BITS Pilani
K K Birla Goa Campus

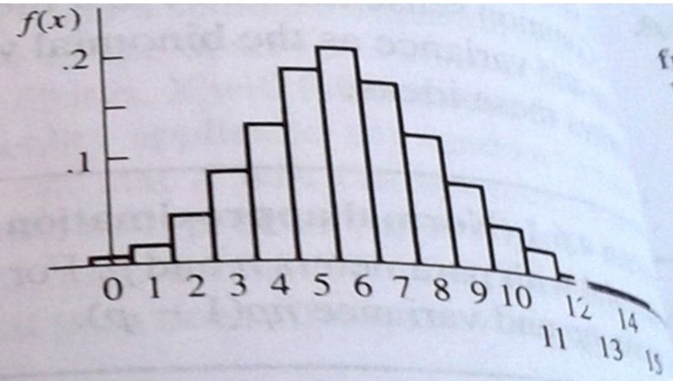
Probability & Statistics,

Dr. Jajati Keshari Sahoo
Department of Mathematics

Normal to binomial distribution

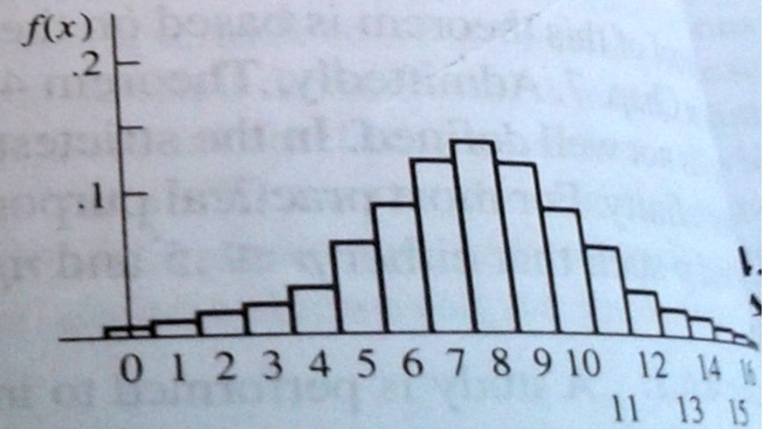
$n = 15$, $p = .4$

x	$f(x)$	x	$f(x)$
0	.0005	8	.1181
1	.0047	9	.0612
2	.0219	10	.0245
3	.0634	11	.0074
4	.1268	12	.0016
5	.1859	13	.0003
6	.2066	14	~ 0
7	.1771	15	~ 0



$n = 20$, $p = .4$

x	$f(x)$	x	$f(x)$
0	~ 0	11	.0710
1	.0005	12	.0355
2	.0031	13	.0145
3	.0124	14	.0049
4	.0350	15	.0013
5	.0746	16	.0003
6	.1244	17	~ 0
7	.1659	18	~ 0
8	.1797	19	~ 0
9	.1597	20	~ 0
10	.1172		



Normal to binomial distribution

Correction for continuity:

We take 0.5(half unit) for correction, i.e.,

Binomial	$\cong$	Normal
$P(X<b)$	$\Leftrightarrow$	$P(X<b-0.5)$
$P(X\leq b)$	$\Leftrightarrow$	$P(X<b+0.5)$
$P(X>a)$	$\Leftrightarrow$	$P(X>a+0.5)$
$P(X\geq a)$	$\Leftrightarrow$	$P(X>a-0.5)$
$P(a<X<b)$	$\Leftrightarrow$	$P(a+0.5<X<b-0.5)$

Normal Approximation to the Binomial Distribution

The normal distribution can be used to approximate the binomial distribution when n is large.

Theorem: Let X be binomial with parameter n and p. For large n, X is approximately normal with mean np and variance $np(1-p)$.

Question: **How to define the largeness of *n* in practice ?**

Normal Approximation to the Binomial Distribution

A good rule of thumb for the normal approximation

For most practical purposes the approximation is acceptable for values of n & p such that

Either $p \leq 0.5$ & $np > 5$

Or $p > 0.5$ & $n(1-p) > 5$.

Normal Approximation to the Binomial Distribution

Problem 1: If a random variable has the binomial distribution with $n = 30$ and $p = 0.60$, use the normal approximation to determine the probabilities that it will take on

(a) a value less than 12;

(b) the value 14;

(c) a value greater than 16.

Solution: $\mu = np = 18$; $\sigma^2 = np(1 - p) = 7.2$; $\sigma = 2.6833$

$$\text{(a)}\; P(X < 12) = P(X < 11.5) \quad \text{(using continuity correction)}$$

$$= F\left(\frac{11.5 - 18}{2.6833}\right) = F(-2.42) = 0.0078.$$

Normal Approximation to the Binomial Distribution

$$(b)\ P(X=14)=P(13.5<X<14.5)\quad \text{(using continuity correction)}$$

$$=F\left(\frac{14.5-18}{2.6833}\right)-F\left(\frac{13.5-18}{2.6833}\right)$$

$$=F(-1.3044)-F(-1.677)$$

$$=0.0961-0.0468=0.0493.$$

$$(c)\ P(X>16)=P(X>16.5)\quad \text{(using continuity correction)}$$

$$=1-P(X\le 16.5)=1-F\left(\frac{16.5-18}{2.6833}\right)$$

$$=1-F(-0.559)=F(0.559)=0.7120.$$

Problem 2

Item produced by a certain manufacturer is independently of acceptable quality with probability 0.95. Find the probability that at most 10 of the next 150 item produced are unacceptable.

Problem 3

One thousand independent rolls of a fair die will be made. Compute an approximation to the probability that number 6 will appear between 150 and 200 times inclusively. If number 5 appears exactly 200 times, find the probability that number 6 will appear less than 150 times.

Problem 4

In 10,000 independent tosses of a coin, the coin landed heads 5800 times. Is it reasonable to assume that the coin is not fair? Explain.

Normal Probability Rule

Theorem: Let X be normally distributed with mean μ and variance σ^2. Then for any $k > 0$, $P(|X - \mu| < k\sigma) = 2F(k) - 1$.

Question: If X is not Normal, then how to get information about these probabilities ?

Chebyshev's Inequality

Statement: Let X be any random variable with mean μ and variance σ^2. Then for any $k > 0$,

$$\boxed{P(|X-\mu| < k\sigma) \ge 1 - \frac{1}{k^2}}.$$

or

$$\boxed{P(|X-\mu| \ge k\sigma) < \frac{1}{k^2}.}$$

Chebyshev's Inequality

Proof: (Case-I: If X is discrete r.v.)

$$\sigma^2 = \sum_x (x-\mu)^2 f(x)$$

$$= \sum_{x=x_1}^{\mu-\sqrt{c}} (\cdot) + \sum_{\mu-\sqrt{c}}^{\mu+\sqrt{c}} (\cdot) + \sum_{\mu+\sqrt{c}} (\cdot)$$

$$\geq \sum_{x=x_1}^{\mu-\sqrt{c}} (\cdot) + \sum_{\mu+\sqrt{c}} (\cdot)$$

Chebyshev's Inequality

Proof: (Cont...)

$$\sigma^2 \geq c\sum_{x=x_1}^{\mu-\sqrt{c}} f(x) + c\sum_{\mu+\sqrt{c}} f(x)$$

$$\geq c\left[P(X \leq \mu - \sqrt{c}) + P(X \geq \mu + \sqrt{c})\right]$$

$$\geq c\left[1 - P(|X - \mu| < \sqrt{c})\right]$$

Chebyshev's Inequality

Proof: (Cont...)

$$P(|X-\mu| < \sqrt{c}) \geq 1 - \frac{\sigma^2}{c}$$

Put $c = k^2\sigma^2$, then

$$\boxed{P(|X-\mu| < k\sigma) \geq 1 - \frac{1}{k^2}.}$$

Problem 5

If *X* is a random variable such that *E(X)=3* and *V(X)=13*.

(a) Determine the lower bound for

$$P(-2 < X < 8).$$

(b) Find the upper bound for

$$P(X \le -2 \text{ and } X \ge 8).$$

Problem 6

Does there exist a random variable X in which

$$P(\mu - 2\sigma < X < \mu + 2\sigma) = 0.6.$$

Problem 7

A fair coin is tossed 1000 times. Find a lower bound of the probability that the proportion of head differs from 0.5 by less than 0.1.