2004-10-07

STATS 3N03/3J04

11-1

BINOMIAL DISTRIBUTION

NOTATION $\quad BIN(n, p)$

EXAMPLE

LET X BE THE NUMBER OF HEADS IN 100 INDEPENDENT TOSSES OF A FAIR COIN.

THEN $X \sim BIN(100, \frac{1}{2})$

" IS DISTRIBUTED AS "

PMF

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

$$x = 0, 1, \dots, n$$

MEAN

$$\mu = E(X) = \sum_{x=0}^{n} x f(x)$$

$$= \sum_{x=0}^{n} x \binom{n}{x} p^x q^{n-x} = \dots = np$$

VARIANCE

$$\sigma^2 = \mathrm{Var}(X) = E\left[(X-\mu)^2\right]$$

$$= \sum_{x=0}^{n} (x-np)^2 \binom{n}{x} p^x q^{n-x}$$

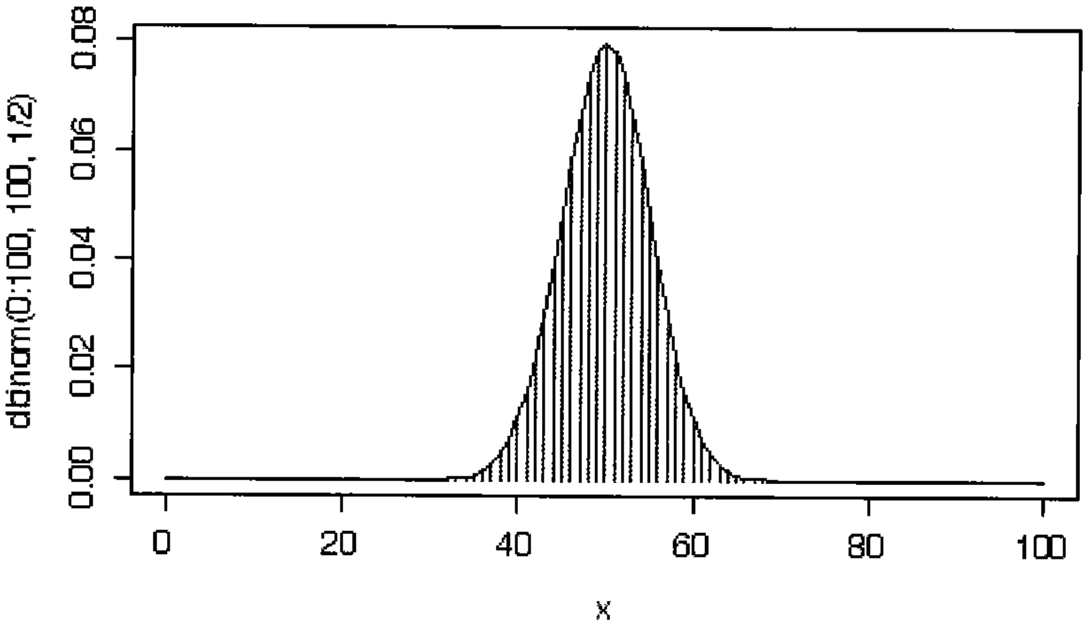
$$= \cdots = npq$$

RELATION TO OTHER DISTRIBUTIONS

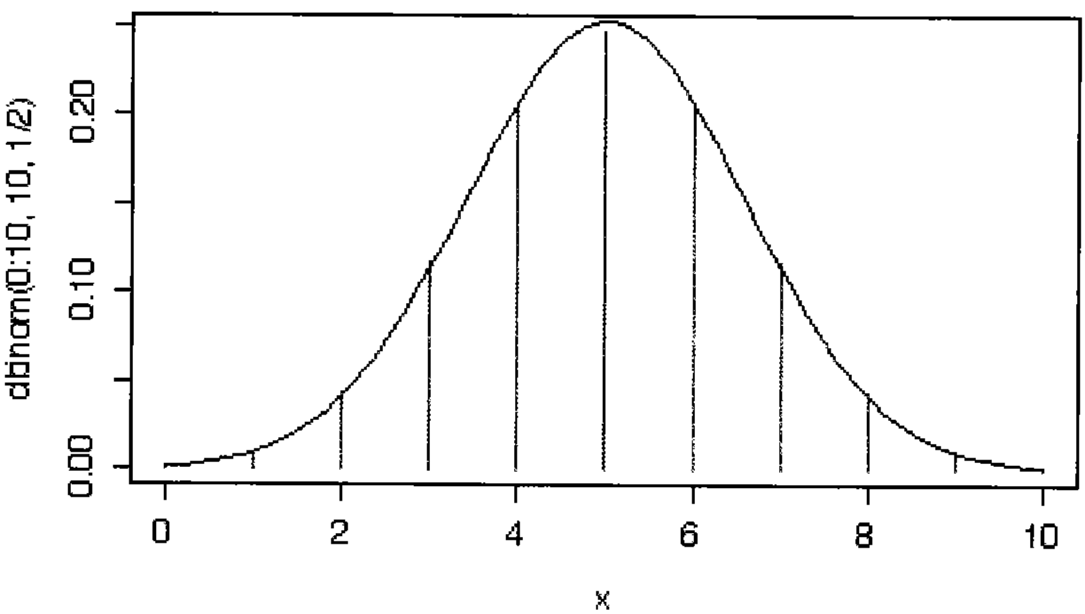
WHEN n IS LARGE AND p IS NOT TOO CLOSE TO 0 OR 1, A $\mathrm{BIN}(n,p)$ DISTRIBUTION CAN BE APPROXIMATED BY A NORMAL DISTRIBUTION WITH MEAN $= np$ AND VARIANCE $= npq$.

```
> plot(0:100, dbinom(0:100,100,1/2), type="h", col="blue",
xlab="x")
> lines(0:100,dnorm(0:100, 50, 5), col="red")
```

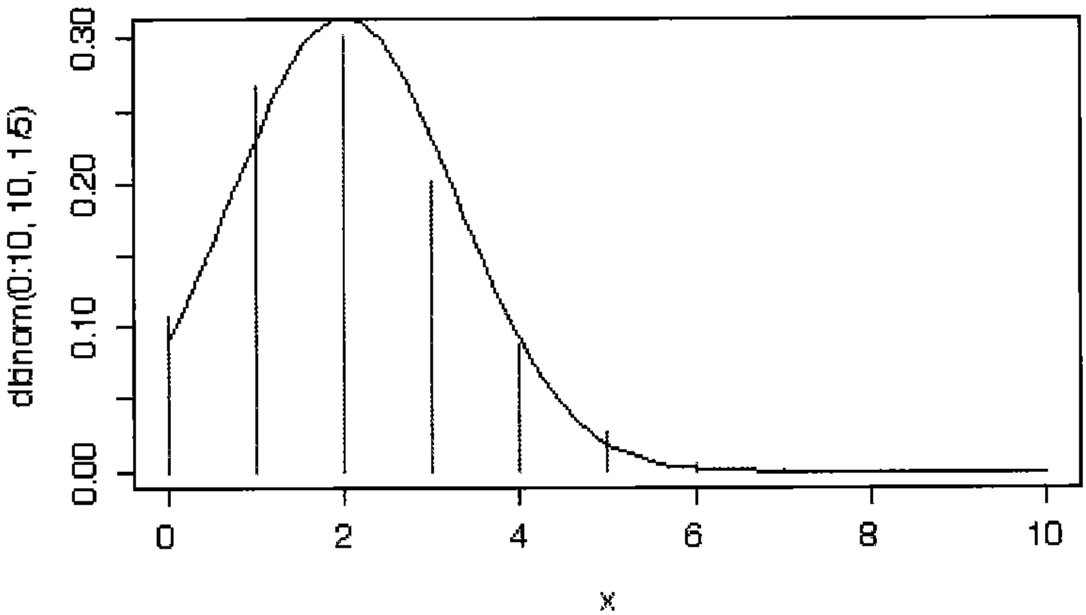
$5 = \sqrt{100 \frac{1}{2} \cdot \frac{1}{2}}$



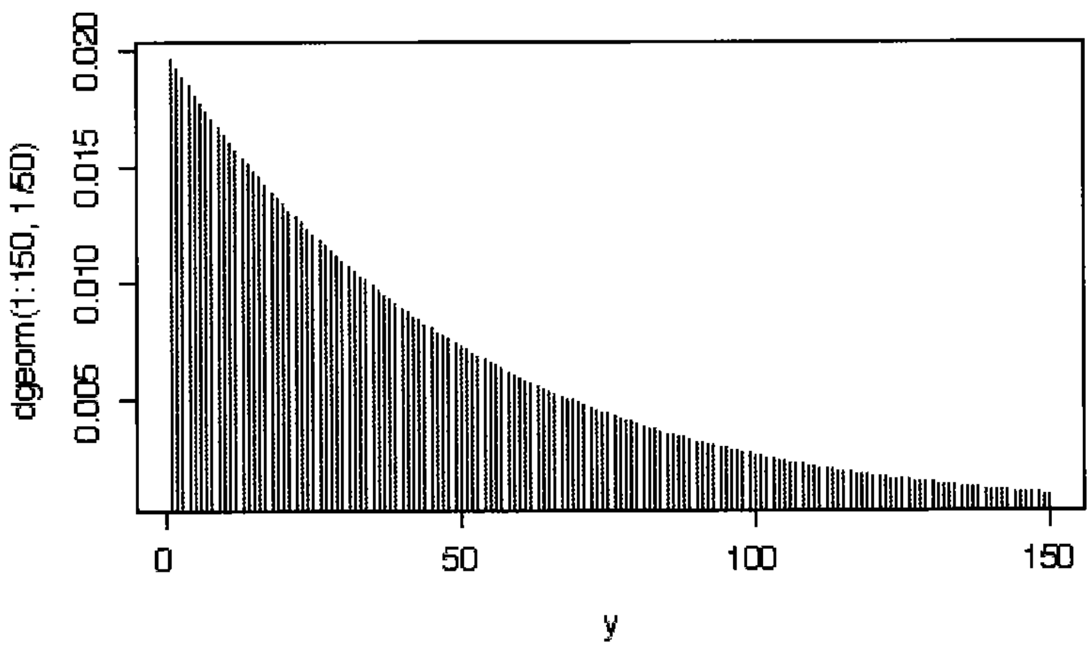

```
> plot(0:10, dbinom(0:10,10,1/2), type="h", col="blue",
xlab="x")
> lines(seq(0,10,len=60), dnorm(seq(0,10,len=60),5,sqrt(2.5)),
col="red")
```



```
> plot(0:10,dbinom(0:10,10,1/5),type="h",col="blue", xlab="x")
> lines(seq(0,10,len=60),dnorm(seq(0,10,len=60), 2, sqrt(1.6)),
col="red")
```




```
> plot(1:150, dgeom(1:150,1/50), type="h", xlab="y", col="blue")
```



GEOMETRIC DISTRIBUTION

SEQUENCE OF INDEPENDENT TRIALS, SAME PROBABILITY p OF "SUCCESS" ON EACH TRIAL.

RANDOM VARIABLE:

Y = NUMBER OF TRIALS UP TO AND INCLUDING THE FIRST "SUCCESS"

PDF

$$f(y) = p\, q^{y-1} \qquad y = 1, 2, \ldots$$

RELATION TO BINOMIAL

BOTH HAVE SEQUENCE OF INDEPENDENT TRIALS, SAME PROBABILITY OF "SUCCESS" ON EACH TRIAL.

BINOMIAL: FIXED NO. OF TRIALS, RANDOM NUMBER OF "SUCCESSES".

GEOMETRIC: FIXED NUMBER OF "SUCCESSES", RANDOM NUMBER OF TRIALS.

MOMENTS:

$$\sum_{y=1}^{\infty} f(y) = p + pq + pq^2 + \cdots$$

$$= p(1 + q + q^2 + \cdots)$$

$$= p(1-q)^{-1} = \frac{p}{p} = 1$$

$$\mu = E(Y) = \sum_{y=1}^{\infty} y f(y)$$

$$= p + 2pq + 3pq^2 + \cdots = \cdots = \frac{1}{p}$$

$$\sigma^2 = Var(Y) = \sum_{y=1}^{\infty} \left(y - \frac{1}{p}\right)^2 pq^{y-1} = \cdots = \frac{q}{p^2}$$

EX RETURN PERIOD (IN YEARS) FOR A "50-YR WIND"

$$\mu = 50 \iff p = \frac{1}{50}$$

11-7

$$\sigma^2 = \frac{q}{p^2} = \frac{\frac{49}{50}}{\left(\frac{1}{50}\right)^2} = (49)(50) = 2450$$

$$\sigma = \sqrt{2450} = 49.497$$

Exercises:

3-65 SHOW THAT $X \sim Bin(50, .1)$

$$\therefore P(X \le 2) = (.9)^{50} + 50(.1)^1(.9)^{49} + \binom{50}{2}(.1)^2(.9)^{48}$$

$$= 0.112$$

dbinom(0:2, 50, .1)

0.00515 0.02863 0.07794

pbinom(2, 50, .1)

0.11173