

3-6 Binomial Distribution

Random experiments and random variables

1. Flip a coin 10 times. Let X = number of heads obtained.
2. A worn machine tool produces 1% defective parts. Let X = number of defective parts in the next 25 parts produced.
3. Each sample of air has a 10% chance of containing a particular rare molecule. Let X = the number of air samples that contain the rare molecule in the next 18 samples analyzed.
4. Of all bits transmitted through a digital transmission channel, 10% are received in error. Let X = the number of bits in error in the next five bits transmitted.

3-6 Binomial Distribution

Random experiments and random variables

5. A multiple choice test contains 10 questions, each with four choices, and you guess at each question. Let X = the number of questions answered correctly.
6. In the next 20 births at a hospital, let X = the number of female births.
7. Of all patients suffering a particular illness, 35% experience improvement from a particular medication. In the next 100 patients administered the medication, let X = the number of patients who experience improvement.

3-6 Binomial Distribution

Definition

A random experiment consists of n Bernoulli trials such that

- (1) The trials are independent
- (2) Each trial results in only two possible outcomes, labeled as “success” and “failure”
- (3) The probability of a success in each trial, denoted as p , remains constant

The random variable X that equals the number of trials that result in a success has a **binomial random variable** with parameters $0 < p < 1$ and $n = 1, 2, \dots$. The probability mass function of X is

$$f(x) = \binom{n}{x} p^x (1 - p)^{n-x} \quad x = 0, 1, \dots, n \quad (3-7)$$

3-6 Binomial Distribution

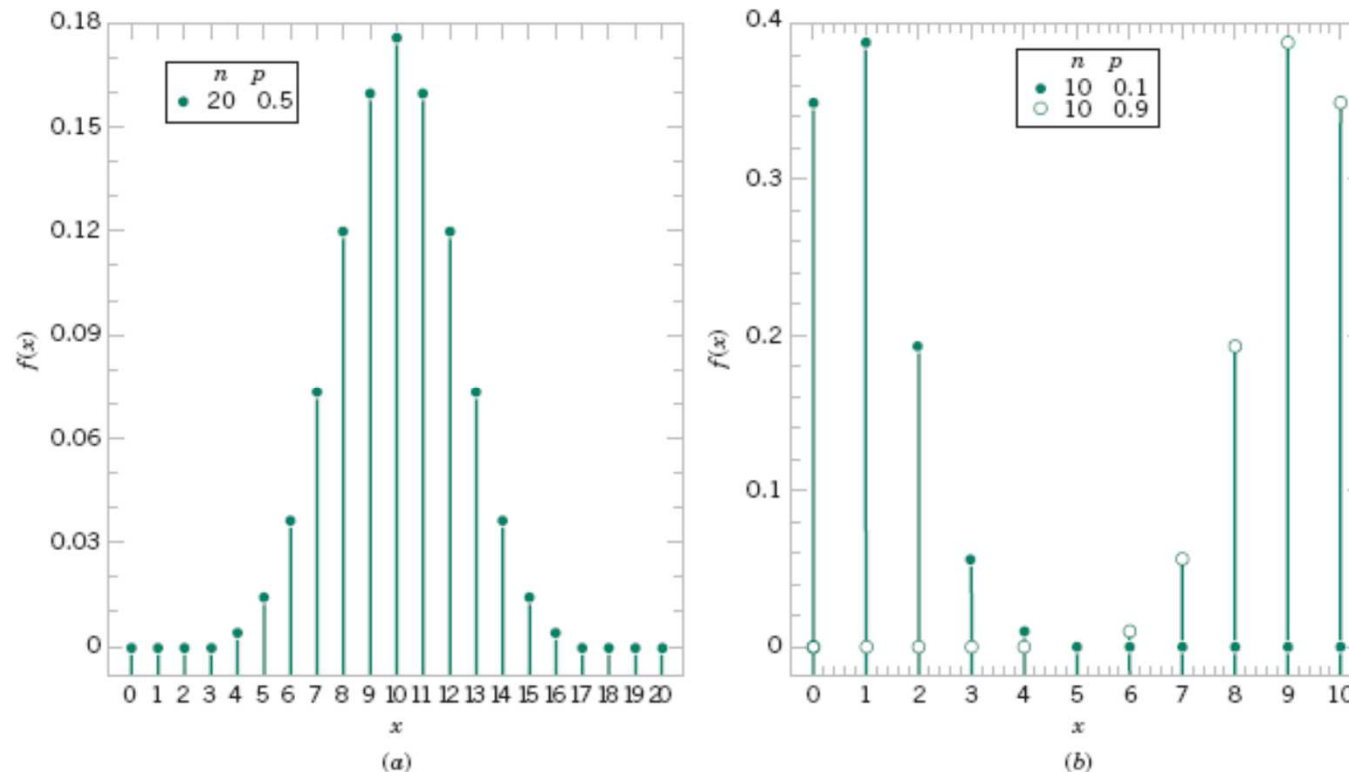


Figure 3-8 Binomial distributions for selected values of n and p .

3-6 Binomial Distribution

Example 3-18

Each sample of water has a 10% chance of containing a particular organic pollutant. Assume that the samples are independent with regard to the presence of the pollutant. Find the probability that in the next 18 samples, exactly 2 contain the pollutant.

Let X = the number of samples that contain the pollutant in the next 18 samples analyzed. Then X is a binomial random variable with $p = 0.1$ and $n = 18$.

Therefore,

$$P(X = 2) = \binom{18}{2} (0.1)^2 (0.9)^{16}$$

Now $\binom{18}{2} = 18!/[2! 16!] = 18(17)/2 = 153$. Therefore,

$$P(X = 2) = 153(0.1)^2(0.9)^{16} = 0.284$$

3-6 Binomial Distribution

Example 3-18

Determine the probability that at least four samples contain the pollutant. The requested probability is

$$P(X \geq 4) = \sum_{x=4}^{18} \binom{18}{x} (0.1)^x (0.9)^{18-x}$$

However, it is easier to use the complementary event,

$$\begin{aligned} P(X \geq 4) &= 1 - P(X < 4) = 1 - \sum_{x=0}^3 \binom{18}{x} (0.1)^x (0.9)^{18-x} \\ &= 1 - [0.150 + 0.300 + 0.284 + 0.168] = 0.098 \end{aligned}$$

Determine the probability that $3 \leq X < 7$. Now

$$\begin{aligned} P(3 \leq X < 7) &= \sum_{x=3}^6 \binom{18}{x} (0.1)^x (0.9)^{18-x} \\ &= 0.168 + 0.070 + 0.022 + 0.005 \\ &= 0.265 \end{aligned}$$

3-6 Binomial Distribution

Mean and Variance

If X is a binomial random variable with parameters p and n ,

$$\mu = E(X) = np \quad \text{and} \quad \sigma^2 = V(X) = np(1 - p) \quad (3-8)$$

3-6 Binomial Distribution

Example 3-19

For the number of transmitted bits received in error in Example 3-16, $n = 4$ and $p = 0.1$, so

$$E(X) = 4(0.1) = 0.4 \quad \text{and} \quad V(X) = 4(0.1)(0.9) = 0.36$$

and these results match those obtained from a direct calculation in Example 3-9.

3-7 Geometric and Negative Binomial Distributions

Example 3-20

The probability that a bit transmitted through a digital transmission channel is received in error is 0.1. Assume the transmissions are independent events, and let the random variable X denote the number of bits transmitted *until* the first error.

Then, $P(X = 5)$ is the probability that the first four bits are transmitted correctly and the fifth bit is in error. This event can be denoted as $\{OOOOE\}$, where O denotes an okay bit. Because the trials are independent and the probability of a correct transmission is 0.9,

$$P(X = 5) = P(OOOOE) = 0.9^4 0.1 = 0.066$$

Note that there is some probability that X will equal any integer value. Also, if the first trial is a success, $X = 1$. Therefore, the range of X is $\{1, 2, 3, \dots\}$, that is, all positive integers.

3-7 Geometric and Negative Binomial Distributions

Definition

In a series of Bernoulli trials (independent trials with constant probability p of a success), let the random variable X denote the number of trials until the first success. Then X is a **geometric random variable** with parameter $0 < p < 1$ and

$$f(x) = (1 - p)^{x-1}p \quad x = 1, 2, \dots \quad (3-9)$$

3-7 Geometric and Negative Binomial Distributions

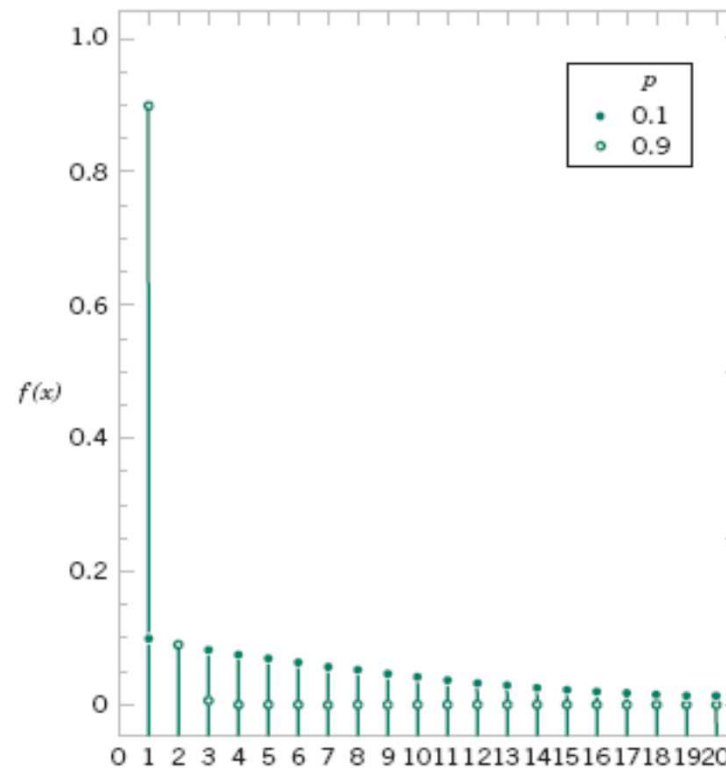


Figure 3-9. Geometric distributions for selected values of the parameter p .

3-7 Geometric and Negative Binomial Distributions

3-7.1 Geometric Distribution

Example 3-21

The probability that a wafer contains a large particle of contamination is 0.01. If it is assumed that the wafers are independent, what is the probability that exactly 125 wafers need to be analyzed before a large particle is detected?

Let X denote the number of samples analyzed until a large particle is detected. Then X is a geometric random variable with $p = 0.01$. The requested probability is

$$P(X = 125) = (0.99)^{124}0.01 = 0.0029$$

3-7 Geometric and Negative Binomial Distributions

Definition

If X is a geometric random variable with parameter p ,

$$\mu = E(X) = 1/p \quad \text{and} \quad \sigma^2 = V(X) = (1 - p)/p^2 \quad (3-10)$$

3-7 Geometric and Negative Binomial Distributions

Lack of Memory Property

A geometric random variable has been defined as the number of trials until the first success.

However, because the trials are independent, the count of the number of trials until the next success can be started at any trial without changing the probability distribution of the random variable. For example, in the transmission of bits, if 100 bits are transmitted, the probability that the first error, after bit 100, occurs on bit 106 is the probability that the next six outcomes are *OOOOOE*. This probability is $(0.9)^5(0.1) = 0.059$, which is identical to the probability that the initial error occurs on bit 6.

The implication of using a geometric model is that the system presumably will not wear out. The probability of an error remains constant for all transmissions. In this sense, the geometric distribution is said to lack any memory. The **lack of memory property** will be discussed again in the context of an exponential random variable in Chapter 4.

3-7 Geometric and Negative Binomial Distributions

3-7.2 Negative Binomial Distribution

A generalization of a geometric distribution in which the random variable is the number of Bernoulli trials required to obtain r successes results in the **negative binomial distribution**.

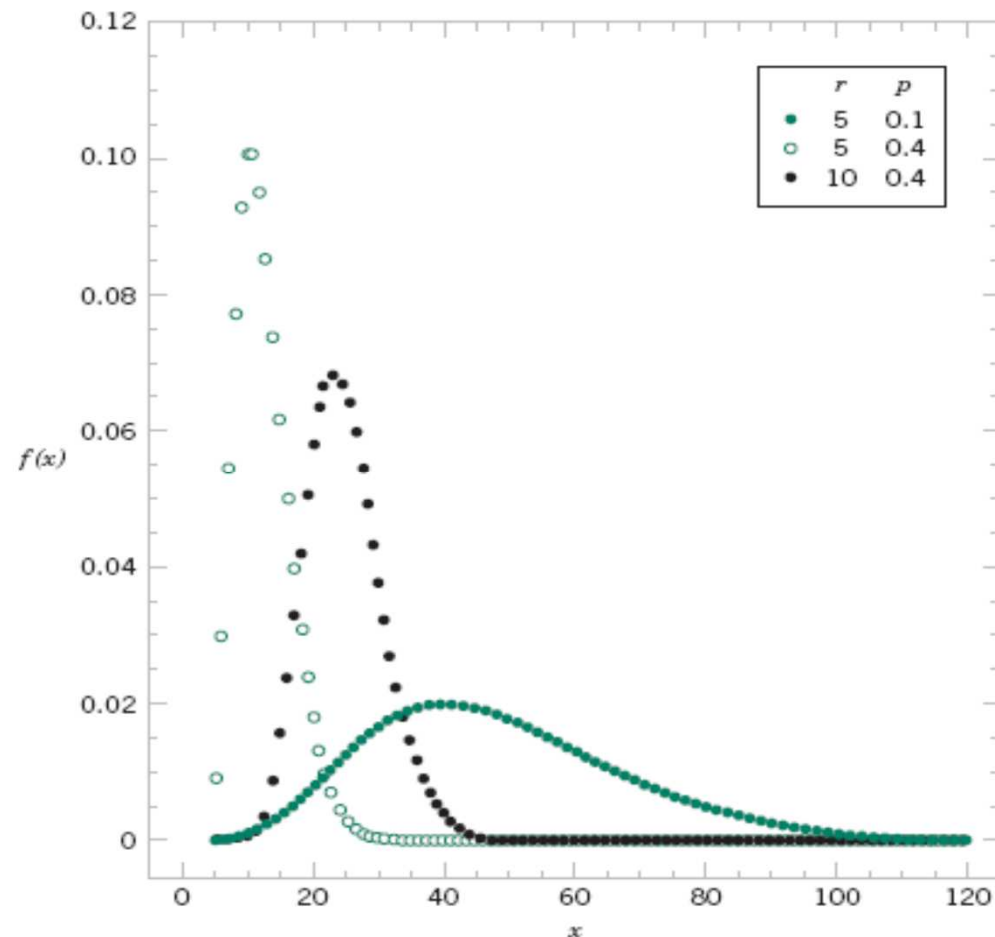
In a series of Bernoulli trials (independent trials with constant probability p of a success), let the random variable X denote the number of trials until r successes occur. Then X is a **negative binomial random variable** with parameters $0 < p < 1$ and $r = 1, 2, 3, \dots$, and

$$f(x) = \binom{x-1}{r-1} (1-p)^{x-r} p^r \quad x = r, r+1, r+2, \dots \quad (3-11)$$

3-7 Geometric and Negative Binomial Distributions

Figure 3-10.

Negative binomial distributions for selected values of the parameters r and p .



3-7 Geometric and Negative Binomial Distributions



Figure 3-11. Negative binomial random variable represented as a sum of geometric random variables.

3-7 Geometric and Negative Binomial Distributions

3-7.2 Negative Binomial Distribution

If X is a negative binomial random variable with parameters p and r ,

$$\mu = E(X) = r/p \quad \text{and} \quad \sigma^2 = V(X) = r(1 - p)/p^2 \quad (3-12)$$

3-7 Geometric and Negative Binomial Distributions

Example 3-25

A Web site contains three identical computer servers. Only one is used to operate the site, and the other two are spares that can be activated in case the primary system fails. The probability of a failure in the primary computer (or any activated spare system) from a request for service is 0.0005. Assuming that each request represents an independent trial, what is the mean number of requests until failure of all three servers?

Let X denote the number of requests until all three servers fail, and let X_1 , X_2 , and X_3 denote the number of requests before a failure of the first, second, and third servers used, respectively. Now, $X = X_1 + X_2 + X_3$. Also, the requests are assumed to comprise independent trials with constant probability of failure $p = 0.0005$. Furthermore, a spare server is not affected by the number of requests before it is activated. Therefore, X has a negative binomial distribution with $p = 0.0005$ and $r = 3$. Consequently,

$$E(X) = 3/0.0005 = 6000 \text{ requests}$$

3-7 Geometric and Negative Binomial Distributions

Example 3-25

What is the probability that all three servers fail within five requests? The probability is $P(X \leq 5)$ and

$$\begin{aligned} P(X \leq 5) &= P(X = 3) + P(X = 4) + P(X = 5) \\ &= 0.0005^3 + \binom{3}{2} 0.0005^3 (0.9995) + \binom{4}{2} 0.0005^3 (0.9995)^2 \\ &= 1.25 \times 10^{-10} + 3.75 \times 10^{-10} + 7.49 \times 10^{-10} \\ &= 1.249 \times 10^{-9} \end{aligned}$$