Unit – 3

Probability Inequalities and Generating Functions

3.1

Probability Inequalities

Inequalities are useful for bounding quanties that might otherwise be hard to compute. They will also be used in the **theory of convergence** and **limit theorems**.

Chebychev's Inequality

When we want to find the probability of an event described by a random variable, its c.d.f or p.d.f. or p.m.f. is required. If it is not known but its *mean* and *variance* are known, we can use **Chebychev's inequality** to find the **upper bound** or **lower bound** for the probability of the event.

Theorem 1: If X is a random variable with mean μ and variance σ^2, then

$$\boldsymbol{P(|X-\mu| \geq \epsilon) \leq \frac{\sigma^2}{\epsilon^2}} \qquad \ldots\ldots\ldots\ldots..(\mathbf{1})$$

where $\epsilon > 0$

Proof: The proof is given for a continuous random variable. Let X be a continuous r.v. with p.d.f. $f(x)$. Then

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

$$= \int_{-\infty}^{\mu-\epsilon} (x-\mu)^2 f(x)dx + \int_{\mu-\epsilon}^{\mu+\epsilon} (x-\mu)^2 f(x)dx + \int_{\mu+\epsilon}^{\infty} (x-\mu)^2 f(x)dx$$

$$\geq \int_{-\infty}^{\mu-\epsilon} (x-\mu)^2 f(x)dx + \int_{\mu+\epsilon}^{\infty} (x-\mu)^2 f(x)dx$$

In the first integral,

$$x \leq \mu - \epsilon \Longrightarrow -x \geq -\mu + \epsilon \Longrightarrow -(x-\mu) \geq \epsilon \Longrightarrow (x-\mu)^2 \geq \epsilon^2$$

In the third integral, $x \geq \mu + \epsilon \Longrightarrow (x-\mu) \geq \varepsilon \Longrightarrow (x-\mu)^2 \geq \epsilon^2$

$$\therefore \sigma^2 \geq \epsilon^2 \left[\int_{-\infty}^{\mu-\epsilon} f(x)dx + \int_{\mu+\epsilon}^{\infty} f(x)dx\right]$$

$$= \epsilon^2[P(X \leq \mu - \epsilon) + P(X \geq \mu + \epsilon)]$$

$$= \epsilon^2 P[\mu - \epsilon \geq X \geq \mu + \epsilon] = \epsilon^2 P[-\epsilon \geq X - \mu \geq \epsilon]$$

$$= \epsilon^2 P[|X-\mu| \geq \epsilon]$$

Thus, $\sigma^2 \geq \epsilon^2 P[|X-\mu| \geq \epsilon] \Longrightarrow P[|X-\mu| \geq \epsilon] \leq \frac{\sigma^2}{\epsilon^2}$

Note: The proof is similar as in the case of d.r.v.X except that integration is replaced by summation.

Alternative forms:

$$P[|X-\mu| < \epsilon] = 1 - P[|X-\mu| \geq \epsilon] \geq 1 - \frac{\sigma^2}{\epsilon^2}$$

$$\Longrightarrow \boldsymbol{P}[|\boldsymbol{X}-\boldsymbol{\mu}| < \epsilon] \geq \mathbf{1} - \frac{\boldsymbol{\sigma}^2}{\boldsymbol{\epsilon}^2} \quad \ldots\ldots\ldots\ldots(2)$$

Let $\epsilon = k\sigma$ for $k > 0$. Then from (1), we have

$$\boldsymbol{P}[|\boldsymbol{X}-\boldsymbol{\mu}| \geq \boldsymbol{k\sigma}] \leq \frac{\mathbf{1}}{\boldsymbol{k}^2} \quad \ldots\ldots\ldots\ldots\ldots(3)$$

and from (2), we have

$$\boldsymbol{P}[|\boldsymbol{X}-\boldsymbol{\mu}| < k\sigma] \geq \mathbf{1} - \frac{\mathbf{1}}{\boldsymbol{k}^2} \quad \ldots\ldots\ldots\ldots(4)$$

Example 1: If a r.v. X has mean 12 and variance 9 and the probability distribution is unknown, then find $P(6 < X < 18)$.

Solution: Since the probability distribution of X is not known, we can't find the value of the required probability. We can find only a lower bound for probability using Chebychev's inequality. We have, for $\epsilon > 0$.

$$P[|X-\mu| < \epsilon] \geq 1 - \frac{\sigma^2}{\epsilon^2}$$

Given $E(X) = \mu = 12$ and $V(X) = \sigma^2 = 9$.

Then $P[|X-12| < \epsilon] \geq 1 - \frac{9}{\epsilon^2} \Rightarrow P[-\epsilon < (X-12) < \epsilon] \geq 1 - \frac{9}{\epsilon^2}$

$$\Rightarrow P[12-\epsilon < X < 12+\epsilon] \geq 1 - \frac{9}{\epsilon^2}$$

Let $\epsilon = 6$. Then $P[6 < X < 18] \geq 1 - \frac{9}{36} = 1 - \frac{1}{4} = 0.75$

$$\Rightarrow P[6 < X < 18] \geq 0.75$$

Thus, the probability of X lying between 6 and 18 is alseast 75%.

Example 2: A d.r.v. X takes the values $-1, 0$ and 1 with probabilities $\frac{1}{8}, \frac{3}{4}$ and $\frac{1}{8}$ respectively. Evaluate $P[|X-\mu| \geq 2\sigma]$ and compare it with the upper bound given by Chebychev's inequality.

Solution: We have,

X	-1	0	1
$p(x)$	$\frac{1}{8}$	$\frac{3}{4}$	$\frac{1}{8}$

Then $E(X) = \mu = \sum xp(x) = -1 \times \frac{1}{8} + 0 \times \frac{3}{4} + 1 \times \frac{1}{8} = 0$

and $E(X^2) = \sum x^2 p(x) = 1 \times \frac{1}{8} + 0 \times \frac{3}{4} + 1 \times \frac{1}{8} = \frac{2}{8} = \frac{1}{4}$

Hence $\sigma^2 = V(X) = E(X^2) - \left(E(X)\right)^2 = \frac{1}{4} - 0 = \frac{1}{4}$

Consider $P[|X-\mu| \geq 2\sigma] = P\left[|X| \geq 2 \cdot \frac{1}{2}\right] = P[|X| \geq 1]$

$$= P(X = -1,1)$$

$$= P(X = -1) + P(X = 1)$$

$$= \frac{1}{8} + \frac{1}{8} = \frac{2}{8} = \frac{1}{4} = 0.25$$

$\Longrightarrow P[|X-\mu| \geq 2\sigma] \leq 0.25$

On the other hand, by Chebychev's inequality,

$$P[|X-\mu| \geq 2\sigma] \leq \frac{1}{2^2} = \frac{1}{4}$$

Note that the two values are same.

Example 3: Use Chebychev's inequality to find how many times must a fair coin be tossed in order that the probability that the ratio of the number of heads to the number of tosses will lie between 0.45 and 0.55 will be at least 0.95.

Solution: Let X denote the number of heads obtained when a fair coin is tossed n times. Then $X \sim B(n,p)$. That is $E(X) = np$ and $V(X) = npq$.

Let $Y = \frac{X}{n}$. Then $E(Y) = E\left(\frac{X}{n}\right) = \frac{1}{n}E(X) = \frac{np}{n} = p$

and $V(Y) = V\left(\frac{X}{n}\right) = E\left(\left(\frac{X}{n}\right)^2\right) - \left(E\left(\frac{X}{n}\right)\right)^2 = \frac{1}{n^2}(E(X^2) - \left(E(X)\right)^2$

$$= \frac{1}{n^2}V(X) = \frac{npq}{n^2} = \frac{pq}{n}.$$

Since $p = \frac{1}{2}$ for a fair coin, $E(Y) = \frac{1}{2}$ and $V(Y) = \frac{1}{4n}$

By Chebychev's inequality for Y

$$P\left[\left|Y - \frac{1}{2}\right| < \epsilon\right] \geq 1 - \frac{\frac{1}{4n}}{\epsilon^2} = 1 - \frac{1}{4n\epsilon^2}$$

$$\Longrightarrow P\left[\frac{1}{2}-\epsilon < Y < \frac{1}{2}+\epsilon\right] \geq 1-\frac{1}{4n\epsilon^2}$$

Notice that, if $\epsilon = 0.05$ then $P(0.45 < Y < 0.55) \geq 1 - \frac{1}{4n\epsilon^2}$

Now, find n when $\epsilon = 0.05$ and $1 - \frac{1}{4n\epsilon^2} = 0.95 \Longrightarrow 1 - \frac{1}{n\times 4\times(0.05)^2} = 0.95$

$\Longrightarrow 1 - \frac{1}{0.01\times n} = 0.95 \Longrightarrow \frac{1}{0.01\times n} = 0.05 \Longrightarrow n = \frac{1}{0.01\times 0.05} = \frac{1}{0.0005} = \frac{10000}{5} = 2000$

Thus, $n = 2000$

Bienayme – Chebychev's inequality

Theorem 3: Let $g(X)$ be a non-negative function of a r.v. X . Then for every $k > 0$, we have

$$\boldsymbol{P[g(X) \geq k] \leq \frac{E[g(X)]}{k}} \ldots\ldots\ldots\ldots (\mathbf{1})$$

Proof: Here we shall prove the theorem for continuous random variable. The proof can be adapted to the case of discrete random variable on replacing integration by summation over the given range of the variable.

Let S be the set of all X for which $g(X) \geq k$.That is, $S = \{x \mid g(x) \geq k\}$.

Now, $E\left[g(X)\right] = \int_S g(x) f(x) dx,$ where $f(x)$ is the p.d.f. of X

$$\geq k\int_S f(x)dx \quad \left(on\ S, g(x) \geq k\right)$$

$$= kP\left[g(X) \geq k\right]$$

$$\Rightarrow P\left[g(X) \geq k\right] \leq \frac{E\left[g(X)\right]}{k}$$

Note:

1. If $g(X) = \left(X - E(X)\right)^2 = (X-\mu)^2$, then $E\left(g(X)\right) = V(X) = \sigma^2$ and replacing k by $\epsilon^2\sigma^2$ in equation (1), we get

$$\boldsymbol{P[(X-\mu)^2 \geq \epsilon^2\sigma^2] \leq \frac{\sigma^2}{\epsilon^2\sigma^2} = \frac{1}{\epsilon^2}}$$

$$\Rightarrow \boldsymbol{P[|X-\mu| \geq \epsilon\sigma] \leq \frac{1}{\epsilon^2}}$$

which is **Chebychev's inequality.**

2. If $g(X) = |X|$ in (1), then we get for any $k > 0$,

$$\boldsymbol{P[|X| \geq k] \leq \frac{E(|X|)}{k}}$$

which is known as **Markov's inequality.**

3. If $g(X) = |X|^r$ in (1), then we get

$$\boldsymbol{P[|X|^r \geq k^r] \leq \frac{E(|X|^r)}{k^r}}$$

which is known as **generalized Markov's inequality.**

Cauchy – Schwartz Inequality

When the j.p.d.f. of X and Y is known, upper bound for expected value of the product of X and Y can be found by using Cauchy – Schwartz inequality when second moments about origin of X and Y are given ($i.e.$, $E(X^2)$ and $E(Y^2)$ are given).

Theorem 2: For any two random variables $\boldsymbol{X}$ and $\boldsymbol{Y}$

$$\boldsymbol{\left(E(XY)\right)^2 \leq E(X^2)E(Y^2)}$$

Proof: Consider $E(X-tY)^2 \geq 0$ for any real number t. That is,

$E(X^2 - 2tXY + t^2Y^2) = E(X^2) - 2tE(XY) + t^2E(Y^2) \geq 0$
which is a quadratic expression in t. This expression is always positive only when t

has complex roots. This is possible only when discriminant of the expression is negative. Thus,

$$4\big(E(XY)\big)^2 - 4E(X^2)E(Y^2) \leq 0$$

$$\Rightarrow \big(E(XY)\big)^2 \leq E(X^2)E(Y^2)$$

Hence the result.

Example 4: The j.p.d.f. of (X, Y) is given by

$$f(x, y) = \frac{x+y}{21} \text{ for } x = 1, 2, 3 \text{ and } y = 1, 2.$$

Verify whether Cauchy-Schwartz inequality.

Solution: The joint and marginal p.m.fs $f_1(x)$ and $f_2(y)$ of X and Y respectively are given in the following table.

$Y \backslash X$	1	2	3	$f_2(y)$
1	$\frac{2}{21}$	$\frac{3}{21}$	$\frac{4}{21}$	$\frac{9}{21}$
2	$\frac{3}{21}$	$\frac{4}{21}$	$\frac{5}{21}$	$\frac{12}{21}$
$f_1(x)$	$\frac{5}{21}$	$\frac{7}{21}$	$\frac{9}{21}$	1

$$E(X) = \sum_{x=1}^{3} x f_1(x) = 1 \times \frac{5}{12} + 2 \times \frac{7}{21} + 3 \times \frac{9}{21} = \frac{46}{21}$$

$$E(X^2) = \sum_{x=1}^{3} x^2 f(x) = 1^2 \times \frac{5}{12} + 2^2 \times \frac{7}{21} + 3^2 \times \frac{9}{21} = \frac{114}{21}$$

$$E(Y) = \sum_{y=1}^{2} y f_2(y) = 1 \times \frac{9}{21} + 2 \times \frac{12}{21} = \frac{33}{21}$$

$$E(Y^2) = \sum_{y=1}^{2} y^2 f_2(y) = 1^2 \times \frac{9}{21} + 2^2 \times \frac{12}{21} = \frac{57}{21}$$

$$E(XY) = \sum_{x=1}^{3}\sum_{y=1}^{2} xyf(x, y)$$

$$= 1 \times 1 \times \frac{2}{21} + 1 \times 2 \times \frac{3}{21} + 1 \times 3 \times \frac{4}{21} + 2 \times 1 \times \frac{3}{21} + 2 \times 2 \times \frac{4}{21} + 2 \times 3 \times \frac{5}{21}$$

$$= \frac{1}{21}(2 + 6 + 12 + 6 + 16 + 30) = \frac{71}{21}$$

Verification of Cauchy-Schwatz inequality:

Here $\left(E(XY)\right)^2 = \left(\frac{71}{21}\right)^2 = 11.755$ and $E(X^2)E(Y^2) = \frac{114}{21} \times \frac{57}{21} = 14.735$

Note that $\left(E(XY)\right)^2 \leq E(X^2)E(Y^2)$.

Example 5: Let X and Y be c.r.vs with j.p.d.f.

$$\boldsymbol{f(x,y) = \begin{cases} x + y, & 0 < x < 1, 0 < y < 1 \\ 0, & otherwise \end{cases}}$$

Verify Cauchy-Schwartz inequality.

Solution: The m.p.d.f. of X is given by

$$f_1(x) = \int_{-\infty}^{\infty} f(x, y)dy = \int_0^1 (x + y)dy = \left[xy + \frac{y^2}{2}\right]_0^1 = x + \frac{1}{2}$$

$$f_1(x) = \begin{cases} x + \frac{1}{2}, & 0 < x < 1 \\ 0, & otherwise \end{cases}$$

Since $f(x, y)$ is symmetric in $x \; and \; y$, the m.p.d.f. of Y is given by

$$f_2(y) = \begin{cases} y + \frac{1}{2}, & 0 < x < 1 \\ 0, & otherwise \end{cases}$$

Now,

$$E(X^2) = \int_0^1 x^2\left(x+\frac{1}{2}\right)dx = \int_0^1 \left(x^3+\frac{x^2}{2}\right)dx = \left[\frac{x^4}{4}+\frac{x^3}{6}\right]_0^1 = \frac{1}{4}+\frac{1}{6} = \frac{10}{24}$$

Similarly, $E(Y^2) = \frac{10}{24}$ .Now,

$$E(XY) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf(x,y)dxdy = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy(x+y)dxdy$$

$$= \int_0^1 \left\{\int_0^1 (x^2y+xy^2)dx\right\}dy = \int_0^1 \left(\frac{x^3y}{3}+\frac{x^2y^2}{2}\right)_0^1 dy$$

$$= \frac{1}{6}\int_0^1 (2y+3y^2)dy = \left[\frac{y^2}{6}+\frac{y^3}{6}\right]_0^1 = \frac{2}{6} = \frac{1}{3}$$

Thus $\left(E(XY)\right)^2 = \left(\frac{1}{3}\right)^2 = \frac{1}{9} = 0.111$, and $E(X^2)E(Y^2) = \frac{10}{24}\times\frac{10}{24} = 0.1736$

Hence $\left(E(XY)\right)^2 \leq\ E(X^2)E(Y^2)$.