MARUDHAR KESARI JAIN COLLEGE FOR WOMEN (AUTONOMOUS)

VANIYAMBADI

PG and Department of Mathematics

II B.Sc. Mathematics – Semester - III

E-Notes (Study Material)

Elective: Mathematical Statistics I	**Code: 23UEMA33**
Unit:3: Moment generating function–Properties of cumulants Chebychev's Inequality-Binomial distribution	
Learning Objectives: To derive certain values incorporated with random variables	
Course Outcome: calculate moments, cumulants, moment generating function and various constants of probability distributions	

Overview:

Unit II

Moment Generating Function. (MGF)

The Moment generating Function of a random variable X about Origin having a probability function. $M_X(t) = E[e^{tx}]$

$$M_X(t) = \sum_{x} e^{tx} p(x)$$ for discreate random variable. $M_X(t) = \int_{-\infty}^{\infty} e^{tx} f(x) dx$ for continuous random variable.

Theorem 1

Find the r^{th} moment x about origin

Statement:

$$M_{CX}(t) = M_X(ct)$$

where,

c being the constant.

Proof:

We know that $M_X(t) = E[e^{tx}]$

LHS:

$$M_{CX}(t) = E[e^{tcx}] \quad -(1)$$

RHS: $M_X(ct) = E[e^{ctx}] - (2)$

From (1) & (2)

$$M_{CX}(t) = M_X(ct)$$

Theorem 2

Addition Theorem for moment generating function.

Statement:

The moment generating function of a Sum of the independent random variable is equal to the product of their respective moment generating function.

Symbolically $x_1, x_2, \dots x_n$ are independent random variable then the moment generating function of the random variables $x_1 + x_2 + \dots + x_n$ is given by

$$M_{x_1+x_2+\dots+x_n}(t) = M_{x_1}(t) \cdot M_{x_2}(t) \dots M_{x_n}(t).$$

Proof:

$$M_{x_1+x_2+\dots+x_n}(t) = E\left[e^{t(x_1+x_2+\dots+x_n)}\right]$$

$$= E\left[e^{tx_1+tx_2+\dots+tx_n}\right]$$

$$= E\left[e^{tx_1} \cdot e^{tx_2} \dots e^{tx_n}\right]$$

$$= E\left[e^{tx_1}\right] E\left[e^{tx_2}\right] \dots E\left[e^{tx_n}\right]$$

$$= M_{x_1}(t) \cdot M_{x_2}(t) \dots M_{x_n}(t)$$

LHS = RHS.

Statement:

Topic: Uniqueness theorem for moment generating function

Statement:

The moment generating function of a distribution, If it uniquily determines the distribution implies that the corresponding to given probability distribution. there is only on probability distribution. Symbolically if X and Y are two random variables then $M_X(t) = M_Y(t)$ which implies X and Y are identical distributed.

Cumulant Theorem:

Cumulant generating function :-

Cumulant Generating function $K(t)$ is defined as $K_X(t) = \log_e M_X(t)$ are r^{th} cumulant.

r^{th} cumulant.

$$K_X(t) = K_1 \frac{t^1}{1!} + K_2 \frac{t^2}{2!} + \dots + K_r \frac{t^r}{r!} \dots$$

$$= \log M_X(t)$$

$$= \log \left[1 + \mu_1' \frac{t}{1!} + \mu_2' \cdot \frac{t^2}{2!} + \dots + \mu_r' \frac{t^r}{r!} \right] \dots$$

Where K_r = coefficient of $\frac{t^r}{r!}$ in $K_X(t)$ is called the r^{th} cumulant.

$$K_1 = \mu_1' = \text{Mean}$$

$$K_2 = \mu_2 = \mu_2' - (\mu_1')^2 = \text{variance}$$

$$K_3 = \mu_3 = \mu_3' - 3\mu_2' \mu_1' + 2\mu_1'^3$$

$$K_4 = \mu_4 = K_3 + 3K_2^2$$

Properties of Cumulant

1) Additive property

Statement:

The r^{th} cumulant of the sum of the independent random variable is equal to the r^{th} cumulant of the individual variables.

$$K_r(x_1 + x_2 + \dots x_n) = K_r(x_1) + K_r(x_2) + \dots + K_r(x_n)$$

where,

x_i, $i = 1, 2, \dots n$ are independent random variable.

Proof:

$$K_X(t) = \log_e M_X(t)$$

$$K_{x_1 + x_2 + \dots + x_n}(t) = \log_e M_{x_1 + x_2 + \dots + x_n}(t)$$

$$= \log\left[M_{x_1}(t) + M_{x_2}(t) + \dots + M_{x_n}(t)\right]$$

$$= \log M_{x_1}(t) + \log M_{x_2}(t) + \dots + \log M_{x_n}(t)$$

$$= K_{x_1}(t) + K_{x_2}(t) + \dots + K_{x_n}(t)$$

$$K_r(x_1 + x_2 + \dots + x_n) = K_r(x_1) + K_r(x_2) + \dots + K_r(x_n)$$

Theorem 5:

Effect of change of origin and scale on moment generating function.

Statement:

Let us transform x to a variable U by changing the both origin and scale in X as follows $U = \frac{x-a}{h}$

where, a and h are constants.

The moment generating function of U is given by

$$M_U(t) = e^{-\frac{at}{h}} \cdot M_x(t/h)$$

proof

We know that

$$M_x(t) = E[e^{tx}]$$

$$M_U(t) = E[e^{tU}]$$

Let $U = \frac{x-a}{h}$

$$M_U(t) = E\left[e^{t\left(\frac{x-a}{h}\right)}\right] = E\left[e^{\frac{tx}{h} - \frac{ta}{h}}\right]$$

$$= E\left[e^{\frac{tx}{h}} \cdot e^{-\frac{at}{h}}\right] = E\left[e^{\frac{tx}{h}}\right] \cdot e^{-at/h}$$

$$M_U(t) = e^{-at/h} \cdot M_x(t/h)$$

Remark:

If $a = E(x) = \mu$ and $h = \sigma_x = \sigma$ then

$$U = \frac{x-a}{h}$$

$$Z = \frac{x - E(x)}{\sigma_x} = \frac{x-\mu}{\sigma}$$, is known as

Standard Normal variant. The moment generating function of Z is given by

$$M_Z(t) = E[e^{tZ}] = E\left[e^{t\left(\frac{x-\mu}{\sigma}\right)}\right]$$

$$= E\left[e^{\frac{tx}{\sigma} - \frac{t\mu}{\sigma}}\right] = E\left[e^{\frac{tx}{\sigma}} \cdot e^{-\frac{t\mu}{\sigma}}\right]$$

$$M_Z(t) = e^{-\frac{t\mu}{\sigma}} \cdot E\left[e^{tx/\sigma}\right]$$

$$M_Z(t) = e^{-t\mu/\sigma} \cdot M_X(t/\sigma)$$

Problems

1) Let the random variables X, assume a value r with the probability $P(X=r) = q^{r-1} \cdot p$, $r = 1, 2, \dots$. find the moment generating function and its mean and variance.

Moment generating function of x

$$M_X(t) = E\left[e^{tx}\right] \qquad (1-x)^{-1} = 1 + x + x^2 + \dots$$

put $X = r$

$$M_r(t) = E\left[e^{tr}\right] = \sum_{r=1}^{\infty} e^{tr} p(r)$$

$$= \sum_{r=1}^{\infty} e^{tr} \cdot q^{r-1} \cdot p = \sum_{r=1}^{\infty} e^{tr} \cdot q^r \cdot q^{-1} \cdot p$$

$$= \frac{p}{q} \sum_{r=1}^{\infty} e^{tr} q^r = \frac{p}{q} \sum_{r=1}^{\infty} (e^t q)^r$$

$$= \frac{p}{q}\left[(qe^t)^1 + (qe^t)^2 + \dots\right]$$

$$= \frac{p}{q} qe^t \left[1 + qe^t + qe^{t^2} + \dots\right]$$

$$M_r(t) = \frac{p}{q} e^t (1 - qe^t)^{-1} = \frac{pe^t}{1-qe^t}$$

Mean :

$$M_X(t) = \left[M'_r(t)\right]_{t=0}$$

$$= \frac{d}{dt}\left[M_r(t)\right] = \frac{d}{dt}\left[\frac{pe^t}{1-qe^t}\right]$$

$$= \frac{(1-qe^t)pe^t - pe^t(-qe^t)}{(1-qe^t)^2}$$ $u = pe^t \quad \frac{du}{dt} = pe^t$

$v = 1-qe^t \quad \frac{dv}{dt} = -qe^t$

$$= \frac{pe^t - pqe^{2t} + pqe^{2t}}{(1-qe^t)^2}$$

$$M_x(t) = \frac{pe^t}{(1-qe^t)^2}$$

when $t = 0$

$$= \frac{pe^0}{(1-qe^0)^2} = \frac{p}{(1-q)^2} = \frac{p}{p^2} = \frac{1}{p}$$

(Mean) $\mu_1' = \frac{1}{p}$

<u>variance</u>:

$$V(x) = E(x^2) - [E(x)]^2$$

$$E(x^2) = \left[M_1''(t)\right]_{t=0} = \frac{d}{dt}\left[\frac{d}{dt}(M_1')\right]_{t=0}$$

$$= \frac{d}{dt}\left[\frac{pe^t}{(1-qe^t)^2}\right]$$

$$= \frac{(1-qe^t)^2 pt - pe^t 2(1-qe^t)(-qe^t)}{(1-qe^t)^4}$$

$$= \frac{(1-qe^t).pe^t + 2qe^t pe^t}{(1-qe^t)^3}$$

$$= \frac{pe^t - pe^t qe^t + 2qe^t pe^t}{(1-qe^t)^3}$$

$$E(x^2) = \frac{pe^t + qe^t pe^t}{(1-qe^t)^3} = \frac{pe^t(1+qe^t)}{(1-qe^t)^3}$$

At $t = 0$

$$E(x^2) = \frac{p(1+q)}{(1-q)^3} = \frac{p(1+q)}{p^3} = \frac{1+q}{p^2}$$

$$V(x) = \frac{1+q}{p^2} - \frac{1}{p^2} = \frac{1+q-1}{p^2} = \frac{q}{p^2}$$

Effect of change of origin and scale on cumulant generating function

If $U = \frac{x-a}{h}$ by moment generating function $M_U(t) = e^{-at/h} M_x(t/h)$

Taking log on b.s

$$\log[M_U(t)] = \log[e^{-at/h} M_x(t/h)]$$

$$= \log[e^{-at/h}] + \log[M_x(t/h)] \quad \log M_x = K_x$$

$$K_U(t) = -\frac{at}{h} + K_x(t/h)$$

Chebychev's inequality

If x is a random variable with Mean (μ_r) and variance (σ^2) then for any positive number K. We have,

$$P\{|x-\mu| \geq K\sigma\} \leq \frac{1}{K^2}$$

$$P\{|x-\mu| < K\sigma\} \geq 1 - \frac{1}{K^2} \qquad E(x) = \int_{-\infty}^{\infty} x f(x)\,dx$$

Proof

Case i: If x is a continuous random variable $\sigma^2 = \sigma_x^2 = E[x - E(x)]^2$

$$E(x) = \mu = \text{Mean}$$

$$\sigma^2 = E[x-\mu]^2$$

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

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

Not defined.

For the first integral,

$$-\infty \le x \le \mu - k\sigma$$

$$x \le \mu - k\sigma$$

$$x - \mu \le -k\sigma$$

Squaring on both side

$$(x-\mu)^2 \le k^2\sigma^2$$

For 3rd integral,

$$\mu + k\sigma \le x \le \infty$$

$$\mu + k\sigma \le x$$

$$k\sigma \le x - \mu$$

Squaring on both sides

$$k^2\sigma^2 \le (x-\mu)^2$$

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

$$\sigma^2 \ge \int_{-\infty}^{\mu-k\sigma} k^2\sigma^2 f(x)\,dx + \int_{\mu+k\sigma}^{\infty} k^2\sigma^2 f(x)\,dx$$

$$\ge k^2\sigma^2 \left[\int_{-\infty}^{\mu-k\sigma} f(x)\,dx + \int_{\mu+k\sigma}^{\infty} f(x)\,dx\right]$$

$$\ge k^2\sigma^2 \left[P(x \le \mu - k\sigma) + P(\mu + k\sigma \le x)\right]$$

$$\ge k^2\sigma^2 \left[P(x - \mu \le -k\sigma) + P(k\sigma \le x - \mu)\right]$$

$$\ge k^2\sigma^2 \left[P(-k\sigma \ge x - \mu \ge k\sigma)\right]$$

$$\sigma^2 \ge k^2\sigma^2 \left[P|x-\mu| \ge k\sigma\right]$$

$$1 \ge k^2 \left[P|x-\mu| \ge k\sigma\right] \rightarrow (*)$$

$$\frac{1}{k^2} \ge P\left[|x-\mu| \ge k\sigma\right]$$

$$P\left[|x-\mu| \ge k\sigma\right] \le \frac{1}{k^2} \rightarrow (1)$$

Also, we know that $p+q=1$

$$P\left[|x-\mu| \ge k\sigma\right] + P\left[|x-\mu| < k\sigma\right] = 1$$

$$P[|x-\mu| < k\sigma] = 1 - P[|x-\mu| \geq k\sigma]$$

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

Case 2:

In case of discrete random variable the proof follow similarly on replacing Integration of Summation.

Remark:

In particular we take $k\sigma = c$ where $c > 0$

Condition 1:

$$P[|x-\mu| \geq k\sigma] \leq \frac{1}{k^2}$$

$$P[|x-\mu| \geq c] \leq \frac{1}{c^2/\sigma^2}$$

$$P[|x-\mu| \geq c] \leq \frac{\sigma^2}{c^2}$$

$$P[|x-\mu| \geq c] \leq \frac{V(x)}{c^2} \quad (\because \sigma^2 \text{ is variance})$$

Condition 2:

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

$$P[|x-\mu| < c] \geq 1 - \frac{1}{c^2/\sigma^2}$$

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

$$P[|x-\mu| < c] \geq 1 - \frac{V(x)}{c^2}$$

Problems

1) If x is a number scored in a throw of a pair die. Show that the chebychev's inequality gives $P[|x-\mu| > 2.5] \leq 0.47$ where μ is a mean of x. where the actual probability is 0.

$P(A) + P(\bar{A}) = 1$

Let

x be a random variable with sample space.

$S = \{1, 2, 3, 4, 5, 6\}$

x :	1	2	3	4	5	6
$P(x)$:	$1/6$	$1/6$	$1/6$	$1/6$	$1/6$	$1/6$

$$E(x) = \sum_{x=1}^{6} x\,p(x) = 1\left(\frac{1}{6}\right) + 2\left(\frac{1}{6}\right) + 3\left(\frac{1}{6}\right) + 4\left(\frac{1}{6}\right) + 5\left(\frac{1}{6}\right) + 6\left(\frac{1}{6}\right)$$

$$= \frac{7}{2}$$

$$E(x^2) = \sum_{x=1}^{6} x^2 p(x)$$

$$= (1)^2\left(\frac{1}{6}\right) + (2)^2\left(\frac{1}{6}\right) + (3)^2\left(\frac{1}{6}\right) + (4)^2\left(\frac{1}{6}\right) + (5)^2\left(\frac{1}{6}\right) + (6)^2\left(\frac{1}{6}\right)$$

$$E(x^2) = \frac{91}{6}$$

$$V(x) = E(x^2) - [E(x)]^2$$

$$V(x) = \frac{91}{6} - \left(\frac{7}{2}\right)^2 = \frac{35}{12}$$

$$P[|x-\mu| > 2.5] \leq 0.47$$

$$P[|x-\mu| < c] \geq 1 - \frac{V(x)}{c^2}$$

$$P[|x-\mu| \geq 2.5] \leq \frac{\frac{35}{12}}{(2.5)^2} \qquad c = 2.5$$

$$P[|x-\mu| \geq 2.5] \leq \frac{7}{15} \qquad = \frac{35/6}{3^2}$$

$$P[|x-\mu| \geq 2.5] \leq 0.467 \qquad |x| \leq 3$$

Acutal probability $\qquad -3 \leq x \leq 3$

$$P[|x-\mu| > 2.5] + P[|x-\mu| \leq 2.5] = 1$$

$$P[|x-\mu| > 2.5] = 1 - P[|x-\mu| \leq 2.5]$$

$$= 1 - P[|x-3.5| \leq 2.5] \quad (\mu = E(x))$$

$$= 1 - P[-2.5 \leq x - 3.5 \leq 2.5]$$

$$= 1 - P[-2.5 + 3.5 \leq x \leq 2.5 + 3.5]$$

$$= 1 - P[1 \leq x \leq 6]$$

$$= 1 - P\left[\frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6}\right]$$

$$= 1 - P(1) = 1 - 1 \qquad P(c) = c$$

$$= 0 \qquad P(1) = 1$$

$$P[|x-\mu| > 2.5] = 0$$

---

2) Two unbaised are thrown to start. If X is the Sum of the numbers shown on the dice. Prove that $P[|x-\underset{E(x)}{7}| \geq 3] \leq \frac{35}{54}$. Compare this with actual probability

$$S = \begin{Bmatrix} (1,1)(1,2)(1,3)(1,4)(1,5)(1,6) \\ (2,1)(2,2)(2,3)(2,4)(2,5)(2,6) \\ (3,1)(3,2)(3,3)(3,4)(3,5)(3,6) \\ (4,1)(4,2)(4,3)(4,4)(4,5)(4,6) \\ (5,1)(5,2)(5,3)(5,4)(5,5)(5,6) \\ (6,1)(6,2)(6,3)(6,4)(6,5)(6,6) \end{Bmatrix}$$

X = Sum of the numbers

X	Favourable Outcomes	
2	(1,1)	1/36
3	(1,2)(2,1)	2/36
4	(1,3)(2,2)(3,1)	3/36
5	(1,4)(2,3)(3,2)(4,1)	4/36
6	(1,5)(2,4)(3,3)(4,2)(5,1)	5/36
7	(1,6)(2,5)(4,3)(3,4)(5,2)(6,1)	6/36
8	(2,6)(3,5)(4,4)(5,3)(6,2)	5/36
9	(3,6)(4,5)(5,4)(6,3)	4/36
10	(4,6)(5,5)(6,4)	3/36
11	(5,6)(6,5)	2/36
12	(6,6)	1/36

$E(x) = 7$

$$E(x^2) = \sum_{x=2}^{12} x^2 f(x)$$

$$= (2)^2\left(\frac{1}{36}\right) + (3)^2\left(\frac{2}{36}\right) + (4)^2\left(\frac{3}{36}\right) + (5)^2\left(\frac{4}{36}\right) +$$

$$(6)^2\left(\frac{5}{36}\right) + (7)^2\left(\frac{6}{36}\right) + (8)^2(5/36) + (9)^2(4/36) +$$

$$(10)^2\left(\frac{3}{36}\right) + (11)^2\left(\frac{2}{36}\right) + (12)^2\left(\frac{1}{36}\right)$$

$E(x^2) = 329/6$

$$V(x) = E(x^2) - [E(x)]^2$$

$$V(x) = \frac{329}{6} - [7]^2 = \frac{35}{6}$$

$$P[|x-7| \geq 3] \leq \frac{35}{54}$$

$$P[|x-\mu| < c] \geq 1 - \frac{V(x)}{c^2}$$

$$P[|x-7| \geq 3] \leq \frac{\frac{35}{6}}{3^2} = \frac{5.833}{9}$$

$$P[|x-7| \geq 3] \leq 0.648148$$

Actual probability

$$P[|x-7| > 3] + P[|x-7| \leq 3] = 1$$

$$P[|x-7| > 3] = 1 - P[|x-7| \leq 3]$$

$$= 1 - P[-3 \leq x-7 \leq 3]$$

$$= 1 - P[-3+7 \leq x \leq 3+7]$$

$$= 1 - P[4 \leq x \leq 10]$$

$$= 1 - P\left[\frac{3}{36} + \frac{4}{36} + \frac{5}{36} + \frac{6}{36} + \frac{5}{36} + \frac{4}{36} + \frac{3}{36}\right]$$

$$P[|x-7| > 3] = 1 - P\left(\frac{5}{6}\right) = 1 - \frac{5}{6} = \frac{1}{6}$$

Binomial Distribution

A random variable X is said to be a binomial distribution, If it assumes only non negative values and its probability mass function is defined by $P(X=x)=p(x)=$

$$\begin{cases} nC_x p^x q^{n-x} & ; x=0,1,2,\dots \\ 0 & ; \text{otherwise} \end{cases}$$

where $q \xrightarrow{\text{Non occurance of probability}} = 1-p \rightarrow$ occurance of probability

The two independent constants n and p in the distribution are known as parameter of the distribution, Sometimes n is also known as the degree of binomial distribution

---

Problems:

1) 10 coins are thrown simultaneously. Find out the probability of getting atleast 7 heads.

Soln)

$$P(x) = nC_x p^x q^{n-x}$$

p = Probability of getting heads (1/2)

q = Probability of not getting heads (1/2)

$$P(x) = nC_x p^x q^{n-x}$$

$$= 10C_x (1/2)^x (1/2)^{10-x} = 10C_x \left(\tfrac{1}{2}\right)^{10}$$

$$= \left(\tfrac{1}{2}\right)^{10}\left[10C_7 + 10C_8 + 10C_9 + 10C_{10}\right]$$

$$= \left(\tfrac{1}{2}\right)^{10}\left[120+45+10+1\right] = \frac{1}{1024}[176]$$

$$\boxed{P(x) = 0.171}$$

*Moment generating function of Binomial distribution.

$$M_x(t) = E\left[e^{tx}\right]$$

$$= \sum_{x=0}^{n} e^{tx} \cdot p(x)$$

$$= \sum_{x=0}^{n} e^{tx}\, nC_x p^x q^{n-x}$$

$$= e^{t(0)}\, nC_0 p^0 q^{n-0} + e^{t}\, nC_1 p^1 q^{n-1} +$$

$$e^{2t}\, nC_2 p^2 q^{n-2} + \ldots$$

$$= q^n + e^t\, nC_1 p q^{n-1} + e^{2t}\, nC_2 p^2 q^{n-2} + \ldots$$

$$M_x(t) = (q + pe^t)^n$$

Characteristic function of Binomial distribution

$$M_x(t) = E\left[e^{itx}\right]$$

$$= \sum_{x=0}^{n} e^{itx} \cdot p(x)$$

$$= \sum_{x=0}^{n} e^{itx}\, nC_x p^x q^{n-x}$$

$$= e^{it(0)}\, nC_0 p^0 q^{n-0} + e^{it}\, nC_1 p^1 q^{n-1} +$$

$$e^{2it}\, nC_2 p^2 q^{n-2} + \ldots$$

$$= q^n + e^{it}\, nC_1 p q^{n-1} + e^{2t}\, nC_2 p^2 q^{n-2} + \ldots$$

$$M_x(t) = (q + pe^{it})^n$$

Probability generating function of Binomial distribution

$$P(s) = E(s)^{x}$$

$$= \sum_{x=0}^{n} s^{x} P(x) = \sum_{x=0}^{n} s^{x} nC_x p^x q^{n-x}$$

$$= s^0 nC_0 p^0 q^{n-0} + s^1 nC_1 p^1 q^{n-1} + s^2 nC_2 p^2 q^{n-2} + \cdots$$

$$= q^n + s\,nC_1 p q^{n-1} + s^2 nC_2 p^2 q^{n-2} + \cdots$$

$$P(s) = (q + ps)^n$$

Additive property of Binomial distribution.

Let $x \sim B(n_1, p_1)$, $y \sim B(n_2, p_2)$ be independent random variable then

$$\left.\begin{aligned} M_x(t) &= (q_1 + p_1 e^t)^{n_1} \\ M_y(t) &= (q_2 + p_2 e^t)^{n_2} \end{aligned}\right\} \rightarrow (1)$$

Since, x and y are independent then the distribution of x + y is

$$M_{x+y}(t) = M_x(t) \cdot M_y(t)$$

$$= (q_1 + p_1 e^t)^{n_1} (q_2 + p_2 e^t)^{n_2} \rightarrow (2)$$

From (2), cannot be expressed in the form of $(q + pe^t)^n$

From the uniqueness Theorem of moment generating function x + y is not a Binomial variate.

Hence, In general, the sum of two independent binomial variate is not a binomial distribution or binomial variate does not process the additive Property

<u>Recurrance relation for the probability of Binomial distribution of fitting of a binomial distribution.</u>

The probability mass function of the binomial distribution is given by

$$P(x) = nC_x p^x q^{n-x} \rightarrow (1)$$

$$P(x+1) = nC_{x+1} p^{x+1} q^{n-(x+1)} \rightarrow (2)$$

$$\frac{(2)}{(1)} \Rightarrow \frac{P(x+1)}{P(x)} = \frac{nC_{x+1} p^{x+1} q^{n-(x+1)}}{nC_x p^x q^{n-x}}$$

$\therefore nC_r = \frac{n!}{r!(n-r)!}$

$nC_{x+1} = \frac{n!}{(x+1)!(n-x-1)!}$

$$= \frac{nC_{x+1} p^x p^1 q^n q^{-x} q^{-1}}{nC_x p^x q^n q^{-x}}$$

$$= \frac{nC_{x+1} p^1 q^{-1}}{nC_x}$$

$$= \frac{\frac{n!}{(x+1)!(n-(x+1))!}}{\frac{n!}{x!(n-(x))!}} pq^{-1}$$

$$= \frac{n!}{(x+1)!(n-(x+1))!} \frac{x!(n-x)!}{n!} pq^{-1}$$

$$= \frac{x!(n-x)! pq^{-1}}{(x+1)!(n-(x+1))!}$$

$$= \frac{x!(n-x)(nx-1)!\,p}{(x+1)(x)!(n-x-1)!\,q}$$

$$\frac{P(x+1)}{P(x)} = \frac{(n-x)p}{(x+1)q}$$

$$P(x+1) = \frac{(n-x)p}{(x+1)q} P(x)$$

which is the required recurrance formula

Since $p(0) = q^n$ where q is calculated from the given data by using $p+q=1$,

$$q = 1-p$$

The remaining probabilities is $P_1, P_2, \ldots$ etc can be obtain from the recurrance formula.

$x=0$ in the recurrance formula

put $x=0$

$$P(1) = \frac{np}{q} P(0)$$

put $x=1$

$$P(2) = \frac{(n-1)p}{2q} P(1)$$

put, $x=2$

$$P(3) = \frac{(n-2)p}{3q} P(2)$$

$\vdots \qquad \vdots$

Mode of Binomial distribution.

The binomial distribution is given by

$$P(x) = nC_x\, p^x q^{n-x} \rightarrow (1)$$

$$P(x-1) = nC_{x-1}\, p^{x-1} q^{n-(x-1)} \rightarrow (2)$$

$$\frac{(1)}{(2)} \Rightarrow \frac{P(x)}{P(x-1)} = \frac{nC_x\, p^x q^{n-x}}{nC_{x-1}\, p^{x-1} q^{n-(x-1)}}$$

$$\frac{P(x)}{P(x-1)} = \frac{nC_x p^x q^n q^{-x}}{nC_{x-1} p^x p^{-1} q^n q^{-x} q^1}$$

$$= \frac{nC_x}{nC_{x-1} p^{-1} q^1} = \frac{p\,nC_x}{nC_{x-1}\, q}$$

$$= \frac{p \dfrac{n!}{x!(n-x)!}}{q \dfrac{n!}{(x-1)!(n-(x-1))!}}$$

$$= p \frac{n!}{x!(n-x)!} \times \frac{(x-1)!(n-x+1)!}{q\, n!}$$

$$x! = x(x-1)!$$

$$(n-x+1)! = (n-x+1)(n-x+1-1)!$$

$$= \frac{p(x-1)!(n-x+1)!}{q\, x!\, (-x+n)!}$$

$$= \frac{p(x-1)!\,(n-x+1)(n-x+1-1)!}{q\, x(x-1)!\,(n-x)!}$$

$$= \frac{p(n-x+1)}{xq}$$

Add & Sub xq in numerator

$$= \frac{np - xp + p + xq - xq}{xq}$$

$$= \frac{xq + p(n-x+1) - xq}{xq}$$

$$= \frac{xq}{xq} + \frac{p(n-x+1) - xq}{xq}$$

$$= 1 + \frac{p(n-x+1) - xq}{xq}$$

$$= 1 + \frac{pn - xp + p - xq}{xq} = 1 + \frac{p(n+1) - x(p+q)}{xq}$$

$$= 1 + \frac{p(n+1) - x}{xq} \rightarrow (3)$$

Mode is the value of x for which $p(x)$ is maximum

case (i) when $p(n+1)$ is not an integer.

Let $(n+1)p = m + f$ where m is an integer, f is fraction, Such that $0 < f < 1$

Sub in (3)

$$(3) \Rightarrow \frac{P(x)}{P(x-1)} = 1 + \frac{P(n+1)-x}{xq}$$

$$\frac{P(x)}{P(x-1)} = 1 + \frac{(m+f)-x}{xq}$$

$$\frac{P(x)}{P(x-1)} > 1 \quad (\because x = 1, 2, 3, \dots m) \rightarrow (*)$$

If $\frac{P(x)}{P(x-1)} < 1$ for $x = m+1, m+2, \dots m$ $\rightarrow$ (A)

Sub $x = 1, 2, \dots m$ in eqn (*)

$$\frac{P(x)}{P(x-1)} > 1$$

$$\frac{P(1)}{P(0)} > 1, \quad \frac{P(2)}{P(1)} > 1, \quad \frac{P(3)}{P(2)} > 1, \dots \frac{P(m)}{P(m-1)} > 1$$

$$P(1) > P(0);\ P(2) > P(1);\ P(3) > P(2) \dots P(m) > P(m-1) \rightarrow (4)$$

Sub $x = m+1, m+2, \dots m$ in (A)

$$(A) \Rightarrow \frac{P(x)}{P(x-1)} < 1$$

$$\frac{P(m+1)}{P(m)} < 1, \quad \frac{P(m+2)}{P(m+1)} < 1, \quad \frac{P(m+3)}{P(m+2)} < 1, \dots$$

$$\frac{P(m)}{P(m-1)}.$$

$P(m+1) < P(m)$, $P(m+2) < P(m+1)$,

$P(m+3) < P(m+2) \ldots P(n) < P(n-1) \rightarrow (5)$

Combaine (4) & (5)

$P(0) < P(1) < P(2) < \ldots < P(m-1) < P(m) > P(m+1)$

$> P(m+2) > \ldots > P(n-1) > P(n)$.

In this case there exist a unit mode value for binomial distribution and it is m, the integral part $(n+1)p$

Case (ii)

When $(n+1)p$ is an integer. Let $(n+1)p = m$ where m is the integer, sub in eqn (3)

$$\frac{P(x)}{P(x-1)} = 1 + \frac{p(n+1) - x}{xq} = 1 + \frac{m-x}{xq}$$

$$\frac{P(x)}{P(x-1)} = \begin{cases} >1 & \text{for } x = 1, 2, \ldots, m-1 \\ =1 & \text{for } x = m \\ <1 & \text{for } x = m+1, m+2, \ldots n \end{cases}$$

$\frac{P(1)}{P(0)} > 1$, $\frac{P(2)}{P(1)} > 1$, $\frac{P(3)}{P(2)} > 1$, $\ldots$, $\frac{P(m-1)}{P(m-2)} > 1$

$P(1) > P(0)$, $P(2) > P(1)$, $P(3) > P(2)$, $\ldots$

$P(m-1) > P(m-2)$

$\frac{P(m+1)}{P(m)} < 1$, $\frac{P(m+2)}{P(m+1)} < 1, \ldots,$ $\frac{P(m-1)}{P(m-2)} < 1$

$P(m+1) < P(m)$, $P(m+2) < P(m+1) \ldots$

$P(n) < P(n-1)$

combine eqn (6) and (7)

$P(0) < P(1) < P(2) < \ldots P(m-1) < P(m) < P(m+1)$

$< P(m+2) \ldots P(n)$

In this case the distribution is bimodel and the two model values are m, m - 1

Problems :

1) command only the following mean of binomial distribution is 3 and variance is 4.

The parameters of binomial distribution is n and p.

$$\text{mean} = np = 3 \rightarrow (1)$$

$$\text{variance} = npq = 4 \rightarrow (2)$$

sub (1) in (2)

$$\text{variance} = nqp = 4$$

$$3q = 4$$

$$q = 4/3 = 1.33371$$

Since $p+q = 1$, which is impossible

Probability can't excist !

Therefore the given statement is wrong.

2) A and B play a game of in which There a chance of winning are in ratio 3:4. Find A's chance of winning atleast 3 game out of five game playing.

Let, p be the Probability that A wins the game

Then , $p = 3/5$

W.K.T $\quad p + q = 1$

$$q = 1 - 3/5 = \frac{5-3}{5} = \frac{2}{5}$$

Hence the binomial probability law the Probability that out of 5 game A wins 'r' games is given by

$$P(x) = {}^nC_x p^x q^{n-x}$$

$$P(r) = {}^nC_r p^r q^{n-r} \quad [\because x = r]$$

$$P(r) = 5C_r \left(\frac{3}{5}\right)^r \left(\frac{2}{5}\right)^{n-r}$$

$$= \sum_{r=3}^{5} 5C_r \left(\frac{3}{5}\right)^r \left(\frac{2}{5}\right)^{5-r}$$

$$= 5C_3 \left(\frac{3}{5}\right)^3 \left(\frac{2}{5}\right)^2 + 5C_4 \left(\frac{3}{5}\right)^4 \left(\frac{2}{5}\right) + 5C_5 \left(\frac{3}{5}\right)^5 \left(\frac{2}{5}\right)^{5-5}$$

$$= 10(0.216)(0.16) + 5(0.1296)(0.4) + (1)(0.07776)(1)$$

$$P(r) = 0.68256$$

Moment generating function of negative binomial distribution (MGFNB)

$$M_x(t) = E[e^{tx}]$$

$$= \sum_{x=0}^{\infty} e^{tx} P(x)$$

$$= \sum_{x=0}^{\infty} e^{tx} \binom{x+r-1}{x} p^r q^x$$

$$= p^r \sum_{x=0}^{\infty} e^{tx} \binom{x+r-1}{x} q^r$$

$$= p^r \sum_{x=0}^{\infty} (qe^t)^x \binom{x+r-1}{x}$$

$$P = \frac{1}{Q}, \quad q = \frac{P}{Q}$$

$$P + q = 1$$

$$\frac{1}{Q}+\frac{P}{Q}=1 \Rightarrow \frac{1}{Q}=1-\frac{P}{Q}$$

$$\frac{1}{Q}=\frac{Q-P}{Q}$$

$$Q-P=1$$

By using MGF of the binomial distribution

$$M_X(t)=(q+pe^t)^n$$

$$M_X(t)=(Q-Pe^t)^{-r}$$

$$=\left(\frac{1}{p}-\frac{q}{p}e^t\right)^{-r}$$

$$=\left(\frac{1}{p}\right)^{-r}(1-qe^t)^{-r}$$

$$M_X(t)=p^r(1-qe^t)^{-r}$$

Cumulants of negative binomial distribution.

$$K_X(t)=\log(M_X(t))=\log(Q-Pe^t)^{-r}$$

$$=-r\log(Q-Pe^t)$$

$$=-r\log\left(Q-P\left(1+\frac{t}{1!}+\frac{t^2}{2!}+\cdots\right)\right)$$

$$=-r\log\left(Q-P-P\left(\frac{t}{1!}+\frac{t^2}{2!}+\cdots\right)\right)$$

$$=-r\log\left[Q-P-P\left(\frac{t}{1!}+\frac{t^2}{2!}+\cdots\right)\right]$$

$$K_X(t)=-r\log\left[1-P\left(\frac{t}{1!}+\frac{t^2}{2!}+\cdots\right)\right] \quad P=\frac{q}{p},\ Q=\frac{1}{p}$$

Additive property of negative binomial distribution.

Let $X_1, X_2, \ldots X_K$ be independent negative binomial i.e, $NB(r_i, P)$ random variable $i=1,\ldots K$ respectively

$X_K=\sum_{i=1}^{K} X_i$ is described as $NB(r_1+r_2+\cdots+r_K, P)$

r, p is a parameter

Proof:

$$P(X=x) = \binom{x+r-1}{x} p^r q^x$$

The MGF of negative binomial distribution is $M_X(t) = p^r(1-qe^t)^{-r}$

By Uniqueness Theorem of MGF

$$M_{X_K}(t) = M_{\sum_{i=1}^{k} x_i}(t)$$

$$= M\left[x_1(t) + x_2(t) + \cdots + x_k(t)\right]$$

$$= M_{x_1}(t) M_{x_2}(t) \cdots M_{x_k}(t)$$

$$= \left[p^{r_1}(1-qe^t)^{-r_1}\right]\left[p^{r_2}(1-qe^t)^{-r_2}\right] \cdots \left[p^{r_k}(1-qe^t)^{-r_k}\right]$$

$$= p^{r_1+r_2+\cdots+r_k} (1-qe^t)^{-(r_1+r_2+\cdots+r_k)}$$

$$= p^{\sum_{i=1}^{k} r_i} (1-qe^t)^{-\sum_{i=1}^{k} r_i}$$

$M_{X_K}(t) = \sum_{i=1}^{k} x_i$ is $NB(r_1+r_2+\cdots+r_k, p)$

Hence the proof.

Theorem

Let X and Y be independent random variable with probability mass function $NB(r_1, p)$ and $NB(r_2, p)$ respectively then the conditional probability mass function of X given $X+Y=t$ is expressed by

$$P\left(\frac{X=x}{x+y=t}\right) = \frac{\binom{x+r_1-1}{x}\binom{t+r_2-x-1}{t-x}}{\binom{t+r_1+r_2-1}{t}}$$

It is particullar $r_1 = r_2 = 1$ conditional distribution is uniform on $\frac{1}{t+1}$ points.

Proof

The additive property of negative binomial distribution $X+Y$ is (r_1+r_2, p)

i.e, $x+y \sim NB(r_1+r_2, p)$.

By conditional probability law,

$$\left(\frac{X=x}{X+Y=y}\right) = \frac{P[(X=x)(Y=t-x)]}{P(X+\cancel{Y}=y)}$$

$$= \frac{\binom{x+r_1-1}{x} p^{r_1} q^{x} \binom{t-x+r_2-1}{t-x} p^{r_2} q^{t-x}}{\binom{x+t-x+r+r_2-1}{x+t-1} p^{r_1+r_2} q^{x+t-x}}$$

$$\left(\frac{X=x}{X+Y=y}\right) = \frac{\binom{x+r_1-1}{x}\binom{t-x+r_2-1}{t-x}}{\binom{t+r_1+r_2-1}{t}}$$

If $r_1 = r_2 = 1$

$$P\left(\frac{X=x}{X+Y=y}\right) = \frac{\binom{x+1-1}{x}\binom{t-x+1-1}{t-x}}{\binom{t+1+1-1}{t}}$$

$$= \frac{\binom{x}{x}\binom{t-x}{t-x}}{\binom{t+1}{t}} = \frac{1}{\binom{t+1}{t}}$$

$$\therefore P\left(\frac{X=x}{X+Y=t}\right) = \frac{t!}{(t+1)!} = \frac{t!}{(t+1)t!} = \frac{1}{t+1}$$

Practice Questions:

2 Marks:

1. Define Moment Generating Function
2. Find r th moment x about origin
3. Write the Characteristic function of binomial distribution
4. 10 coins are thrown simultaneously, find out the probability of getting atleast 7 heads.
5. Define Binomial distribution

5 Marks:

1. Addition Theorem for Moment Generating function
2. State and Prove Uniqueness theorem for Moment generating function
3. Explain the effect of change of origin and scale on moment generating function
4. Explain Mode of a binomial distribution
5. Additive property of negative binomial distribution

10 Marks:

1. State and prove Chebyshev's Inequality
2. Explain Recurrence relation for the probability of binomial distribution of fitting of a binomial distribution

References:

S.C. Gupta and V.K. Kapoor, Fundamentals of Mathematical Statistics, Sultan Chand & Co, New Delhi, Reprint 2019.