(L6) Probability & Statistics

Bernoulli Distribution.

$S = \{ \text{success}, \text{failure} \}$.

$X(s) = 1 , \quad X(f) = 0$

The pmf is $f(x) = p^x (1-p)^{1-x}, \; x \in \{0, 1\}$

$\mu = p , \quad \sigma^2 = pq = p(1-p)$.

Binomial Distribution.

'n' - independent Bernoulli trials.

$$\boxed{X(\cdots\cdots) = X(\underbrace{1\,0\,0\,1 \cdots 1}_{n \text{ times}})}$$

We are interested to study total number of successes.

Q. What is the probability of having 'x' successes. $x = 1, 2, 3 \ldots$
$x \leq n$

The pmf is

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

$$x = 0, 1, 2, \ldots$$

↓

binomial distribution

denoted by $B(n, p)$ or $b(n, p)$.

<u>Exp</u>. Suppose that the probability of germination of a beet seed is 0.8 and the germination of a seed is called a success. Suppose we plant 10 seeds and we assume that the germination of one seed is independent of the germination of another seed.

Let X denote the # of seeds germinated.

Q. Find $P(X \leq 6) = ?$

<u>Ans</u>. 0.1209

Q. Find Mgf of $B(n,p)$.

$$M(t) = E(e^{tX})$$
$$= \sum_{x=0}^{n} e^{tx} \binom{n}{x} p^x (1-p)^{n-x}$$
$$= \sum_{x=0}^{n} \binom{n}{x} (pe^t)^x (1-p)^{n-x}$$
$$= [(1-p) + pe^t]^n, \quad -\infty < t < \infty.$$

$$\mu = M'(0) = np$$
$$\sigma^2 = M''(0) - [M'(0)]^2$$
$$= np(1-p).$$

Setting $n=1$, $\underbrace{\mu = p, \ \sigma^2 = p(1-p)}_{\text{for Bernoulli.}}$

Another explanation of $B(n,p)$

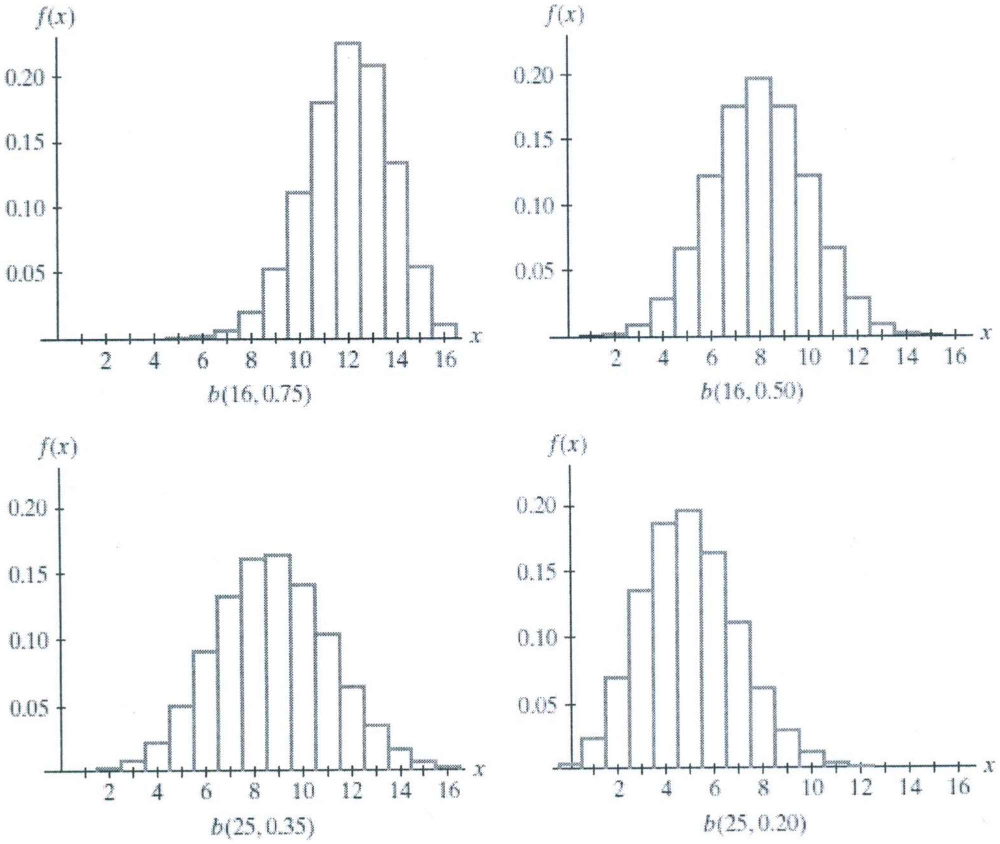
$$p = \frac{N_1}{N_1 + N_2}$$

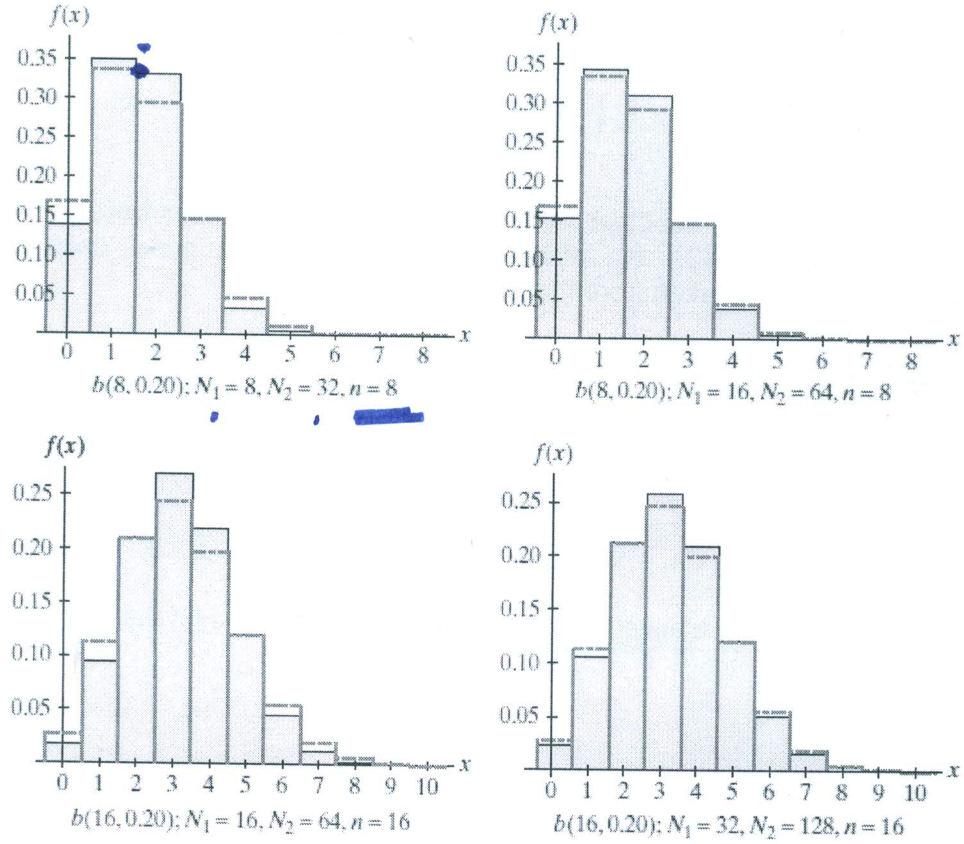
X = the # of success balls in a random sample of size n that is taken from this urn.

Then X follows hypergeometric distribution with pmf

$$f(x) = \frac{\binom{N_1}{x}\binom{N_2}{n-x}}{\binom{N_1+N_2}{n}}.$$

when the balls are picked "without replacement"


f(x)
0.20
0.15
0.10
0.05
2 4 6 8 10 12 14 16 x
b(16, 0.75)
b(16, 0.50)
b(25, 0.35)
b(25, 0.20)



Binomial and hypergeometric (shaded) probability histograms

When balls are picked with replacement then.

$X \approx B(n, p)$.

However binomial is close to hypergeometric when n is small and $N = N_1 + N_2$ is large.

~~Negative bi~~

Negative binomial

Observe a sequence of independent Bernoulli trials untill exactly r successes occur, where 'r' is a fixed positive number.

Let X be a random variable which denotes the # of trials needed to observe the rth success.

The pmf is

$$f(x) = \binom{x-1}{r-1} p^r q^{x-r}, \quad x = r, r+1, \ldots$$

mgf of -ve binomial:

$$M(t) = E(e^{tX}).$$

$$= \sum_{x=r}^{\infty} e^{tx} \binom{x-1}{r-1} p^r (1-p)^{x-r}$$

$$= \frac{(pe^t)^r}{[1-(1-p)e^t]^r}$$ where $(1-p)e^t < 1$.

$$\mu = \frac{r}{p}$$

$$\sigma^2 = \frac{r(r+1-p)}{p^2} - \frac{r^2}{p^2} = \frac{r(1-p)}{p^2}$$

<u>why is it called '-ve'</u> binomial?

Consider the function

$$h(b) = (1-b)^{-r}, \quad -1 < b < 1$$

Then $$\frac{1}{(1-b)^r} = \sum_{k=0}^{\infty} \frac{h^{(k)}(0)}{k!} b^k$$

$$(1-b)^{-r} = \sum_{k=0}^{\infty} \binom{r+k-1}{r-1} b^k .$$

If we set $x = k + \cancel{r}$

i.e. $k = x - r$

Then $$(1-b)^{-r} = \sum_{\cancel{x}=r}^{\infty} \binom{r+x-r-1}{r-1} b^{x-r}$$

$$= \sum_{x=r}^{\infty} \binom{x-1}{r-1} b^{x-r} \quad (1)$$

except the factor p^r, the -ve binomial (pmf) looks like (1) when $b = 1-p$.

Connection to geometric distribution.

Setting $r = 1$,
we obtain

$$f(x) = p(1-p)^{x-1}, \quad x = 1, 2, 3, \ldots$$

Quiz. Suppose that during practice a basketball player can make a free throw 80% of the time. Furthermore, assume that a sequence of free-throw shooting can be thought of as independent Bernoulli trials.

Let X equal the minimum # of free-throws that this player must attempt to make a total of 10 shots.

Find $P(X = 12)$.

<u>Poisson Process</u>

Let the # of occurences of some event in a given continuous interval be counted. Then we have a Poisson process with parameter $\lambda > 0$ if the following are satisfied.

a) the # of occurrences in nonoverlapping subintervals are independent.

b) The probability of exactly one occurrence in a sufficiently short sub-interval of length h is approx. λh

c) the prob. of two or more occurrences in a sufficiently short interval is zero.

$$P(X=x) = \binom{n}{x}\left(\frac{\lambda}{n}\right)^{x}\left(1-\frac{\lambda}{n}\right)^{n-x}$$

$$\lim_{n\to\infty} P(X=x) = \frac{\lambda^{x}e^{-\lambda}}{x!}$$

(H.W.) $\cancel{= \frac{n!}{x!(n-x)!}}$

Poisson distribution.

where we partition the unit interval into n subintervals of equal length $\frac{1}{n}$. Then consider that n is large.

- the prob. of one occurence in any one subinterval of length $1/n$ is $\lambda \cdot \frac{1}{n}$
- the prob. of two or more occurences in any one subinterval is zero.