Chapter 2 Discrete Distributions

- Random Variables of the Discrete Type
  - Uniform Distribution
  - Hypergeometric Distribution
- Mathematical Expectation
- Moment Generating Function
- Bernoulli Trials and the Binomial Distribution
- Geometric and Negative Binomial Distribution
- The Poisson Distribution

The Moment-Generating Function

Definition 2.3-1

Let X be a random variable of the discrete type with pmf $f(x)$ and space S. If there is a positive number h such that

$$E(e^{tX}) = \sum_{x \in S} e^{tx} f(x)$$

exists and is finite for $-h < t < h$, then the function defined by

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

is called the **moment-generating function of** X (or of the distribution of X). This function is often abbreviated as mgf.

- Generating Property
- Uniqueness Property

$$M(t) = E(e^{tX}) = \sum_{x \in S} e^{tx} f(x)$$

The Moment-Generating Function

- Let X be a discrete random variable with p.m.f $P_X(x)$ and space S.
- If there is a positive number h such that

$$E\left[e^{tX}\right] = \sum_{x_i \in S} e^{t\,x_i} \; P_X(x_i)$$

exists and is finite for -h < t < h.

- Then, the function of t defined by $M(t) \equiv E[e^{tX}]$ is called the moment-generating function of X, and often abbreviated as m.g.f.

The Moment-Generating Function

- Let X and Y be two discrete random variables defined on the same space S.

If $E\left[e^{tX}\right] = E\left[e^{tY}\right]$,

then the probability mass functions of X and Y are equal.

- Insight of the argument above：

Assume that S = { s_1, s_2, … } contains only positive integers.

Then, we have

$$P_X(s_1)e^{ts_1} + P_X(s_2)e^{ts_2} + \ldots$$
$$= P_Y(s_1)e^{ts_1} + P_Y(s_2)e^{ts_2} + \ldots \quad .$$

Therefore, $P_X(s_1) = P_Y(s_1)$, i.e. X and Y have the same p.m.f.

(by mathematical transform theory)

Example 2.3-5 If X has the m.g.f.

$$M(t) = e^t\left(\frac{3}{6}\right) + e^{2t}\left(\frac{2}{6}\right) + e^{3t}\left(\frac{1}{6}\right),$$

then the probabilities are

$$P(X = 1) = \frac{3}{6}, \qquad P(X = 2) = \frac{2}{6}, \qquad P(X = 3) = \frac{1}{6}.$$

We can write this, if we choose to do so, by saying X has the p.m.f.

$$f(x) = \frac{4 - x}{6}, \qquad x = 1, 2, 3.$$

▲

$$M(t) = E(e^{tX}) = \sum_{x \in S} e^{tx} f(x)$$

$$M'(t) = \sum_{x \in S} x e^{tx} f(x),$$

$$M''(t) = \sum_{x \in S} x^2 e^{tx} f(x), \dashrightarrow M^{(r)}(t) = \sum_{x \in S} x^r e^{tx} f(x).$$

Setting $t = 0$, we see that

$$M'(0) = \sum_{x \in S} x f(x) = E(X),$$

$$M''(0) = \sum_{x \in S} x^2 f(x) = E(X^2),$$

and, in general,

$$M^{(r)}(0) = \sum_{x \in S} x^r f(x) = E(X^r).$$

Moment-Generating Function

- Let $M_X(t)$ be the m.g.f of a discrete random variable X.

$$\frac{d^k M_X(t)}{dt^K} = \sum_{x_i \in S} x_i^{\,k} e^{tx_i} P_X(x_i).$$

Furthermore,

$$\frac{d^k M_X(0)}{dt^K} = \sum_{x_i \in S} x_i^{\,k} P_X(x_i). = E\left[X^k\right].$$

- In particular,

$$\boxed{\mu_X = M_X{}'(0)\, and\; \sigma^2 = M_X{}''(0) - \left[M_X{}'(0)\right]^2.}$$

Chapter 2 Discrete Distributions

- Random Variables of the Discrete Type
  - Uniform Distribution
  - Hypergeometric Distribution
- Mathematical Expectation
- Moment Generating Function
- Bernoulli Trials and the Binomial Distribution
- Geometric and Negative Binomial Distribution
- The Poisson Distribution

Hypergeometric distribution （超幾何分布）

-- select n objects from N_1+N_2 objects
(e.g., red chips and blue chips)

$\boxed{x \le n,\ x \le N_1, \text{ and } n - x \le N_2}$

the probability of selecting exactly x red chips is

$$f(x) = P(X = x) = \frac{\binom{N_1}{x}\binom{N_2}{n-x}}{\binom{N}{n}}$$

若n=1，<u>超幾何分布</u>可以簡化為<u>伯努利分布</u>

Bernoulli Distribution

- Bernoulli experiment
  - a random experiment
  - the outcome of which can be classified in one of two mutually exclusive and exhaustive ways, say, success and failure.
- A sequence of Bernoulli trials occurs
  - when a Bernoulli experiment is performed several independent times
  - and the probability of success, say p, remains the same from trial to trial.

Definition ***Bernoulli Random Variable:*** X is a *Bernoulli* random variable if the PMF of X has the form

$$P_X(x) = \begin{cases} 1-p & x=0 \\ p & x=1 \\ 0 & \text{otherwise} \end{cases}$$

where the parameter p is in the range $0 < p < 1$.

Bernoulli Distribution

$$f(x) = p^x(1-p)^{1-x}, \qquad x = 0, 1,$$

Example 2.4-1 Suppose that the probability of germination of a beet seed is 0.8 and the germination of a seed is called a success. If we plant 10 seeds and can assume that the germination of one seed is independent of the germination of another seed, this would correspond to 10 Bernoulli trials with $p = 0.8$.

- Let X be a Bernoulli random variable.
  - The p.m.f of X can be written as

$$P_X(k) = p^k (1-p)^{1-k},$$

where $k = 0$ or 1 and p is the probability of success.

- The expected value of X is

$$\sum_{k=0}^{1} x p^k (1-p)^{1-k} = p.$$

- The variance of X is

$$\sum_{k=0}^{1} (k-p)^2 p^k (1-p)^{1-k} = p(1-p).$$

X has a **Bernoulli distribution**. The expected value of X is

$$\mu = E(X) = \sum_{x=0}^{1} x\, p^x (1-p)^{1-x} = (0)(1-p) + (1)(p) = p,$$

and the variance of X is

$$\sigma^2 = \mathrm{Var}(X) = \sum_{x=0}^{1} (x-p)^2 p^x (1-p)^{1-x}$$

$$= (0-p)^2(1-p) + (1-p)^2 p = p(1-p) = pq.$$

It follows that the standard deviation of X is

$$\sigma = \sqrt{p(1-p)} = \sqrt{pq}.$$

A sequence of Bernoulli trials

Example 2.4-3 Out of millions of instant lottery tickets, suppose that 20% are winners. If five such tickets are purchased, then $(0, 0, 0, 1, 0)$ is a possible observed sequence in which the fourth ticket is a winner and the other four are losers. Assuming independence among winning and losing tickets, we observe that the probability of this outcome is

X_i

$$(0.8)(0.8)(0.8)(0.2)(0.8) = (0.2)(0.8)^4.$$

Example 2.4-4 If five beet seeds are planted in a row, a possible observed sequence would be $(1, 0, 1, 0, 1)$ in which the first, third, and fifth seeds germinated and the other two did not. If the probability of germination is $p = 0.8$, the probability of this outcome is, assuming independence,

Y_i

$$(0.8)(0.2)(0.8)(0.2)(0.8) = (0.8)^3(0.2)^2.$$

Binomial Distribution

- Let X be the random variable corresponding to the number of successes in a sequence of Bernoulli trials.
- Then,

$$P_X(k) = \text{Prob}(X = k) = C_k^n p^k (1-p)^{n-k},$$

where n is the number of Bernoulli trials and p is the probability of success in one trial.

- X is said to have a binomial distribution and is normally denoted by b(n, p).

a binomial experiment satisfies the following properties:

1. A Bernoulli (success-failure) experiment is performed n times.
2. The trials are independent. (with replacement)
3. The probability of success on each trial is a constant p; the probability of failure is $q = 1 - p$.
4. The random variable X equals the number of successes in the n trials.

Example 2.4-5 In the instant lottery with 20% winning tickets, if X is equal to the number of winning tickets among $n = 8$ that are purchased, then the probability of purchasing two winning tickets is

$$f(2) = P(X = 2) = \binom{8}{2}(0.2)^2(0.8)^6 = 0.2936.$$

The distribution of the random variable X is $b(8, 0.2)$. ■

Example 2.4-6 In order to obtain a better feeling for the effect of the parameters n and p on the distribution of probabilities, four probability histograms are displayed in Figure 2.4-1. ■

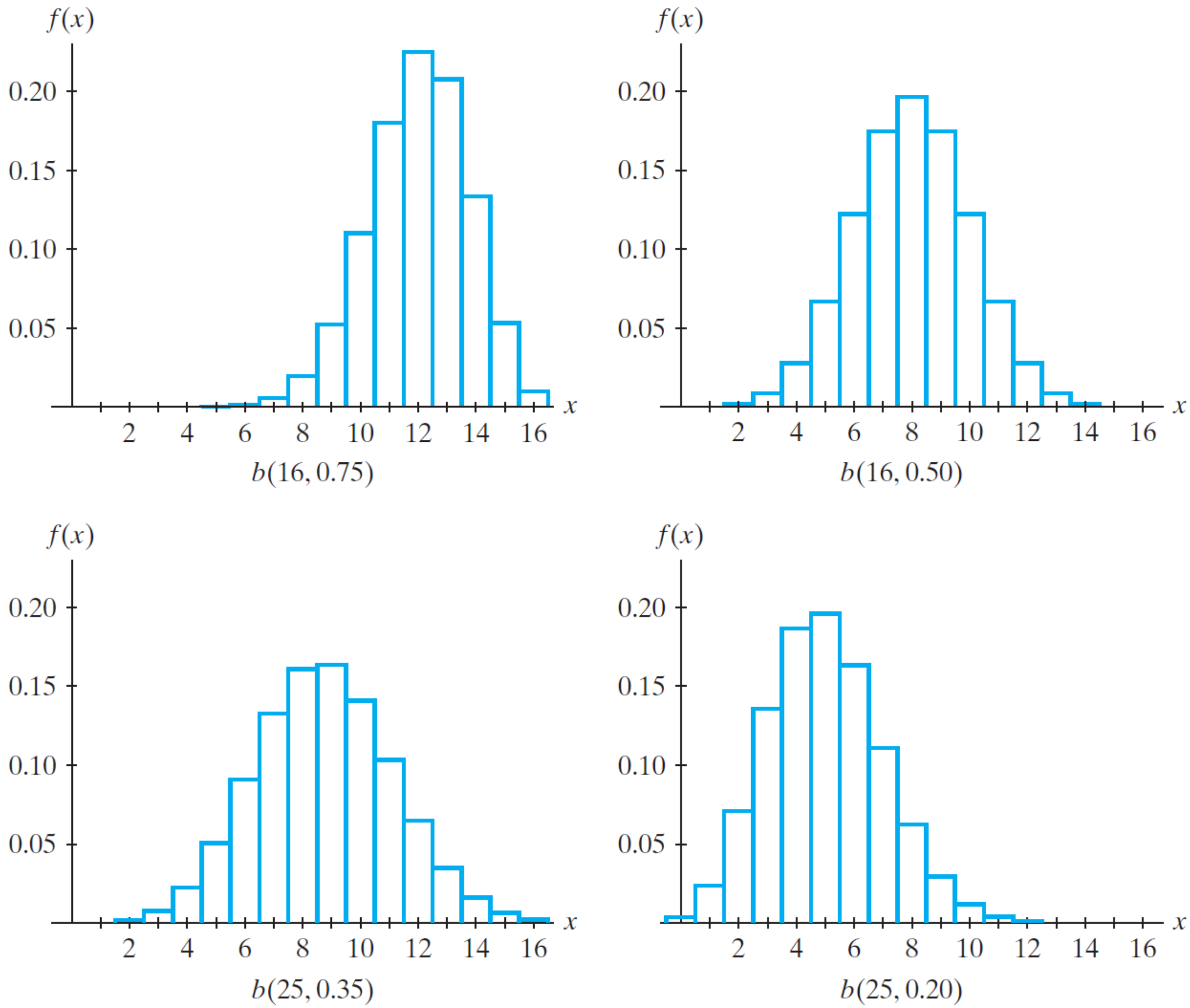


Figure 2.4-1 Binomial probability histograms

Example 2.4-1 Suppose that the probability of germination of a beet seed is 0.8 and the germination of a seed is called a success. If we plant 10 seeds and can assume that the germination of one seed is independent of the germination of another seed, this would correspond to 10 Bernoulli trials with $p = 0.8$. ■

Example 2.4-7 In Example 2.4-1, the number X of seeds that germinate in $n = 10$ independent trials is $b(10, 0.8)$; that is,

$$f(x) = \binom{10}{x}(0.8)^x(0.2)^{10-x}, \qquad x = 0, 1, 2, \ldots, 10.$$

In particular,

$$\begin{aligned} P(X \le 8) &= 1 - P(X = 9) - P(X = 10) \\ &= 1 - 10(0.8)^9(0.2) - (0.8)^{10} = 0.6242. \end{aligned}$$

Also, with a little more work, we could compute

$$P(X \le 6) = \sum_{x=0}^{6}\binom{10}{x}(0.8)^x(0.2)^{10-x} = 0.1209.$$ ■

Properties of Probability Mass Function

$$P_X(k) \equiv Prob(X = k) = \sum_{q \in Q_k} Prob(q),$$

- The p.m.f. of a random variable X satisfies the following three properties:

$$(1)\ P_X(x) > 0\ ,\ x \in S\ :\ the\ space\ of\ X.$$

$$(2)\ \sum_{x_i \in S} P_X(x_i) = 1\ .$$

$$(3)\ Prob(A) = \sum_{x_j \in A} P_X(x_j)\ ,\ where\ A \subseteq S.$$

Def: Cumulative Distribution Function

(Probability Distribution Function)

(Distribution Function)

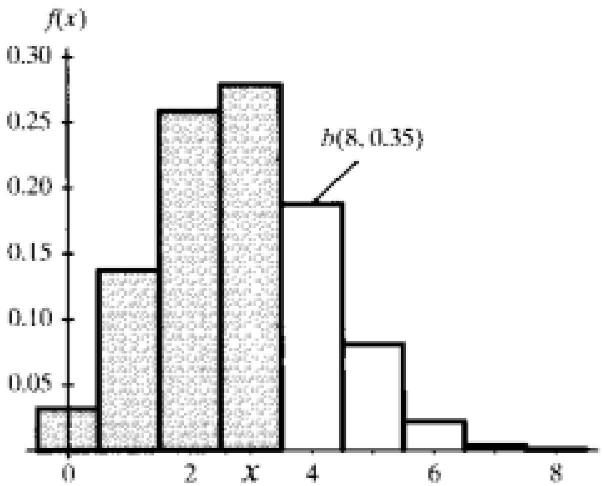
- For a random variable X, we define its cumulative distribution function F as

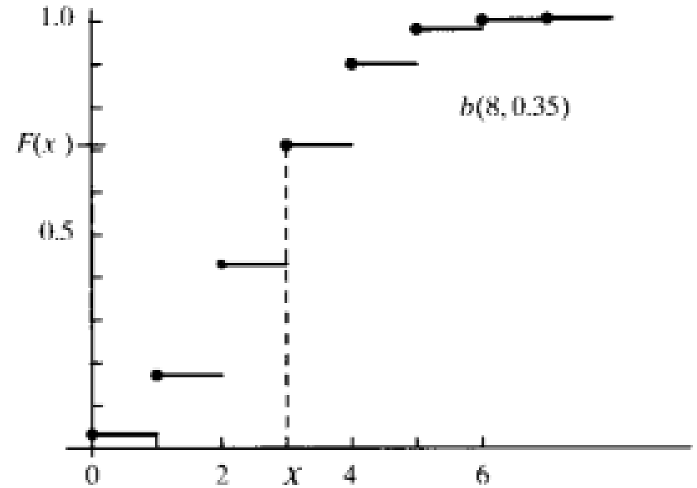
$$F_X(t) = Prob(X \le t)$$

-- Cumulative Distribution Function, CDF

Cumulative Distribution Function (Distribution Function)

Table II The Binomial Distribution





$$F(x) = P(X \le x) = \sum_{k=0}^{x} \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k}$$

For the binomial distribution given in Example 2.4-7, namely, the $b(10, 0.8)$ distribution, the distribution function is defined by

$$F(x) = P(X \le x) = \sum_{y=0}^{\lfloor x \rfloor} \binom{10}{y} (0.8)^y (0.2)^{10-y},$$

where $\lfloor x \rfloor$ is the greatest integer in x. A graph of this cdf is shown in Figure 2.4-2. Note that the vertical jumps at the integers in this step function are equal to the probabilities associated with those respective integers.

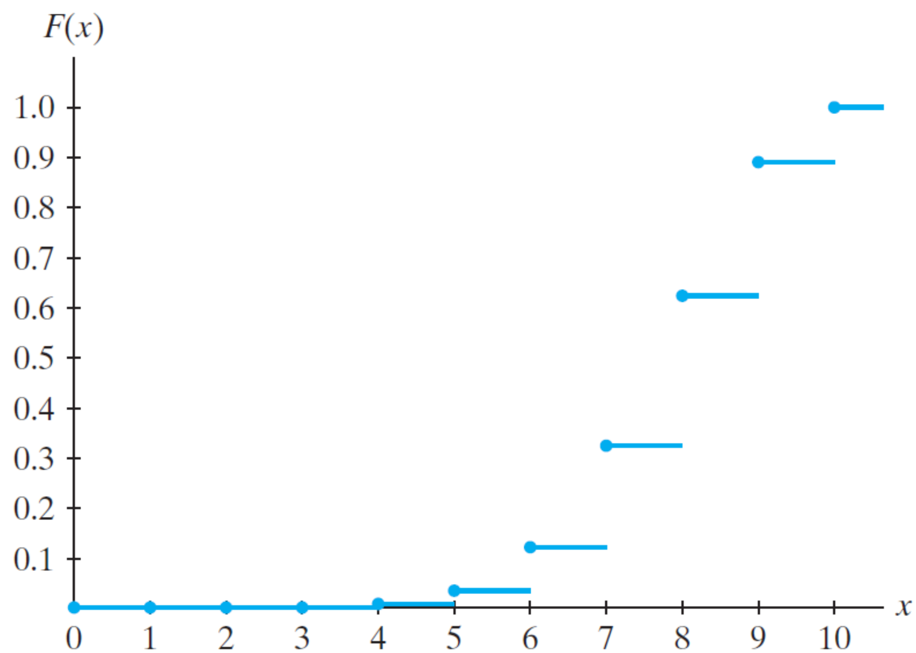


Figure 2.4-2 Distribution function for the $b(10, 0.8)$ distribution

$$F_X(t) = Prob(X \le t)$$

$$P_X(k) = \text{Prob}(X = k) = C_k^n p^k (1-p)^{n-k},$$

Table II *continued*

		p									
n	x	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
10	0	0.5987	0.3487	0.1969	0.1074	0.0563	0.0282	0.0135	0.0060	0.0025	0.0010
	1	0.9139	0.7361	0.5443	0.3758	0.2440	0.1493	0.0860	0.0464	0.0233	0.0107
	2	0.9885	0.9298	0.8202	0.6778	0.5256	0.3828	0.2616	0.1673	0.0996	0.0547
	3	0.9990	0.9872	0.9500	0.8791	0.7759	0.6496	0.5138	0.3823	0.2660	0.1719
	4	0.9999	0.9984	0.9901	0.9672	0.9219	0.8497	0.7515	0.6331	0.5044	0.3770
	5	1.0000	0.9999	0.9986	0.9936	0.9803	0.9527	0.9051	0.8338	0.7384	0.6230
	6	1.0000	1.0000	0.9999	0.9991	0.9965	0.9894	0.9740	0.9452	0.8980	0.8281
	7	1.0000	1.0000	1.0000	0.9999	0.9996	0.9984	0.9952	0.9877	0.9726	0.9453
	8	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9995	0.9983	0.9955	0.9893
	9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9997	0.9990
	10	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

p.494

$$P(X \leqq 5) = ?$$

Example 2.4-8 Leghorn chickens are raised for laying eggs. Let $p = 0.5$ be the probability that a newly hatched chick is a female. Assuming independence, let X equal the number of female chicks out of 10 newly hatched chicks selected at random. Then the distribution of X is $b(10, 0.5)$. From Table II in Appendix B, the probability of 5 or fewer female chicks is

$$P(X \le 5) = 0.6230.$$

The probability of exactly 6 female chicks is

$$\begin{aligned} P(X = 6) &= \binom{10}{6}\left(\frac{1}{2}\right)^6\left(\frac{1}{2}\right)^4 \\ &= P(X \le 6) - P(X \le 5) \\ &= \boxed{0.8281} - \boxed{0.6230} = 0.2051, \end{aligned}$$

since $P(X \le 6) = 0.8281$. The probability of at least 6 female chicks is

$$P(X \ge 6) = 1 - P(X \le 5) = 1 - 0.6230 = 0.3770.$$

- Table for p > 0.5 ?

Example 2.4-9 Suppose that we are in one of those rare times when 65% of the American public approve of the way the president of the United States is handling the job. Take a random sample of $n = 8$ Americans and let Y equal the number who give approval. Then, to a very good approximation, the distribution of Y is $b(8, 0.65)$. (Y would have the stated distribution exactly if the sampling were done with replacement, but most public opinion polling uses sampling without replacement.) To find $P(Y \geq 6)$, note that

$$P(Y \geq 6) = P(8 - Y \leq 8 - 6) = P(X \leq 2),$$

where $X = 8 - Y$ counts the number who disapprove. Since $q = 1 - p = 0.35$ equals the probability of disapproval by each person selected, the distribution of X is $b(8, 0.35)$. (See Figure 2.4-3.) From Table II in Appendix B, since $P(X \leq 2) = 0.4278$, it follows that $P(Y \geq 6) = 0.4278$.

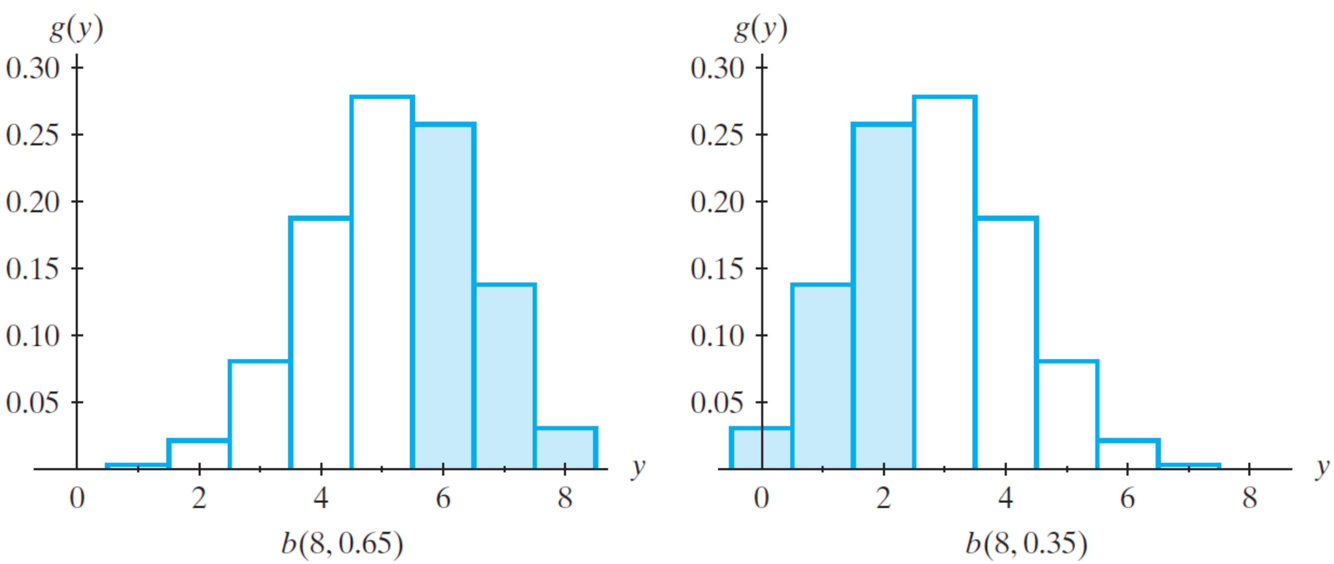


Figure 2.4-3 Presidential approval histogram

Similarly,

$$
\begin{aligned}
P(Y \le 5) &= P(8 - Y \ge 8 - 5) \\
&= P(X \ge 3) \;=\; 1 - P(X \le 2) \\
&= 1 - 0.4278 \;=\; 0.5722
\end{aligned}
$$

and

$$
\begin{aligned}
P(Y = 5) &= P(8 - Y = 8 - 5) \\
&= P(X = 3) \;=\; P(X \le 3) - P(X \le 2) \\
&= 0.7064 - 0.4278 \;=\; 0.2786.
\end{aligned}
$$

Recall that if n is a positive integer, then

$$(a+b)^n = \sum_{x=0}^{n} \binom{n}{x} b^x a^{n-x}.$$

Thus the sum of the binomial probabilities, if we use the above binomial expansion with $b = p$ and $a = 1 - p$, is

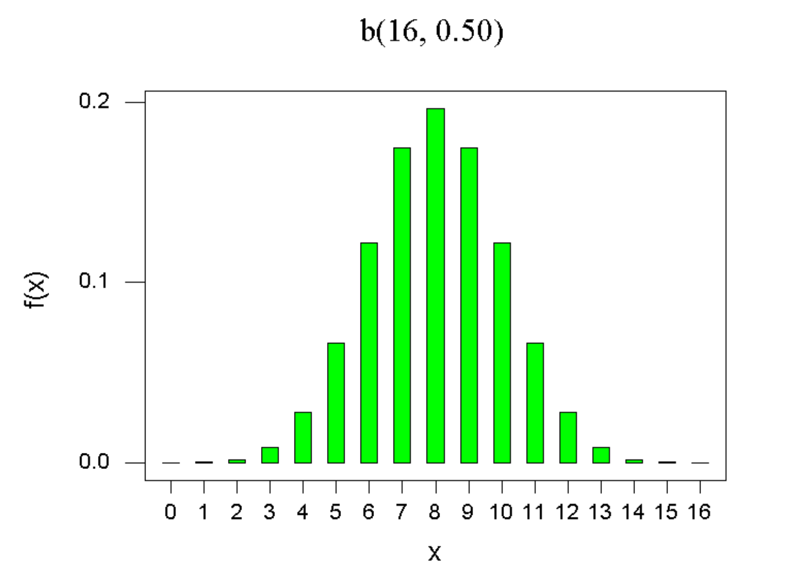
$$\sum_{x=0}^{n} \binom{n}{x} p^x (1-p)^{n-x} = [(1-p)+p]^n = 1,$$

a result that had to follow from the fact that $f(x)$ is a p.m.f.

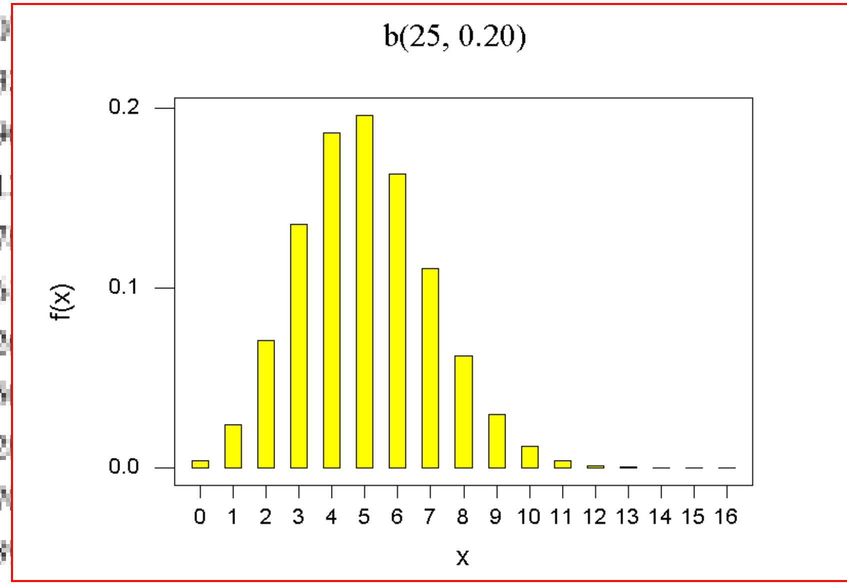
$\Rightarrow$ Satisfy the properties of a p.m.f.

Table II *continued*

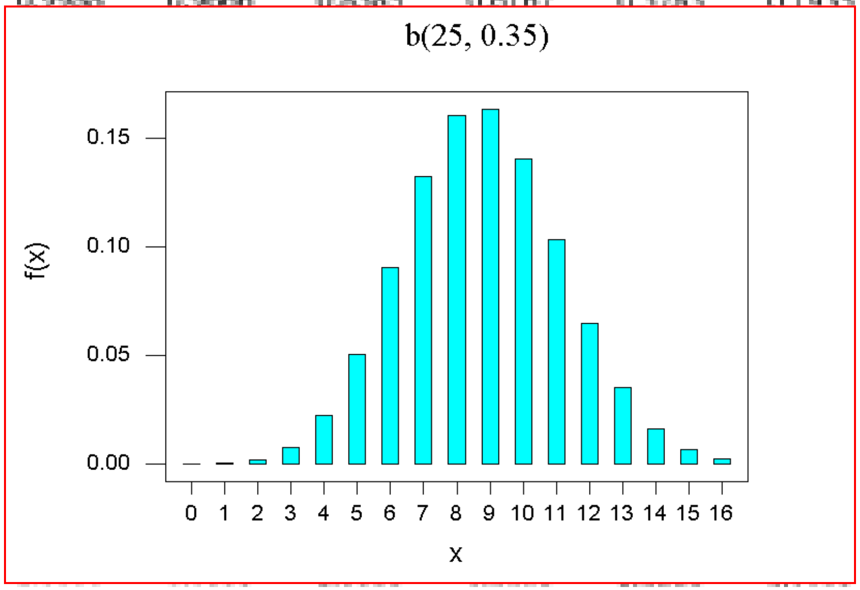
n	x	p = 0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
16	0	0.4401	0.1853	0.0743	0.0281	0.0100	0.0033	0.0010	0.0003	0.0001	0.0000
	1	0.8108	0.5147	0.2839	0.1407	0.0635	0.0261	0.0098	0.0033	0.0010	0.0003
	2	0.9571	0.7892						0.0183	0.0066	0.0021
	3	0.9930	0.9316						0.0651	0.0281	0.0106
	4	0.9991	0.9830						0.1666	0.0853	0.0384
	5	0.9999	0.9967						0.3288	0.1976	0.1051
	6	1.0000	0.9995						0.5272	0.3660	0.2272
	7	1.0000	0.9999						0.7161	0.5629	0.4018
	8	1.0000	1.0000						0.8577	0.7441	0.5982
	9	1.0000	1.0000						0.9417	0.8759	0.7728
	10	1.0000	1.0000						0.9809	0.9514	0.8949
	11	1.0000	1.0000						0.9951	0.9851	0.9616
	12	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9998	0.9991	0.9965	0.9894
	13	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9994	0.9979
	14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9997
	15	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	16	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000



25	0	0.2774	0.0718	0.0172	0.0038	0.0008	0.0001	0.0000	0.0000	0.0000	0.0000
	1	0.6424	0.2712	0.0931	0.0274	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	2	0.8729	0.5371	0.2537	0.0982	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	3	0.9659	0.7636	0.4711	0.2340	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	4	0.9928	0.9020	0.6821	0.4207	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	5	0.9988	0.9666	0.8385	0.6167	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	6	0.9998	0.9905	0.9305	0.7800	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	7	1.0000	0.9977	0.9745	0.8909	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	8	1.0000	0.9995	0.9920	0.9532	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	9	1.0000	0.9999	0.9979	0.9827	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	10	1.0000	1.0000	0.9995	0.9944	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	11	1.0000	1.0000	0.9999	0.9985	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	12	1.0000	1.0000	1.0000	0.9996	0.9966	0.9825	0.9396	0.8462	0.6937	0.5000
	13	1.0000	1.0000	1.0000	0.9999	0.9991	0.9940	0.9745	0.9222	0.8173	0.6550
	14	1.0000	1.0000	1.0000	1.0000	0.9998	0.9982	0.9907	0.9656	0.9040	0.7878
	15	1.0000	1.0000	1.0000	1.0000	1.0000	0.9995	0.9971	0.9868	0.9560	0.8852
	16	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9992	0.9957	0.9826	0.9461
	17	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9998	0.9988	0.9942	0.9784
	18	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9997	0.9984	0.9927
	19	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9996	0.9980
	20	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9995
	21	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999
	22	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	23	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	24	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	25	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000



25	0	0.2774	0.0718	0.0172	0.0038	0.0008	0.0001	0.0000	0.0000	0.0000	0.0000
	1	0.6424	0.2712	0.0931	0.0274	0.0070	0.0016	0.0003	0.0001	0.0000	0.0000
	2	0.8729	0.5371	0.2537	0.0982	0.0321	0.0090	0.0021	0.0004	0.0001	0.0000
	3	0.9659	0.7636	0.4711	0.2340	0.0962	0.0332	0.0097	0.0024	0.0005	0.0001
	4	0.9928	0.9020	0.6821	0.4207	0.2137	0.0905	0.0320	0.0095	0.0023	0.0005
	5	0.9988	0.9666	0.8385	0.6167	0.3783	0.1935	0.0826	0.0294	0.0086	0.0020
	6							0.1734	0.0736	0.0258	0.0073
	7							0.3061	0.1536	0.0639	0.0216
	8							0.4668	0.2735	0.1340	0.0539
	9							0.6303	0.4246	0.2424	0.1148
	10							0.7712	0.5858	0.3843	0.2122
	11							0.8746	0.7323	0.5426	0.3450
	12							0.9396	0.8462	0.6937	0.5000
	13							0.9745	0.9222	0.8173	0.6550
	14							0.9907	0.9656	0.9040	0.7878
	15							0.9971	0.9868	0.9560	0.8852
	16							0.9992	0.9957	0.9826	0.9461
	17	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9998	0.9988	0.9942	0.9784
	18	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9997	0.9984	0.9927
	19	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9996	0.9980
	20	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9995
	21	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999
	22	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	23	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	24	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	25	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000



Mean and Variance of the Binomial Distribution ?

- Let X be b(n , p).

$$E[X] = \sum_{k=0}^{n} k C_k^n p^k (1-p)^{n-k}$$

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

$$E[X^2] = \sum_{k=0}^{n} k^2 C_k^n p^k (1-p)^{n-k}$$

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

are both difficult to compute. *(but can be solved iteratively, as in [Ross], p.144.)*

Example 3.3-11 Consider the simple experiment of flipping a fair coin five independent times. If X equals the number of heads that is observed, then X is $b(5, 0.5)$, $\mu = 2.5$, $\sigma^2 = 1.25$, and $\sigma = 1.118$. This experiment was simulated 100 times, yielding the following data:

How?

2	3	2	4	1	2	1	1	4	2	4	2	0	4	4	2	4	4	3	4
2	2	4	4	1	1	3	3	1	4	2	3	1	2	4	1	2	5	3	2
4	3	2	2	2	3	5	2	0	3	2	1	3	4	2	2	4	0	2	1
3	3	2	3	2	1	3	2	2	2	1	1	3	3	1	1	4	2	1	5
3	2	3	0	3	5	3	2	4	3	3	5	2	3	3	1	3	2	1	1

For these data $\bar{x} = 2.47$, $s^2 = 1.5243$, and $s = 1.235$. In Figure 3.3-3 the probability histogram and the relative frequency histogram (shaded) are given. ▲

$$\mu = np \qquad \text{and} \qquad \sigma^2 = np(1-p) = npq.$$

How?

(use mgf)

$$\sigma = \sqrt{np(1-p)}.$$

Figure 3.3-1: Binomial Probability Histograms, $b(n, 0.25)$, $n = 2, 4, \ldots, 32$.

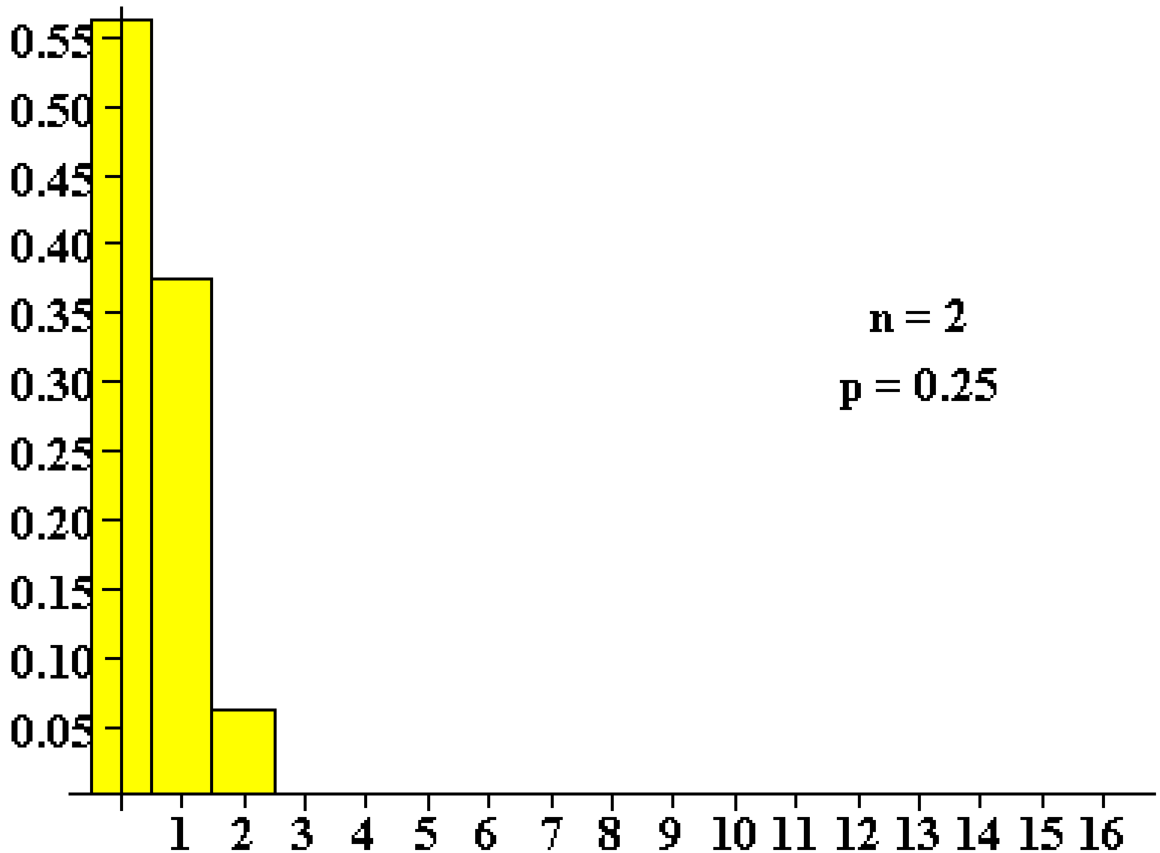


Figure 3.3-1: Binomial Probability Histograms, $b(10, p)$, $p = 0.05, 0.10, \ldots, 0.95$.

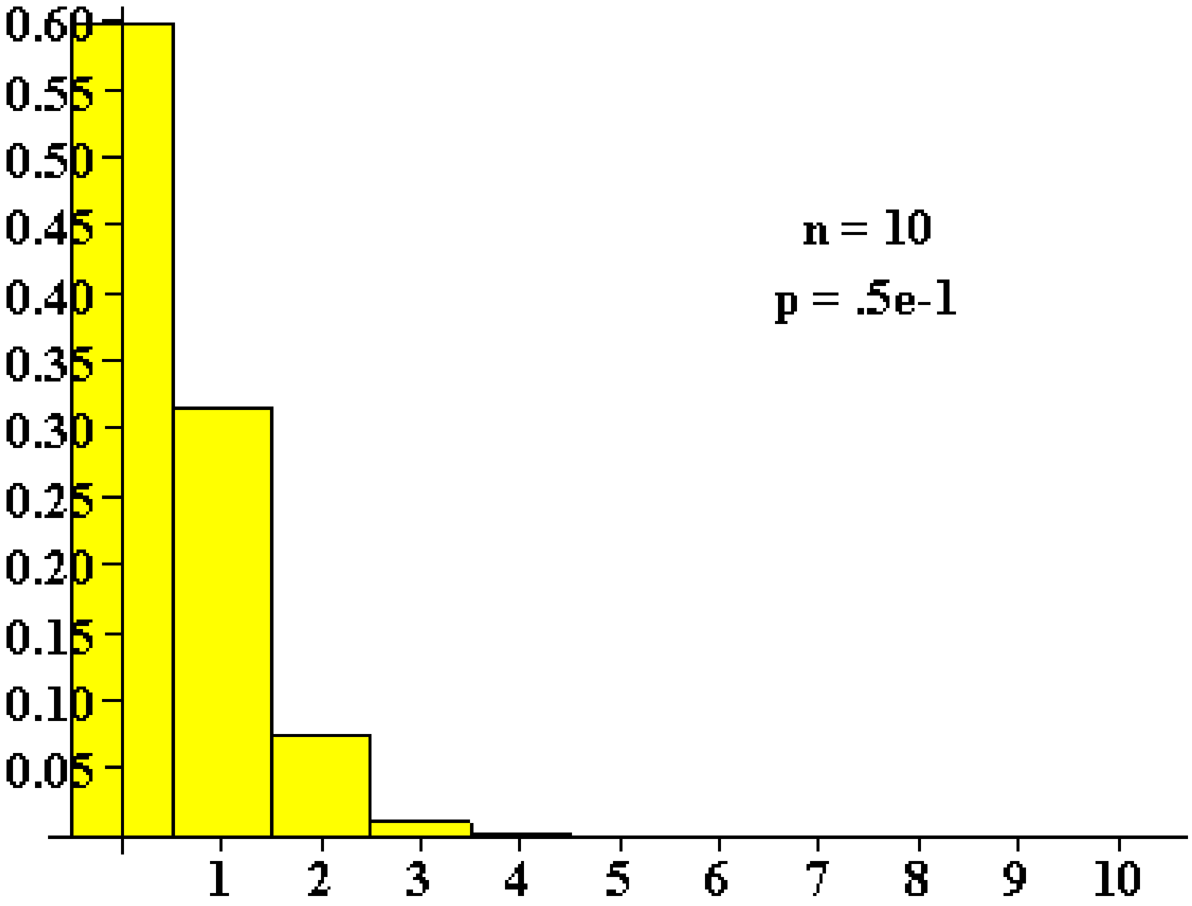


Table II *continued*

		p									
n	x	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
10	0	0.5987	0.3487	0.1969	0.1074	0.0563	0.0282	0.0135	0.0060	0.0025	0.0010
	1	0.9139	0.7361	0.5443	0.3758	0.2440	0.1493	0.0860	0.0464	0.0233	0.0107
	2	0.9885	0.9298	0.8202	0.6778	0.5256	0.3828	0.2616	0.1673	0.0996	0.0547
	3	0.9990	0.9872	0.9500	0.8791	0.7759	0.6496	0.5138	0.3823	0.2660	0.1719
	4	0.9999	0.9984	0.9901	0.9672	0.9219	0.8497	0.7515	0.6331	0.5044	0.3770
	5	1.0000	0.9999	0.9986	0.9936	0.9803	0.9527	0.9051	0.8338	0.7384	0.6230
	6	1.0000	1.0000	0.9999	0.9991	0.9965	0.9894	0.9740	0.9452	0.8980	0.8281
	7	1.0000	1.0000	1.0000	0.9999	0.9996	0.9984	0.9952	0.9877	0.9726	0.9453
	8	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9995	0.9983	0.9955	0.9893
	9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9997	0.9990
	10	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Moment-Generating Function of Binomial Distribution

- Let X be b（n , p）.

$$E[X]=\sum_{k=0}^{n} kC_k^n p^k(1-p)^{n-k}$$

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

$$E[X^2]=\sum_{k=0}^{n} k^2C_k^n p^k(1-p)^{n-k}$$

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

are both difficult to compute.

On the other hand, we can easily derive the m.g.f. of a binomial distribution.

$$M_X(t) = E\left[e^{tX}\right] = \sum_{k=0}^{n} e^{tk} C_k^n p^k (1-p)^{n-k}$$

$$= \sum_{k=0}^{n} C_k^n \left(pe^t\right)^k (1-p)^{n-k} = \left[pe^t + (1-p)\right]^n.$$

Moment-Generating Function of the Binomial Distribution

$$M_X'(t) = n\left[pe^t + (1-p)\right]^{n-1} pe^t$$

$$M_X''(t) = n(n-1)\left[pe^t + (1-p)\right]^{n-2}\left(pe^t\right)^2 + npe^t\left[pe^t + (1-p)\right]^{n-1}$$

$$M_X'(0) = np$$

$$M_X''(0) = n(n-1)p^2 + np.$$

Therefore,

$$\mu_X = M_X'(0) = np$$

$$\sigma_X^{\ 2} = M_X''(0) - \left[M_X'(0)\right]^2$$

$$= n^2p^2 - np^2 + np - n^2p^2$$

$$= np(1\text{-}p).$$

Binomial Distribution

$$P_X(k) = \text{Prob}(X = k) = C_k^n p^k (1-p)^{n-k},$$

What is its relation with hypergeometric distribution?

?

Suppose that an urn contains N_1 success balls and N_2 failure balls. Let $p = N_1/(N_1 + N_2)$, and let X equal the number of success balls in a random sample of size n that is taken from this urn. If the sampling is done one at a time with replacement, then the distribution of X is $b(n, p)$; if the sampling is done without replacement, then X has a hypergeometric distribution with pmf

$$f(x) = \frac{\dbinom{N_1}{x}\dbinom{N_2}{n-x}}{\dbinom{N_1+N_2}{n}},$$

where x is a nonnegative integer such that $x \le n$, $x \le N_1$, and $n - x \le N_2$. When $N_1 + N_2$ is large and n is relatively small, it makes little difference if the sampling is done with or without replacement. In Figure 2.4-4, the probability histograms are compared for different combinations of n, N_1, and N_2.

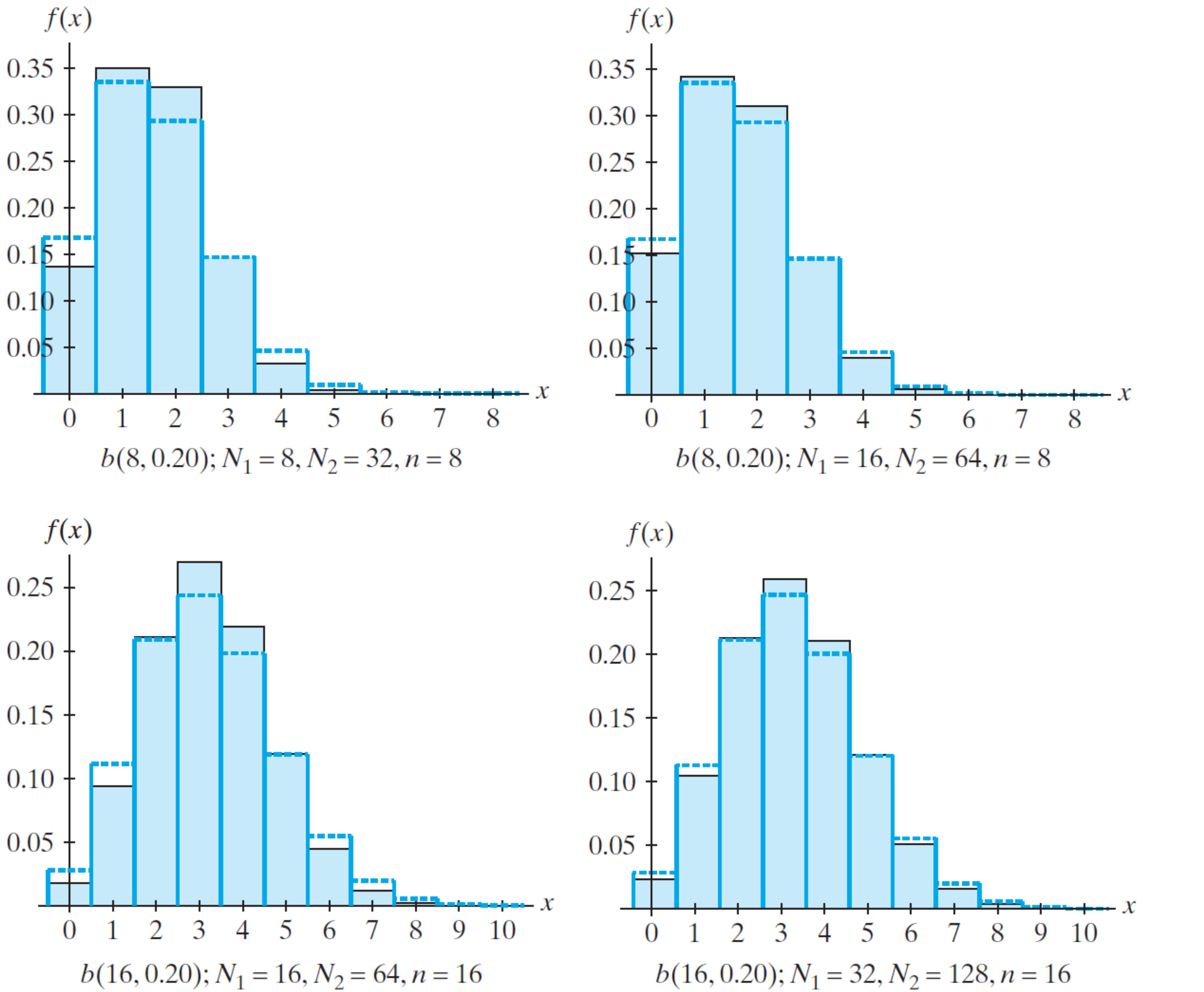


Figure 2.4-4 Binomial and hypergeometric (shaded) probability histograms

Chapter 2 Discrete Distributions

- Random Variables of the Discrete Type
  - Uniform Distribution
  - Hypergeometric Distribution
- Mathematical Expectation
- Moment Generating Function
- Bernoulli Trials and the Binomial Distribution
- Geometric and Negative Binomial Distribution
- The Poisson Distribution

From the sum of a geometric series we also note that, when k is an integer,

$$P(X > k) = \sum_{x=k+1}^{\infty} (1-p)^{x-1} p = \frac{(1-p)^k p}{1-(1-p)} = (1-p)^k = q^k, \text{?}$$

and thus the value of the distribution function at a positive integer k is

$$P(X \le k) = \sum_{x=1}^{k} (1-p)^{x-1} p = 1 - P(X > k) = 1 - (1-p)^k = 1 - q^k.$$

$$P(X > k) = \sum_{x=k+1}^{\infty} (1-p)^{x-1} p = \frac{(1-p)^k p}{1-(1-p)} = (1-p)^k = q^k,$$

$$P(X \le k) = \sum_{x=1}^{k} (1-p)^{x-1} p = 1 - P(X > k) = 1 - (1-p)^k = 1 - q^k.$$

Example 3.4-4 Some biology students were checking the eye color for a large number of fruit flies. For the individual fly, suppose that the probability of white eyes is 1/4 and the probability of red eyes is 3/4, and that we may treat these observations as having independent Bernoulli trials. The probability that at least four flies have to be checked for eye color to observe a white-eyed fly is given by

$$P(X \ge 4) = P(X > 3) = q^3 = \left(\frac{3}{4}\right)^3 = \frac{27}{64} = 0.4219.$$

The probability that at most four flies have to be checked for eye color to observe a white-eyed fly is given by

$$P(X \le 4) = 1 - q^4 = 1 - \left(\frac{3}{4}\right)^4 = \frac{175}{256} = 0.6836.$$

$$P(X \le k) = \sum_{x=1}^{k} (1-p)^{x-1} p = 1 - P(X > k) = 1 - (1-p)^k = 1 - q^k.$$

The probability that the first fly with white eyes is the fourth fly considered is

$$P(X = 4) = q^{4-1} p = \left(\frac{3}{4}\right)^3 \left(\frac{1}{4}\right) = \frac{27}{256} = 0.1055.$$

It is also true that

$$\begin{aligned} P(X = 4) &= P(X \le 4) - P(X \le 3) \\ &= [1 - (3/4)^4] - [1 - (3/4)^3] \\ &= \left(\frac{3}{4}\right)^3 \left(\frac{1}{4}\right). \end{aligned}$$

Negative Binomial Distribution

Let the r.v. ***X*** denote the number of trials needed to observe the r^{th} success

X: the trial number on which the r^{th} success is observed

The p.m.f. of ***X*** equals the product of the following two probabilities:

1. the probability of obtaining exactly *r-1* successes in the first *x-1* trials

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

2. the probability *p* of a success on the r^{th} trial

Thus, the p.m.f. of ***X*** is

$$\boxed{g(x) = \binom{x-1}{r-1} p^{r}(1-p)^{x-r} = \binom{x-1}{r-1} p^{r} q^{x-r}, \; x = r, r+1, \ldots .}$$

$$\sum_{x=r}^{\infty} g(x) \stackrel{?}{=} 1$$

The reason for calling the more general distribution the negative binomial distribution is the following. Consider $h(w) = (1-w)^{-r}$, the binomial $(1-w)$ with the negative exponent $-r$. Using Maclaurin's series expansion (Appendix D.4), we have

$$\boxed{(1-w)^{-r} = \sum_{k=0}^{\infty} \frac{h^{(k)}(0)}{k!} w^k = \sum_{k=0}^{\infty} \binom{r+k-1}{r-1} w^k, \qquad -1 < w < 1.}$$

?

Appendix D.4 Infinite Series

Taylor's series

A function $f(x)$ possessing derivatives of all orders at $x = b$ can be expanded in the following **Taylor series**:

$$f(x) = f(b) + \frac{f'(b)}{1!}(x-b) + \frac{f''(b)}{2!}(x-b)^2 + \frac{f'''(b)}{3!}(x-b)^3 + \cdots .$$

Maclaurin's series

If $b = 0$, we obtain the special case that is often called the **Maclaurin series**;

$$f(x) = f(0) + \frac{f'(0)}{1!}x + \frac{f''(0)}{2!}x^2 + \frac{f'''(0)}{3!}x^3 + \cdots .$$

Maclaurin's series

If $b = 0$, we obtain the special case that is often called the **Maclaurin series**;

$$f(x) = f(0) + \frac{f'(0)}{1!}x + \frac{f''(0)}{2!}x^2 + \frac{f'''(0)}{3!}x^3 + \cdots .$$

For example, if $f(x) = e^x$, so that all derivatives of $f(x) = e^x$ are $f^{(r)}(x) = e^x$, then $f^{(r)}(0) = 1$, for $r = 1, 2, 3, \ldots$. Thus, the Maclaurin series expansion of $f(x) = e^x$ is

$$e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots .$$

For examples, $$e = 1 + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \cdots$$

$$e^{-1} = 1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \cdots + \frac{(-1)^n}{n!} + \cdots .$$

Maclaurin's series

If $b = 0$, we obtain the special case that is often called the **Maclaurin series**;

$$f(x) = f(0) + \frac{f'(0)}{1!}x + \frac{f''(0)}{2!}x^2 + \frac{f'''(0)}{3!}x^3 + \cdots .$$

the binomial (1-*w*) with the negative exponent -*r*

As another example, consider the **negative binomial**

$$h(w) = (1-w)^{-r},$$

where r is a positive integer. Here

$$h'(w) = r(1-w)^{-(r+1)},$$

$$h''(w) = (r)(r+1)(1-w)^{-(r+2)},$$

$$h'''(w) = (r)(r+1)(r+2)(1-w)^{-(r+3)},$$

$$\vdots .$$

$$h'(w) = r(1-w)^{-(r+1)},$$

$$h''(w) = (r)(r+1)(1-w)^{-(r+2)},$$

$$h'''(w) = (r)(r+1)(r+2)(1-w)^{-(r+3)}$$

$$\vdots\ .$$

In general, $h^{(k)}(0) = (r)(r+1)\cdots(r+k-1) = (r+k-1)!/(r-1)!$. Thus,

$$\begin{aligned}(1-w)^{-r} &= 1 + \frac{(r+1-1)!}{(r-1)!\,1!}w + \frac{(r+2-1)!}{(r-1)!\,2!}w^2 + \cdots + \frac{(r+k-1)!}{(r-1)!\,k!}w^k + \cdots \\ &= \sum_{k=0}^{\infty}\binom{r+k-1}{r-1}w^k.\end{aligned}$$

The reason for calling the more general distribution the negative binomial distribution is the following. Consider $h(w) = (1-w)^{-r}$, the binomial $(1-w)$ with the negative exponent $-r$. Using Maclaurin's series expansion (Appendix D.4), we have

$$\boxed{(1-w)^{-r} = \sum_{k=0}^{\infty} \frac{h^{(k)}(0)}{k!} w^k = \sum_{k=0}^{\infty} \binom{r+k-1}{r-1} w^k, \qquad -1 < w < 1.}$$

#

If we let $x = k + r$ in the summation, then $k = x - r$ and

$$(1-w)^{-r} = \sum_{x=r}^{\infty} \binom{r+x-r-1}{r-1} w^{x-r} = \sum_{x=r}^{\infty} \binom{x-1}{r-1} w^{x-r},$$

the summand of which is, except for the factor p^r, the negative binomial probability when $w = q$. In particular, we see that the sum of the probabilities for the negative binomial distribution is 1 because

$$\sum_{x=r}^{\infty} g(x) = \sum_{x=r}^{\infty} \binom{x-1}{r-1} p^r q^{x-r} = p^r (1-q)^{-r} = 1.$$

→ Satisfy the properties of a p.m.f.

Next, show that the mean and the variance of X are

$$\mu = E(X) = \frac{r}{p} \quad \text{and} \quad \sigma^2 = \frac{rq}{p^2} = \frac{r(1-p)}{p^2}.$$

In particular, if $r = 1$ so that X has a geometric distribution, then

$$\mu = \frac{1}{p} \quad \text{and} \quad \sigma^2 = \frac{q}{p^2} = \frac{1-p}{p^2}.$$

The mean $\mu = 1/p$ agrees with our intuition. For illustration, if $p = 1/6$, then we would expect, on the average, $1/(1/6) = 6$ trials before the first success.

- *In the textbook, it is shown how to compute the mean and variance of the Negative Binomial distribution as follows.*

To find these moments, we determine the mgf of the negative binomial distribution. It is

$$
\begin{aligned}
M(t) &= \sum_{x=r}^{\infty} e^{tx} \binom{x-1}{r-1} p^r (1-p)^{x-r} \\
&= (pe^t)^r \sum_{x=r}^{\infty} \binom{x-1}{r-1} [(1-p)e^t]^{x-r} \\
&= \frac{(pe^t)^r}{[1-(1-p)e^t]^r}, \qquad \text{where } (1-p)e^t < 1
\end{aligned}
$$

(or, equivalently, when $t < -\ln(1-p)$). Thus,

$$
\begin{aligned}
M'(t) &= (pe^t)^r(-r)[1-(1-p)e^t]^{-r-1}[-(1-p)e^t] \\
&\quad + r(pe^t)^{r-1}(pe^t)[1-(1-p)e^t]^{-r} \\
&= r(pe^t)^r[1-(1-p)e^t]^{-r-1}
\end{aligned}
$$

and

$$
\begin{aligned}
M''(t) &= r(pe^t)^r(-r-1)[1-(1-p)e^t]^{-r-2}[-(1-p)e^t] \\
&\quad + r^2(pe^t)^{r-1}(pe^t)[1-(1-p)e^t]^{-r-1}.
\end{aligned}
$$

$$\boxed{M'(0) = E(X) = \mu \qquad \text{and} \qquad M''(0) - [M'(0)]^2 = E(X^2) - [E(X)]^2 = \sigma^2.}$$

Accordingly,

$$M'(0) = rp^r p^{-r-1} = rp^{-1}$$

and

$$\begin{aligned} M''(0) &= r(r+1)p^r p^{-r-2}(1-p) + r^2 p^r p^{-r-1} \\ &= rp^{-2}[(1-p)(r+1) + rp] \ = \ rp^{-2}(r+1-p). \end{aligned}$$

Hence, we have

$$\mu = \frac{r}{p} \qquad \text{and} \qquad \sigma^2 = \frac{r(r+1-p)}{p^2} - \frac{r^2}{p^2} = \frac{r(1-p)}{p^2}.$$

- *Exercises 2.5-5 and 2.5-6 show a somewhat easier way for finding the mean and variance of the four distributions: Bernoulli, Binomial, Geometric, and Negative Binomial.*

Example 2.5-3 shows the effect of p and r on the negative binomial distribution

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

$$\mu = E(X) = \frac{r}{p}$$

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

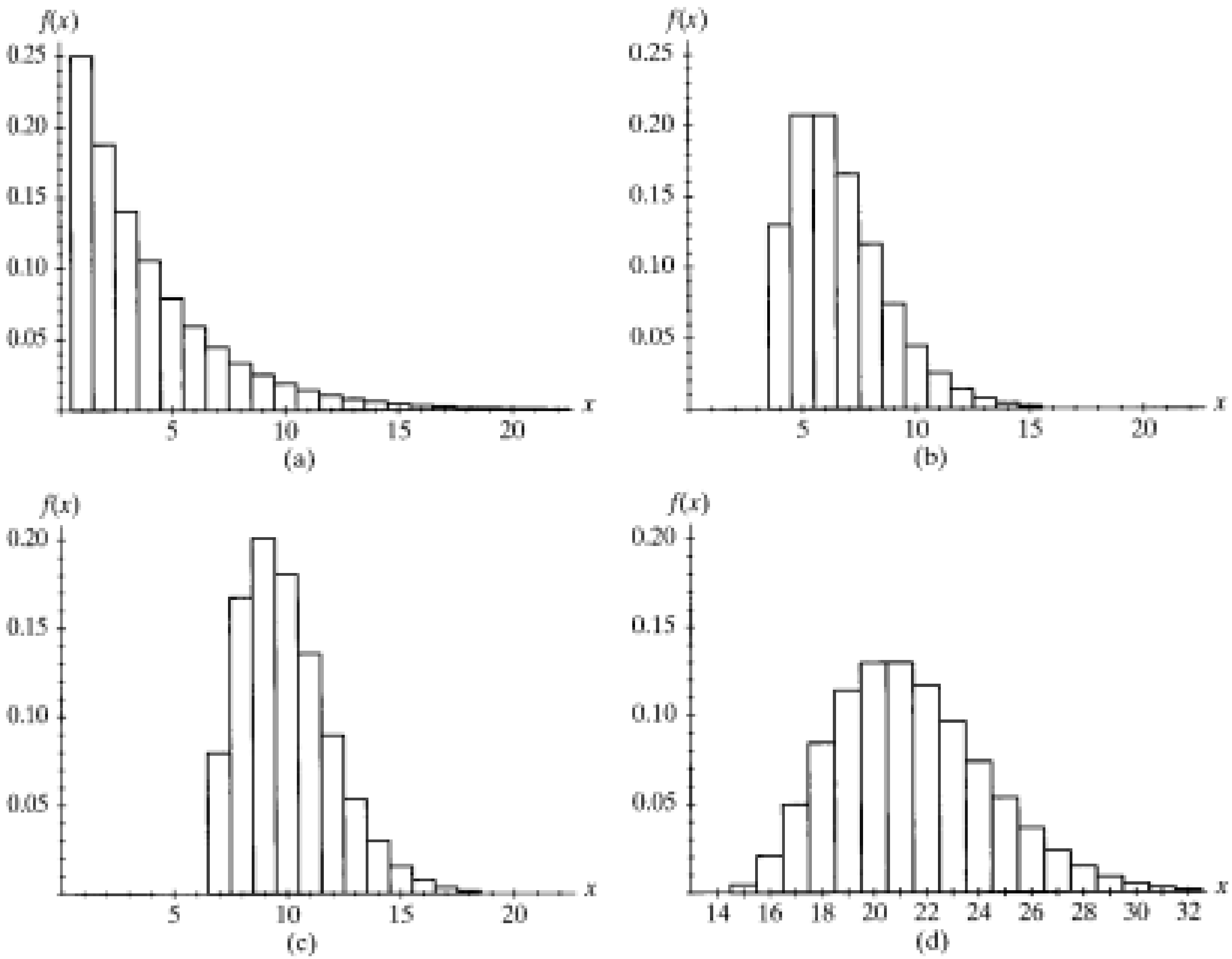


Figure 3.4-1 Negative binomial probability histograms: (a) $r = 1$, $p = 0.25$, (b) $r = 4$, $p = 0.6$, (c) $r = 7$, $p = 0.7$, (d) $r = 15$, $p = 0.7$