7. Probability Distributions

7. Probability Distributions

7.1 Introduction

In this section the concept of random variables will be discussed and some typical probability distributions will be introduced.

7.2 Random Variables

7.2.1 Introduction

There are many cases in which each outcome of an experiment can be adequately represented by a single numerical value. In such cases a function may be used to associate a real value with each outcome in the sample space. A function like this is called a random variable.

7.2.2 Random Variables and Range Space

A *random variable* is a real-value function whose domain is a sample space. Random variables are usually denoted by upper-case letters such as X, Y and Z.

The set of all possible values of a random variable is called its *range space*, usually denoted by R.

Example 7.2-1

Suppose 3 fair coins are tossed. The sample space of this experiment can be written as

$$S = \{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\}$$

Let the random variable X be the number of tails facing upwards. Then

$$\begin{aligned} X(HHH) &= 0 \\ X(HHT) &= X(HTH) = X(THH) = 1 \\ X(HTT) &= X(THT) = X(TTH) = 2 \\ X(TTT) &= 3 \end{aligned}$$

The range space $R = \{0, 1, 2, 3\}$.

Example 7.2-2

Suppose we record the time interval between the arrival of two customers in a shop. A suitable sample space for this experiment is

$$S = \{t : t \geq 0\}.$$

Then a possible random variable is $X(s) = r$.
The corresponding range space is $R = \{t : t \geq 0\}$.

In Example 7.2-1, R is a countable set and X is called a *discrete random variable*.
In Example 7.2-2, R is uncountable and X is called a *continuous random variable*.

7.2.3 Probability Distributions

The probability P(X = x) is the probability of the equivalent event

$$\{s \in S : X(s) = x\}.$$

P(X = x) is often written as p(x).

In Example 7.2-1,

$$p(0) = P(X = 0) = P(HHH) = \frac{1}{8}$$

$$p(1) = P(X = 1) = P(HHT, HTH, THH) = \frac{3}{8}$$

$$p(2) = P(X = 2) = P(HTT, THT, TTH) = \frac{3}{8}$$

$$p(3) = P(X = 3) = P(TTT) = \frac{1}{8}$$

For a discrete random variable X, the set of pairs $\{(x, p(x)) : x \in R\}$ is called the *probability distribution* (p.d) of X.

For a discrete random variable, the probability distribution must satisfy the following conditions:

(i) $0 \le p(x) \le 1$ for any $x \in R$

(ii) $\sum_{x \in R} p(x) = 1$

A probability distribution is often written in tabular form. For example, the p.d. of X in Example 1 may be written as

x	0	1	2	3
p(x) = P(X = x)	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

Example 7.2-3

Suppose two fair dice are tossed. Let the random variable X be the total score on the two dices. The range space is

$$R = \{2, 3, 4, \ldots, 12\}.$$

The probability distribution of X is

x	2	3	4	5	6	7	8	9	10	11	12
p(x)	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Example 7.2-4

A salesman visits 3 customers A, B and C to sell a product. The probabilities that A, B and C will order the product are 0.2, 0.3 and 0.5 respectively. Let X be the number of customers that will order the product. Find the probability distribution of X.

Let A, B and C be the events that customer A, B and C will order the product.

$$\begin{aligned} P(X = 0) &= P(\bar{A}\bar{B}\bar{C}) \\ &= (1 - 0.2)(1 - 0.3)(1 - 0.5) \\ &= 0.28 \end{aligned}$$

$$\begin{aligned} P(X = 1) &= P(A\bar{B}\bar{C} \text{ or } \bar{A}B\bar{C} \text{ or } \bar{A}\bar{B}C) \\ &= (0.2)(0.7)(0.5) + (0.8)(0.3)(0.5) + (0.8)(0.7)(0.5) \\ &= 0.47 \end{aligned}$$

$$\begin{aligned} P(X = 2) &= P(AB\bar{C} \text{ or } A\bar{B}C \text{ or } \bar{A}BC) \\ &= (0.2)(0.3)(0.5) + (0.2)(0.7)(0.5) + (0.8)(0.3)(0.5) \\ &= 0.22 \end{aligned}$$

$$\begin{aligned} P(X = 3) &= P(ABC) \\ &= (0.2)(0.3)(0.5) \\ &= 0.03 \end{aligned}$$

Example 7.2-5

In a test paper, there are five true-or-false questions. Two marks are awarded for each correct answer and one mark is deducted for each wrong answer. Suppose a student answers each question by choosing T or F randomly and let X be the total marks he gets.

The sample space can be wriiten as

$$S = \{WWWWW, WWWWC, WWWCW, \ldots, CCCCC\}$$

where, for example, the outcome WWWWC means that the first 4 answers are wrong and the fifth answer is correct.

Number of outcomes in S = $2^5 = 32$

Number of outcomes with k C's = ${}_5C_k$, with k = 0, 1, 2, 3, 4, 5

Obviously $X = 2k - (5 - k)$ where k is the number of correct answers. The range space of X is

$$R = \{-5, -2, 1, 4, 7, 10\}.$$

The probability distribution of X is shown below

x	-5	-2	1	4	7	10
p(x)	$\frac{1}{32}$	$\frac{5}{32}$	$\frac{10}{32}$	$\frac{10}{32}$	$\frac{5}{32}$	$\frac{1}{32}$

7.3 Mathematical Expectation

7.3.1 Expected Value

In each trial of an experiment, a random variable X may assume any value in the range space R. In the long-run, an average may be associated with X. This long run average is called the expected value of X. Similarly, a variance may be associated with X. The variance indicates the spread of values taken by the random variable.

Example 7.3-1

The length in meters, X, of metal bars produced in a machine is such that

$$P(X < 4.5) = 0.1$$
$$P(4.5 \le X \le 5) = 0.7$$
$$P(X > 5) = 0.2.$$

If $4.5 \le X \le 5$, a bar will be accepted by the customers with a profit of \$52.

If $X < 4.5$, a bar will be discarded with a loss of \$20.

If $X > 5$, a bar will be cut to the specified length and the profit will be reduced to \$48. Find the expected profit.

Consider N bars produced and let

A be the number of bars accepted,
R be the number of bars rejected, and
C be the number of bars that have to be cut.

$$\text{Average profit } P = \frac{52A + (-20)R + 48C}{N}$$
$$= 52\frac{A}{N} - 20\frac{R}{N} + 48\frac{C}{N}$$

$$\therefore \lim_{N\to\infty} P = 52\lim_{N\to\infty}\frac{A}{N} - 20\lim_{N\to\infty}\frac{R}{N} + 48\lim_{N\to\infty}\frac{C}{N}$$
$$= 52(0.7) - 20(0.1) + 48(0.2)$$
$$= 44$$

The expected profit is therefore \$44.

In general, let X be a discrete random variable with possible values $x_1, x_2, \ldots, x_n$ and corresponding probabilities $p_1, p_2, \ldots, p_n$. The *expectation* of X is defined as

$$E(X) = \sum_x xp(x)$$
$$= x_1p_1 + x_2p_2 + \ldots + x_np_n .$$

The expectation of X is also called the *expected value* of X.

Example 7.3-2

The random variable X has the distribution shown below

x	1	2	3	4
P(X = x)	0.3	0.2	0.4	0.1

The expected value of X
= E(X)
= (1)(0.3) + (2)(0.2) + (3)(0.4) + (4)(0.1)
= 2.3

Example 7.3-3

The random variable X a probability distribution given by

$$P(X = x) = \frac{c}{x}, \qquad \text{where } x = 1, 2, 3, 4, 5.$$

(a) Find c.
(b) Find E(X)

(a) $$c\left(1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5}\right) = 1$$

$$c = \frac{60}{137}$$

(b) $$E(X) = (c)(1) + (\frac{c}{2})(2) + (\frac{c}{3})(3) + (\frac{c}{4})(4) + (\frac{c}{5})(5)$$

$$= 5c$$

$$= \frac{300}{137}$$

Example 7.3-4

In a game, a player tosses 3 fair coins. He wins $10 if 3 heads occur, $5 if 2 heads occur, $2 if only 1 head occurs and losses $15 if no heads occur. What is his expected gain?

His expected gain
= $10(1/8) + 5(3/8) + 2(3/8) - 15(1/8)$
= 2 (dollars)

7.3.2 Function of a Random Variable

If X is a discrete random variable with probability distribution p(x) and if g(x) is a real-valued function of X, then the expected value of g(X) is defined as

$$E[g(X)] = \sum_x g(x)p(x)$$

Example 7.3-5

A salesman is employed by a computer manufacturer to sell PC's. The salary he gets in a day is calculated by the formula

$$g(x) = 90 + 60x$$

where x is the number of PC's he sells in that day.

Assume that in each day he may sell zero to four PC's with probabilities listed in the table below:

Number of PC's, x	0	1	2	3	4
Probability, p(x)	0.05	0.2	0.4	0.2	0.15

Let X be the number of PC's sold in a day.

His expected daily salary is

$$\begin{aligned} E[g(X)] &= g(0)\,p(0) + g(1)\,p(1) + g(2)\,p(2) + g(3)\,p(3) + g(4)\,p(4) \\ &= (90)(0.05) + (90+60)(0.2) + (90+120)(0.4) + (90+180)(0.2) \\ &\quad + (90+240)(0.15) \\ &= 222 \text{ (dollars)} \end{aligned}$$

It can be proved that, for any real numbers a and b,

$$E(aX + b) = aE(X) + b .$$

Example 7.3-6

In Example 7.3-6, expected daily salary

$= E(90+60X)$

$= 90 + 60E(X)$

$= 90 + 60[(0)(0.05) + (1)(0.2) + + (2)(0.4) + (3)(0.2) + (4)(0.15)]$

$= 222$ (dollars)

7.3.3 Variance of a Random Variable

The *variance* of a random variable X is defined as

$$V(X) = \sigma^2 = E[(X - \mu)^2] ,$$

where $\mu = E(X)$.

The *standard deviation*, σ, of a random variable X is defined

$$\sigma = \sqrt{E[(X-\mu)^2]} .$$

Example 7.3-7

The probability distribution of the random variable X is shown below:

x	1	2	3
P(X = x)	0.1	0.5	0.4

Find E(X) and V(X).

$$E(X) = \sum_x x\,p(x)$$
$$= 1(0.1) + 2(0.5) + 3(0.4)$$
$$= 2.3$$

$$V(X) = E[(X - \mu)^2]$$
$$= \sum_x (x - \mu)^2 p(x)$$
$$= (1 - 2.3)^2 (0.1) + (2 - 2.3)^2 (0.5) + (3 - 2.3)^2 (0.4)$$
$$= 0.41$$

Example 7.3-8

Daily sales records for a shop selling electric appliances show that it will sell zero, one, two or three air-conditioners with the probabilities:

Number of Sales	0	1	2	3
Probability	0.5	0.3	0.15	0.05

Calculate the expected value, variance and standard variation for daily sales.

Expected value
$= (0)(0.5) + (1)(0.3) + (2)(0.15) + (3)(0.05)$
$= 0.75$

Variance
$= (0 - 0.75)^2(0.5) + (1 - 0.75)^2(0.3) + (2 - 0.75)^2(0.15) + (3 - 0.75)^2(0.05)$
$= 0.7875$

Standard deviation
$= \sqrt{0.7875}$
$= 0.8874$

7.4 Binomial Distribution

7.4.1 Introducing Binomial Distribution

Consider the following random variables:

(a) X is the number of "6" obtained in 10 rolls of a fair die.
(b) X is the number of tails obtained in 100 tosses of a fair coin.
(c) X is the number of defective light bulb in a batch of 1000.
(d) X is the number of boys in a family of 5 children.

In each case, a basic experiment is repeated a number of times. For example, the basic experiment in case (a) is rolling the die once.

The following are common characteristics of the random variables in cases (a) to (d):

(1) The number of trials n of the basic experiment is fixed in advance.
(2) Each trial has two possible outcomes which may be called "success" and "failure".
(3) The trials are independent.
(4) The probability of success is fixed.

A random variable X defined to be the number of successes among n trials called a *binomial random variable* if the properties (1) to (4) are satisfied.

Mathematically, we write

$X \sim \text{Bin}(n, p)$, where n = number of trials, and
p = probability of success.

The following tables shows a summary for the cases (a) to (d):

Case	"success"	P(success), i.e. p	n	Distribution
(a)	Getting a "6"	1/6	10	X ~ Bin(10, 1/6)
(b)	Getting a tail	1/2	100	X ~ Bin(100,1/2)
(c)	Defective bulb	p (not given)	1000	X ~ Bin(1000, p)
(d)	Child is a boy	0.5	5	X ~ Bin(5, 0.5)

7.4.2 The Probability Distribution of a Binomial Random Variable

If $X \sim \text{Bin}(n, p)$, then $p(x) = P(X = x) = {}_nC_x\, p^x (1 - p)^{n-x}$ where $x = 0, 1, 2, \ldots, n$.

A sketch of the proof for this result is given below.
Consider x successes among n trials.
The probability that all the x successes occur in specified trials (for example, in the first x trials) is $p^x (1 - p)^{n-x}$.
The number of outcomes in which there are exactly x successes is ${}_nC_x$.
Therefore $P(X = x) = {}_nC_x\, p^x (1 - p)^{n-x}$.

Definition: If p is the probability of success of an event in one trial, then the probability of x successes in n trials is ${}_nC_x\, p^x (1 - p)^{n-x}$.

Example 7.4-1

A fair coin is tossed 8 times. Find the probability of obtaining 5 heads.

Let X be the number of heads obtained in 8 tosses. Then $X \sim \text{Bin}(8, \frac{1}{2})$.

$$\begin{aligned} P(5 \text{ heads}) &= {}_8C_5\left(\frac{1}{2}\right)^5\left(1-\frac{1}{2}\right)^{8-5} \\ &= \frac{56}{256} \\ &= \frac{7}{32} \end{aligned}$$

Example 7.4-2

There are 10 multiple-choice questions in a test and each question has 5 options. Suppose a student answers all 10 questions by randomly picking an option in each question. Find the probability that

(a) he will answer six questions correctly,

(b) he will get at least 3 correct answers.

Let X be the number of correct answers he will get. Then $X \sim \text{Bin}(10, 0.2)$.

(a) $P(X = 6) = {}_{10}C_6\,(0.2)^6\,(1 - 0.2)^{10-6}$
$= 0.00551$

(b) P(at least 3 correct answers)
$= 1 - P(X = 0) - P(X = 1) - P(X = 2)$
$= 1 - {}_{10}C_0\,(0.8)^{10} - {}_{10}C_1\,(0.2)\,(0.8)^9 - {}_{10}C_2\,(0.2)^2\,(0.8)^8$
$= 0.322$

Example 7.4-3

Binary digits 0 and 1 are transmitted along a data channel in which the presence of noise results in the fact that each digit may be wrongly received with a probabilty of 0.00002. Each message is transmitted in blocks of 2000 digits.

(a) What is the probabilty that at least one digit in a block is wrongly received?

(b) If a certain message has a length of 20 blocks, find the probability that 2 or more blocks are wrongly received.

(a) Let X be the number of digits wrongly received in a block of 2000 digits. Then $X \sim \text{Bin}(2000, 0.00002)$.

$$\begin{aligned} P(X \geq 1) &= 1 - P(X = 0) \\ &= 1 - (1 - 0.00002)^{2000} \\ &= 0.0392 \end{aligned}$$

(b) Let Y be the number of block that are wrongly received among the 20 blocks. Then $Y \sim \text{Bin}(20, 0.0392)$.

$$\begin{aligned} P(Y \geq 2) &= 1 - P(Y = 0) - P(Y = 1) \\ &= 1 - (1 - 0.0392)^{20} - {}_{20}C_1\,(0.0392)\,(1 - 0.0392)^{19} \\ &= 0.184 \end{aligned}$$

7.5 Poisson Distribution

7.5.1 Introducing Poisson Processes and Random Variables

A Poisson process is a process that generates sequences of isolated events occurring in time or space with the following characteristics:

(a) the events that occur in non-overlapping intervals of time or space are independent,
(b) the probability that two or more events occur in a small enough interval is negligible,
(c) events are occuring at a constant average rate per unit time (or length or area or volume).

If a random variable X is defined to be the number of events from a Poisson process that occur in a fixed length of time, then X is said to be a Poisson random variable, or

$$X \sim P_o(\theta)$$

where θ is the average number of events that occur in a time interval of this length.

The probability distribution for a Poisson random variable X is

$$p(x) = P(X = x) = \frac{e^{-\theta}\theta^{x}}{x!}$$

where x = 0, 1, 2, ….

7.5.2 Examples of Poisson Random Variables

In each of the following examples, X may be taken as a Poisson random variable:

(a) X is the number of cars passing a junction in one minute,
(b) X is the number of air bubbles in a 100cm x 100cm area of a glass plate,
(c) X is the number of customers arriving at a shop in a period of 5 minutes.

Example 7.5-1

A manufacturer of a certain type of LCD-screen finds that the number of "dead spots" on a screen is a $P_o(0.8)$ random variable. A screen with more than 3 "dead spots" has to be sold at a discount. Find the proportion of screens that are sold at a discount.

Let X be the number of "dead spots" on a screen. Then $X \sim P_o(0.8)$.

The proportion of screens that are sold at a discount
$= P(X > 3)$
$= 1 - p(0) - p(1) - p(2) - p(3)$
$$= 1 - e^{-0.8} - e^{-0.8}(0.8) - \frac{e^{-0.8}(0.8)^2}{2!} - \frac{e^{-0.8}(0.8)^3}{3!}$$
$= 0.00908$

Example 7.5-2

A radioactive source is emitting α-particles at an average of 2.6 particles per minute. Find the probability that the number of particles emitted in one minute is

(a) 0
(b) 1
(c) more than 1

Let X be the number of particles emitted in one minute.

Then $X \sim P_o(2.6)$ and $p(x) = \dfrac{e^{-2.6}(2.6)^x}{x!}$.

(a) $P(X = 0) = e^{-2.6} = 0.0743$

(b) $P(X = 1) = e^{-2.6}(2.6) = 0.193$

(c) $P(X > 1) = 1 - 0.0743 - 0.193 = 0.7327$

7.5.3 Poisson Approximation to Binomial Distribution

When n is large and p is small, the binomial distribution Bin(n, p) can be approximated by the Poisson distribution $P_o(np)$, i.e.

$$p(x) = \frac{e^{-np}(np)^x}{x!}$$

A rule of thumb for the conditions under which this approximation may be used is

$n \geq 100$, $np < 10$.

Example 7.5-3

A company ships memory chips in boxes of 200. It is known that 0.001 of all the chips are defective. Find the probability that less than 3 chips in a box are defective.

Let X be the number of defective chips in a box. Then $X \sim \text{Bin}(200, 0.001)$.

The average number of defective chips in a box
= 200 x 0.001
= 0.2

The distribution of X may be approximated by $P_o(0.2)$.

$$\therefore P(X < 3) = p(0) + p(1) + p(2)$$
$$= e^{-0.2} + e^{-0.2}(0.2) + \frac{e^{-0.2}(0.2)^2}{2!}$$
$$= 0.999$$

7.6 Normal Distribution

7.6.1 Normal Distribution

Many continuous random variables have been found to follow Normal distributions.

A random variable X is said to have a Normal distribution, with parameters μ and σ^2 if it can take any real value and has p.d.f.

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right], \quad -\infty < x < \infty \ .$$

In this case, we write $X \sim N(\mu, \sigma^2)$. It can be shown that μ is the expected value of X and σ^2 is the variance.

$$P(a < X < b) = \int_a^b f(x)dx$$

$$= \int_a^b \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]dx$$

This integral cannot be done algebraically and its value has to be found by numerical methods of integration. The value of $P(a < X < b)$ can also be viewed as the area under the curve of f(x) from x = a to x = b.

7.6.2 Characteristics of the Normal Probability Distribution

1. The curve has the bell shape.
2. It is symmetric about the center.
3. The mean, mode median are the same value and lies at the center.
4. The two tails of the normal distribution extend indefinitely and never touch the horizontal axis. (Graphically, of course, this is impossible to show.)

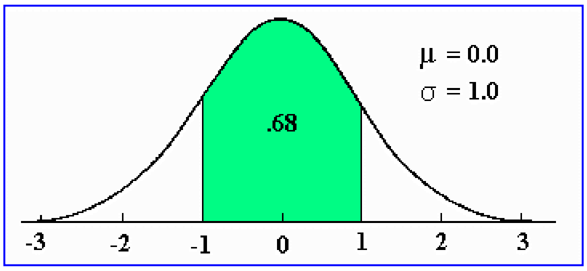


Diagram of the Standard Normal Curve

7.6.3 Standard Normal Distribution

A standard normal distribution is the normal distribution with $\mu = 0$ and $\sigma^2 = 1$, i.e. the N(0,1) distribution. A standard normal random variable is often denoted by Z.

If $Z \sim N(0,1)$, its c.d.f. is usually written as

$$\Phi(z) = P(Z \le z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} \exp(-\frac{x^2}{2})dx$$

Values of $\Phi(z)$ can be looked up in the standard normal table.

An interactive Normal table can be found in the following web site:
http://davidmlane.com/hyperstat/normal_distribution.html or
http://www.stat.berkeley.edu/~stark/Java/NormHiLite.htm

Note:

(1) $\Phi(z)$ may be interpreted as the area to the left of z under the standard normal curve.

(2) $P(Z < -z) = P(Z > z)$, since the standard normal curve is symmetrical about the line $Z = 0$.

(3) The area of the normal distribution lies between Z = -1 and +1 is 68.3% ; Z = -2 and +2 is 95.4%; Z = -3 and +3 is 99.7%.

Example 7.6-1

Given Z is the normal distribution with $\mu = 0$ and $\sigma^2 = 1$, i.e. $Z \sim N(0,1)$, find the following probabilities using the standard normal table.

(a) $P(Z \leq 1.25)$

(b) $P(Z > 2.33)$

(c) $P(0.5 < Z < 1.5)$

(d) $P(Z < -1.25)$

(e) $P(-1.5 < Z < -0.5)$

(a) $P(Z \leq 1.25)$

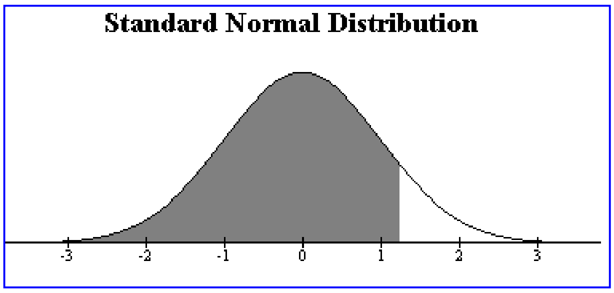
$= 0.8943$

(b) $P(Z > 2.33)$

$= 1 - \Phi(2.33)$

$= 1 – 0.9901$

$= 0.0099$

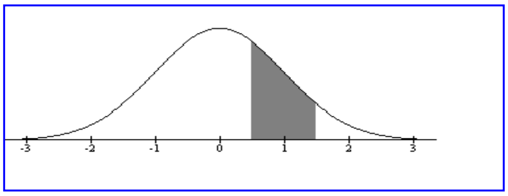


(c) $P(0.5 < Z < 1.5)$

$= \Phi(1.5) - \Phi(0.5)$

$= 0.9332 – 0.6915$

$= 0.2417$



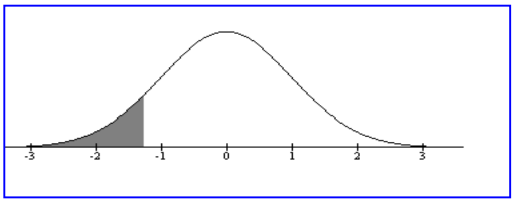
(d) $P(Z < -1.25)$

$= P(z > 1.25)$

$= 1 - \Phi(1.25)$

$= 1 – 0.8944$

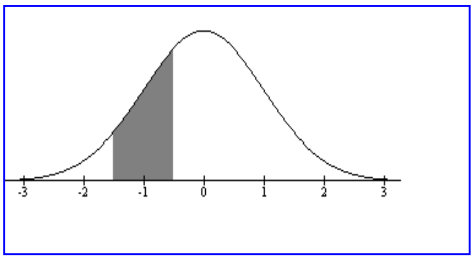
$= 0.1056$



(e) $P(-1.5 < Z < -0.5)$

$= P(0.5 < Z < 1.5)$

$= 0.2417$



Exercise

Calculate these areas under the normal curve:

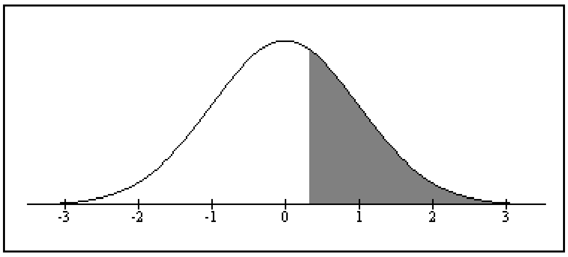
a. $Z < 1.45$
b. between $Z = 0$ and $Z = 1.2$
c. between $Z = 0$ and $Z = -0.9$
d. $Z > 1.6$
e. between $Z = 0.3$ and $Z = 1.56$
f. between $Z = 0.2$ and $Z = -0.2$
g. between $Z = -0.5$ and $Z = -1.2$
h. between $Z = -1.04$ and $Z = 1.95$

Example 7.6-2

Given Z is the normal distribution with $\mu = 0$ and $\sigma^2 = 1$, i.e. $Z \sim N(0,1)$, find the value of c if

(a) $P(Z \le c) = 0.8888$
(b) $P(Z > c) = 0.37$
(c) $P(Z < c) = 0.025$
(d) $P(0 \le Z \le c) = 0.4924$

(a) $\Phi(c) = 0.8888$
$c = 1.22$



(b) $1 - \Phi(c) = 0.37$
$\Phi(c) = 0.63$
$c = 0.34$

(c) Obviously c is negative and the standard normal table cannot be used directly. Recall that $P(Z < -z) = P(Z > z)$.

$$\begin{aligned} \therefore\ 0.025 &= P(Z < c) \\ &= P(Z > -c) \\ &= 1 - \Phi(-c) \\ \text{i.e. } \Phi(-c) &= 0.975 \\ -c &= 1.96 \\ c &= -1.96 \end{aligned}$$

(d)
$$\begin{aligned} P(0 \le Z \le c) &= \Phi(c) - 0.5 \\ &= 0.4924 \\ \Phi(c) &= 0.9924 \\ c &= 2.43 \end{aligned}$$

7.6.4 Standardization of Normal Random Variables

If $X \sim N(\mu, \sigma^2)$, then it can be shown that

$$Z = \frac{X-\mu}{\sigma} \sim N(0,1) .$$

In other words $\frac{X-\mu}{\sigma}$ is the standard normal random variable.

The process of transforming X to Z is called *standardization* of the random variable X.

Example 7.6-3

Suppose $X \sim N(10, 6.25)$. Find the following probabilities

(a) $P(X < 13)$
(b) $P(X > 5)$
(c) $P(8 < X < 15)$

Let $Z = \frac{X-10}{\sqrt{6.25}} = \frac{X-10}{2.5}$.

(a) $$\begin{aligned} P(X < 13) &= P(Z < \frac{13-10}{2.5}) \\ &= P(Z < 1.2) \\ &= 0.8849 \end{aligned}$$

(b) $$\begin{aligned} P(X > 5) &= P(Z > \frac{5-10}{2.5}) \\ &= P(Z > -2) \\ &= P(Z < 2) \\ &= 0.9773 \end{aligned}$$

(c) $$\begin{aligned} P(8 < X < 15) &= P(\frac{8-10}{2.5} < Z < \frac{15-10}{2.5}) \\ &= P(-0.8 < Z < 2) \\ &= \Phi(2) - \Phi(-0.8) \\ &= \Phi(2) - (1 - \Phi(0.8)) \\ &= 0.9772 - (1 - 0.7881) \\ &= 0.7653 \end{aligned}$$

Example 7.6-4

The lifetime of light bulbs produced in a certain factory is normally distributed with mean 3000 hours and standard deviation 50 hours. Find the probability that a randomly chosen light bulb has a lifetime less than 2900 hours.

Let X be the lifetime of a randomly chosen light bulb.
Then $X \sim N(3000, 50^2)$.

Let $Z = \frac{X - 3000}{50}$.

$$\begin{aligned} P(X < 2900) &= P(Z < -2) \\ &= 1 - 0.9772 \\ &= 0.0228 \end{aligned}$$

Example 7.6-5

If a random variable has the normal distribution with $\mu = 40$ and $\sigma = 2.4$, find the probabilities that it will take a value

(a) less than 43.6
(b) greater than 38.2
(c) between 40.6 and 43.0
(d) between 35.8 and 44.2.

Let X be the random variable.

(a)
$$\begin{aligned} P(X < 43.6) &= P(Z < \frac{43.6 - 40}{2.4}) \\ &= P(Z < 1.5) \\ &= 0.9332 \end{aligned}$$

(b)
$$\begin{aligned} P(X > 38.2) &= P(Z > \frac{38.2 - 40}{2.4}) \\ &= P(Z > -0.75) \\ &= P(Z < 0.75) \\ &= 0.7734 \end{aligned}$$

(c)
$$\begin{aligned} P(40.6 < X < 43.0) &= P(\frac{40.6 - 40}{2.4} < Z < \frac{43.0 - 40}{2.4}) \\ &= P(0.25 < Z < 1.25) \\ &= 0.2957 \end{aligned}$$

(d)
$$\begin{aligned} P(35.8 < X < 44.2) &= P(\frac{35.8 - 40}{2.4} < Z < \frac{44.2 - 40}{2.4}) \\ &= P(-1.75 < Z < 1.75) \\ &= 0.9198 \end{aligned}$$

7.7 Refernece

Mendendall, W., Breaver, R.J., and Beaver B.M., *Introduction to Probability and Statistics*, 10th ed., Duxbury Press, 1998.

Freund, J.E. and Simon, G.A., *Modern Elementary Statistics*, 9th ed., Prentic Hall, 1997.

Statistics Textbooks on the web:
http://davidmlane.com/hyperstat/index.html
http://www.anu.edu.au/nceph/surfstat/surfstat-home/surfstat.html

Area Under the Standard Normal Curve

Note: An entry in the table is the area under the curve to the left of z, i.e. Φ(z).

Z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
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	0.7703	0.7734	0.7764	0.7793	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	0.8906	0.8925	0.8943	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
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998