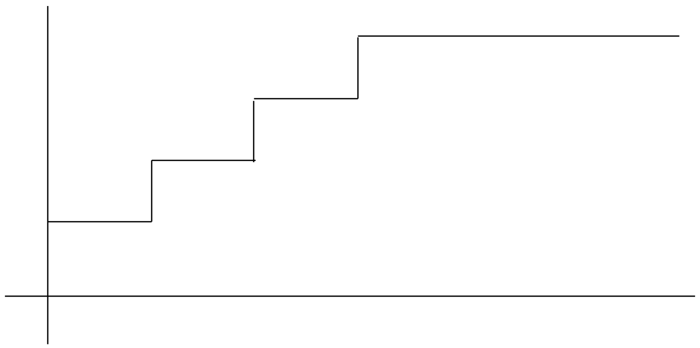
Discrete RV : if F(x) is constant except for a finite number of jump discontinuities (piecewise constant/step)

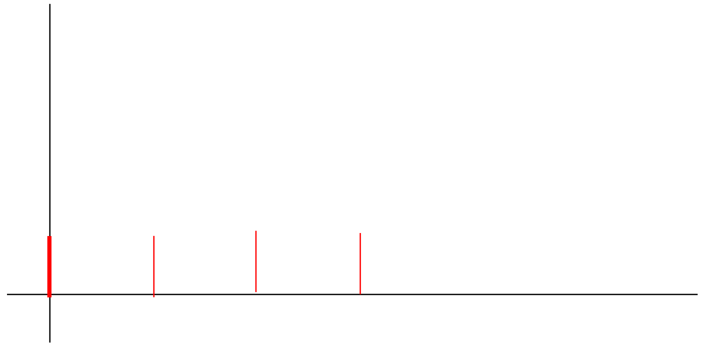


Probability density function:
Derivative of distribution function F(x)

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

As F(x) is monotonically increasing, so f(x)>=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 $f_X(x) = \sum_i p_i \delta(X - x_i)$, xi: jump discontinuities of F(x)
probability mass function (pmf)





pmf

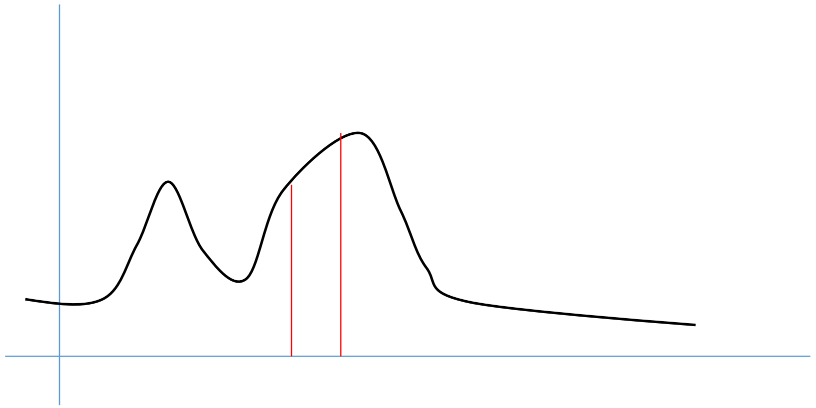
From definition,

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

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

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

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

P{0.4<X<=0.6} =?

2. X have the triangular pdf

$$f(x)=\begin{cases} x, & 0<x<=1 \\ 2-x, & 1<=x<=2 \\ 0, & \text{otherwise} \end{cases}$$

Check f(x) is pdf

P{0.3<X<=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,

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

$= \int_{0.3}^{1.5} f(x)\,dx$

$= \int_{0.3}^{1} x\,dx + \int_1^{1.5} (2-x)\,dx$

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

$= \cancel{0.8705}\ 0.83$

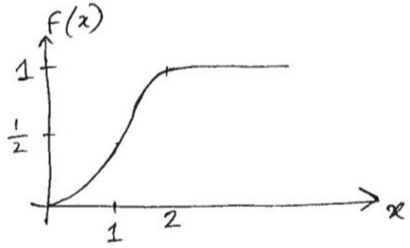
other way,

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

$= 2.1.5 - \frac{1.5^2}{2} - 1 - \frac{0.3^2}{2}$

$= 0.83$

If we plot F(x), it would be



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

DF is F(x)=P(X<=x)=sum (pi) for xi<=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.

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

$$\left.F(x)=P(X<=x)=\begin{cases} 0 & , x<0 \\ 1-p & , 0<=x<1 \\ 1 & , x>=1 \end{cases}\right\} \text{DF}$$

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, ……, x_n with probabilities p_1, p_2,……, p_n respectively. Expectation or expected value of X is

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

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.

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

Mean – of a RV X is $E(X)=\mu$

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 μ= E(X)

As per definition, μ_0’ = μ_0 = 1, μ_1’ = E(X) = μ, μ_1 = 0, μ_2 = σ^2

Central moments can be obtained from non-central moments as

$$\mu_2 = E(X-\mu)^2 = E(X)^2 - \left(E(X)\right)^2 = \mu_2' - (\mu_1')^2$$

$$\mu_3 = \mu_3' - 3\mu_2'\mu_2' + 2(\mu_1')^3$$

$$\mu_4 = \mu_4' - 4\mu_3'\mu_1' + 6\mu_2'(\mu_1')^2 - 3(\mu_1')^4$$

Expectation – mean, variance, moments

Let, a continuous RV X assumes the values between –inf to +inf with pdf f(x)

Expectation or expected value of X is

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

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

$$M(s) = \int_{-\infty}^{\infty} \left(1 + sx + \frac{s^2x^2}{2!} + \cdots\right) f(x)dx$$

$$M(s) = 1 + s\int_{-\infty}^{\infty} xf(x)dx + \frac{s^2}{2!}\int_{-\infty}^{\infty} x^2 f(x)dx + \cdots$$

$$M'(0) = \int_{-\infty}^{\infty} xf(x)dx = E(x)$$

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

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

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

Similarly, third order moment : measure of skewness

Fourth order moment : measure of kurtosis

Example-17 Let, X have the PDF as

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

$\therefore$ MGF is $M(s) = \int_0^\infty e^{sx} \frac{1}{2} e^{-x/2} \, dx$

$$= \int_0^\infty \frac{1}{2} e^{(s-\frac{1}{2})x} \, dx = \left[\frac{1}{2} \frac{e^{(s-\frac{1}{2})x}}{s-\frac{1}{2}} \right]_0^\infty$$

$$= \frac{1}{1-2s} \quad , s < \frac{1}{2}$$

$$M'(s) = \frac{d}{ds}\left(\frac{1}{1-2s}\right) = \frac{-1(-2)}{(1-2s)^2} = \frac{2}{(1-2s)^2}$$

$$M''(s) = \frac{d^2}{ds^2}\left(\frac{1}{1-2s}\right) = \frac{-2.2(1-2s).(-2)}{(1-2s)^4} = \frac{8}{(1-2s)^3}$$

now, $M'(0) = E(X) = 2$

$M''(0) = E(X^2) = 8$

$\therefore \mu = 2$

$\sigma^2 = 8 - 2^2 = 4$

now, to check these results,

$$E(X) = \int_0^\infty x \cdot \frac{1}{2} e^{-x/2} \, dx = \frac{1}{2} \int_0^\infty x e^{-x/2} \, dx$$

$$= \frac{1}{2} \left[x \int e^{-x/2} dx - \int \left(1 . \int e^{-x/2} dx\right) dx \right]_0^\infty$$

$$= \frac{1}{2} \left[-2x e^{-x/2} - 4 e^{-x/2} \right]_0^\infty$$

$$= \frac{1}{2} \cdot 4 = 2$$

$$E(X^2) = \int_0^\infty x^2 \frac{1}{2} e^{-x/2}\, dx = \frac{1}{2}\left[x^2(-2e^{-x/2}) - \int 2x(-2e^{-x/2})\, dx\right]_0^\infty$$

$$= \frac{1}{2}\left[-2x^2 e^{-x/2} + 4\int x e^{-x/2}\, dx\right]_0^\infty$$

(from E (X) ~~ot~~ solution)

$$= \left[-x^2 e^{-x/2}\right]_0^\infty + 2.4$$

$$= 8$$

$$\therefore \sigma^2 = 8 - 2^2 = 4$$ (hence verified)

Suppose, X is a discrete RV taking values xi with probability pi

$$M(s) = \sum_{i} p_i e^{s x_i}$$

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,

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

With z=1,

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

So,

$$\Gamma'(1) = E(X)$$

$$\Gamma''(1) = E\big(X(X-1)\big) = E(X^2 - X) = E(X^2) - E(X)$$

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

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

$$= pz + q$$

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

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

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

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

$$\& ,\ E(X^2) = E(X) = p$$

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

$$= pq$$

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)>=0,

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

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

If jw 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 (μ-ε, μ+ε) is negligible if the ratio σ/ε is sufficiently small. This result is fundamental and known as Chebyshev inequality; i.e., for any ε>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=μ with probability 1.

2. For specific densities , the bound is too high. Eg. If X is Normal RV, then P{|X-μ|>=3σ} = 0.0027 but this inequality gives <=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:

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

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

Some moment inequalities:

Theorem 1: Let $h(X)$ be a non negative function of RV X. If $E[h(X)]$ exists, then for every $\varepsilon>0$,

$P\{h(X) \geq \varepsilon\} \leq E[h(X)]/\varepsilon$

Corollary 1: If $h(X)=X$

$P\{X \geq \varepsilon\} \leq E(X)/\varepsilon$: Markov inequality

2: if $h(X)=(X-\mu)^2$ and $\varepsilon=k^2\sigma^2$

$P\{(X-\mu)^2 \geq k^2\sigma^2\} \leq \sigma^2/(k^2\sigma^2)=1/k^2$

$P\{|X-\mu| \geq k\sigma\} \leq 1/k^2$: Chebyshev inequality

$P\{|X-\mu| \leq k\sigma\} \geq 1-1/k^2$: lower bound for probability

Special RV:

1. Bernoulli RV: $X=0,1$ and $P(X=1)=p$, $P(X=0)=1-p$; $0<p<1$

   $E(X)=p$, $Var(X)=p(1-p)$

2. Binomial RV: n no. of independent Bernoulli trials and X is the total number of successes in n trials

$P(X=k)={}^{n}C_{k}\, p^{k} (1-p)^{n-k}$ $k=0,1,2,\ldots, n$; $0<=p<=1$

$X \sim B(n,p)$

$E(X)=np$, $Var(X)=npq$,

Mode: unimodal ($[(n+1)p]$) or bimodal ($(n+1)p$, $(n+1)p-1$)

Theorem 1: Let, Xi ($i=1, 2, \ldots, k$) be independent binomial RVs with $Xi \sim B(ni, p)$

$S=X1+X2+\ldots Xk$

$S \sim B(ni+n2+\ldots nk, p)$

3. Poisson RV: X takes values 0, 1, 2,….. Inf with

$P(X=k)=e^{-\lambda} \lambda^{k}/k!$, $\lambda>0$

$X \sim P(\lambda)$