Chapter-5
Chapter-6

Some Discrete Probability Distribution
Some Continuous Probability Distribution

5-Discrete probability distributions.

Binomial probability distribution
Hypergeometric probability distribution
Multinomial probability distribution
Poisson probability distribution

6-Continuous probability distributions.

Normal probability distribution
Student's t distribution
Chi-square distribution
F distribution

The Bernoulli Process

Bernoulli process must possess the following properties:

1. The experiment consists of repeated trials.
2. Each trial results in an outcome that may be classified as a success or a failure.
3. The probability of success, denoted by p, remains constant from trial to trial.
4. The repeated trials are independent.

Binomial Distribution A Bernoulli trial can result in a success with probability p and a failure with probability $q = 1 - p$. Then the probability distribution of the binomial random variable X, the number of successes in n independent trials, is

$$b(x; n, p) = \binom{n}{x} p^x q^{n-x}, \quad x = 0, 1, 2, \ldots, n.$$

Theorem 5.1: The mean and variance of the binomial distribution $b(x; n, p)$ are

$$\mu = np \text{ and } \sigma^2 = npq.$$

Many types of probability problems have only two outcomes or can be reduced to two outcomes. For example, when a coin is tossed, it can land heads or tails. When a baby is born, it will be either male or female. In a basketball game, a team either wins or loses. A true/false item can be answered in only two ways, true or false. Other situations can be reduced to two outcomes. For example, a medical treatment can be classified as effective or ineffective, depending on the results. A person can be classified as having normal or abnormal blood pressure, depending on the measure of the blood pressure gauge. A multiple-choice question, even though there are four or five answer choices, can be classified as correct or incorrect. Situations like these are called *binomial experiments*.

Tossing Coins

A coin is tossed 3 times. Find the probability of getting exactly two heads.

Example

The probability that a certain kind of component will survive a shock test is $3/4$. Find the probability that exactly 2 of the next 4 components tested survive.

Solution:

$$b\left(2; 4, \frac{3}{4}\right) = \binom{4}{2}\left(\frac{3}{4}\right)^2\left(\frac{1}{4}\right)^2 = \left(\frac{4!}{2!\ 2!}\right)\left(\frac{3^2}{4^4}\right) = \frac{27}{128}.$$

Destination Weddings Twenty-six percent of couples who plan to marry this year are planning destination weddings. In a random sample of 12 couples who plan to marry, find the probability that

a. Exactly 6 couples will have a destination wedding
b. At least 6 couples will have a destination wedding
c. Fewer than 5 couples will have a destination wedding

a. 0.047 *b.* 0.065 *c.* 0.821

Source: *Time* magazine.

Driving While Intoxicated

A report from the Secretary of Health and Human Services stated that 70% of single-vehicle traffic fatalities that occur at night on weekends involve an intoxicated driver. If a sample of 15 single-vehicle traffic fatalities that occur at night on a weekend is selected, find the probability that exactly 12 involve a driver who is intoxicated.

Source: *100% American* by Daniel Evan Weiss.

0.17.

Survey on Employment

A survey from Teenage Research Unlimited (Northbrook, Illinois) found that 30% of teenage consumers receive their spending money from part-time jobs. If 5 teenagers are selected at random, find the probability that at least 3 of them will have part-time jobs.

0.162

Example 5.2: The probability that a patient recovers from a rare blood disease is 0.4. If 15 people are known to have contracted this disease, what is the probability that (a) at least 10 survive, (b) from 3 to 8 survive, and (c) exactly 5 survive?

Find the mean and variance of the binomial random variable

***Solution*:** Let X be the number of people who survive.

(a) $$P(X \geq 10) = 1 - P(X < 10) = 1 - \sum_{x=0}^{9} b(x; 15, 0.4) = 1 - 0.9662$$
$$= 0.0338$$

(b) $$P(3 \leq X \leq 8) = \sum_{x=3}^{8} b(x; 15, 0.4) = \sum_{x=0}^{8} b(x; 15, 0.4) - \sum_{x=0}^{2} b(x; 15, 0.4)$$
$$= 0.9050 - 0.0271 = 0.8779$$

(c) $$P(X = 5) = b(5; 15, 0.4) = \sum_{x=0}^{5} b(x; 15, 0.4) - \sum_{x=0}^{4} b(x; 15, 0.4)$$
$$= 0.4032 - 0.2173 = 0.1859$$

Table A.1 (continued) Binomial Probability Sums $\sum_{x=0}^{r} b(x; n, p)$

		p									
n	r	0.10	0.20	0.25	0.30	0.40	0.50	0.60	0.70	0.80	0.90
15	0	0.2059	0.0352	0.0134	0.0047	0.0005	0.0000				
	1	0.5490	0.1671	0.0802	0.0353	0.0052	0.0005	0.0000			
	2	0.8159	0.3980	0.2361	0.1268	0.0271	0.0037	0.0003	0.0000		
	3	0.9444	0.6482	0.4613	0.2969	0.0905	0.0176	0.0019	0.0001		
	4	0.9873	0.8358	0.6865	0.5155	0.2173	0.0592	0.0093	0.0007	0.0000	
	5	0.9978	0.9389	0.8516	0.7216	0.4032	0.1509	0.0338	0.0037	0.0001	
	6	0.9997	0.9819	0.9434	0.8689	0.6098	0.3036	0.0950	0.0152	0.0008	
	7	1.0000	0.9958	0.9827	0.9500	0.7869	0.5000	0.2131	0.0500	0.0042	0.0000
	8		0.9992	0.9958	0.9848	0.9050	0.6964	0.3902	0.1311	0.0181	0.0003
	9		0.9999	0.9992	0.9963	0.9662	0.8491	0.5968	0.2784	0.0611	0.0022
	10		1.0000	0.9999	0.9993	0.9907	0.9408	0.7827	0.4845	0.1642	0.0127
	11			1.0000	0.9999	0.9981	0.9824	0.9095	0.7031	0.3518	0.0556
	12				1.0000	0.9997	0.9963	0.9729	0.8732	0.6020	0.1841
	13					1.0000	0.9995	0.9948	0.9647	0.8329	0.4510
	14						1.0000	0.9995	0.9953	0.9648	0.7941
	15							1.0000	1.0000	1.0000	1.0000

5.7 One prominent physician claims that 70% of those with lung cancer are chain smokers. If his assertion is correct,

(a) find the probability that of 10 such patients recently admitted to a hospital, fewer than half are chain smokers;

(b) find the probability that of 20 such patients recently admitted to a hospital, fewer than half are chain smokers.

Solution: $p = 0.7$.

(a) For $n = 10$, $P(X < 5) = P(X \leq 4) = 0.0474$.

(b) For $n = 20$, $P(X < 10) = P(X \leq 9) = 0.0171$.

Table A.1 (continued) Binomial Probability Sums $\sum_{x=0}^{r} b(x; n, p)$

		p									
n	r	**0.10**	**0.20**	**0.25**	**0.30**	**0.40**	**0.50**	**0.60**	**0.70**	**0.80**	**0.90**
8	**0**	0.4305	0.1678	0.1001	0.0576	0.0168	0.0039	0.0007	0.0001	0.0000	
	1	0.8131	0.5033	0.3671	0.2553	0.1064	0.0352	0.0085	0.0013	0.0001	
	2	0.9619	0.7969	0.6785	0.5518	0.3154	0.1445	0.0498	0.0113	0.0012	0.0000
	3	0.9950	0.9437	0.8862	0.8059	0.5941	0.3633	0.1737	0.0580	0.0104	0.0004
	4	0.9996	0.9896	0.9727	0.9420	0.8263	0.6367	0.4059	0.1941	0.0563	0.0050
	5	1.0000	0.9988	0.9958	0.9887	0.9502	0.8555	0.6846	0.4482	0.2031	0.0381
	6		0.9999	0.9996	0.9987	0.9915	0.9648	0.8936	0.7447	0.4967	0.1869
	7		1.0000	1.0000	0.9999	0.9993	0.9961	0.9832	0.9424	0.8322	0.5695
	8				1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
9	**0**	0.3874	0.1342	0.0751	0.0404	0.0101	0.0020	0.0003	0.0000		
	1	0.7748	0.4362	0.3003	0.1960	0.0705	0.0195	0.0038	0.0004	0.0000	
	2	0.9470	0.7382	0.6007	0.4628	0.2318	0.0898	0.0250	0.0043	0.0003	0.0000
	3	0.9917	0.9144	0.8343	0.7297	0.4826	0.2539	0.0994	0.0253	0.0031	0.0001
	4	0.9991	0.9804	0.9511	0.9012	0.7334	0.5000	0.2666	0.0988	0.0196	0.0009
	5	0.9999	0.9969	0.9900	0.9747	0.9006	0.7461	0.5174	0.2703	0.0856	0.0083
	6	1.0000	0.9997	0.9987	0.9957	0.9750	0.9102	0.7682	0.5372	0.2618	0.0530
	7		1.0000	0.9999	0.9996	0.9962	0.9805	0.9295	0.8040	0.5638	0.2252
	8			1.0000	1.0000	0.9997	0.9980	0.9899	0.9596	0.8658	0.6126
	9					1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
10	**0**	0.3487	0.1074	0.0563	0.0282	0.0060	0.0010	0.0001	0.0000		
	1	0.7361	0.3758	0.2440	0.1493	0.0464	0.0107	0.0017	0.0001	0.0000	
	2	0.9298	0.6778	0.5256	0.3828	0.1673	0.0547	0.0123	0.0016	0.0001	
	3	0.9872	0.8791	0.7759	0.6496	0.3823	0.1719	0.0548	0.0106	0.0009	0.0000
	4	0.9984	0.9672	0.9219	0.8497	0.6331	0.3770	0.1662	0.0473	0.0064	0.0001
	5	0.9999	0.9936	0.9803	0.9527	0.8338	0.6230	0.3669	0.1503	0.0328	0.0016
	6	1.0000	0.9991	0.9965	0.9894	0.9452	0.8281	0.6177	0.3504	0.1209	0.0128
	7		0.9999	0.9996	0.9984	0.9877	0.9453	0.8327	0.6172	0.3222	0.0702
	8		1.0000	1.0000	0.9999	0.9983	0.9893	0.9536	0.8507	0.6242	0.2639
	9				1.0000	0.9999	0.9990	0.9940	0.9718	0.8926	0.6513

5.9 In testing a certain kind of truck tire over rugged terrain, it is found that 25% of the trucks fail to complete the test run without a blowout. Of the next 15 trucks tested, find the probability that

(a) from 3 to 6 have blowouts;

(b) fewer than 4 have blowouts;

(c) more than 5 have blowouts.

Solution:

For $n = 15$ and $p = 0.25$, we have

(a) $P(3 \leq X \leq 6) = P(X \leq 6) - P(X \leq 2) = 0.9434 - 0.2361 = 0.7073.$

(b) $P(X < 4) = P(X \leq 3) = 0.4613.$

(c) $P(X > 5) = 1 - P(X \leq 5) = 1 - 0.8516 = 0.1484.$

5.26 Assuming that 6 in 10 automobile accidents are due mainly to a speed violation, find the probability that among 8 automobile accidents, 6 will be due mainly to a speed violation

(a) by using the formula for the binomial distribution;

(b) by using Table A.1.

Home work

5.11 The probability that a patient recovers from a delicate heart operation is 0.9. What is the probability that exactly 5 of the next 7 patients having this operation survive?

Home work

5.12 A traffic control engineer reports that 75% of the vehicles passing through a checkpoint are from within the state. What is the probability that fewer than 4 of the next 9 vehicles are from out of state?

Solution

5.11 From Table A.1 with $n = 7$ and $p = 0.9$, we have

$$P(X = 5) = P(X \leq 5) - P(X \leq 4) = 0.1497 - 0.0257 = 0.1240.$$

5.12 From Table A.1 with $n = 9$ and $p = 0.25$, we have $P(X < 4) = 0.8343$.

Multinomial Distribution

If a given trial can result in the k outcomes $E_1, E_2, \ldots, E_k$ with probabilities $p_1, p_2, \ldots, p_k$, then the probability distribution of the random variables $X_1, X_2, \ldots, X_k$, representing the number of occurrences for $E_1, E_2, \ldots, E_k$ in n independent trials, is

$$f(x_1, x_2, \ldots, x_k; p_1, p_2, \ldots, p_k, n) = \binom{n}{x_1, x_2, \ldots, x_k} p_1^{x_1} p_2^{x_2} \cdots p_k^{x_k},$$

with

$$\sum_{i=1}^{k} x_i = n \text{ and } \sum_{i=1}^{k} p_i = 1.$$

5.7: The complexity of arrivals and departures of planes at an airport is such that computer simulation is often used to model the "ideal" conditions. For a certain airport with three runways, it is known that in the ideal setting the following are the probabilities that the individual runways are accessed by a randomly arriving commercial jet:

Runway 1: $p_1 = 2/9$,
Runway 2: $p_2 = 1/6$,
Runway 3: $p_3 = 11/18$.

What is the probability that 6 randomly arriving airplanes are distributed in the following fashion?

Runway 1: 2 airplanes,
Runway 2: 1 airplane,
Runway 3: 3 airplanes

5.23 The probabilities are 0.4, 0.2, 0.3, and 0.1, respectively, that a delegate to a certain convention arrived by air, bus, automobile, or train. What is the probability that among 9 delegates randomly selected at this convention, 3 arrived by air, 3 arrived by bus, 1 arrived by automobile, and 2 arrived by train?

Solution

5.23 Using the multinomial distribution, we have

$$f(x_1, x_2, \ldots, x_k; p_1, p_2, \ldots, p_k, n) = \binom{n}{x_1, x_2, \ldots, x_k} p_1^{x_1} p_2^{x_2} \cdots p_k^{x_k},$$

$$\binom{9}{3, 3, 1, 2}(0.4)^3(0.2)^3(0.3)(0.2)^2 = 0.0077.$$

Hypergeometric Distribution

as a **hypergeometric experiment**, that is, one that possesses the following two properties:

1. A random sample of size n is selected without replacement from N items.
2. Of the N items, k may be classified as successes and $N-k$ are classified as failures.

Hypergeometric Distribution The probability distribution of the hypergeometric random variable X, the number of successes in a random sample of size n selected from N items of which k are labeled success and $N-k$ labeled failure, is

$$h(x; N, n, k) = \frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}},$$

5.9: Lots of 40 components each are deemed unacceptable if they contain 3 or more defectives. The procedure for sampling a lot is to select 5 components at random and to reject the lot if a defective is found. What is the probability that exactly 1 defective is found in the sample if there are 3 defectives in the entire lot?

Solution:

Using the hypergeometric distribution with

$$h(x; N, n, k) = \frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}},$$

$n = 5,\ N = 40,\ k = 3,$ and $x = 1$

$$h(1; 40, 5, 3) = \frac{\binom{3}{1}\binom{37}{4}}{\binom{40}{5}} = 0.3011.$$

Theorem 5.2:

The mean and variance of the hypergeometric distribution $h(x; N, n, k)$ are

$$\mu = \frac{nk}{N} \text{ and } \sigma^2 = \frac{N-n}{N-1} \cdot n \cdot \frac{k}{N}\left(1 - \frac{k}{N}\right).$$

Example 5.11:

Find the mean and variance of the random variable of Example 5.9 and then use Chebyshev's theorem to interpret the interval $\mu \pm 2\sigma$.

Solution:

Since Example 5.9 was a hypergeometric experiment with $N = 40$, $n = 5$, and $k = 3$, by Theorem 5.2, we have

$$\mu = \frac{(5)(3)}{40} = \frac{3}{8} = 0.375,$$

and

$$\sigma^2 = \left(\frac{40-5}{39}\right)(5)\left(\frac{3}{40}\right)\left(1 - \frac{3}{40}\right) = 0.3113.$$

Class Activity

5.29 A homeowner plants 6 bulbs selected at random from a box containing 5 tulip bulbs and 4 daffodil bulbs. What is the probability that he planted 2 daffodil bulbs and 4 tulip bulbs?

$$h(2; 9, 6, 4) = \frac{\binom{4}{2}\binom{5}{4}}{\binom{9}{6}} = \frac{5}{14}.$$

5.31 A random committee of size 3 is selected from 4 doctors and 2 nurses. Write a formula for the probability distribution of the random variable X representing the number of doctors on the committee. Find $P(2 \leq X \leq 3)$.

$$h(x; 6, 3, 4) = \frac{\binom{4}{x}\binom{2}{3-x}}{\binom{6}{3}}, \text{ for } x = \quad 2, 3.$$

5.32 From a lot of 10 missiles, 4 are selected at random and fired. If the lot contains 3 defective missiles that will not fire, what is the probability that

(a) all 4 will fire?

(b) at most 2 will not fire?

$$h(4; 10, 4, 7) = \frac{1}{6}.$$

$$= \sum_{x=0}^{2} h(x; 10, 4, 3) = \frac{29}{30}.$$

5.33 If 7 cards are dealt from an ordinary deck of 52 playing cards, what is the probability that

(a) exactly 2 of them will be face cards?

(b) at least 1 of them will be a queen?

(a) $\frac{\binom{12}{2}\binom{40}{5}}{\binom{52}{7}} = 0.3246.$

(b) $1 - \frac{\binom{48}{7}}{\binom{52}{7}} = 0.4496.$

Geometric Distribution

If repeated independent trials can result in a success with probability p and a failure with probability $q = 1 - p$, then the probability distribution of the random variable X, the number of the trial on which the first success occurs, is

$$g(x; p) = pq^{x-1}, \quad x = 1, 2, 3, \ldots .$$

Example:

For a certain manufacturing process, it is known that, on the average, 1 in every 100 items is defective. What is the probability that the fifth item inspected is the first defective item found?

with $x = 5$ and $p = 0.01$,

$$g(5; 0.01) = (0.01)(0.99)^4 = 0.0096.$$

Experiments yielding numerical values of a random variable X, the number of outcomes occurring during a given time interval or in a specified region, are called **Poisson experiments.** The given time interval may be of any length, such as a minute, a day, a week, a month, or even a year. For example, a Poisson experiment

Poisson Probability Formula

Probabilities for a random variable X that has a Poisson distribution are given by the formula

$$P(X = x) = e^{-\lambda}\frac{\lambda^x}{x!}, \qquad x = 0, 1, 2, \ldots,$$

where λ is a positive real number and $e \approx 2.718$. (Most calculators have an e key.) The random variable X is called a **Poisson random variable** and is said to have the **Poisson distribution** with parameter λ.

Mean and Standard Deviation of a Poisson Random Variable

The mean and standard deviation of a Poisson random variable with parameter λ are

$$\mu = \lambda \quad \text{and} \quad \sigma = \sqrt{\lambda},$$

respectively.

Example:

5.17: During a laboratory experiment, the average number of radioactive particles passing through a counter in 1 millisecond is 4. What is the probability that 6 particles enter the counter in a given millisecond?

$$p(6;4) = \frac{e^{-4}4^6}{6!}$$

$$= \sum_{x=0}^{6} p(x;4) - \sum_{x=0}^{5} p(x;4)$$

$$= 0.8893 - 0.7851 = 0.1042.$$

IMR in Finland The infant mortality rate (IMR) is the number of deaths of children under 1 year old per 1000 live births during a calendar year. From the *World Factbook*, the Central Intelligence Agency's most popular publication, we found that the IMR in Finland is 3.5. Use the Poisson approximation to determine the probability that, of 500 randomly selected live births in Finland, there are

a. no infant deaths.
b. at most three infant deaths.

Table A.2 Poisson Probability Sums $\sum_{x=0}^{r} p(x;\mu)$

	μ								
r	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0	0.9048	0.8187	0.7408	0.6703	0.6065	0.5488	0.4966	0.4493	0.4066
1	0.9953	0.9825	0.9631	0.9384	0.9098	0.8781	0.8442	0.8088	0.7725
2	0.9998	0.9989	0.9964	0.9921	0.9856	0.9769	0.9659	0.9526	0.9371
3	1.0000	0.9999	0.9997	0.9992	0.9982	0.9966	0.9942	0.9909	0.9865
4		1.0000	1.0000	0.9999	0.9998	0.9996	0.9992	0.9986	0.9977
5				1.0000	1.0000	1.0000	0.9999	0.9998	0.9997
6							1.0000	1.0000	1.0000

	μ								
r	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
0	0.3679	0.2231	0.1353	0.0821	0.0498	0.0302	0.0183	0.0111	0.0067
1	0.7358	0.5578	0.4060	0.2873	0.1991	0.1359	0.0916	0.0611	0.0404
2	0.9197	0.8088	0.6767	0.5438	0.4232	0.3208	0.2381	0.1736	0.1247
3	0.9810	0.9344	0.8571	0.7576	0.6472	0.5366	0.4335	0.3423	0.2650
4	0.9963	0.9814	0.9473	0.8912	0.8153	0.7254	0.6288	0.5321	0.4405
5	0.9994	0.9955	0.9834	0.9580	0.9161	0.8576	0.7851	0.7029	0.6160
6	0.9999	0.9991	0.9955	0.9858	0.9665	0.9347	0.8893	0.8311	0.7622
7	1.0000	0.9998	0.9989	0.9958	0.9881	0.9733	0.9489	0.9134	0.8666
8		1.0000	0.9998	0.9989	0.9962	0.9901	0.9786	0.9597	0.9319
9			1.0000	0.9997	0.9989	0.9967	0.9919	0.9829	0.9682
10				0.9999	0.9997	0.9990	0.9972	0.9933	0.9863
11				1.0000	0.9999	0.9997	0.9991	0.9976	0.9945
12					1.0000	0.9999	0.9997	0.9992	0.9980
13						1.0000	0.9999	0.9997	0.9993
14							1.0000	0.9999	0.9998
15								1.0000	0.9999
16									1.0000

Example:

Emergency Room Traffic Desert Samaritan Hospital keeps records of emergency room (ER) traffic. Those records indicate that the number of patients arriving between 6:00 P.M. and 7:00 P.M. has a Poisson distribution with parameter $\lambda = 6.9$. Determine the probability that, on a given day, the number of patients who arrive at the emergency room between 6:00 P.M. and 7:00 P.M. will be

a. exactly 4.
b. at most 2.
c. between 4 and 10, inclusive.
d. Obtain a table of probabilities for the random variable X, the number of patients arriving between 6:00 P.M. and 7:00 P.M. Stop when the probabilities become zero to three decimal places.
e. Use part (d) to construct a (partial) probability histogram for X.
f. Identify the shape of the probability distribution of X.

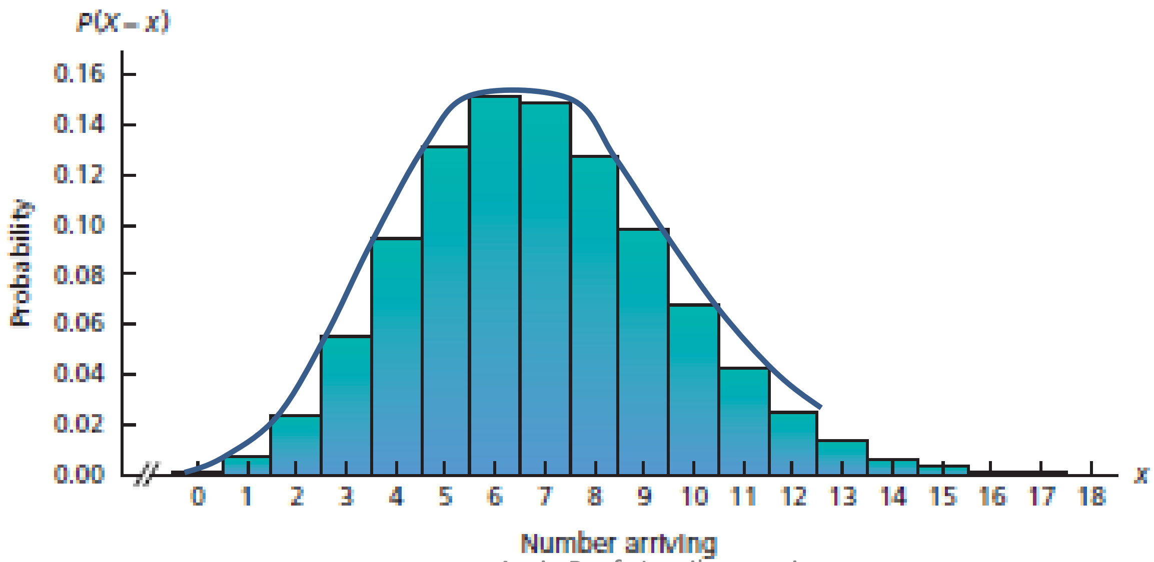
Answers:

a) 0.095
b)0.032
c)0.821

d)

Number arriving x	Probability $P(X = x)$	Number arriving x	Probability $P(X = x)$
0	0.001	10	0.068
1	0.007	11	0.043
2	0.024	12	0.025
3	0.055	13	0.013
4	0.095	14	0.006
5	0.131	15	0.003
6	0.151	16	0.001
7	0.149	17	0.001
8	0.128	18	0.000
9	0.098		

e)



Poisson Approximation to the Binomial Distribution

Recall that the binomial probability formula is

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

To Approximate Binomial Probabilities by Using a Poisson Probability Formula

Step 1 Find n, the number of trials, and p, the success probability.

Step 2 Continue only if $n \geq 100$ and $np \leq 10$.

Step 3 Approximate the binomial probabilities by using the Poisson probability formula

$$P(X = x) = e^{-np} \frac{(np)^x}{x!}.$$

Theorem 5.5:

Let X be a binomial random variable with probability distribution $b(x; n, p)$. When $n \to \infty$, $p \to 0$, and $np \overset{n\to\infty}{\longrightarrow} \mu$ remains constant,

$$b(x; n, p) \overset{n\to\infty}{\longrightarrow} p(x; \mu).$$

Example 5.19:

In a certain industrial facility, accidents occur infrequently. It is known that the probability of an accident on any given day is 0.005 and accidents are independent of each other.

(a) What is the probability that in any given period of 400 days there will be an accident on one day?

(b) What is the probability that there are at most three days with an accident?

***Solution*:**

$$n = 400 \text{ and } p = 0.005. \text{ Thus, } np = 2.$$

(a) $P(X = 1) = e^{-2}2^1 = 0.271$ and

(b) $P(X \leq 3) = \sum_{x=0}^{3} e^{-2}2^x/x! = 0.857.$

Class Activity

5.55 The probability that a student pilot passes the written test for a private pilot's license is 0.7. Find the probability that a given student will pass the test

(a) on the third try;

(b) before the fourth try.

(a) $P(X = 3) = g(3; 0.7) = (0.7)(0.3)^2 = 0.0630.$

(b) $P(X < 4) = \sum_{x=1}^{3} g(x; 0.7) = \sum_{x=1}^{3} (0.7)(0.3)^{x-1} = 0.9730.$

5.56 On average, 3 traffic accidents per month occur at a certain intersection. What is the probability that in any given month at this intersection

(a) exactly 5 accidents will occur?

(b) fewer than 3 accidents will occur?

(c) at least 2 accidents will occur?

(a) $p(5; 3) = \sum_{x=0}^{5} p(x; 3) - \sum_{x=0}^{4} p(x; 3) = 0.1008.$

(b) $P(X < 3) = P(X \leq 2) = 0.4232.$

(c) $P(X \geq 2) = 1 - P(X \leq 1) = 0.8009.$

5.57 On average, a textbook author makes two word-processing errors per page on the first draft of her textbook. What is the probability that on the next page she will make

(a) 4 or more errors?

(b) no errors?

(a) $P(X \geq 4) = 1 - P(X \leq 3) = 0.1429.$

(b) $P(X = 0) = p(0; 2) = 0.1353.$

5.61 Suppose that, on average, 1 person in 1000 makes a numerical error in preparing his or her income tax return. If 10,000 returns are selected at random and examined, find the probability that 6, 7, or 8 of them contain an error.

$\mu = np = (10000)(0.001) = 10$, so

$P(6 \leq X \leq 8) \approx 0.8622 - 0.5246 = 0.3376.$

Continuous probability distributions.

- ✓ Normal probability distribution

Student's t distribution

Chi-square distribution

F distribution

Continuous Prob.Distribution

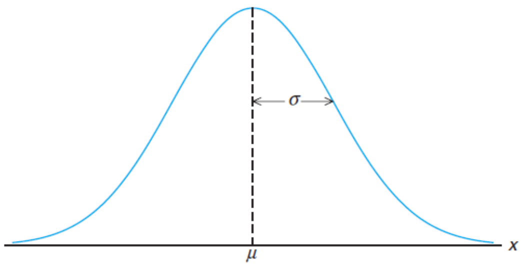


Figure 6.2: The normal curve.

The most important continuous probability distribution in the entire field of statistics is the **normal distribution**. Its graph, called the **normal curve**, is the

Normal Distribution The density of the normal random variable X, with mean μ and variance σ^2, is

$$n(x;\mu,\sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}, \quad -\infty < x < \infty,$$

where $\pi = 3.14159\ldots$ and $e = 2.71828\ldots$.

Theorem 6.2: The mean and variance of $n(x;\mu,\sigma)$ are μ and σ^2, respectively. Hence, the standard deviation is σ.

Area under Normal Curve

$$P(x_1 < X < x_2) = \int_{x_1}^{x_2} n(x;\mu,\sigma)\,dx = \frac{1}{\sqrt{2\pi}\sigma}\int_{x_1}^{x_2} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}dx$$

is represented by the area of the shaded region.

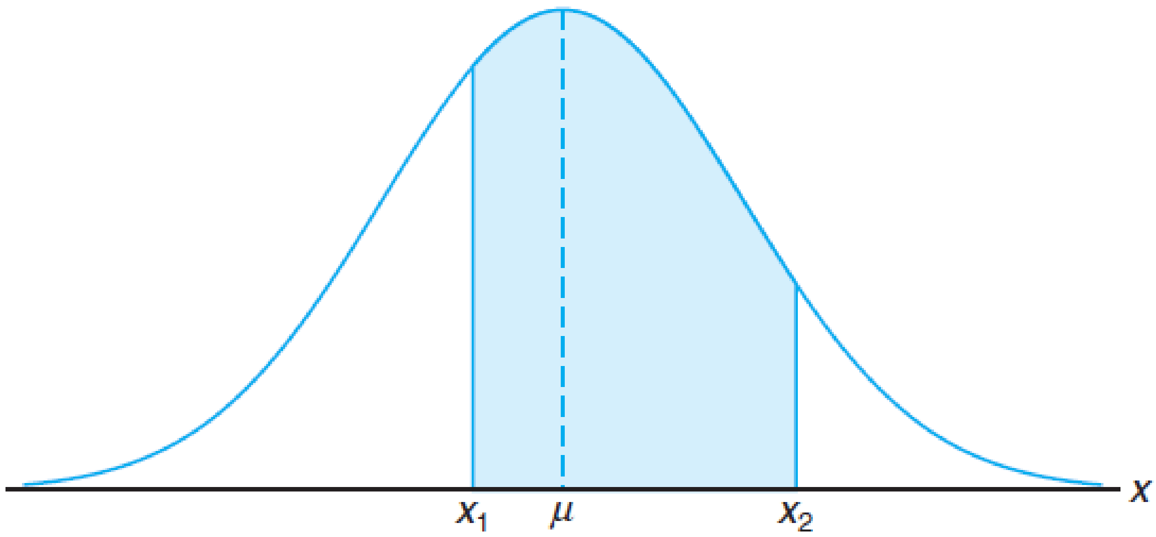


Figure 6.6: $P(x_1 < X < x_2) =$ area of the shaded region.

Z-curves

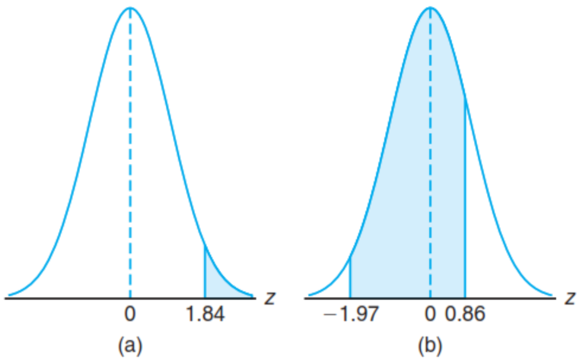
$$Z = \frac{X - \mu}{\sigma}.$$

Definition 6.1: The distribution of a normal random variable with mean 0 and variance 1 is called a standard normal distribution.

6.2: Given a standard normal distribution, find the area under the curve that lies

(a) to the right of $z = 1.84$ and

(b) between $z = -1.97$ and $z = 0.86$.



Use table

(a) $$1 - 0.9671 = 0.0329.$$

(b) $$0.8051 - 0.0244 = 0.7807.$$

6.4: Given a random variable X having a normal distribution with $\mu = 50$ and $\sigma = 10$, find the probability that X assumes a value between 45 and 62.

$$P(45 < X < 62) = P(-0.5 < Z < 1.2)$$

Use table

$$= P(Z < 1.2) - P(Z < -0.5)$$

$$= 0.8849 - 0.3085 = 0.5764.$$

6.5: Given that X has a normal distribution with $\mu = 300$ and $\sigma = 50$, find the probability that X assumes a value greater than 362.

$$z = \frac{362 - 300}{50} = 1.24.$$

$$P(X > 362) = P(Z > 1.24)$$

$$= 1 - P(Z < 1.24) = 1 - 0.8925$$

$$= 0.1075.$$

Table A.3 Areas under the Normal Curve

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
−3.4	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0002
−3.3	0.0005	0.0005	0.0005	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0003
−3.2	0.0007	0.0007	0.0006	0.0006	0.0006	0.0006	0.0006	0.0005	0.0005	0.0005
−3.1	0.0010	0.0009	0.0009	0.0009	0.0008	0.0008	0.0008	0.0008	0.0007	0.0007
−3.0	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010
−2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
−2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
−2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
−2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
−2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
−2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
−2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
−2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
−2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
−2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
−1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
−1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
−1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
−1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
−1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
−1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681
−1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
−1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
−1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
−1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
−0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
−0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
−0.7	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148
−0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
−0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
−0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
−0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
−0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
−0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
−0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641

Table A.3 (continued) Areas under the Normal Curve

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998

Applications of the Normal Distribution

Example 6.7:

A certain type of storage battery lasts, on average, 3.0 years with a standard deviation of 0.5 year. Assuming that battery life is normally distributed, find the probability that a given battery will last less than 2.3 years.

Solution:

$$P(X < 2.3) = P(Z < -1.4) = 0.0808.$$

Example 6.8:

An electrical firm manufactures light bulbs that have a life, before burn-out, that is normally distributed with mean equal to 800 hours and a standard deviation of 40 hours. Find the probability that a bulb burns between 778 and 834 hours.

Solution:

$$P(778 < X < 834) = P(-0.55 < Z < 0.85) = P(Z < 0.85) - P(Z < -0.55)$$
$$= 0.8023 - 0.2912 = 0.5111.$$

Theorem 6.3:

If X is a binomial random variable with mean $\mu = np$ and variance $\sigma^2 = npq$, then the limiting form of the distribution of

$$Z = \frac{X - np}{\sqrt{npq}},$$

as $n \to \infty$, is the standard normal distribution $n(z; 0, 1)$.

Example:

$$P(X = 4) = b(4; 15, 0.4) = 0.1268,$$

Binomial

$$z_1 = \frac{3.5 - 6}{1.897} = -1.32 \quad \text{and} \quad z_2 = \frac{4.5 - 6}{1.897} = -0.79.$$

If X is a binomial random variable and Z a standard normal variable, then

$$P(X = 4) = b(4; 15, 0.4) \approx P(-1.32 < Z < -0.79)$$

Normal

$$= P(Z < -0.79) - P(Z < -1.32) = 0.2148 - 0.0934 = 0.1214.$$

This agrees very closely with the exact value of 0.1268.

The normal approximation is most useful in calculating binomial sums for large values of n.

Example:

$$P(7 \le X \le 9) = \sum_{x=0}^{9} b(x;15,0.4) - \sum_{x=0}^{6} b(x;15,0.4)$$

Binomial

$$= 0.9662 - 0.6098 = 0.3564,$$

$$z_1 = \frac{6.5-6}{1.897} = 0.26 \quad \text{and} \quad z_2 = \frac{9.5-6}{1.897} = 1.85.$$

Now,

Normal

$$P(7 \le X \le 9) \approx P(0.26 < Z < 1.85) = P(Z < 1.85) - P(Z < 0.26)$$
$$= 0.9678 - 0.6026 = 0.3652.$$

Applications of the Normal Distribution

Example 6.15:

The probability that a patient recovers from a rare blood disease is 0.4. If 100 people are known to have contracted this disease, what is the probability that fewer than 30 survive?

Solution:

$$\mu = np = (100)(0.4) = 40 \text{ and } \sigma = \sqrt{npq} = \sqrt{(100)(0.4)(0.6)} = 4.899.$$

$$z = \frac{29.5 - 40}{4.899} = -2.14,$$

$$\boxed{P(X < 30) \approx P(Z < -2.14) = 0.0162.}$$

Example 6.16:

A multiple-choice quiz has 200 questions, each with 4 possible answers of which only 1 is correct. What is the probability that sheer guesswork yields from 25 to 30 correct answers for the 80 of the 200 problems about which the student has no knowledge?

Solution:

$$P(25 \le X \le 30) = \sum_{x=25}^{30} b(x; 80, 1/4).$$

Using the normal curve approximation with

$$\mu = np = (80)\left(\frac{1}{4}\right) = 20$$

and

$$\sigma = \sqrt{npq} = \sqrt{(80)(1/4)(3/4)} = 3.873,$$

$$z_1 = \frac{24.5 - 20}{3.873} = 1.16 \text{ and } z_2 = \frac{30.5 - 20}{3.873} = 2.71.$$

$$\boxed{\begin{aligned} P(25 \le X \le 30) &= \sum_{x=25}^{30} b(x; 80, 0.25) \approx P(1.16 < Z < 2.71) \\ &= P(Z < 2.71) - P(Z < 1.16) = 0.9966 - 0.8770 = 0.1196. \end{aligned}}$$

Home Activity

6.13 A research scientist reports that mice will live an average of 40 months when their diets are sharply restricted and then enriched with vitamins and proteins. Assuming that the lifetimes of such mice are normally distributed with a standard deviation of 6.3 months, find the probability that a given mouse will live

(a) more than 32 months;

(b) less than 28 months;

(c) between 37 and 49 months.

6.18 The heights of 1000 students are normally distributed with a mean of 174.5 centimeters and a standard deviation of 6.9 centimeters. Assuming that the heights are recorded to the nearest half-centimeter, how many of these students would you expect to have heights

(a) less than 160.0 centimeters?

(b) between 171.5 and 182.0 centimeters inclusive?

(c) equal to 175.0 centimeters?

(d) greater than or equal to 188.0 centimeters?

6.20 The weights of a large number of miniature poodles are approximately normally distributed with a mean of 8 kilograms and a standard deviation of 0.9 kilogram. If measurements are recorded to the nearest tenth of a kilogram, find the fraction of these poodles with weights

(a) over 9.5 kilograms;

(b) of at most 8.6 kilograms;

(c) between 7.3 and 9.1 kilograms inclusive.

6.34 A pair of dice is rolled 180 times. What is the probability that a total of 7 occurs

(a) at least 25 times?

(b) between 33 and 41 times inclusive?

(c) exactly 30 times?

6.27 The probability that a patient recovers from a delicate heart operation is 0.9. Of the next 100 patients having this operation, what is the probability that

(a) between 84 and 95 inclusive survive?

(b) fewer than 86 survive?

Summary

Binomial distribution

$$\begin{aligned} n &= \text{number of trials} \\ p &= \text{probability of success} \\ P(x) &= \binom{n}{x} p^x q^{n-x} \\ \mathbf{E}(X) &= np \\ \mathrm{Var}(X) &= npq \end{aligned}$$

Geometric distribution

$$\begin{aligned} p &= \text{probability of success} \\ P(x) &= (1-p)^{x-1}p, \quad x = 1, 2, \ldots \\ \mathbf{E}(X) &= \frac{1}{p} \\ \mathrm{Var}(X) &= \frac{1-p}{p^2} \end{aligned}$$

Summary

Poisson distribution

$$\begin{aligned}
\lambda &= \text{frequency, average number of events} \\
P(x) &= e^{-\lambda}\frac{\lambda^x}{x!}, \quad x = 0, 1, 2, \ldots \\
\mathbf{E}(X) &= \lambda \\
\mathrm{Var}(X) &= \lambda
\end{aligned}$$

Normal distribution

$$\begin{aligned}
\mu &= \text{expectation, location parameter} \\
\sigma &= \text{standard deviation, scale parameter} \\
f(x) &= \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{\frac{-(x-\mu)^2}{2\sigma^2}\right\}, \quad -\infty < x < \infty \\
\mathbf{E}(X) &= \mu \\
\mathrm{Var}(X) &= \sigma^2
\end{aligned}$$

ANY Questions?