Foundation of
Data Science and Analytics

Probability Distribution - 2

Arun K. Timalsina

Material Adaptation : Applied Statistics and Probability for Engineers, by Montgomery and Runger. John Wiley & Sons,

Example 3-1: Voice Lines

- A voice communication system for a business contains 48 external lines. At a particular time, the system is observed, and some of the lines are being used.
- Let X denote the number of lines in use. Then, X can assume any of the integer values 0 through 48.
- The system is observed at a random point in time. If 10 lines are in use, then $x = 10$.

Example 3-2: Wafers

In a semiconductor manufacturing process, 2 wafers from a lot are sampled. Each wafer is classified as *pass* or *fail*. Assume that the probability that a wafer passes is 0.8, and that wafers are independent.

The sample space for the experiment and associated probabilities are shown in Table 3-1. The probability that the 1st wafer passes and the 2nd fails, denoted as *pf* is $P(pf) = 0.8 * 0.2 = 0.16$.

The random variable X is defined as the number of wafers that pass.

Table 3-1 Wafer Tests

Outcome			
Wafer #			
1	2	Probability	x
Pass	Pass	0.64	2
Fail	Pass	0.16	1
Pass	Fail	0.16	1
Fail	Fail	0.04	0
		1.00	

Example 3-3: Particles on Wafers

- Define the random variable *X* to be the number of contamination particles on a wafer. Although wafers possess a number of characteristics, the random variable *X* summarizes the wafer only in terms of the number of particles. The possible values of *X* are the integers 0 through a very large number, so we write $x \geq 0$.
- We can also describe the random variable *Y* as the number of chips made from a wafer that fail the final test. If there can be 12 chips made from a wafer, then we write $0 \leq y \leq 12$. (changed)

Probability Distributions

- A random variable X associates the outcomes of a random experiment to a number on the number line.
- The probability distribution of the random variable X is a description of the probabilities with the possible numerical values of X.
- A probability distribution of a discrete random variable can be:
  1. A list of the possible values along with their probabilities.
  2. A formula that is used to calculate the probability in response to an input of the random variable's value.

Example 3-4: Digital Channel

- There is a chance that a bit transmitted through a digital transmission channel is received in error.
- Let *X* equal the number of bits received in error of the next 4 transmitted.
- The associated probability distribution of *X* is shown as a graph and as a table.

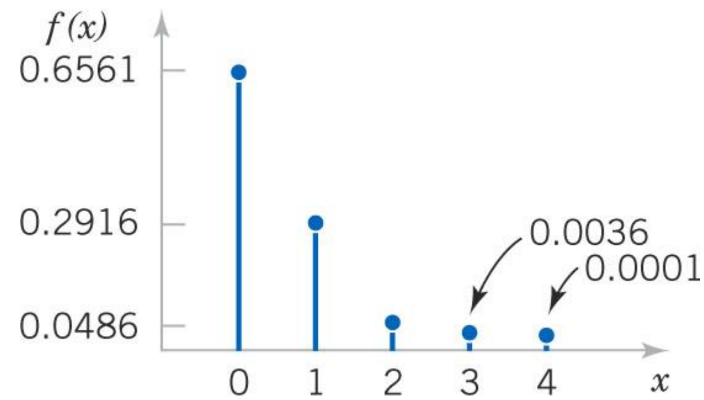


Figure 3-1 Probability distribution for bits in error.

$P(X=0) =$	0.6561
$P(X=1) =$	0.2916
$P(X=2) =$	0.0486
$P(X=3) =$	0.0036
$P(X=4) =$	0.0001
	1.0000

Probability Mass Function

Suppose a loading on a long, thin beam places mass only at discrete points. This represents a probability distribution where the beam is the number line over the range of x and the probabilities represent the mass. That's why it is called a probability mass function.

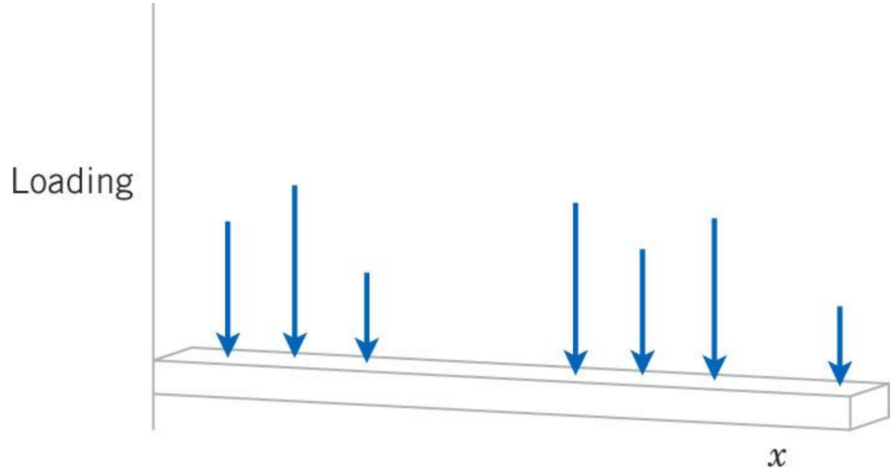


Figure 3-2 Loading at discrete points on a long, thin beam.

Probability Mass Function Properties

For a discrete random variable X with possible values $x_1, x_2, \ldots x_n$, a probability mass function is a function such that:

(1) $f(x_i) \geq 0$

(2) $\sum_{i=1}^{n} f(x_i) = 1$

(3) $f(x_i) = P(X = x_i)$

Example 3-5: Wafer Contamination

- Let the random variable *X* denote the number of wafers that need to be analyzed to detect a large particle. Assume that the probability that a wafer contains a large particle is 0.1, and that the wafers are independent. Determine the probability distribution of *X*.
- Let *p* denote a wafer for which a large particle is **p**resent & let *a* denote a wafer in which it is **a**bsent.
- The sample space is: $S = \{p, ap, aap, aaap, \ldots\}$
- The range of the values of *X* is: x = 1, 2, 3, 4, …

$P(X=1) =$	0.1	0.1
$P(X=2) =$	$(0.9)*0.1$	0.09
$P(X=3) =$	$(0.9)^2*0.1$	0.081
$P(X=4) =$	$(0.9)^3*0.1$	0.0729
		0.3439

Cumulative Distribution Functions

- Example 3-6: From Example 3.4, we can express the probability of three or fewer bits being in error, denoted as $P(X \leq 3)$.
- The event $(X \leq 3)$ is the union of the mutually exclusive events: $(X=0)$, $(X=1)$, $(X=2)$, $(X=3)$.
- From the table:

x	$P(X=x)$	$P(X \leq x)$
0	0.6561	0.6561
1	0.2916	0.9477
2	0.0486	0.9963
3	0.0036	0.9999
4	0.0001	1.0000
	1.0000	

$$P(X \leq 3) = P(X=0) + P(X=1) + P(X=2) + P(X=3) = 0.9999$$

$$P(X = 3) = P(X \leq 3) - P(X \leq 2) = 0.0036$$

Cumulative Distribution Function Properties

The cumulative distribution function is built from the probability mass function and vice versa.

The cumulative distribution function of a discrete random variable X, denoted as $F(x)$, is:

$$F\left(x\right)=F\left(X\ =x\right)=\sum_{x_i\leq x}x_i$$

For a discrete random variable X, $F\left(x\right)$ satisfies the following properties:

(1) $F\left(x\right)=P\left(X\ \leq x\right)=\sum_{x_i\leq x}f\left(x_i\right)$

(2) $0\leq F\left(x\right)\leq 1$

(3) If $x\leq y$, then $F\left(x\right)\leq F\left(y\right)$

Example 3-7: Cumulative Distribution Function

- Determine the probability mass function of X from this cumulative distribution function:

$F(x) =$	0.0		$x < -2$
	0.2		$-2 \le x < 0$
	0.7		$0 \le x < 2$
	1.0		$2 \le x$

PMF	
$f(2) =$	0.2
$f(0) =$	0.5
$f(2) =$	0.3

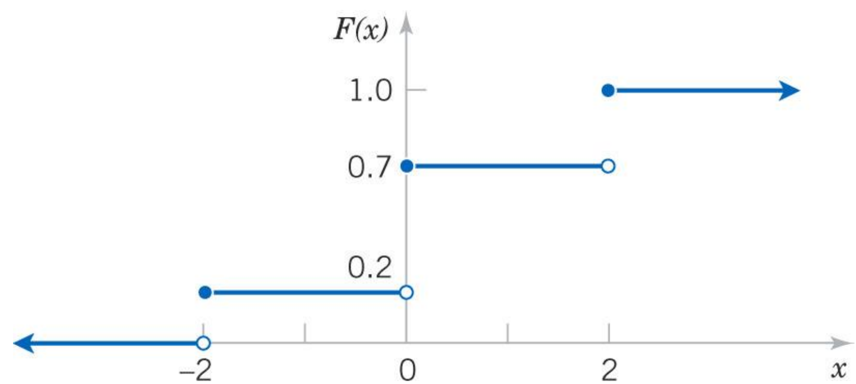


Figure 3-3 Graph of the CDF

Example 3-8: Sampling without Replacement

A day's production of 850 parts contains 50 defective parts. Two parts are selected at random without replacement. Let the random variable *X* equal the number of defective parts in the sample. Create the CDF of *X*.

$$P(X=0)=\tfrac{800}{850}\cdot\tfrac{799}{849}=0.886$$

$$P(X=1)=2\cdot\tfrac{800}{850}\cdot\tfrac{50}{849}=0.111$$

$$P(X=2)=\tfrac{50}{850}\cdot\tfrac{49}{849}=0.003$$

Therefore,

$$F(0)=P(X\le 0)=0.886$$

$$F(1)=P(X\le 1)=0.997$$

$$F(2)=P(X\le 2)=1.000$$

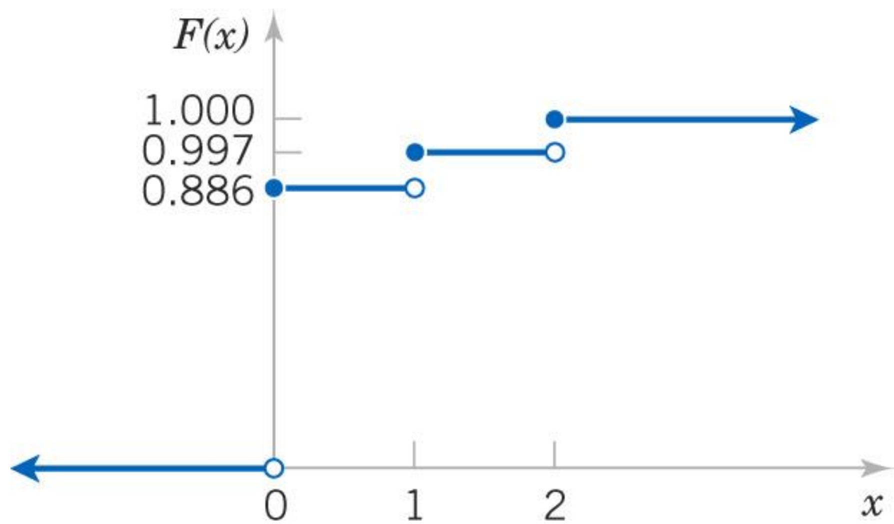


Figure 3-4 CDF. Note that F(x) is defined for all x, $-\infty < x < \infty$, not just 0, 1 and 2.

Summary Numbers of a Probability Distribution

- The mean is a measure of the center of a probability distribution.
- The variance is a measure of the dispersion or variability of a probability distribution.
- The standard deviation is another measure of the dispersion. It is the square root of the variance.

Mean Defined

The mean or expected value of the discrete random variable X, denoted as μ or $E(X)$, is

$$\mu = E(X) = \sum_x x \cdot f(x)$$

- The mean is the weighted average of the possible values of *X*, the weights being the probabilities where the beam balances. It represents the center of the distribution. It is also called the arithmetic mean.
- If *f*(*x*) is the probability mass function representing the loading on a long, thin beam, then *E*(*X*) is the fulcrum or point of balance for the beam.
- The mean value may, or may not, be a given value of *x*.

Variance Defined

The variance of X, denoted as σ^2 or $V(X)$, is

$$\sigma^2 = V(X) = E(X-\mu)^2 = \sum_x (x-\mu)^2 \cdot f(x) = \sum_x x^2 \cdot f(x) - \mu^2$$

- The variance is the measure of dispersion or scatter in the possible values for X.
- It is the average of the squared deviations from the distribution mean.

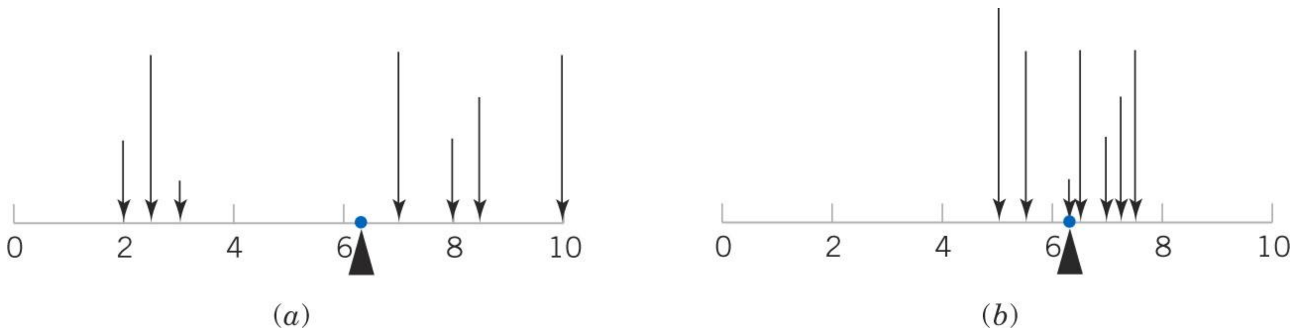


Figure 3-5 The mean is the balance point. Distributions (a) & (b) have equal mean, but (a) has a larger variance.

Variance Formula Derivations

$$V(X) = \sum_x (x-\mu)^2 f(x) \text{ is the definitional formula}$$

$$= \sum_x \left(x^2 - 2\mu x + \mu^2\right) f(x)$$

$$= \sum_x x^2 f(x) - 2\mu \sum x f(x) + \mu^2 \sum_x f(x)$$

$$= \sum_x x^2 f(x) - 2\mu^2 + \mu^2$$

$$= \sum_x x^2 f(x) - \mu^2 \text{ is the computational formula}$$

The computational formula is easier to calculate manually.

Different Distributions Have Same Measures

These measures do not uniquely identify a probability distribution – different distributions could have the same mean & variance.

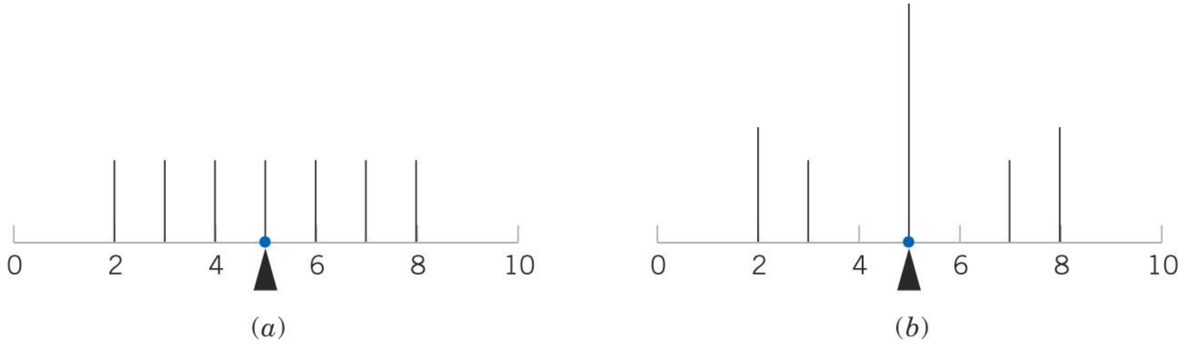


Figure 3-6 These probability distributions have the same mean and variance measures, but are very different in shape.

Exercise 3-9: Digital Channel

In Exercise 3-4, there is a chance that a bit transmitted through a digital transmission channel is an error. *X* is the number of bits received in error of the next 4 transmitted. Use table to calculate the mean & variance.

			Definitional formula		
x	$f(x)$	$x*f(x)$	$(x-0.4)^2$	$(x-0.4)^2*f(x)$	$x^2*f(x)$
0	0.6561	0.0000	0.160	0.1050	0.0000
1	0.2916	0.2916	0.360	0.1050	0.2916
2	0.0486	0.0972	2.560	0.1244	0.1944
3	0.0036	0.0108	6.760	0.0243	0.0324
4	0.0001	0.0004	12.960	0.0013	0.0016
	Totals =	0.4000		0.3600	0.5200
		= Mean		= Variance (σ^2)	= $E(x^2)$
		= μ		$\sigma^2 = E(x^2) - \mu^2 =$	0.3600
				Computational formula	

Exercise 3-10 Marketing

- Two new product designs are to be compared on the basis of revenue potential. Revenue from Design A is predicted to be \$3 million. But for Design B, the revenue could be \$7 million with probability 0.3 or only \$2 million with probability 0.7. Which design is preferable?
- Answer:
  - Let X & Y represent the revenues for products A & B.
  - $E(X)$ = \$3 million. $V(X)$ = 0 because x is certain.
  - $E(Y)$ = \$3.5 million = 7*0.3 + 2*0.7 = 2.1 + 1.4
  - $V(X)$ = 5.25 million dollars^2 or $(7-3.5)^2$*.3 + $(2-3.5)^2$*.7 = 3.675 + 1.575
  - SD(X) = 2.29 million dollars , the square root of the variance.
  - Standard deviation has the same units as the mean, not the squared units of the variance.

Exercise 3-11: Messages

The number of messages sent per hour over a computer network has the following distribution. Find the mean & standard deviation of the number of messages sent per hour.

x	$f(x)$
10	0.08
11	0.15
12	0.30
13	0.20
14	0.20
15	0.07
	1.00

$x*f(x)$	$x^2*f(x)$
0.80	8
1.65	18.15
3.60	43.2
2.60	33.8
2.80	39.2
1.05	15.75
12.50	158.10
$= E(X)$	$= E(X^2)$

Mean = 12.5

Variance = $158.10^2 - 12.5^2 = 1.85$

Standard deviation = 1.36

Note that: $E(X^2) \neq [E(X)]^2$

A Function of a Random Variable

If X is a discrete random variable with probability mass function $f(x)$,

$$E\left[h(X)\right] = \sum_x h(x)\, f(x) \qquad (3\text{-}4)$$

If $h(x) = (X - \mu)^2$, then its expectation is the variance of X.

Example 3-12: Digital Channel

In Example 3-9, X is the number of bits in error in the next four bits transmitted. What is the expected value of the square of the number of bits in error?

x	$f(x)$
0	0.6561
1	0.2916
2	0.0486
3	0.0036
4	0.0001
	1.0000

$x^2 * f(x)$
0.0000
0.2916
0.1944
0.0324
0.0016
0.5200
$= E(x^2)$

Discrete Uniform Distribution

- Simplest discrete distribution.
- The random variable X assumes only a finite number of values, each with equal probability.
- A random variable X has a discrete uniform distribution if each of the n values in its range, say $x_1, x_2, \ldots, x_n$, has equal probability.

$$f(x_i) = 1/n \qquad (3\text{-}5)$$

Example 3-13: Discrete Uniform Random Variable

The first digit of a part's serial number is equally likely to be the digits 0 through 9. If one part is selected from a large batch & X is the 1^{st} digit of the serial number, then X has a discrete uniform distribution as shown.

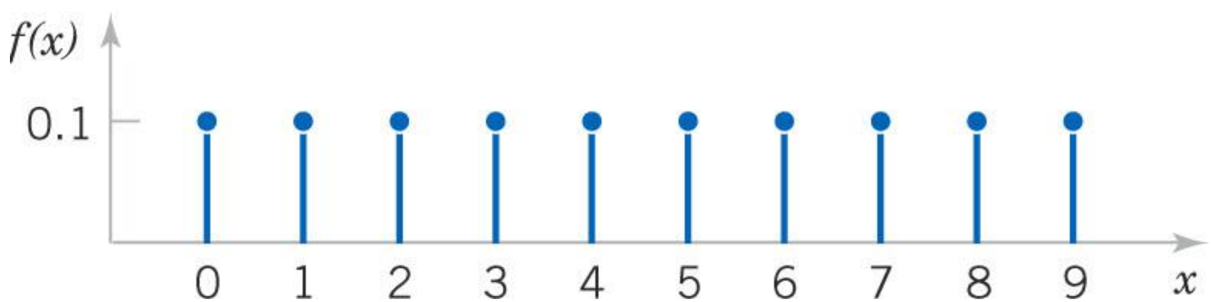


Figure 3-7 Probability mass function, $f(x) = 1/10$ for $x = 0, 1, 2, \ldots, 9$

General Discrete Uniform Distribution

- Let X be a discrete uniform random variable from a to b for $a < b$. There are b – (a-1) values in the inclusive interval. Therefore:

$$f(x) = 1/(b\text{-}a+1)$$

- Its measures are:

$$\mu = E(x) = 1/(b\text{-}a)$$

$$\sigma^2 = V(x) = [(b\text{-}a+1)^2 - 1]/12 \qquad (3\text{-}6)$$

Note that the mean is the midpoint of a & b.

Example 3-14: Number of Voice Lines

Per Example 3-1, let the random variable *X* denote the number of the 48 voice lines that are in use at a particular time. Assume that *X* is a discrete uniform random variable with a range of 0 to 48. Find *E*(*X*) & *SD*(*X*).

Answer:

$$\mu = \frac{48+0}{2} = 24$$

$$\sigma_X = \sqrt{\frac{(48-0+1)^2-1}{12}} = \sqrt{\frac{2400}{12}} = 14.142$$

Example 3-15 Proportion of Voice Lines

Let the random variable Y denote the proportion of the 48 voice line that are in use at a particular time & X as defined in the prior example. Then $Y = X/48$ is a proportion. Find $E(Y)$ & $V(Y)$.

Answer:

$$E(Y) = E(X) / 48 = 24/48 = 0.5$$

$$V(Y) = V(X) / 48^2 = 14.142^2 / 2304 = 0.0868$$

Examples of Binomial Random Variables

1. Flip a coin 10 times. X = # heads obtained.
2. A worn tool produces 1% defective parts. X = # defective parts in the next 25 parts produced.
3. A multiple-choice test contains 10 questions, each with 4 choices, and you guess. X = # of correct answers.
4. Of the next 20 births, let X = # females.

These are binomial experiments having the following characteristics:

1. Fixed number of trials (n).
2. Each trial is termed a success or failure. X is the # of successes.
3. The probability of success in each trial is constant (p).
4. The outcomes of successive trials are independent.

Example 3-16: Digital Channel

The chance that a bit transmitted through a digital transmission channel is received in error is 0.1. Assume that the transmission trials are independent. Let X = the number of bits in error in the next 4 bits transmitted. Find $P(X=2)$.

Answer:

Let E denote a bit in error
Let O denote an OK bit.
Sample space & x listed in table.
6 outcomes where $x = 2$.
Prob of each is $0.1^2 * 0.9^2 = 0.0081$
Prob(X=2) = 6*0.0081 = 0.0486

$$P(X=2) = C_2^4 (0.1)^2 (0.9)^2$$

Outcome	*x*	Outcome	*x*
OOOO	0	EOOO	1
OOOE	1	EOOE	2
OOEO	1	EOEO	2
OOEE	2	EOEE	3
OEOO	1	EEOO	2
OEOE	2	EEOE	3
OEEO	2	EEEO	3
OEEE	3	EEEE	4

Binomial Distribution Definition

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

$$f(x) = C_x^n p^x (1-p)^{n-x} \quad \text{for } x = 0, 1, \ldots n \qquad (3\text{-}7)$$

- Based on the binomial expansion:

$$(a+b)^n = \sum_{k=0}^{n} C_k^n a^k b^{n-k}$$

Binomial Distribution Shapes

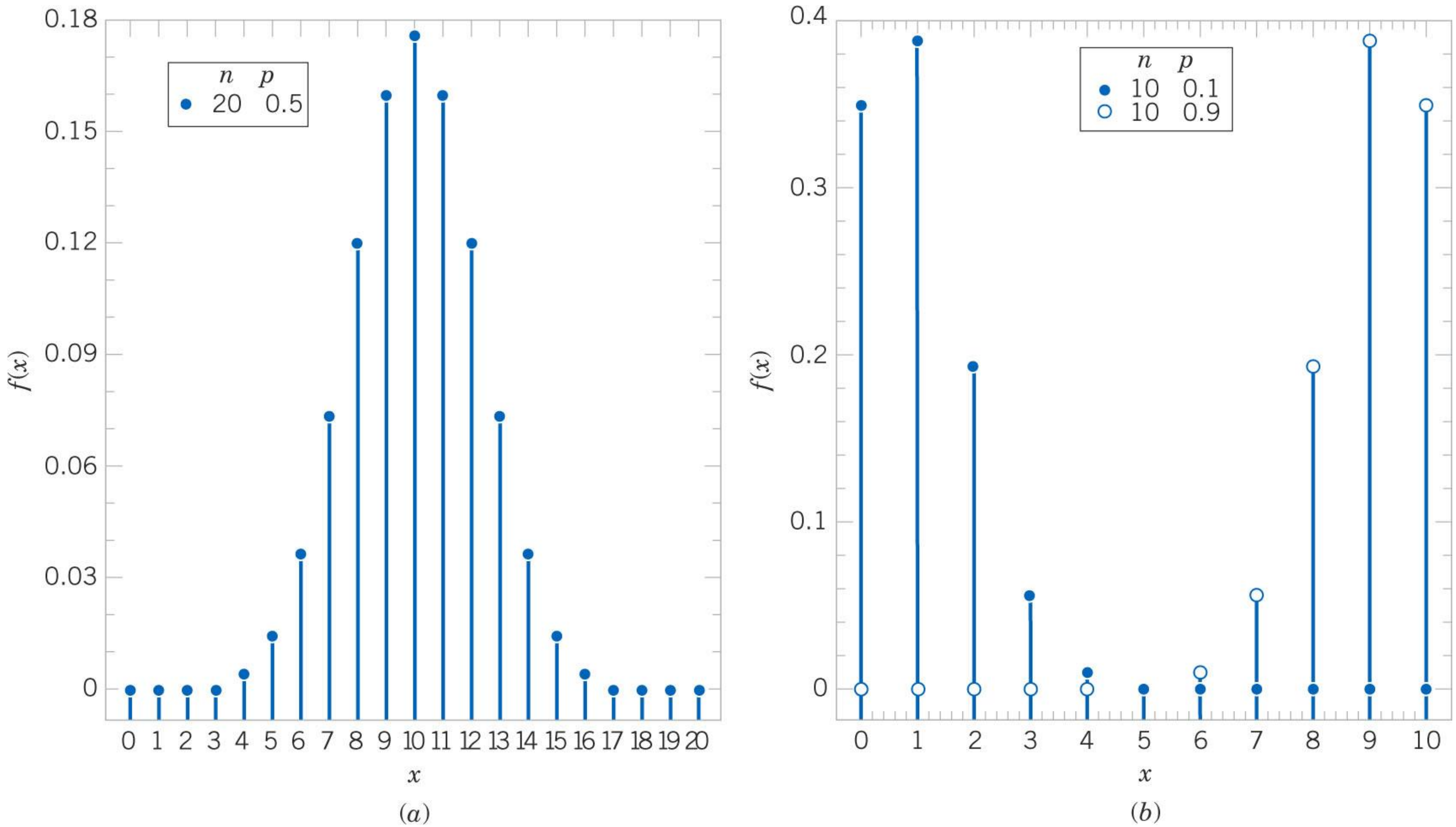


Figure 3-8 Binomial Distributions for selected values of *n* and *p*. Distribution (a) is symmetrical, while distributions (b) are skewed. The skew is right if *p* is small.

Example 3-17: Binomial Coefficients

Exercises in binomial coefficient calculation:

$$C_3^{10} = \frac{10!}{3!7!} = \frac{10 \cdot 9 \cdot 8 \cdot 7!}{3 \cdot 2 \cdot 1 \cdot 7!} = 120$$

$$C_{10}^{15} = \frac{15!}{10!5!} = \frac{15 \cdot 14 \cdot 13 \cdot 12 \cdot 11}{5 \cdot 4 \cdot 3 \cdot 2 \cdot 1} = 3,003$$

$$C_4^{100} = \frac{100!}{4!96!} = \frac{100 \cdot 99 \cdot 98 \cdot 97}{4 \cdot 3 \cdot 2 \cdot 1} 3,921,225$$

Exercise 3-18: Organic Pollution-1

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.

Answer: Let X denote 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

$$P(X=2)=C_2^{18}(0.1)^2(0.9)^{16}=153(0.1)^2(0.9)^{16}=0.2835$$

0.2835	= BINOMDIST(2,18,0.1,FALSE)

Exercise 3-18: Organic Pollution-2

Determine the probability that at least 4 samples contain the pollutant.

Answer:

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

$$= 1 - P(X < 4)$$

$$= 1 - \sum_{x=0}^{3} C_x^{18} (0.1)^x (0.9)^{18-x}$$

$$= 0.098$$

0.0982	= 1 - BINOMDIST(3,18,0.1,TRUE)

Exercise 3-18: Organic Pollution-3

Now determine the probability that $3 \le X \le 7$.

Answer:

$$P(3 \le X \le 7) = \sum_{x=3}^{7} C_x^{18} (0.1)^x (0.9)^{18-x} = 0.265$$

$$P(X \le 7) - P(X \le 2)$$

0.2660	= BINOMDIST(7,18,0.1,TRUE) - BINOMDIST(2,18,0.1,TRUE)

Appendix A, Table II (pg. 705) is a cumulative binomial table for selected values of *p* and *n*.

Binomial Mean and Variance

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

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

Example 3-19:

For the number of transmitted bit received in error in Example 3-16, n = 4 and p = 0.1. Find the mean and variance of the binomial random variable.

Answer:

$$\mu = E(X) = np = 4*0.1 = 0{,}4$$

$$\sigma^2 = V(X) = np(1-p) = 4*0.1*0.9 = 3.6$$

$$\sigma = SD(X) = 1.9$$

Example 3-20: New Idea

The probability that a bit, sent through a digital transmission channel, is received in error is 0.1. Assume that the transmissions are independent. Let *X* denote the number of bits transmitted until the 1st error.

P(*X*=5) is the probability that the 1st four bits are transmitted correctly and the 5th bit is in error.

$P(X=5) = P(OOOOE) = 0.9^4 0.1 = 0.0656.$

x is the total number of bits sent.

This illustrates the geometric distribution.

Geometric Distribution

- Similar to the binomial distribution – a series of Bernoulli trials with fixed parameter p.
- Binomial distribution has:
  - Fixed number of trials.
  - Random number of successes.
- Geometric distribution has reversed roles:
  - Random number of trials.
  - Fixed number of successes, in this case 1.

- $f(x) = p(1-p)^{x-1}$ where: (3-9)
  - $x = 1, 2, \ldots \infty$, the number of failures until the 1st success.
  - $0 < p < 1$, the probability of success.

Geometric Graphs

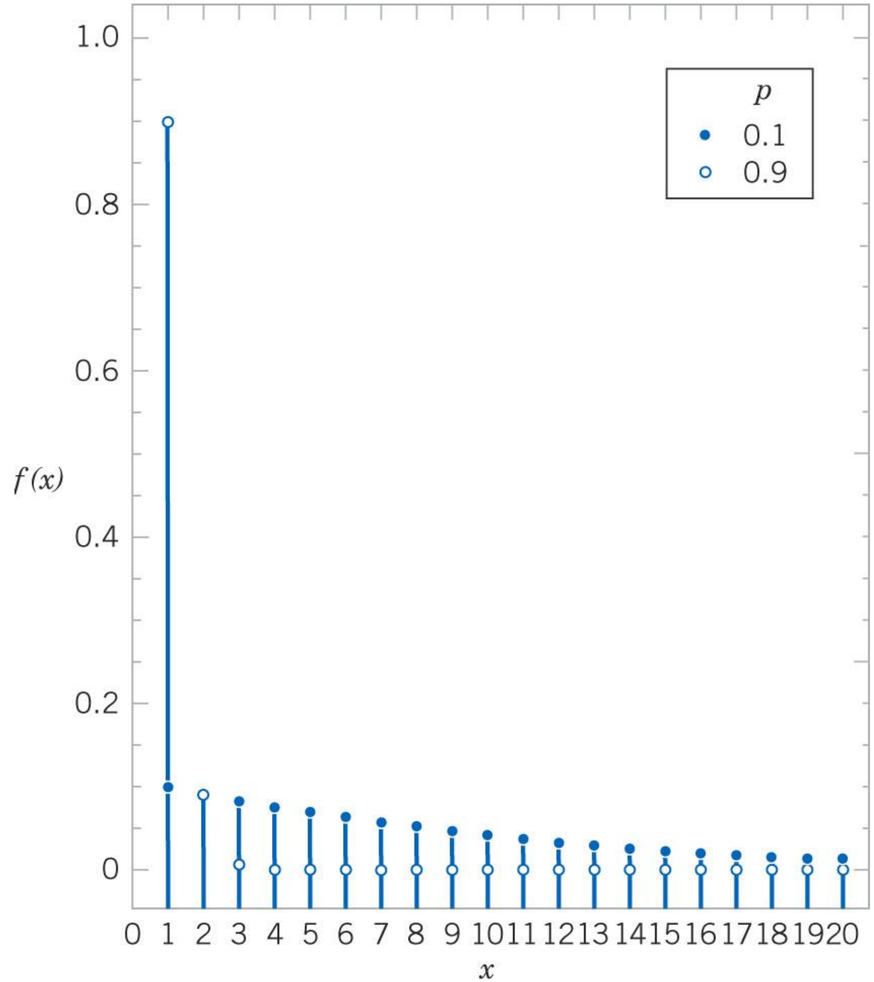


Figure 3-9 Geometric distributions for parameter p values of 0.1 and 0.9. The graphs coincide at $x = 2$.

Example 3.21: Geometric Problem

The probability that a wafer contains a large particle of contamination is 0.01. Assume that the wafers are independent. What is the probability that exactly 125 wafers need to be analyzed before a particle is detected?

Answer:

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

$$P(X=125) = (0.99)^{124}(0.01) = 0.00288.$$

Geometric Mean & Variance

- If X is a geometric random variable with parameter p,

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

Exercise 3-22: Geometric Problem

Consider the transmission of bits in Exercise 3-20. Here, $p = 0.1$. Find the mean and standard deviation.

Answer:

Mean = $\mu = E(X) = 1 / p = 1 / 0.1 = 10$

Variance = $\sigma^2 = V(X) = (1-p) / p^2 = 0.9 / 0.01 = 90$

Standard deviation = sqrt(99) = 9.487

Lack of Memory Property

- For a geometric random variable, the trials are independent. Thus the count of the number of trials until the next success can be started at any trial without changing the probability.
- The probability that the next bit error will occur on bit 106, given that 100 bits have been transmitted, is the same as it was for bit 006.
- Implies that the system does not wear out!

Example 3-23: Lack of Memory

In Example 3-20, the probability that a bit is transmitted in error is 0.1. Suppose 50 bits have been transmitted. What is the mean number of bits transmitted until the next error?

Answer:

The mean number of bits transmitted until the next error, after 50 bits have already been transmitted, is $1 / 0.1 = 10$.

Example 3-24: New Idea

The probability that a bit, sent through a digital transmission channel, is received in error is 0.1. Assume that the transmissions are independent. Let X denote the number of bits transmitted until the 4th error.

P(X=10) is the probability that 3 errors occur over the first 9 trials, then the 4th success occurs on the 10th trial.

$$\text{3 errors occur over the first 9 trials} = C_3^9 p^3 (1-p)^6$$

$$\text{4th error occurs on the 10th trial} = C_3^9 p^4 (1-p)^6$$

In general, probabilities for X can be determined as follows. Here $P(X = x)$ implies that $r-1$ successes occur in the first $x-1$ trials and the rth success occurs on trial x. The probability that $r-1$ successes occur in the first $x-1$ trials is obtained from the binomial distribution to be

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

for $r \le x$. The probability that trial x is a success is p. Because the trials are independent, these probabilities are multiplied so that

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

This leads to the following result.

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

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

Negative Binomial Definition

- In a series of independent trials with constant probability of 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, ...
- The probability mass function is:

$$f(x) = C_{r-1}^{x-1} p^r (1-p)^{x-r} \text{ for } x = r, r+1, r+2 \ldots \qquad (3\text{-}11)$$

- From the prior example for $f(X=10|r=4)$:
  - x-1 = 9
  - r-1 = 3

Negative Binomial Graphs

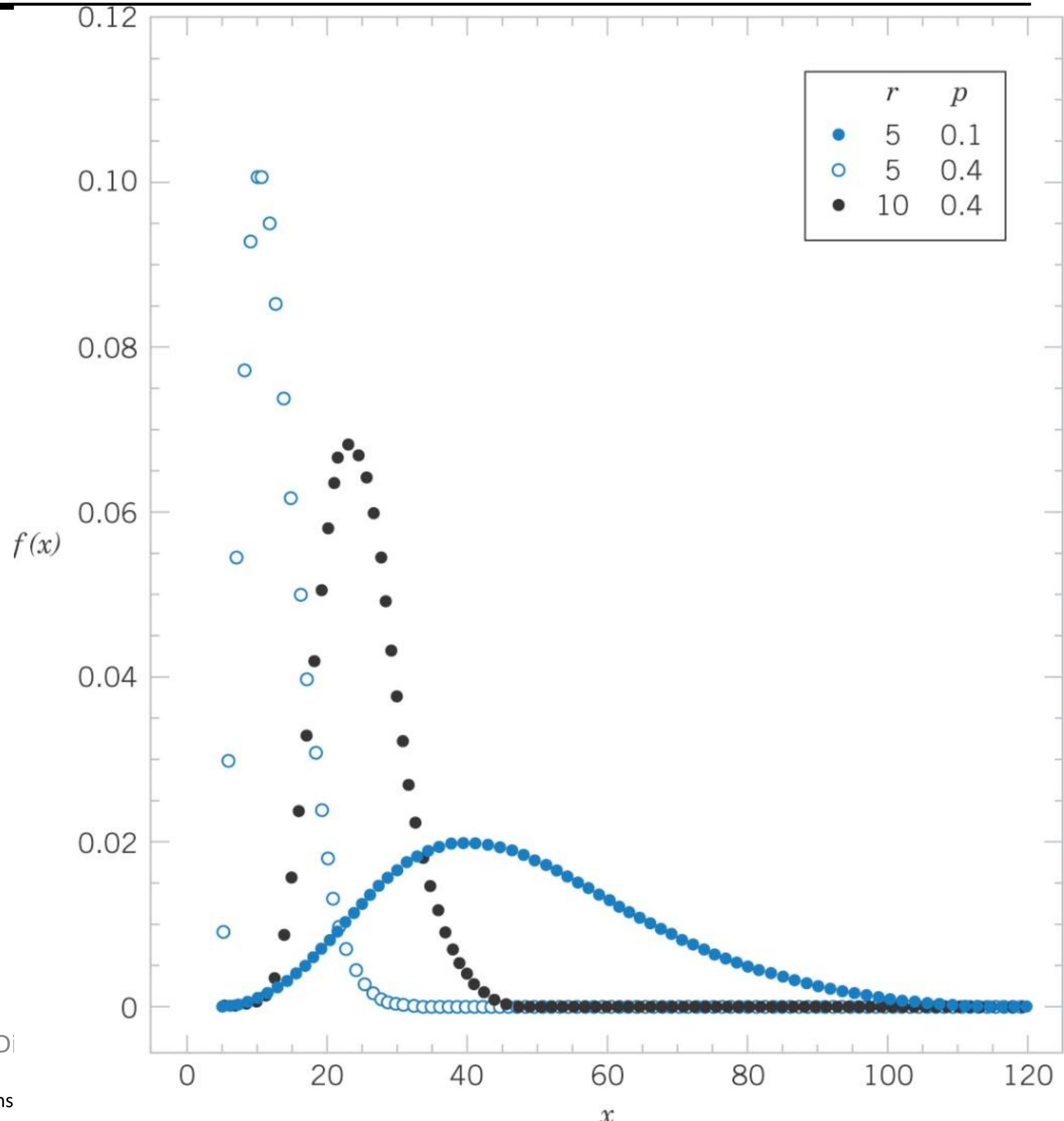
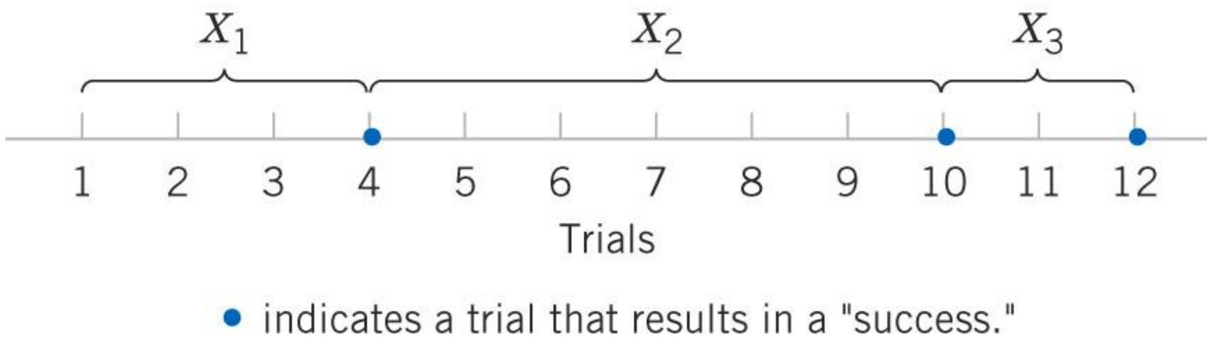


Figure 3-10 Negative binomial distributions for 3 different parameter combinations.

Lack of Memory Property

$$X = X_1 + X_2 + X_3$$



- Let X_1 denote the number of trials to the 1st success.
- Let X_2 denote the number of trials to the 2nd success, since the 1st success.
- Let X_3 denote the number of trials to the 3rd success, since the 2nd success.
- Let the X_i be geometric random variables – independent, so without memory.
- Then $X = X_1 + X_2 + X_3$
- Therefore, X is a negative binomial random variable, a sum of three geometric rv's.

Negative Binomial Mean & Variance

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

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

What's In A Name?

- Binomial distribution:
  - Fixed number of trials (n).
  - Random number of successes (x).
- Negative binomial distribution:
  - Random number of trials (x).
  - Fixed number of successes (r).
- Because of the reversed roles, a negative binomial can be considered the opposite or negative of the binomial.

Example 3-25: Web Servers-1

A Web site contains 3 identical computer servers. Only one is used to operate the site, and the other 2 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) from a request for service is 0.0005. Assume that each request represents an independent trial. What is the mean number of requests until failure of all 3 servers?

Answer:

- Let X denote the number of requests until all three servers fail.
- Let $r = 3$ and $p=0.0005 = 1/2000$
- Then $\mu = 3 / 0.0005 = 6{,}000$ requests

Example 3-25: Web Servers-2

What is the probability that all 3 servers fail within 5 requests? (X = 5)

Answer:

$$P(X \le 5) = P(X = 3) + P(X = 4) + P(X = 5)$$

$$= 0.005^3 + C_2^3 0.0005^3 0.9995 + C_2^4 0.0005^3 0.9995^2$$

In Excel	
1.250E-10	= 0.0005^3
3.748E-10	= NEGBINOMDIST(1, 3, 0.0005)
7.493E-10	= NEGBINOMDIST(2, 3, 0.0005)
1.249E-09	

Note that Excel uses a different definition of *X*; # of failures before the r^{th} success, not # of trials.

Hypergeometric Distribution

- Applies to sampling without replacement.
- Trials are not independent & a tree diagram used.
- A set of *N* objects contains:
  - *K* objects classified as success
  - *N* - *K* objects classified as failures
- A sample of size *n* objects is selected without replacement from the *N* objects, where:
  - $K \le N$ *and* $n \le N$
- Let the random variable *X* denote the number of successes in the sample. Then *X* is a hypergeometric random variable.

$$f(x) = \frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}} \text{ where } x = \max(0, n+K-N) \text{ to } \min(K, n) \qquad (3\text{-}13)$$

Hypergeometric Graphs

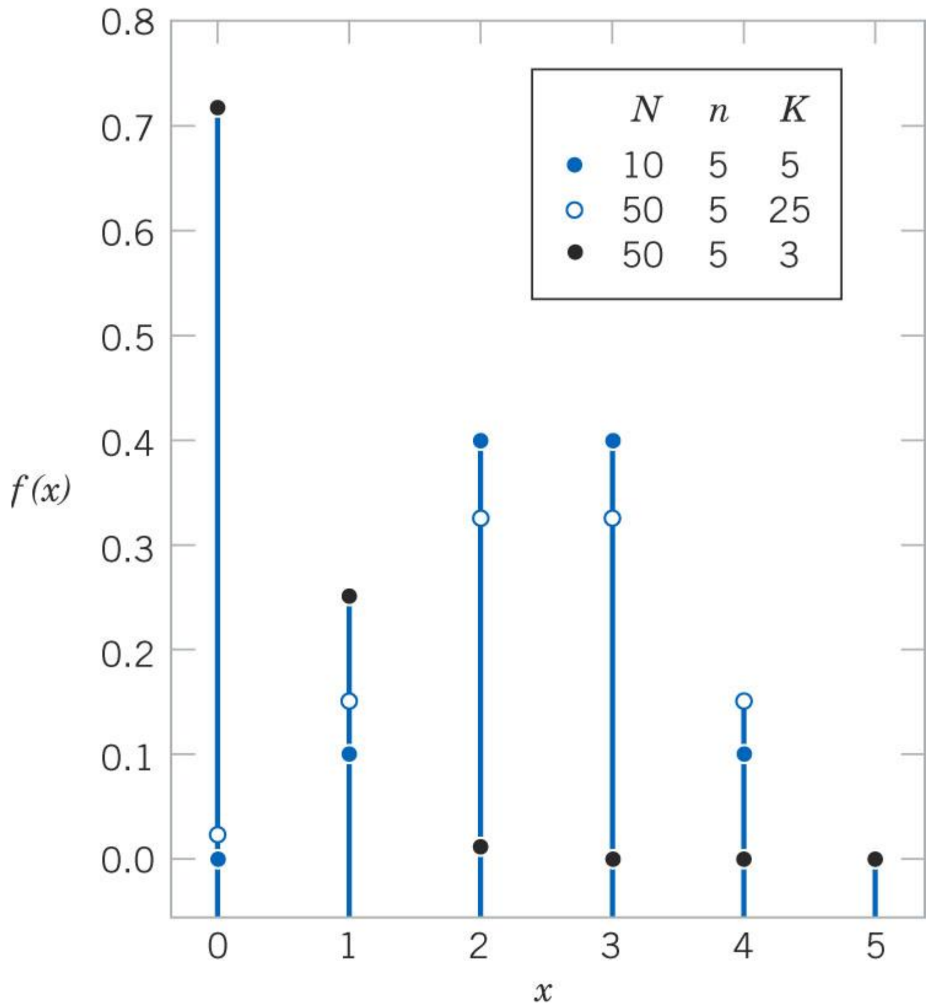


Figure 3-12 Hypergeometric distributions for 3 parameter sets of *N*, *K*, and *n*.

Example 3-26: Sampling without Replacement

From an earlier example, 50 parts are defective on a lot of 850. Two are sampled. Let *X* denote the number of defectives in the sample. Use the hypergeometric distribution to find the probability distribution.

Answer:

In Excel	
0.8857	= HYPGEOMDIST(0,2,50,850)
0.1109	= HYPGEOMDIST(1,2,50,850)
0.0034	= HYPGEOMDIST(2,2,50,850)

$$P(X=0)=\frac{\binom{50}{0}\binom{800}{2}}{\binom{850}{2}}=\frac{319,660}{360,825}=0.886$$

$$P(X=1)=\frac{\binom{50}{1}\binom{800}{1}}{\binom{850}{2}}=\frac{40,000}{360,825}=0.111$$

$$P(X=2)=\frac{\binom{50}{2}\binom{800}{0}}{\binom{850}{2}}=\frac{1,225}{360,825}=0.003$$

Example 3-27: Parts from Suppliers-1

A batch of parts contains 100 parts from supplier A and 200 parts from Supplier B. If 4 parts are selected randomly, without replacement, what is the probability that they are all from Supplier A?

Answer:

Let X equal the number of parts in the sample from Supplier A.

$$P(X=4)=\frac{\binom{100}{4}\binom{200}{0}}{\binom{300}{4}}=0.0119$$

In Excel	
0.01185	= HYPGEOMDIST(4,100,4,300)

Example 3-27: Parts from Suppliers-2

What is the probability that two or more parts are from Supplier A?

Answer:

$$P(X \geq 2) = P(X = 2) + P(X = 3) + P(X = 4)$$

$$= \frac{\binom{100}{2}\binom{200}{2}}{\binom{300}{4}} + \frac{\binom{100}{3}\binom{200}{1}}{\binom{300}{4}} + \frac{\binom{100}{4}\binom{200}{1}}{\binom{300}{4}}$$

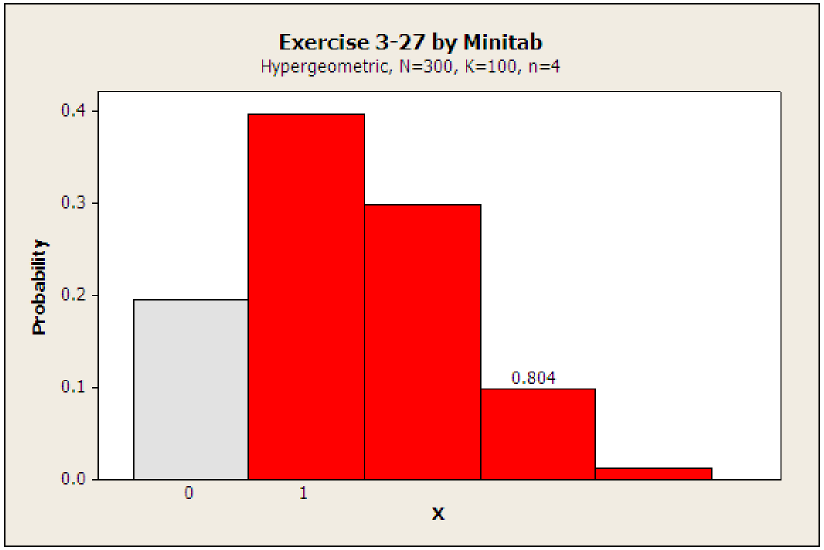
$$= 0.298 + 0.098 + 0.0119 = 0.408$$

In Excel	
0.40741	= HYPGEOMDIST(2,100,4,300)
	+ HYPGEOMDIST(3,100,4,300)
	+ HYPGEOMDIST(4,100,4,300)

Example 3-27: Parts from Suppliers-3

What is the probability that at least one part is from Supplier A?

Answer:



$$P(X \geq 1) = 1 - P(X = 0) = 1 - \frac{\binom{100}{0}\binom{200}{4}}{\binom{300}{4}} = 0.804$$

In Excel	
0.80445	= 1 - HYPGEOMDIST(0,100,4,300)

Hypergeometric Mean & Variance

- If X is a hypergeometric random variable with parameters N, K, and n, then

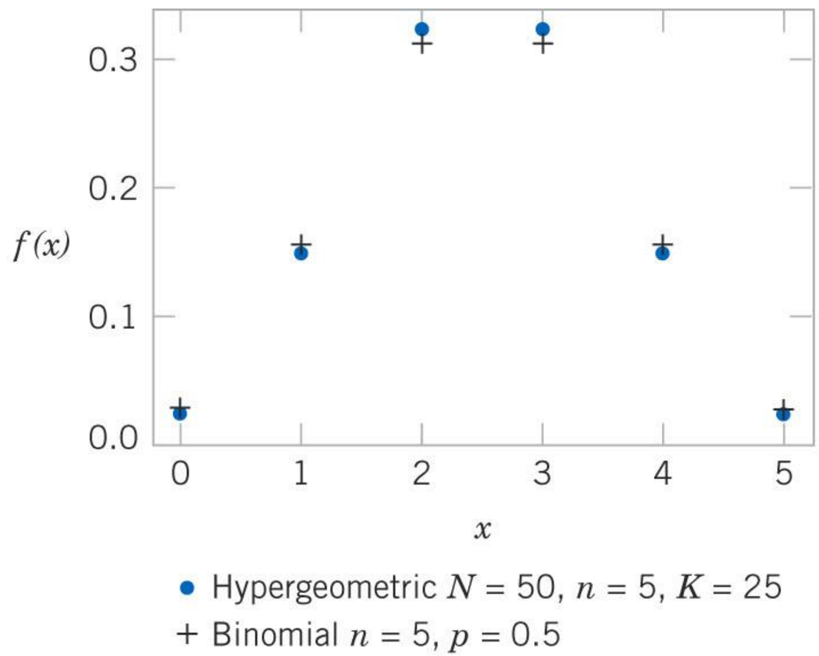
$$\mu = E(X) = np \quad \text{and} \quad \sigma^2 = V(X) = np(1-p)\left(\frac{N-n}{N-1}\right) \qquad (3\text{-}14)$$

where $p = K/N$

and $\left(\frac{N-n}{N-1}\right)$ is the finite population correction factor.

σ^2 approaches the binomial variance as n/N becomes small.

Hypergeometric & Binomial Graphs



	0	1	2	3	4	5
Hypergeometric probability	0.025	0.149	0.326	0.326	0.149	0.025
Binomial probability	0.031	0.156	0.312	0.312	0.156	0.031

Figure 3-13 Comparison of hypergeometric and binomial distributions.

Example 3-29: Customer Sample-1

A listing of customer accounts at a large corporation contains 1,000 accounts. Of these, 700 have purchased at least one of the company's products in the last 3 months. To evaluate a new product, 50 customers are sampled at random from the listing. What is the probability that more than 45 of the sampled customers have purchased in the last 3 months?

Let X denote the number of customers in the sample who have purchased from the company in the last 3 months. Then X is a hypergeometric random variable with *N* = 1,000, *K* = 700, *n* = 50. This a lengthy problem! ☹

$$P(X > 45) = \sum_{x=46}^{50} \frac{\binom{700}{x}\binom{300}{50-x}}{\binom{1,000}{50}}$$

Example 3-29: Customer Sample-2

Since n/N is small, the binomial will be used to approximate the hypergeometric. Let $p = K/N = 0.7$

$$P(X > 45) = \sum_{x=46}^{50} \binom{50}{x} 0.7^x (1-0.7)^{50-x} = 0.00017$$

In Excel	
0.000172	= 1 - BINOMDIST(45, 50, 0.7, TRUE)

The hypergeometric value is 0.00013. The absolute error is 0.00004, but the percent error in using the approximation is (17-13)/13 = 31%.

Poisson Distribution

As the number of trials (*n*) in a binomial experiment increases to infinity while the binomial mean (*np*) remains constant, the binomial distribution becomes the Poisson distribution.

Example 3-30:

$$\text{Let } \lambda = np = E(x), \text{ so } p = \lambda/n$$

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

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

$$= \frac{e^{-\lambda}\lambda^x}{x!}$$

Example 3-31: Wire Flaws

Flaws occur at random along the length of a thin copper wire. Let X denote the random variable that counts the number of flaws in a length of L mm of wire. Suppose the average number of flaws in L is λ.

Partition L into n subintervals (1 μm) each. If the subinterval is small enough, the probability that more than one flaw occurs is negligible.

Assume that the:

- Flaws occur at random, implying that each subinterval has the same probability of containing a flaw.
- Probability that a subinterval contains a flaw is independent of other subintervals.

X is now binomial. $E(X) = np = \lambda$ and $p = \lambda/n$

As n becomes large, p becomes small and a Poisson process is created.

Examples of Poisson Processes

In general, the Poisson random variable X is the number of events (counts) per interval.

1. Particles of contamination per wafer.
2. Flaws per roll of textile.
3. Calls at a telephone exchange per hour.
4. Power outages per year.
5. Atomic particles emitted from a specimen per second.
6. Flaws per unit length of copper wire.

Poisson Distribution Definition

- The random variable X that equals the number of events in a Poisson process is a Poisson random variable with parameter $\lambda > 0$, and the probability mass function is:

$$f(x) = \frac{e^{-\lambda}\lambda^{x}}{x!} \quad \text{for} \quad x = 0, 1, 2, 3, \ldots\infty \qquad (3\text{-}16)$$

Poisson Graphs

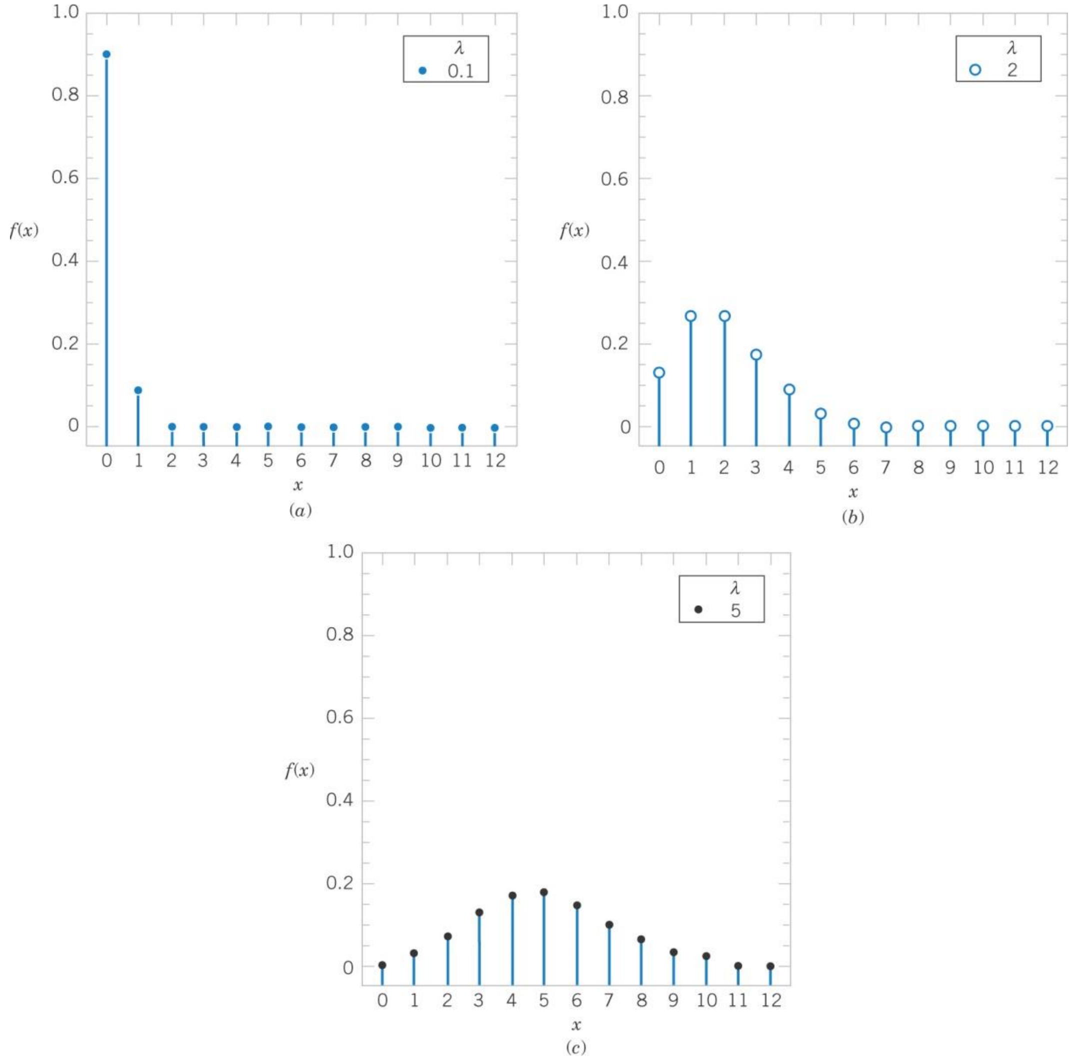


Figure 3-14 Poisson distributions for λ = 0.1, 2, 5.

Poisson Requires Consistent Units

It is important to use consistent units in the calculation of Poisson:

– Probabilities
– Means
– Variances

- Example of unit conversions:
  – Average # of flaws per mm of wire is 3.4.
  – Average # of flaws per 10 mm of wire is 34.
  – Average # of flaws per 20 mm of wire is 68.

Example 3-32: Calculations for Wire Flaws-1

For the case of the thin copper wire, suppose that the number of flaws follows a Poisson distribution of 2.3 flaws per mm. Let *X* denote the number of flaws in 1 mm of wire. Find the probability of exactly 2 flaws in 1 mm of wire.

Answer:

$$P(X=2)=\frac{e^{-2.3}2.3^2}{2!}=0.265$$

In Excel	
0.26518	= POISSON(2, 2.3, FALSE)

Example 3-32: Calculations for Wire Flaws-2

Determine the probability of 10 flaws in 5 mm of wire. Now let X denote the number of flaws in 5 mm of wire.

Answer:

$$E(X) = \lambda = 5 \text{ mm} \cdot \ 2.3 \text{ flaws/mm} = 11.5 \text{ flaws}$$

$$\mathrm{P}(X = 10) = e^{-11.5} \frac{11.5^{10}}{10!} = 0.113$$

In Excel	
0.1129	= POISSON(10, 11.5, FALSE)

Example 3-32: Calculations for Wire Flaws-3

Determine the probability of at least 1 flaw in 2 mm of wire. Now let X denote the number of flaws in 2 mm of wire. Note that $P(X \geq 1)$ requires ∞ terms. ☹

Answer:

$$E(X) = \lambda = 2 \text{ mm} \cdot \ 2.3 \text{ flaws/mm} = 4.6 \text{ flaws}$$

$$\mathrm{P}(X \geq 1) = 1 - P(X = 0) = 1 - e^{-4.6} \frac{4.6^0}{0!} = 0.9899$$

In Excel	
0.989948	= 1 - POISSON(0, 4.6, FALSE)

Example 3-33: CDs-1

Contamination is a problem in the manufacture of optical storage disks (CDs). The number of particles of contamination that occur on a CD has a Poisson distribution. The average number of particles per square cm of media is 0.1. The area of a disk under study is 100 cm^2. Let X denote the number of particles of a disk. Find $P(X = 12)$.

Answer:

$$E(X) = \lambda = 100 \text{ cm}^2 \cdot 0.1 \text{ particles/cm}^2 = 10 \text{ particles}$$

$$\mathrm{P}(X = 12) = e^{-10} \frac{10^{12}}{12!} = 0.095$$

In Excel	
0.0948	= POISSON(12, 10, FALSE)

Example 3-33: CDs-2

Find the probability that zero particles occur on the disk. Recall that λ = 10 particles.

Answer:

$$P(X = 0) = e^{-10}\frac{10^0}{0!} = 4.54 \cdot 10^{-5}$$

In Excel	
4.540E-05	= POISSON(0, 10, FALSE)

Example 3-33: CDs-3

Determine the probability that 12 or fewer particles occur on the disk. That will require 13 terms in the sum of probabilities.☹ Recall that λ = 10 particles.

Answer:

$$\mathrm{P}(X \le 12) = P(X = 0) + P(X = 1) + \ldots + P(X = 12)$$

$$= \sum_{x=0}^{12} e^{-10} \frac{10^x}{x!} = 0.792$$

In Excel	
0.7916	= POISSON(12, 10, TRUE)

Poisson Mean & Variance

If X is a Poisson random variable with parameter λ, then:

$$\mu = E(X) = \lambda \quad \text{and} \quad \sigma^2 = V(X) = \lambda \qquad (3\text{-}17)$$

The mean and variance of the Poisson model are the same. If the mean and variance of a data set are not about the same, then the Poisson model would not be a good representation of that set.

The derivation of the mean and variance is shown in the text.

Discrete Random Variables

Many physical systems can be modeled by the same or similar random experiments and random variables. The distribution of the random variable involved in each of these common systems can be analyzed, and the results can be used in different applications and examples.

In this chapter, we present the analysis of several random experiments and discrete random variables that frequently arise in applications.

We often omit a discussion of the underlying sample space of the random experiment and directly describe the distribution of a particular random variable.

Important Terms & Concepts of Chapter 3

Bernoulli trial

Binomial distribution

Cumulative probability distribution – discrete random variable

Discrete uniform distribution

Expected value of a function of a random variable

Finite population correction factor

Geometric distribution

Hypergeometric distribution

Lack of memory property – discrete random variable

Mean – discrete random variable

Mean – function of a discrete random variable

Negative binominal distribution

Poisson distribution

Poisson process

Probability distribution – discrete random variable

Probability mass function

Standard deviation – discrete random variable

Variance – discrete random variable

3 Discrete Random Variables and Probability Distributions

CHAPTER OUTLINE