3.4 The Binomial probability distribution

~~There are many~~ There are many experiments that conform either exactly or ~~appropriately~~ approximately to the following list of requirements:

1. The experiment consists of a sequence of n smaller experiments called <u>trials</u>, where n is fixed in advance of the experiment.
2. Each trial can result in one of the same two possible outcomes, which we ~~generally~~ generically denote by success (S) and failure (F).
3. The trials are independent, so that the outcome on any particular trial does not influence the outcome on any other trial.
4. The ~~prob~~ probability of success $P(S)$ is constant from trial to trial; we denote this probability by p.

<u>Defn.</u> An experiment for which conditions 1 – 4 are satisfied is called a <u>binomial experiment.</u>

~~Defn~~: <u>The Binomial Random variable and Distribution</u>.

<u>Defn</u>: The <u>binomial random variable X</u> associated with a binomial experiment consisting of n trials is defined as

$X =$ The number of S's among the n trials.

Notation: Because the pmf of a binomial rv X depends on the two parameters n and p, we denote the pmf by $b(x; n, p)$

Defn: pmf of binomial (n, p) distribution is given by

$$b(x; n,p) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x} & , x = 0, 1, 2, 3, 4, \ldots n. \\ 0 & \text{otherwise} \end{cases}$$

The Mean and variance of the rv X having binomial distribution with parameters n & p.

$$E(X) = np$$

$$V(X) = np(1-p)$$

$$S.d = \sigma_X = \sqrt{np(1-p)} \quad (>0).$$

<u>Proposition</u> If $X \sim Bin(n, p)$, Then

$E(X) = np$, $V(X) = np(1-p) = npq$,

and $\sigma_X = \sqrt{npq}$ (where $q = 1-p$).

<u>Exercises Section 3.4.</u>

46. a. $b(3; 8, .35) = \binom{8}{3} (.35)^3 (1 - .35)^{8-3}$

$$= \frac{8 \times 7 \times 6}{2 \times 3} (.35)^3 (.65)^5$$

$$= .2786$$

b. $.2787$,

c. $P(3 \le X \le 5) = 0.7451$,

d. $P(1 \le X) = 0.6126$.

50. Let X = number of incoming calls that involve fax messages.
By given condition, $n = 25$, $p = .25$ and
$X \sim b(x; 25, .25)$.

a. $$P(X \le 6) = \sum_{x=0}^{6} b(x; 25, .25)$$
$$= 0.5611$$

b. $$b(6; 25, .25) = \binom{25}{6} (.25)^6 (.75)^{19}$$
$$= 0.1828.$$

c. $$P(X \geq 6) = 1 - P(X < 6)$$
$$= 1 - P(X \le 5)$$
$$= 1 - 0.3783$$
$$= 0.6217$$

d. $$P(X > 6) = 1 - P(X \le 6) = 1 - 0.5611$$
$$= 0.4389.$$

51. a. $E(X) = np = 25 \times \frac{1}{4} = 6.25$

b. $V(X) = np(1-p) = 6.25 \times .75 = 4.6875$
$\sigma_x = \sqrt{4.6875} = 2.165.$

c. To exceed the expected number by more than 2 standard deviations means that
$$X > E(X) + 2.\sigma_x,$$
$$P(X > E(X) + 2.\sigma_x) = P(X > 10.58)$$

$= 1 - P(x \le 10.58)$

$= 1 - P(x \le 10) = 1 - 0.97 = 0.03.$

59. Denote

$x =$ number of houses with a fire detector.

$x \sim Bin(25, p).$

a. Denote

$R = \{$ rejecting claim when $p = .8\}$

We are given that we reject the claim when $x \le 15$, we actually need to calculate probability that $x \le 15$ when $p = .8$, Therefore

$$P(R) = P(x \le 15 \text{ and } p = .8)$$

$$= \sum_{x=0}^{15} \binom{25}{x} (.8)^x (.2)^{25-x}$$

$$= B(15; 25, .8) = 0.017.$$

b. $P(x > 15$ and $p = .7)$

$$= 1 - B(15; 25, .7)$$

$$= 1 - 0.189 = 0.811$$

$P(x > 15$ and $p = .6)$

$$= 1 - 0.575 = 0.425.$$

c. $P(x \le 14$ and $p = .8) = .006.$

$P(x > 14$ and $p = .7) = .902$

$P(x > 14$ and $p = .6) = .586.$

62. a. $V(X) = np(1-p)$

$\therefore V(X) = 0$ if $p = 0$ or if $p = 1$.

In first case, when $p=0$, implies that every trial fails, so the variability in X is zero.

When $p=1$, every trial succeeds, so again there is no variability in X.

b. $$\frac{d}{dp}V(X) = \frac{d}{dp}(np - np^2)$$

$$= n - 2np = n(1-2p).$$

Since, $\frac{d^2}{dp^2}V(X) = -2n < 0$

the value p for which

$$n(1-2p) = 0$$

is the ~~maximum~~ maxima.

$$n(1-2p) = 0 \Rightarrow p = 1/2 = .5 \; (n > 0).$$

So for $p = .5$, $V(X)$ is maximum.

Prob (64) Show that $E(X) = np$, $V(X) = np(1-p)$, when X is a binomial random variable.

Ans: $$E(X) = \sum_{x=0}^{n} x\, b(x; n, p)$$

$$= \sum_{x=0}^{n} x \binom{n}{x} p^x (1-p)^{n-x}$$

$$= \sum_{x=1}^{n} x \frac{n!}{x!\,(n-x)!} p^x (1-p)^{n-x}$$

$$= \sum_{x=1}^{n} \frac{np\,(n-1)!}{(x-1)!\,\{(n-1)-(x-1)\}!}\, p^{x-1} (1-p)^{\{(n-1)-(x-1)\}}$$

$[\because x \geq 1]$

$$= np \sum_{k=0}^{n_1} \frac{n_1!}{k!\,(n_1-k)!}\, p^{k} (1-p)^{n_1-k}$$

Put $x-1 = k$, $n-1 = n_1$

$$= np\,(p + 1-p)^{n_1}$$

[using binomial expansion taking $a = p$ and $b = 1-p$]

$$= np \cdot 1 = np \text{ (Proved).}$$

Now, to find $V(X)$ we use the following ~~fo~~ shortcut formula :

$$V(X) = E(X(X-1)) - E(X)\,(E(X)-1).$$

$$E(X(X-1)) = \sum_{x=0}^{n} x(x-1) \binom{n}{x} p^{x} (1-p)^{n-x}$$

$$= \sum_{x=2}^{n} x(x-1) \frac{n!}{x!\,(n-x)!}\, p^{2} \cdot p^{\cancel{x-2}} (1-p)^{n-x}$$

$$= \sum_{x=2}^{n} \frac{n(n-1)\,(n-2)!}{(x-2)!\,((n-2)-(x-2))!} \cdot p^{2} \cdot p^{x-2} \cdot (1-p)^{n-x}$$

[Put $n-2 = n_2$, $x-2 = k$]

$$= \sum_{k=0}^{n_2} n(n-1)p^2 \frac{n_2!}{k!\,(n_2-k)!}\, p^{k} (1-p)^{n_2-k}$$

$$= n(n-1)p^2 \sum_{k=0}^{n_2} \frac{n_2!}{k!\,(n_2-k)!}\, p^{k} (1-p)^{n_2-k}$$

$[\because n_2 - x = n_2 - k]$

$$= n(n-1)p^2 \quad (p+1-p)^{n-2} \quad \text{[Using binomial expansion]}$$

$$= n(n-1)p^2$$

$$\therefore V(X) = n(n-1)p^2 - np(np-1)$$

$$= \cancel{n^2p^2} - np^2 - \cancel{n^2p^2} + np$$

$$= np - np^2$$

$$= np(1-p).$$

~~67. Refer to Chebyshev's inequality given in Exercise 44. Calculate~~ $\cancel{P(|X-\mu| \geq K\sigma)}$

A result called Chebyshev's inequality states that for any probability distribution of an r.v. X and any number K that is atleast 1, $P(|X-\mu| \geq K\sigma) \leq \frac{1}{K^2}$.

Putting $K\sigma = t$, we get $P(|X-\mu| \geq t) \leq \sigma^2/t^2$.

Ex. The r.v. Y has the prob.^{ty} distribution as given in the follow table.

y	45	46	47	48	49	50	51	52	53	54	55
$p(y)$	0.05	0.10	0.12	0.14	0.25	0.17	.06	.05	.03	.02	.01

Find $E(Y)$, $\sigma = SD(Y)$ and $P(|Y-\mu| \geq t)$ for $t = 2\sigma, 3\sigma$.

Sol: $$E(Y) = \sum_{y=45}^{55} y\,p(y) = 48.84.$$

$$V(Y) = E(Y^2) - (E(Y))^2$$

$$= \sum_{y=45}^{55} y^2 p(y) - (48.84)^2$$

$$= 2389.84 - (48.84)^2 = 4.4944 = \sigma^2$$

$$\Rightarrow \sigma = 2.12 \ (>0).$$

By Chebyshev's inequality, we have

$$P(|Y-\mu| \geq t) \leq \frac{\sigma^2}{t^2}.$$

for $t = 2\sigma$,

a) $$P[\,|Y-\mu| \geq 2\sigma\,] \leq \frac{\sigma^2}{(2\sigma)^2} = \frac{1}{4}$$

$$P[\,|Y-\mu| \geq 2\sigma\,]$$

$$= P[\,Y-\mu \geq 2\sigma\,] + P(Y-\mu \leq -2\sigma)$$

$$= P[\,Y \geq \mu + 2\sigma\,] + P(Y \leq \mu - 2\sigma)$$

$$= P[\,Y \geq 53.08\,] + P[\,Y \leq 44.60\,]$$

$$= P(Y=54) + P(Y=55) + 0$$

$$= 0.02 + 0.01 + 0 = 0.03.$$

(b) Similarly find $P[\,|Y-\mu| \geq 3\sigma\,]$ [Task].