STA 2200 PROBABILITY AND STATISTICS II

Purpose At the end of the course the student should be able to handle problems involving probability distributions of a discrete or a continuous random variable.

Objectives

By the end of this course the student should be able to;

(1) Define the probability mass, density and distribution functions, and to use these to determine expectation, variance, percentiles and mode for a given distribution

(2) Appreciate the form of the probability mass functions for the binomial, geometric, hypergeometic and Poisson distributions, and the probability density functions for the uniform, exponential gamma , beta and normal, functions, and their applications

(3) Apply the moment generating function and transformation of variable techniques

(4) Apply the principles of statistical inference for one sample problems.

DESCRIPTION

Random variables: discrete and continuous, probability mass, density and distribution functions, expectation, variance, percentiles and mode. Moments and moment generating function. Moment generating function and transformation Change of variable technique for univariate distribution. Probability distributions: hypergeometric, binomial, Poisson, uniform, normal, beta and gamma. Statistical inference including one sample normal and t tests.

Pre-Requisites: STA 2100 Probability and Statistics I, SMA 2104 Mathematics for Science

Course Text Books

1) RV Hogg, JW McKean & AT Craig *Introduction to Mathematical Statistics*, 6th ed., Prentice Hall, 2003 ISBN 0-13-177698-3
2) J Crawshaw & J Chambers *A Concise Course in A-Level statistics, with worked examples*, 3rd ed. Stanley Thornes, 1994 ISBN 0-534- 42362-0

Course Journals:

1) Journal of Applied Statistics (J. Appl. Stat.) [0266-4763; 1360-0532]
2) Statistics (Statistics) [0233-1888]

Further Reference Text Books And Journals:

a) HJ Larson *Introduction to Probability Theory and Statistical Inference(Probability and Mathematical Statistics)* 3rd ed., Wiley, 1982
b) Uppal, S. M. , Odhiambo, R. O. & Humphreys, H. M. *Introduction to Probability and Statistics*. JKUAT Press, 2005
c) I Miller & M Miller *John E Freund's Mathematical Statistics with Applications,* 7th ed., Pearsons Education, Prentice Hall, New Jersey, 2003 ISBN: 0131246461
d) Statistical Science (Stat. Sci.) [0883-4237]
e) Journal of Mathematical Sciences
f) The Annals of Applied Probability

1. RANDOM VARIABLES

1.1 Introduction

In application of probability, we are often interested in a number associated with the outcome of a random experiment. Such a quantity whose value is determined by the outcome of a random experiment is called a **random variable**. It can also be defined as any quantity or attribute whose value varies from one unit of the population to another.
A **discrete** random variable is function whose range is finite and/or countable, Ie it can only assume values in a finite or countably infinite set of values. A **continuous** random variable is one that can take any value in an interval of real numbers. (There are *uncountably* many real numbers in an interval of positive length.)

1.2 Discrete Random Variables and Probability Mass Function

Consider the experiment of flipping a fair coin three times. The number of tails that appear is noted as a discrete random variable. *X*= number of tails that appear in 3 flips of a fair coin.
There are 8 possible outcomes of the experiment: namely the sample space consists of

$$S = \{\text{HHH}, \text{HHT}, \text{HTH}, \text{HTT}, \text{THH}, \text{THT}, \text{TTH}, \text{TTT}\}$$
$$\text{X} = \{\ 0 \quad 1, \quad 1, \quad 2, \quad 1, \quad 2, \quad 2, \quad 3\}$$

are the corresponding values taken by the random variable *X*.
Now, what are the possible values that *X* takes on and what are the probabilities of *X* taking a particular value?
From the above we see that the possible values of *X* are the 4 values

$$\text{X} = \{0, 1, 2, 3\}$$

Ie the sample space is a disjoint union of the 4 events $\{X = j\}$ for *j*=0,1,2,3
Specifically in our example:

$\{\text{X}=0\} = \{\text{HHH}\}$ $\qquad$ $\{\text{X}=1\} = \{\text{HHT, HTH, THH}\}$

$\{\text{X}=2\} = \{\text{TTH, HTT, THT}\}$ $\qquad$ $\{\text{X}=3\} = \{\text{TTT}\}$

Since for a fair coin we assume that each element of the sample space is equally likely (with probability $\frac{1}{8}$, we find that the probabilities for the various values of *X,* called the *probability distribution* of *X* or the *probability mass function (pmf)*. can be summarized in the following table listing the possible values beside the probability of that value

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

Note: The probability that *X* takes on the value *x,* ie $\text{p}(\text{X} = \text{x})$, is defined as the sum of the probabilities of all points in *S* that are assigned the value *x*.
We can say that this pmf places mass $\frac{3}{8}$ on the value $\text{X} = 2$.
The "masses" (or probabilities) for a pmf should be between 0 and 1.
The total mass (i.e. total probability) must add up to 1.

Definition: The **probability mass function** of a discrete variable is a graph, table, or formula that specifies the proportion (or probabilities) associated with each possible value the random variable can take. The mass function $\text{P}(\text{X} = \text{x})$ (or just p(x) has the following properties:

$$0 \le \text{p}(\text{x}) \le 1 \text{ and } \sum_{\text{all x}} \text{p}(\text{x}) = 1$$

More generally, let X have the following properties

i) It is a discrete variable that can only assume values $x_1, x_2, \ldots. x_n$

ii) The probabilities associated with these values are
$P(X = x_1) = p_1$, $P(X = x_2) = p_2 \ldots\ldots P(X = x_n) = p_n$

Then X is a discrete random variable if $0 \le p_i \le 1$ and $\sum_{i=1}^{n} p_i = 1$

Remark: We denote random variables with capital letters while realized or particular values are denoted by lower case letters.

Example 1
Two tetrahedral dice are rolled together once and the sum of the scores facing down was noted. Find the pmf of the random variable 'the sum of the scores facing down.'
Solution

+	1	2	3	4
1	2	3	4	5
2	3	4	5	6
3	4	5	6	7
4	5	6	7	8

$X = \{1, 2, 3, 4, 5, 6, 7, 8\}$

Therefore t is given the pmf by the table below

x	2	3	4	5	6	7	8
P(X=x)	$\frac{1}{16}$	$\frac{1}{8}$	$\frac{3}{16}$	$\frac{1}{4}$	$\frac{3}{16}$	$\frac{1}{8}$	$\frac{1}{16}$

This can also be written as a function

$$P(X = x) = \begin{cases} \frac{x-1}{16} \text{ for } x = 2, 3, 4, 5 \\ \frac{9-x}{16} \text{ for } x = 6, 7, 8 \end{cases}$$

Example 2
The pmf of a discrete random variable W is given by the table below

w	-3	-2	-1	0	1
P(W=w)	0.1	0.25	0.3	0.15	d

Find the value of the constant d, $P(-3 \le \text{w} < 0)$, $P(\text{w} > \text{-}1)$ and $P(-1 < \text{w} < 1)$
Solution

$$\sum_{\text{all w}} \text{p(W} = \text{w}) = 1 \quad \Rightarrow \quad 0.1 + 0.25 + 0.3 + 0.15 + d = 1 \quad \Rightarrow \quad d = 0.2$$

$$P(-3 \le \text{w} < 0) = P(W = -3) + P(W = -2) + P(W = -1) = 0.65$$

$$P(\text{w} > \text{-}1) = P(\text{w} = 0) + P(\text{w} = 1) = 0.15 + 0.2 = 0.35$$

$$P(-1 < \text{w} < 1) = P(W = 0) = 0.15$$

Example 3
A discrete random variable Y has a pmf given by the table below

y	0	1	2	3	4
P(Y=y)	c	2c	5c	10c	17c

Find the value of the constant c hence computes $P(1 \le \text{Y} < 3)$
Solution

$$\sum_{\text{ally}} \text{p(Y} = \text{y}) = 1 \quad \Rightarrow \quad c(1 + 2 + 5 + 10 + 17) = 1 \quad \Rightarrow \quad c = \frac{1}{35}$$

$$P(1 \le \text{Y} < 3) = P(Y = 1) + P(Y = 2) = \frac{2}{35} + \frac{5}{35} = \frac{1}{5}$$

Exercise
1. A die is loaded such that the probability of a face showing up is proportional to the face number. Determine the probability of each sample point.
2. Roll a fair die and let X be the square of the score that show up. Wtrie down the probability distribution of X hence compute $P(X < 15)$ and $P(3 \le \text{X} < 30)$

3. Let X be the random variable the number of fours observed when two dice are rolled together once. Show that X is a discrete random variable.
4. The pmf of a discrete random variable X is given by $P(X = x) = kx$ for $x = 1,2,3,4,5,6$ Find the value of the constant k, $P(X < 4)$ and $P(3 \le X < 6)$
5. A fair coin is flip until a head appears. Let N represent the number of tosses required to realize a head. Find the pmf of N
6. A discrete random variable Y has a pmf given by $P(Y = y) = c\left(\frac{3}{4}\right)^x$ for $y = 0,1,2,.....$ Find the value of the constant c and $P(X < 3)$
7. Verify that $f(x) = \dfrac{2x}{k(k+1)}$ for $x = 0,1,2,.....k$ can serve as a pmf of a random variable X.
8. For each of the following determine c so that the function can serve as a pmf of a random variable X.
   a) $f(x) = cx$ for $x = 1,2,3,4,5$
   b) $f(x) = cx^2$ for $x = 0,1,2,.....k$
   c) $f(x) = c\left(\frac{1}{6}\right)^x$ for $x = 0,1,2,3.....$
   d) $f(x) = c2^{-x}$ for $x =$ for $x = 0,1,2,....$
9. A coin is loaded so that heads is three times as likely as the tails. For 3 independent tosses of the coin find the pmf of the total number of heads realized and the probability of realizing at most 2 heads.

1.3 Continuous Random Variables and Probability Density Function

A **continuous** random variable can assume any value in an interval on the real line or in a collection of intervals. The sample space is uncountable. For instance, suppose an experiment involves observing the arrival of cars at a certain period of time along a highway on a particular day. Let T denote the time that lapses before the 1st arrival, the T is a continuous random variable that assumes values in the interval $[0,\infty)$

Definition: A random variable *X* is *continuous* if there exists a nonnegative function f so that, for every interval *B*, $P(X \in B) = \int_B f(x)\,dx$. The function $f = f(x)$ is called the *probability density function* of *X*.

Definition: Let X be a continuous random variable that assumes values in the interval $(-\infty,\infty)$, The f(x) is said to be a probability density function (pdf) of X if it satisfies the following conditions

$$f(x) \ge 0 \text{ for all x },\ p(a \le x \le b) = \int_a^b f(x)\,dx = 1 \text{ and } \int_{-\infty}^{\infty} f(x)\,dx = 1$$

The support of a continuous random variable is the smallest interval containing all values of x where f(x) >= 0.

Remark A crucial property is that, for any real number x, we have $P(X = x) = 0$ (implying there is no difference between $P(X \le x)$ and $P(X < x)$); that is it is not possible to talk about the probability of the random variable assuming a particular value. Instead, we talk about the probability of the random variable assuming a value within a given interval. The probability of the random variable assuming a value within some given interval from $x = a$ to $x = b$ is defined to be the area under the graph of the probability density function between $x = a$ and $x = b$.

Example 1

Let X be a continuous random variable. Show that the function

$$\text{f(x)} = \begin{cases} \frac{1}{2}x, & 0 \le x \le 2 \\ 0, & elsewhere \end{cases}$$ is a pdf of X hence compute $P(0 \le X < 1)$ and $P(-1 < X < 1)$

Solution

$\text{f(x)} \ge 0$ for all x in the interval $0 \le x \le 2$ and $\int_0^2 \frac{1}{2}\text{xdx} = \frac{1}{4}\left[x^2\right]_0^2 = 1$. Therefore f(x)is indeed a pdf of X.

Now $P(0 \le X < 1) = \int_0^1 \frac{1}{2}\text{xdx} = \frac{1}{4}\left[x^2\right]_0^1 = \frac{1}{4}$ and

$P(-1 < X < 1) = P(-1 < X < 0) + P(0 < X < 1) = 0 + \frac{1}{4} = \frac{1}{4}$

Example 2

The time X, in hours, between computer failures is a continuous random variable with density

$$\text{f(x)} = \begin{cases} \lambda e^{-0.01x} & \text{for } x > 0 \\ 0, & elsewhere \end{cases}$$ Find λ hence compute $P(50 \le X < 150)$ and $P(X < 100)$

Solution

$\text{f(x)} \ge 0$ for all x in $0 \le x < \infty$ Thus $1 = \lambda\int_0^\infty e^{-0.01x}dx = -100\lambda\left[e^{-0.01x}\right]_0^\infty = 100\lambda \Rightarrow \lambda = 0.01$..

Now $P(50 \le X < 150) = 0.01\int_{50}^{150} e^{-0.01x}dx = \left[-e^{-0.01x}\right]_{50}^{150} = e^{-0.5} - e^{-1.5} \approx 0.3834005$ and

$P(X < 100) = 0.01\int_0^{100} e^{-0.01x}dx = \left[-e^{-0.01x}\right]_0^{100} = 1 - e^{-1} \approx 0.6321206$

Example 3

A continuous random variable X has a probability density function given by

$$\text{f(x)} = \begin{cases} 0.25, & 0 < x < 2 \\ 0.5x + c, & 2 < x < 3 \\ 0, & elsewhere \end{cases}$$ Find c hence compute $P(1 \le X < 2.5)$.

Solution

$\int_{\text{all x}} \text{f(x)dx} = 1 = \int_0^2 \frac{1}{4}dx + \int_2^3 \left(\frac{1}{2}x + c\right)dx = \frac{1}{2} + \frac{5}{4} + c \Rightarrow c = -\frac{3}{4}$

$P(1 \le X < 2.5) = P(1 \le X < 2) + P(2 \le X < 2.5) = \int_1^2 \frac{1}{4}dx + \int_2^{2.5}\left(\frac{1}{2}x - \frac{3}{4}\right)dx = \frac{1}{4} + \left[\frac{x^2 - 3x}{4}\right]_2^{2.5} = \frac{1}{4} + \frac{3}{16} = \frac{7}{16}$

Exercise

1) Suppose that the random variable X has p.d.f. given by $\text{f(x)} = \begin{cases} cx, & 0 \le x \le 1 \\ 0, & elsewhere \end{cases}$ Find the value of the constant c hence determine m so that $P(X \le m) = \frac{1}{2}$
2) Let X be a continuous random variable with pdf $\text{f(x)} = \begin{cases} \frac{x}{5} + k, & 0 \le x \le 3 \\ 0, & elsewhere \end{cases}$ Find the value of the constant k hence compute $P(1 < X < 3)$
3) A continuous random variable Y has the pdf given by $\text{f(y)} = \begin{cases} k(1+y), & 4 \le x \le 7 \\ 0, & elsewhere \end{cases}$ Find the value of the constant k hence compute $P(Y < 5)$ and $P(5 < Y < 6)$

4) A continuous random variable X has the pdf given by $f(x) = \begin{cases} k(1+x)^2, & -2 \le x \le 0 \\ 4k & 0 < x \le \frac{4}{3} \\ 0, & elsewhere \end{cases}$ Find the value of the constant k hence compute $P(X > 1)$ and $P(-1 < X < 1)$

5) A continuous random variable X has the pdf given by $f(x) = \begin{cases} kx, & 0 \le x \le 2 \\ k(4-x) & 2 < x \le 4 \\ 0, & elsewhere \end{cases}$ Find the value of the constant k hence compute $P(X > 3)$ and $P(1 < X < 3)$

1.4 Distribution Function of a Random Variables

Definition: For any random variable X, we define the **cumulative distribution function (CDF)**, *F(x)* as $F(x) = P(X \le x) = \begin{cases} \sum_{t=-\infty}^{x} f(t) \text{ If X is discrete} \\ \int_{-\infty}^{x} f(t)dt \text{ If X is continuous} \end{cases}$ for every x.

\\ Again here t is introduced to facilitate summation /integration//

Properties of any cumulative distribution function

- $\lim_{x\to\infty} F(x) = 1$ and $\lim_{x\to-\infty} F(x) = 0$
- F(x) is a non-decreasing function.
- F (x) is a right continuous function of *x*. In other words $\lim_{t\to x} F(t) = F(x)$

Reminder If the c.d.f. of X is F(x) and the p.d.f. is f(x), then differentiate F(x) to get f(x), and integrate f(x) to get F(x);

Theorem: For any random variable X and real values a < b, $P(a \le X \le b) = F(b) - F(a)$

Example 1

Let X be a discrete random variable with pmf given by $f(x) = \begin{cases} \frac{1}{20}(1+x) & \text{for } x = 1,2,3,4,5 \\ 0, & elsewhere \end{cases}$.

Determine the cdf of X hence compute $P(X > 3)$

Solution

$$F(x) = \sum_{t=-\infty}^{x} f(t) = \frac{1}{20}\sum_{t=1}^{x}(x+1) = \frac{1}{20}(2+3+....+x) = \frac{1}{20}\left\{\frac{x}{2}\left[4+(x-1)\right]\right\} = \frac{x(x+3)}{40}$$

$$F(x) = \begin{cases} 0 & \text{for } x < 1 \\ \frac{x(x+3)}{40} & \text{for } x = 1,2,3,4,5 \\ 1 & \text{for } x > 5 \end{cases} \quad \text{Recall for an AP } S_n = \frac{n}{2}\left[2a+(n-1)d\right]$$

$$P(X > 3) = 1 - P(X \le 3) = 1 - \frac{3\times 6}{40} = \frac{11}{20}$$

Example 2

Suppose X is a continuous random variable whose pdf f(x) is given by

$f(x) = \begin{cases} \frac{1}{2}x, & 0 \le x \le 2 \\ 0, & elsewhere \end{cases}$. Obtain the cdf of X hence compute $P(X > \frac{2}{3})$

Solution

$$F(x) = \int_{-\infty}^{x} f(t)dt = \int_{0}^{x} \tfrac{1}{2}t\,dt = \tfrac{1}{4}\left[t^2\right]_0^x = \tfrac{1}{4}x^2 \quad \text{thus } F(x) = \begin{cases} 0, & x < 0 \\ \frac{x^2}{4}, & 0 \le x \le 2 \\ 1, & x > 2 \end{cases}$$

$$P\left(X > \tfrac{2}{3}\right) = 1 - P\left(X \le \tfrac{2}{3}\right) = 1 - \tfrac{1}{4}\left(\tfrac{2}{3}\right)^2 = \tfrac{8}{9}$$

Example 3

A continuous random variable X has a probability density function given by

$$f(x) = \begin{cases} 0.25, & 0 < x < 2 \\ 0.5x + c, & 2 < x < 3 \\ 0, & elsewhere \end{cases}$$ Find the cdf of X hence compute $P(1.5 \le X > 2.5)$.

Solution

$$F(x) = \int_{-\infty}^{x} f(t)dt = \begin{cases} \int_0^x \frac{1}{4}dt = \frac{x}{4} \\ \int_2^x \left(\frac{1}{2}t - \frac{3}{4}\right)dt + k = \frac{x^2 - 3x}{4} + \frac{1}{2} + k \end{cases}$$

Under the two levels, F(2) must be the same. (reason for introducing k)

$$\Rightarrow F(2) = \tfrac{1}{2} = k \quad \therefore \quad F(x) = \begin{cases} 0 \text{ for } x < 0 \\ \frac{x}{4} \text{ for } 0 \le x \le 2 \\ \frac{x^2 - 3x}{4} + 1 \text{ for } 2 \le x \le 3 \\ 1, \text{ for } x > 2 \end{cases}$$

Exercise

1. 1. The pdf of a continuous random variable X is given by $f(x) = \begin{cases} C/\sqrt{x}, & 0 \le x \le 4 \\ 0, & elsewhere \end{cases}$ Find the value of the constant C, the cdf of X and $P(X \ge 1)$
2. 2. The pdf of a random variable X is given by $g(x) = \begin{cases} kx(1-x), & 0 \le x \le 1 \\ 0, & elsewhere \end{cases}$ Find the value of the constant k, the cdf of X and the value of m such that $G(x) = \tfrac{1}{2}$
3. 3. Find the cdf of a random variable Y whose pdf is given by;

   a) $f(x) = \begin{cases} 1/3, & 0 \le x \le 1 \\ 1/3, & 2 \le x \le 4 \\ 0, & elsewhere \end{cases}$ b) $f(x) = \begin{cases} x/2, & 0 \le x \le 1 \\ 1/2, & 1 \le x \le 2 \\ (3-x)/2, & 2 \le x \le 3 \\ 0, & elsewhere \end{cases}$
4. 4. If the cdf of a random variable Y is given by $F(x) = 1 - \dfrac{9}{y^2}$ for $Y \ge 3$ and $F(x) = 0$ for $Y < 3$, find $P(X \le 5)$, $P(X > 8)$ and the pdf of X

1.5 Derived Random Variables

Give the pdf of a random variable say X, we can obtain the distribution of a second random variable say Y provided that we know some functional relationship between X and Y say $Y = u(X)$. Eg $Y = 2X + 3$, $Y = X^3$, $Y = \sqrt{X} + c$ etc

For a 1-1 relationship between X and Y eg $\mathrm{Y} = 2X + 3,$ $\mathrm{f(x)}$ and $\mathrm{g(y)}$ yields exactltly the same probabilities only the random variable and the set of values it can assume changes.

Example 1

Give the pmf of a random variable X as $\mathrm{f(x)} = \begin{cases} x/6 & \text{for } x = 1, 2, 3 \\ 0, & elsewhere \end{cases}$ find the pmf of $\mathrm{Y} = X^2$

Solution

The only values of Y with non zero probabilities are $\mathrm{Y} = 1,$ $\mathrm{Y} = 4$ and $\mathrm{Y} = 9$. Now

$P(\mathrm{Y} = 1) = P(\mathrm{X}^2 = 1) = P(\mathrm{X} = 1) = \tfrac{1}{6}$ $\quad$ $P(\mathrm{Y} = 4) = P(\mathrm{X}^2 = 4) = P(\mathrm{X} = 2) = \tfrac{1}{3}$ and

$P(\mathrm{Y} = 9) = P(\mathrm{X}^2 = 9) = P(\mathrm{X} = 3) = \tfrac{1}{2}$

In some cases several values of X will give rise to the same value of Y. The procedure is just the same as above but it is necessary to add the several probabilities that are associated with each value x that provides a unique value y.

Example 2

Give the pmf of a r.v X as $\mathrm{f(x)} = \begin{cases} \dfrac{x+1}{15} & \text{for } x = 0, 1, 2, 3, 4 \\ 0, & elsewhere \end{cases}$ find the pmf of $\mathrm{Y} = (X - 2)^2$

Solution

x	0	1	2	3	4
y	4	1	0	1	4

$P(\mathrm{Y} = 0) = P(\mathrm{X} = 2) = \tfrac{1}{5}$ $\quad$ $P(\mathrm{Y} = 1) = P(\mathrm{X} = 1) + P(\mathrm{X} = 3) = \tfrac{2}{15} + \tfrac{4}{15} = \tfrac{2}{5}$

$P(\mathrm{Y} = 1) = P(\mathrm{X} = 0) = P(\mathrm{X} = 4) = \tfrac{1}{15} + \tfrac{1}{3} = \tfrac{2}{5}$ Therefore the pmf of Y can be written as

y	0	1	4
$P(\mathrm{Y} = \mathrm{y})$	$\tfrac{2}{5}$	$\tfrac{2}{5}$	$\tfrac{2}{5}$

\

Exercise

1. Suppose the pmf of a r.v X is given by $\mathrm{f(x)} = \begin{cases} 1/6 & \text{for } x = 1, 2, 3, 4, 5, 6 \\ 0, & elsewhere \end{cases}$, Obtain the pmf of $\mathrm{Y} = 2X^2$ and $\mathrm{Z} = X - 3$
2. Let the pmf of a r.v X be given by $\mathrm{f(x)} = \begin{cases} \dfrac{x^2+1}{18} & \text{for } x = 0, 1, 2, 3 \\ 0, & elsewhere \end{cases}$, determine the pmf of $\mathrm{Y} = X^2 + 1$
3. Suppose the pmf of a r.v X is given by $\mathrm{f(x)} = \begin{cases} x/10 & \text{for } x = 0, 1, 2, 3, 4 \\ 0, & elsewhere \end{cases}$, Obtain the pmf of $\mathrm{Y} = |X - 2|$

4. Let the pmf of a r.v X be given by $f(x)=\begin{cases}\left(\frac{1}{2}\right)^x & \text{for } x=1,2,3,....\\ 0, & elsewhere\end{cases}$, determine the pmf of $Y=\begin{cases}1 \text{ if } x \text{ is odd}\\ 0 \text{ if } x \text{ is even}\end{cases}$

5. Suppose the r.v X has a pmf given by $f(x)=\begin{cases}\frac{1}{6}\left(\frac{5}{6}\right)^x & \text{for } x=0,1,2,3,....\\ 0, & elsewhere\end{cases}$, Obtain the pmf of $Y=X-1$

1.6 Change of Variable Technique

.Let X be c. r.v with pdf f(x) and let Y be a function of X, then the pdf of Y $g(y)=f(x)\left|\frac{dx}{dy}\right|$

NB If $F(x)$ is the cdf of a r.v X, then a r.v $Y=F(x)$ has a uniform distribution over $[0,1]$

Example 1

A continuous r.v X has a pdf given by $f(x)=\begin{cases}5x^4, 0\le x\le 1\\ 0 \ \ otherwise\end{cases}$. Determine the pdf of $Y=x^3$

Solution

$$Y=x^3 \quad\Rightarrow\quad X=Y^{1/3} \quad\Rightarrow\quad \frac{dx}{dy}=\frac{1}{3}Y^{-2/3} \quad g(y)=f(x)\left|\frac{dx}{dy}\right|=5\left(y^{1/3}\right)^4\times\frac{1}{3}y^{-2/3}=\begin{cases}\frac{5}{3}y^{2/3}, 0\le x\le 1\\ 0 \ \ otherwise\end{cases}$$

Example 2

A r.v X has pdf $f(x)=\begin{cases}24x^2, 0\le x\le \frac{1}{2}\\ 0 \ \ otherwise\end{cases}$ determine the pdf of $Y=8X^3$

Solution

$$Y=8x^3\Rightarrow \quad X=\tfrac{1}{2}Y^{1/3} \quad\Rightarrow\quad \frac{dx}{dy}=\tfrac{1}{6}y^{-2/3} \quad g(y)=f(x)\left|\frac{dx}{dy}\right|=24\left(\tfrac{1}{2}y^{1/3}\right)^2\times\tfrac{1}{6}y^{-2/3}=\begin{cases}1, \ 0\le x\le 1\\ 0, \ \ elsewhere\end{cases}$$

NB $Y=8X^3$ is the cdf of X

Exercise:

1. For a r.v X with $f(x)=\begin{cases}5x^4, 0\le x\le 1\\ 0 \ \ otherwise\end{cases}$, determine the pdf of $Y=2\ln x$ and its range.

2. A r.v X has pdf $f(x)=\begin{cases}e^{-x}, x\ge 0\\ 0 \ \ otherwise\end{cases}$ determine the pdf of $Y=X^4$

3. The probability density function of X is given by $f(x)=\begin{cases}\dfrac{2}{\pi(x^2+4)} & -\infty\le X\le\infty\\ 0 & otherwise\end{cases}$

   Obtain the probability density function of $Y=\tan^{-1}(\frac{x}{2})$

4. Which transformation will change a r.v X with pdf is as below to a uniform R.V Y whose range is $0 \le x \le 1$ a) $f(x) = \begin{cases} 2e^{-2x}, & x \ge 0 \\ 0 & otherwise \end{cases}$ b) $f(x) = \begin{cases} \frac{1}{2}(x-3), & 3 \le x \le 5 \\ 0 & otherwise \end{cases}$

1.7 Expectation and Variance of a Random Variable

1.7.1 Expected Values

One of the most important things we'd like to know about a random variable is: what value does it take on average? What is the average price of a computer? What is the average value of a number that rolls on a die? The value is found as the average of all possible values, weighted by how often they occur (i.e. probability)

Definition: Let X be a random variable with probability distribution $p(X = x)$. Then the **expected value** of X, denoted $E(X)$ or μ, is given by;

$$E(x) = \mu = \begin{cases} \sum_{x=-\infty}^{\infty} xp(X = x) & \text{if X is discrete} \\ \int_{-\infty}^{\infty} xp(X = x)dx & \text{if X is continuous} \end{cases}.$$

Theorem: Let X be a r.v. with probability distribution $p(X=x)$ and let $g(x)$ be a real-valued function of X. ie $g: \mathbb{R} \to \mathbb{R}$, then the expected value of $g(x)$ is given by

$$E[g(x)] = \begin{cases} \sum_{x=-\infty}^{\infty} g(x)p(X = x) & \text{if X is discrete} \\ \int_{-\infty}^{\infty} g(x)p(X = x)dx & \text{if X is continuous} \end{cases}.$$

Theorem: Let X be a r.v. with probability distribution $p(X = x)$. Then

(i) $E(c) = c$, where c is an arbitrary constant.

(ii) $E[ax + b] == a\mu + b$ where a and b are arbitrary constants

(iii) $E[kg(x)] = kE[g(x)]$ *where* $g(x)$ is a function of X

(iv) $E[ag_1(x) \pm bg_2(x)] = aE[g_1(x)] \pm bE[g_2(x)]$ and in general $E\left[\sum_{i=1}^{n} c_i g_i(x)\right] = \sum_{i=1}^{n} c_i E[g_i(x)]$ are functions of X. This property of expectation is called *linearity property*

Proof

We will sketch the proof using a continuous random variable since the proof using the discrete random variable is similar and was also discussed in probability and statistics I.

(i) $E[c] = \int_{All\ x} cP(X = x)dx = c \int_{All\ x} P(X = x)dx = c(1) = c$

(ii) $E[ax + b] = \int_{All\ x} (ax + b)P(X = x)dx = \int_{All\ x} axP(X = x)dx + \int_{All\ x} bP(X = x)dx$

$$= a \int_{All\ x} xP(X = x)dx + b \int_{All\ x} P(X = x)dx = a\mu + b$$

(iii) $E[kg(x)] = \int_{All\ x} kg(x)P(X = x)dx = k \int_{All\ x} g(x)P(X = x)dx = kE[g(x)]$

(iv) $E[ag_1(x) \pm bg_2(x)] = \int_{All\ x} [ag_1(x) \pm bg_2(x)] P(X=x)dx = \int_{All\ x} ag_1(x) P(X=x)dx \pm \int_{All\ x} bg_2(x) P(X=x)dx$

$$= E[ag_1(x)] \pm E[bg_2(x)] = aE[g_1(x)] \pm bE[g_2(x)] \text{ ie from part iii}$$

1.7.2 Variance and Standard Deviation

Definition: Let X be a r.v with mean $E(X)=\mu$, the **variance** of X, denoted σ^2 or Var(X), is given by $Var(X)=\sigma^2=E(X-\mu)^2$. The units for variance are square units. The quantity that has the correct units is **standard deviation**, denoted σ. It's actually the positive square root of Var(X).

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

Theorem: $Var(X)=E(X-\mu)^2=E(X)^2-\mu^2$

Proof:

$Var(X)=E(X-\mu)^2=E\left(X^2-2X\mu+\mu^2\right)=E(X)^2-2\mu E(X)+\mu^2=E(X)^2-\mu^2$ Since $E(X)=\mu$

Theorem: $Var(aX+b)=a^2\,\mathrm{var}(X)$

Proof:

Recall that $E[aX+b]=a\mu+b$ therefore

$Var(aX+b)=E[(aX+b)-(a\mu+b)]^2=E[a(X-\mu)]^2==E\left[a^2(X-\mu)^2\right]=a^2E\left[(X-\mu)^2\right]=a^2\,\mathrm{var}(X)$

Remark

(i) The expected value of X always lies between the smallest and largest values of X.

(ii) In computations, bear in mind that variance cannot be negative!

Example 1

Given a probability distribution of X as below, find the mean and standard deviation of X.

x	0	1	2	3
P(X=x)	1/8	1/4	3/8	1/4

Solution

x	0	1	2	3	total
$p(X=x)$	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{3}{8}$	$\frac{1}{4}$	1
$xp(X=x)$	0	$\frac{1}{4}$	$\frac{3}{4}$	$\frac{3}{4}$	$\frac{7}{4}$
$x^2p(X=x)$	0	$\frac{1}{4}$	$\frac{3}{2}$	$\frac{9}{4}$	4

$E(X)=\mu=\sum_{x=0}^{3} xp(X=x)=1.75$ and

standard deviation

$\sigma=\sqrt{E(X^2)-\mu^2}=\sqrt{4-1.75^2}\approx 0.968246$

Example 2

The probability distribution of a r.v X is as shown below, find the mean and standard deviation of; a) X b) $Y=12X+6$.

x	0	1	2
P(X=x)	1/6	1/2	1/3

Solution

x	0	1	2	total
$p(X=x)$	$\frac{1}{6}$	$\frac{1}{2}$	$\frac{1}{3}$	1
$xp(X=x)$	0	$\frac{1}{2}$	$\frac{2}{3}$	$\frac{7}{6}$
$x^2p(X=x)$	0	$\frac{1}{2}$	$\frac{4}{3}$	$\frac{11}{6}$

$E(X)=\mu=\sum_{x=0}^{2} xp(X=x)=\frac{7}{6}$ and

$E(X^2)=\sum_{x=0}^{2} x^2p(X=x)=\frac{11}{6}$

Standard deviation $\sigma = \sqrt{E(X^2) - \mu^2} = \sqrt{11/6 - (7/6)^2} = \sqrt{17/6} \approx 1.6833$

Now $E(Y) = 12E(X) + 6 = 12(7/6) + 6 = 20$

$Var(Y) = Var(12X + 6) = 12^2 \times Var(X) = 144 \times \sqrt{17/6} \approx 242.38812$

Example 3

A continuous random variable X has a pdf given by $f(x) = \begin{cases} \frac{1}{2}x, & 0 \le x \le 2 \\ 0, & elsewhere \end{cases}$, find the mean and standard of X

Solution

$$E(x) = \int_{-\infty}^{\infty} xf(x)dx = \int_0^2 \tfrac{1}{2}x^2 dx = \left[\frac{x^3}{6}\right]_0^2 = \frac{4}{3} \text{ and } E(x^2) = \int_{-\infty}^{\infty} x^2 f(x)dx = \int_0^2 \tfrac{1}{2}x^3 dx = \left[\frac{x^4}{8}\right]_0^2 = 2$$

Standard deviation $\sigma = \sqrt{E(X^2) - \mu^2} = \sqrt{2 - (4/3)^2} = \frac{\sqrt{2}}{3}$

Exercise

1. Suppose X has a probability mass function given by the table below

x	2	3	4	5	6
P(X=x)	0.01	0.25	0.4	0.3	0.04

   Find the mean and variance of; X

2. Suppose X has a probability mass function given by the table below

x	11	12	13	14	15
P(X=x)	0.4	0.2	0.2	0.1	0.1

   Find the mean and variance of; X

3. Let X be a random variable with $P(X = 1) = 0.2$, $P(X = 2) = 0.3$, and $P(X = 3) = 0.5$. What is the expected value and standard deviation of; a)X *b)* $Y = 5X - 10$?
4. A random variable W has the probability distribution shown below,

w	0	1	2	3
P(W=w)	2d	0.3	d	0.1

   Find the values of the constant d hence determine the mean and variance of W. Also find the mean and variance of $Y = 10X + 25$
5. A random variable X has the probability distribution shown below,

x	1	2	3	4	5
P(X=x)	7c	5c	4c	3c	c

   Find the values of the constant c hence determine the mean and variance of X.
6. The random variable Z has the probability distribution shown below,

z	2	3	5	7	11
P(Z=z)	$1/6$	$1/3$	$1/4$	x	y

   If $E(Z) = 4\frac{2}{3}$, find the values of x and y hence determine the variance of Z
7. A discrete random variable M has the probability distribution $f(m) = \begin{cases} \frac{m}{36}, & m = 1, 2, 3, ..., 8 \\ 0, & elsewhere \end{cases}$, find the mean and variance of M
8. For a discrete random variable Y the probability distribution is $f(y) = \begin{cases} \frac{5-y}{10}, & y = 1, 2, 3, 4 \\ 0, & elsewhere \end{cases}$, calculate $E(Y)$ and var(Y)
9. Suppose X has a pmf given by $f(x) = \begin{cases} kx & \text{for } x = 1, 2, 3, 4 \\ 0, & elsewhere \end{cases}$, find the value of the constant k hence obtain the mean and variance of X

10. A team of 3 is to be chosen from 4 girl and 6 boys. If X is the number of girls in the team, find the probability distribution of X hence determine the mean and variance of X
11. A fair six sided die has; '1' on one face, '2' on two of it's faces and '3' on the remaining three faces. The die is rolled twice. If T is the total score write down the probability distribution of T hence determine; a) the probability that T is more than 4. b) the mean and variance of T
12. The pdf of a continuous r.v R is given by $f(r) = \begin{cases} kr \text{ for } 0 \le r \le 4 \\ 0, \; elsewhere \end{cases}$, (a) Determine c. hence Compute $P(10 \le r \le 2)$, $E(X)$ and $Var(X)$.
13. A continuous r.v M has the pdf given by $f(m) = \begin{cases} k(1 - \frac{m}{10}) \text{ for } M \le 10 \\ 0, \; elsewhere \end{cases}$, find the value of the constant k, the mean and the variance of X
14. A continuous r.v X has the pdf given by $f(x) = \begin{cases} k(1-x) \text{ for } 0 \le x \le 1 \\ 0, \; elsewhere \end{cases}$, findt the value of the constant k. Also find the mean and the variance of X
15. The lifetime of new bus engines, T years, has continuous pdf $f(t) = \begin{cases} {}^{d}\!/_{t^2} \text{ if } x \ge 1 \\ 0, \text{ if } x < 1 \end{cases}$ find the value of the constant d hence determine the mean and standard deviation of T
16. An archer shoots an arrow at a target. The distance of the arrow from the centre of the target is a random variable X whose p.d.f. is given by $f(x) = \begin{cases} k(3 + 2x - x^2) \;\; if \; x \le 3 \\ 0, \;\; if \; x > 3 \end{cases}$ find the value of the constant k. Also find the mean and standard deviation of X
17. A continuous r.v X has the pdf given by $f(x) = \begin{cases} k(1+x), \; -1 \le x < 0 \\ 2k(1-x), \; 0 \le x \le 1 \\ 0, \; elsewhere \end{cases}$, find the value of the constant k. Also find the mean and the variance of X
18. A continuous r.v X has the pdf given by $f(x) = \begin{cases} e^{-x} \text{ for } x > 0 \\ 0, \; elsewhere \end{cases}$, find the mean and standard deviation of; a) X b) $Y = e^{\frac{3}{4}x}$

1.8 Mode, Median, Quartiles and Percentiles

Another measure commonly used to summarize random variables are the mode and median; Mode is the value of x that maximizes the pdf. That is the value of x for which $f^{\prime}(x) = 0$. Median is the value m such that "half of the distribution lies to the left of m and half to the right". More formally, m should satisfy $F_X(m) = 0.5$.

Note: If there is a value m such that the graph of y= f(x) is symmetric about x=m, then both the expected value and the median of X are equal to m.

The lower quartile Q_1 and the upper quartile Q_3 are similarly defined by

$$F_X(Q_1) = 0.25 \text{ and } F_X(Q_3) = 0.75$$

Thus, the probability that X lies between Q_1 and Q_3 is $0.75 - 0.25 = 0.5$, so the quartiles give an estimate of how spread-out the distribution is. More generally, we define the nth percentile of X to be the value of x_n such that $F_X(x_n) = 0.01n$ or $n/_{100}$, that is, the probability that X is smaller than x_n is n%.

Example A random variable X has the pdf given by $f(x) = \begin{cases} 2x, & 0 \le x \le 1 \\ 0, & elsewhere \end{cases}$ Find the lower, middle and upper quartiles.

Solution

On the interval $0 \le x \le 1$, the cdf of X is given by $F(x) = x^2$ thus

a) At lower quartile Q_1, $F(Q_1) = Q_1^{\ 2} = 0.25 \Rightarrow Q_1 = \sqrt{0.25} = 0.5$

b) At median m, $F(m) = m^2 = 0.5 \Rightarrow m = \sqrt{0.5} = \frac{1}{\sqrt{2}}$

c) At upper quartile u $F(Q_3) = Q_3^{\ 2} = 0.75 \Rightarrow Q_3 = \sqrt{0.75} = \frac{\sqrt{3}}{2}$

Qn Find the 64th percentile of the pdf in the above example.

A continuous r.v X has the pdf given by $f(x) = \begin{cases} k(1-x) & \text{for } 0 \le x \le 1 \\ 0, & elsewhere \end{cases}$, find the of mode.

2. PROBABILITY DISTRIBUTION

2.1 Discrete Distribution

Among the discrete distributions that we will look at includes the Bernoulli, binomial, Poisson, geometric and hyper-geometric

2.1.1 Bernoulli distribution

Definition: A Bernoulli trial is a random experiment in which there are only two possible outcomes - success and failure. Eg

- Tossing a coin and considering heads as success and tails as failure.
- Checking items from a production line: success = not defective, failure = defective.
- Phoning a call centre: success = operator free; failure = no operator free.

A Bernoulli random variable X takes the values 0 and 1 and $P(X = 1) = p$ and $P(X = 0) = 1 - p$

Definition: A r.v X is said to be a real Bernoulli distribution if it's pmf is given by;

$$P(X = x) = \begin{cases} p^x(1-p)^{1-x} & for \ \ x = 0,1 \\ 0 & otherwise \end{cases}$$

We abbreviate this as $X \sim B(p)$ ie p is the only parameter here. It can be easily checked that the mean and variance of a Bernoulli random variable are $\mu = p$ and $\sigma^2 = p(1-p1)$

2.1.2 Binomial Distribution

Consider a sequence of *n* independent, Bernoulli trials, with each trial having two possible outcomes, *success* or *failure*. Let *p* be the probability of a success for any single trial. Let*X* denote the number of successes on *n* trials. The random variable *X* is said to have a **binomial distribution** and has probability mass function

$$P(X = x) = {}_nC_r \times p^x(1-p)^{n-x} \quad for \ \ x = 0,1,2.....n$$

We abbreviate this as $X \sim \text{Bin}(n, p)$ read as “X follows a binomial distribution with parameters n and p”. ${}_nC_r$ Counts the number of outcomes that include exactly x successes and $n-x$ failures.

The mean and variance of a Binomial random variable are respectively given by;

$$\mu = np \quad \text{and} \quad \sigma^2 = np(1-p1)$$

Let’s check to make sure that if X has a binomial distribution, then $\sum_{x=0}^{n} P(X = x) = 1$. We will need the binomial expansion for any polynomial:

$$(p+q)^n = \sum_{x=0}^{n} {}_nC_r \times p^x q^{n-x} \text{ therefore } \sum_{x=0}^{n} {}_nC_x \times p^x (1-p)^{n-x} = [p+(1-p)]^n = 1^n = 1 \text{ so}$$

Example 1

A biased coin is tossed 6 times. The probability of heads on any toss is 0:3. Let X denote the number of heads that come up. Calculate: (i) $P(X = 2)$ (ii) $P(X = 3)$ (iii) $P(1 < X < 5)$

Solution

If we call heads a success then X has a binomial distribution with parameters n=6 and p=0:3.

(i) $P(X = 2) = {}_6C_2 \times (0.3)^2 (0.7)^4 = 0.324135$

(ii) $P(X = 3) = {}_6C_3 \times (0.3)^3 (0.7)^3 = 0.18522$

(iii) $$\begin{aligned} P(1 < X \le 5) &= P(X = 2) + P(X = 3) + P(X = 4) + P(X = 5) \\ &= 0.324 + 0.185 + 0.059 + 0.01 = 0.578 \end{aligned}$$

Example 2

A quality control engineer is in charge of testing whether or not 90% of the DVD players produced by his company conform to specifications. To do this, the engineer randomly selects a batch of 12 DVD players from each day's production. The day's production is acceptable provided no more than 1 DVD player fails to meet specifications’. Otherwise, the entire day's production has to be tested.

a) What is the probability that the engineer incorrectly passes a day's production as acceptable if only 80% of the day's DVD players actually conform to speciication?
b) What is the probability that the engineer unnecessarily requires the entire day's production to be tested if in fact 90% of the DVD players conform to speciffications?

Solution

a) Let X denote the number of DVD players in the sample that fail to meet speciffications. In part (i) we want $P(X \le 1)$ with binomial parameters $n = 12$ and $p = 0.2$

$$\begin{aligned} P(X \le 1) &= P(X = 0) + P(X = 1) = {}_{12}C_0 \times (0.2)^0 (0.8)^{12} + {}_{12}C_1 \times (0.2)^1 (0.8)^{11} \\ &= 0.069 + 0.206 = 0.275 \end{aligned}$$

b) We now want $P(X > 1)$ with parameters $n = 12$ and $p = 0.1$.

$$P(X \le 1) = P(X = 0) + P(X = 1) = {}_{12}C_0 \times (0.1)^0 (0.9)^{12} + {}_{12}C_1 \times (0.1)^1 (0.9)^{11} = 0.659$$

So $P(X > 1) = 0.34$

Example 3

Bits are sent over a communications channel in packets of 12. If the probability of a bit being corrupted over this channel is 0:1 and such errors are independent, what is the probability that no more than 2 bits in a packet are corrupted?

If 6 packets are sent over the channel, what is the probability that at least one packet will contain 3 or more corrupted bits?

Let X denote the number of packets containing 3 or more corrupted bits. What is the probability that X will exceed its mean by more than 2 standard deviations?

Solution

Let C denote the number of corrupted bits in a packet. Then in the first question, we want

$P(C \le 2) = P(C=0) + P(C=1) + P(C=2)$

$= {}_{12}C_0(0.1)^0(0.9)^{12} + {}_{12}C_1(0.1)^1(0.9)^{11} + {}_{12}C_2(0.1)^2(0.9)^{10}$

$= 0.282 + 0.377 + 0.23 = 0.889.$

Therefore the probability of a packet containing 3 or more corrupted bits is

$P(C \ge 3) = 1 - P(C \le 2) = 1 - 0.889 = 0.111.$

Let X be the number of packets containing 3 or more corrupted bits. X can be modelled with a binomial distribution with parameters $n = 6$ and $\text{p} = 0.111$. The probability that at least one packet will contain 3 or more corrupted bits is:

$P(X \ge 1) = 1 - P(X = 0) = 1 - {}_6C_0 \times (0.111)^0(0.889)^6 = 0.494.$

The mean of X is $\text{E(X)} = 6(0.111) = 0.666$ and its standard deviation is

$= \sqrt{6(0.111)(0.889)} = 0.77$

So the probability that X exceeds its mean by more than 2 standard deviations is

$P(X > \mu + 2\sigma) = P(X > 2.2) = P(X \ge 3)$ since X is discrete.

Now $P(X \ge 3) = 1 - P(X \le 2) = 1 - \{P(X=0) + P(X=1) + P(X=2)\}$

$= 1 - \left[{}_6C_0 \times (0.111)^0(0.889)^6 + {}_6C_1 \times (0.111)^1(0.889)^5 + {}_6C_2 \times (0.111)^2(0.889)^4\right]$

$= 1 - (0.4936 + 0.3698 + 0.1026) = 0.032$

Exercise

1. A fair coin is tossed 10 times. What is the probability that exactly 6 heads will occur.
2. If 3% of the electric bulbs manufactured by a company are defective find the probability that in a sample of 100 bulbs exactly 5 bulbs are defective.
3. An oil exploration firm is formed with enough capital to finance 10 explorations. The probability of a particular exploration being successful is 0.1. Find mean and variance of the number of successful explorations.
4. Emily hits 60% of her free throws in basketball games. She had 25 free throws in last week's game.
   a) What is the expected number and the standard deviation of Emily's hit ?
   b) Suppose Emily had 7 free throws in yesterday's game.What is the probability that she made at least 5 hits?
5. A coin is loaded so that heads has 60% chance of showing up. This coin is tossed 3 times.
   a) What are the mean and the standard deviation of the number of heads that turned out?
   b) What is the probability that the head turns out at least twice?
   c) What is the probability that an odd number of heads turn out in 3 flips?
6. According to the 2009 current Population Survey conducted by the U.S. Census Bureau, 40% of the U.S. population 25 years old and above have completed a bachelor's degree or more. Given a random sample of 50 people 25 years old or above, what is expected number of people and the standard deviation of the number of people who have completed a bachelor's degree.
7. Joe throws a fair die six times and face number 3 appeared twice. It he incredibly lucky or unusual?
8. If the probability of being a smoker among a group of cases with lung cancer is .6, what's the probability that in a group of 8 cases you have; (a) less than 2 smokers? (b0 More than 5? (c) What are the expected value and variance of the number of smokers?
9. The manufacturer of the disk drives in one of the well-known brands of microcomputers expects 2% of the disk drives to malfunction during the microcomputer's warranty

period. Calculate the probability that in a sample of 100 disk drives, that not more than three will malfunction

10. Suppose 90% of the cars on Thika super highways does over 17 km per litre.
    a) What is the expected number and the standard deviation of cars on Thika super highways that will do over 17 km per litre.in a random sample of 15 cars ?
    b) What is the probability that in a random sample of 15 cars exactly 10 of these will do over 17 km per litre?

2.1.3 Poisson distribution

Named after the French mathematician Simeon Poisson, the distribution is used to model the number of events, (such as the number of telephone calls at a business, number of customers in waiting lines, number of defects in a given surface area, airplane arrivals, or the number of accidents at an intersection), occurring within a given time interval. Other such random events where Poisson distribution can apply includes;

- the number of hits to your web site in a day
- the number of calls that arrive in each day on your mobile phone
- the rate of job submissions in a busy computer centre per minute.
- the number of messages arriving to a computer server in any one hour.

Poisson probabilities are useful when there are a large number of independent trials with a small probability of success on a single trial and the variables occur over a period of time. It can also be used when a density of items is distributed over a given area or volume. The formula for the Poisson probability mass function is $P(X=x)=\frac{\lambda^x e^{-\lambda}}{x!},\ x=0,1,2,....$ This is abbreviated as $X \sim Po(\lambda)$. λ is the shape parameter which indicates the average number of events in the given time interval. The mean and variance of this distribution are equal ie $\mu=\sigma^2=\lambda$

Let's check to make sure that if *X* has a poisson distribution, then $\sum_{x=0}^{\infty} P(X=x)=1$. We will need to recall that $e^{\lambda}=1+\frac{\lambda}{1!}+\frac{\lambda^2}{2!}+\frac{\lambda^3}{3!}+\frac{\lambda^4}{4!}+.....$ Now

$$\sum_{x=0}^{\infty} P(X=x)=\sum_{x=0}^{\infty}\frac{\lambda^x e^{-\lambda}}{x!}=e^{-\lambda}\sum_{x=0}^{\infty}\frac{\lambda^x}{x!}=e^{-\lambda}e^{\lambda}=e^0=1$$

Remark The major difference between Poisson and Binomial distributions is that the Poisson does not have a fixed number of trials. Instead, it uses the fixed interval of time or space in which the number of successes is recorded.

Example 1

Consider a computer system with Poisson job-arrival stream at an average of 2 per minute. Determine the probability that in any one-minute interval there will be

a) 0 jobs; b) exactly 2 jobs; c) at most 3 arrivals. d) more than 3 arrivals

Solution

Job Arrivals with λ=2

a) No job arrivals: $P(X=0)=e^{-2}=0.1353353$

b) Exactly 3 job arrivals: $P(X=3)=\frac{2^3 e^{-2}}{3!}=0.1804470$

c) At most 3 arrivals

d) $P(X \le 3) = P(X=0)+P(X=1)+P(X=2)+P(X=3) = \left(1+\frac{2}{1}+\frac{2^2}{2}+\frac{2^3}{3!}\right)e^{-2} = 0.8571$

e) more than 3 arrivals $P(X > 3) = 1 - P(X \le 3) == 1 - 0.8571 = 0.1429$

Example 2

If there are 500 customers per eight-hour day in a check-out lane, what is the probability that there will be exactly 3 in line during any five-minute period?

Solution

The expected value during any one five minute period would be 500 / 96 = 5.2083333. The 96 is because there are 96 five-minute periods in eight hours. So, you expect about 5.2 customers in 5 minutes and want to know the probability of getting exactly 3.

$$P(X=3) = \frac{(-500/96)^{3} e^{--500/96}}{3!} = 0.1288\,(\text{approx})$$

Example 3

If new cases of West Nile in New England are occurring at a rate of about 2 per month, then what's the probability that exactly 4 cases will occur in the next 3 months?

Solution

X ~ Poisson (λ=2/month)

$$\text{P(X} = 4 \text{ in 3 months)} = \frac{(2*3)^4 e^{-(2*3)}}{4!} = \frac{6^4 e^{-(6)}}{4!} = 13.4\%$$

Exactly 6 cases?

$$\text{P(X} = 6 \text{ in 3 months)} = \frac{(2*3)^6 e^{-(2*3)}}{6!} = \frac{6^6 e^{-(6)}}{6!} = 16\%$$

Exercise

1. Calculate the Poisson distribution whose λ (Average Rate of Success)) is 3 & X (Poisson Random Variable) is 6.
2. Customers arrive at a checkout counter according to a Poisson distribution at an average of 7 per hour. During a given hour, what are the probabilities that
   a) No more than 3 customers arrive?
   b) At least 2 customers arrive?
   c) Exactly 5 customers arrive?
3. Manufacturer of television set knows that on an average 5% of their product is defective. They sells television sets in consignment of 100 and guarantees that not more than 2 set will be defective. What is the probability that the TV set will fail to meet the guaranteed quality?
4. It is known from the past experience that in a certain plant there are on the average of 4 industrial accidents per month. Find the probability that in a given year will be less that 3 accidents.
5. Suppose that the change of an individual coal miner being killed in a mining accident during a year is 1.1499. Use the Poisson distribution to calculate the probability that in the mine employing 350 miners- there will be at least one accident in a year.
6. The number of road construction projects that take place at any one time in a certain city follows a Poisson distribution with a mean of 3. Find the probability that exactly five road construction projects are currently taking place in this city. (0.100819)
7. The number of road construction projects that take place at any one time in a certain city follows a Poisson distribution with a mean of 7. Find the probability that more than four road construction projects are currently taking place in the city. (0.827008)

8. The number of traffic accidents that occur on a particular stretch of road during a month follows a Poisson distribution with a mean of 7.6. Find the probability that less than three accidents will occur next month on this stretch of road. (0.018757)
9. The number of traffic accidents that occur on a particular stretch of road during a month follows a Poisson distribution with a mean of 7. Find the probability of observing exactly three accidents on this stretch of road next month. (0.052129)
10. The number of traffic accidents that occur on a particular stretch of road during a month follows a Poisson distribution with a mean of 6.8. Find the probability that the next two months will both result in four accidents each occurring on this stretch of road. (0.00985)
11. Suppose the number of babies born during an 8-hour shift at a hospital's maternity wing follows a Poisson distribution with a mean of 6 an hour. Find the probability that five babies are born during a particular 1-hour period in this maternity wing. (0.160623)
12. The university policy department must write, on average, five tickets per day to keep department revenues at budgeted levels. Suppose the number of tickets written per day follows a Poisson distribution with a mean of 8.8 tickets per day. Find the probability that less than six tickets are written on a randomly selected day from this distribution. (0.128387)
13. A taxi firm has two cars which it hires out day by day. The number of demands for a car on each day is distributed as Poisson distribution with mean 1.5. Calculate the proportion of days on which neither car is used and the proportion of days on which some demands is refused
14. If calls to your cell phone are a Poisson process with a constant rate λ=0.5 calls per hour, what's the probability that, if you forget to turn your phone off in a 3 hour lecture, your phone rings during that time? How many phone calls do you expect to get during this lecture?
15. The average number of defects per wafer (defect density) is 3. The redundancy built into the design allows for up to 4 defects per wafer. What is the probability that the redundancy will not be sufficient if the defects follow a Poisson distribution?
16. The mean number of errors due to a particular bug occurring in a minute is 0.0001
    a) What is the probability that no error will occur in 20 minutes?
    b) How long would the program need to run to ensure that there will be a 99.95% chance that an error wills showup to highlight this bug?

Properties of Poisson

- The mean and variance are both equal to λ.
- The sum of independent Poisson variables is a further Poisson variable with mean equal to the sum of the individual means.
- As well as cropping up in the situations already mentioned, the Poisson distribution provides an approximation for the Binomial distribution.

2.1.4 Geometric Distribution

Suppose a Bernoulli trial with success probability *p* is performed repeatedly until the first success appears we want to find the probability that the first success occurs on the yth trial. ie let *Y* denote the number of trials needed to obtain the first success. The sample space S={*s;fs;ffs, fffs ffffs …}*. This is an *infinite* sample space (though it is still discrete). What is the probability of a sample point, say $P(fffs) = P(Y = 4)$ *)*? Since successive trials are independent (this is implicit in the statement of the problem), we have

$P(fffs) = P(Y = 4) = q^3p$ where $q = 1 - p$ and $0 \le p \le 1$

Definition: A r.v. *Y* is said to have a **geometric probability distribution** if and only if

$$P(Y = y) = \begin{cases} pq^{y-1} \text{for } y = 1, 2, 3 ..... \text{ where } q = 1 - p \\ 0 \quad \textit{otherwise} \end{cases}.$$

This is abbreviated as $X \sim Geo(p)$.

The only parameter for this geometric distribution is p (ie the probability of success in each trial). To be sure everything is consistent; we should check that the probabilities of all the sample points add up to 1. Now

$$\sum_{y=1}^{\infty} P(Y = y) = \sum_{y=1}^{\infty} pq^{y-1} = \frac{p}{1-q} = 1$$

Recall sum to infinity of a convergent G.P is $s = \frac{a}{1-r}$

The cdf of a geometric distributions is given by

$$F(y) = P(Y \leq y) = P(Y = 1) + P(Y = 2) + P(Y = 3) + ... + P(Y = y)$$

$$= p + pq + pq^2 + ...pq^{y-1} = \frac{p(1-q^y)}{1-q} = 1 - q^y$$

Let $Y \sim$ Geo(p), then $\mu = E(Y) = \frac{1}{p}$ and $Var(X) = \sigma^2 = \frac{q}{p^2}$ Show?

Example 1

A sharpshooter normally hits the target 70% of the time.

a) Find the probability that her first hit is on the second shot
b) Find the mean and standard deviation of the number of shots required to realize the 1^{st} hit

Solution

Let X be the random variable 'the number of shoots required to realize the 1^{st} hit'

$x \sim Geo(0.7)$ and $P(X = x) = 0.7(1 - 0.7)^{x-1},\ x = 1, 2, 3, ...$

a) $P(X = 2) = p(1 - \rho) = 0.7(0.3) = 0.21$
b) $\mu = \frac{1}{\rho} = \frac{1}{0.7} = 1.428571$ and $\sigma = \frac{\sqrt{1-p}}{p} = \frac{\sqrt{1-0.7}}{0.7} \approx 0.78$

Example 2

The State Department is trying to identify an individual who speaks Farsi to fill a foreign embassy position. They have determined that 4% of the applicant pool are fluent in Farsi.

a) If applicants are contacted randomly, how many individuals can they expect to interview in order to find one who is fluent in Farsi?
b) What is the probability that they will have to interview more than 25 until they find one who speaks Farsi?

Solution

a) $\mu = \frac{1}{\rho} = \frac{1}{0.04} = 25$
b) $P(X \leq 25) = (1 - \rho)^n = 1 - (0.04)^{25} = 1 \ \Rightarrow\ P(X > 25) = 1 - P(X \leq 25) = 0$

Example 3

From past experience it is known that 3% of accounts in a large accounting population are in error. What is the probability that 5 accounts are audited before an account in error is found? What is the probability that the first account in error occurs in the first five accounts audited?

Solution

$P(Y = 5) = 0.03(0.97)^4 = 0.02655878 \qquad P(Y \leq 5) = 1 - 0.97^5 = 0.14126597$

Exercise

1. Over a very long period of time, it has been noted that on Friday's 25% of the customers at the drive-in window at the bank make deposits. What is the probability that it takes 4 customers at the drive-in window before the first one makes a deposit.
2. It is estimated that 45% of people in Fast-Food restaurants order a diet drink with their lunch. Find the probability that the fourth person orders a diet drink. Also find the probability that the first diet drinker of th e day occurs before the 5th person.
3. What is the probability of rolling a sum of seven in fewer than three rolls of a pair of dice? Hint (The random variable, X, is the number of rolls before a sum of 7.)
4. In New York City at rush hour, the chance that a taxicab passes someone and is available is 15%. a) How many cabs can you expect to pass you for you to find one that is free and b) what is the probability that more than 10 cabs pass you before you find one that is free.
5. An urn contains N white and M black balls. Balls are randomly selected, one at a time, until a black ball is obtained. If we assume that each selected ball is replaced before the next one is drawn, what is;
   a) the probability that exactly n draws are needed?
   b) the probability that at least k draws are needed?
   c) the expected value and Variance of the number of balls drawn?
6. In a gambling game a player tosses a coin until a head appears. He then receives $\$2n$, where n is the number of tosses.
   a) What is the probability that the player receives $8.00 in one play of the game?
   b) If the player must pay $5.00 to play, what is the win/loss per game?
7. An oil prospector will drill a succession of holes in a given area to find a productive well. The probability of success is 0.2.
   a) What is the probability that the 3rd hole drilled is the first to yield a productive well?
   b) If the prospector can afford to drill at most 10 well, what is the probability that he will fail to find a productive well?
8. A well-travelled highway has itstraffic lights green for 82% of the time. If a person travelling the road goes through 8 traffic intersections, complete the chart to find a) the probability that the first red light occur on the nth traffic light and b) the cumulative probability that the person will hit the red light on or before the nth traffic light.
9. An oil prospector will drill a succession of holes in a given area to find a productive well. The probability of success is 0.2.
   a) What is the probability that the 3rd hole drilled is the first to yield a productive well?
   b) If the prospector can afford to drill at most 10 well, what is the probability that he will fail to find a productive well?

2.1.5 The negative binomial distribution

Suppose a Bernoulli trial is performed until the t^{th} success is realized. Then the random variable "the number of trials until the t^{th} success is realized" has a negative binomial distribution

Definition: A random variable X has the negative binomial distribution, also called the Pascal distribution, denoted $X \sim NB(r, p)$, if there exists an integer $n \geq 1$and a real number $p \in (0,1)$ such that $P(X = r + x) = {}_{r+x-1}C_x \times p^r (1-p)^x = 1, 2, 3, \ldots$

If r=1 the negative binomial distribution reduces to a geometric distribution.

2.1.6 Hyper geometric Distribution

Hyper geometric experiments occur when the trials are not independent of each other and occur due to sampling without replacement hyper-geometric probabilities involve the multiplication of two combinations together and then division by the total number of combinations

Suppose we have a population of N elements that possess one of two characteristics, e.g. D of them are defective and $N-D$ are non defective. A sample of n elements is randomly selected from the population. The r.v. of interest, Y, is the number of defective elements in the sample.

Definition: A r.v. Y is said to have a **hyper geometric probability distribution** if and only if

$$P(Y=y)=\frac{{}_rC_y \times {}_{N-D}C_{n-y}}{{}_NC_n} \text{ for } y=1,2,3,\ldots \text{n} \quad \text{y} \le \text{n and n - y} \le N-r$$

Theorem: If Y is a r.v with a hyper geometric distribution, then $\mu = E(Y) = \dfrac{nr}{N}$

Example 1

Boxes contain 2000 items of which 10% are defective. Find the probability that no more than 2 defectives will be obtained in a sample of size 10 drawn Without Replacement

Solution

Let Y be the number of defectives

$$P(Y \le 2) = P(Y=0)+P(Y=1)+P(Y=2) = \frac{{}_{180}C_{10}}{{}_{200}C_{10}} + \frac{{}_{20}C_1 \times {}_{180}C_9}{{}_{200}C_{10}} + \frac{{}_{20}C_2 \times {}_{180}C_8}{{}_{200}C_{10}}$$

$$= 0.3398 + 0.3974 + 0.1975 = 0.9347 \quad \Rightarrow \quad P(Y>2) = 0.0653$$

Example 2

How many ways can 3 men and 4 women be selected from a group of 7 men and 10 women?

Solution

The answer is $\dfrac{{}_7C_2 \times {}_{10}C_4}{{}_{17}C_7} = \dfrac{7350}{19448} = 0.3779\,(\text{approx})$

Note that the sum of the numbers in the numerator are the numbers used in the combination in the denominator.

This can be extended to more than two groups and called an extended hypergeometric problem.

Exercise

1. A bottle contains 4 laxative and 5 aspirin tablets. 3 tablets are drawn at random from the bottle. Find the probability that; a) exactly one, b) at most 1 c) at least 2 are laxative tablet.
2. Want is the probability of getting at most 2 diamonds in the5 selected without replacement from a well shuffled deck?
3. A massager has to deliver 10 out of 16 letters to computing department the rest to statistics department. She mixed up the letters and delivered 10 letters at random to computing department. What is the probability that, only 6 letters for computing department actually got there?
4. In a class there are 20 students. 6 are compulsive smokers and they always keep cigarette in their lockers. One day prefects checked at random on 10 lockers. What is the probability that they find cigarette in at most 2 lockers?
5. A box holds 8 green, 4 white and 8 red beads. 6 beads are drawn at random without replacement from the box. What is the probability that 3 red, 2 green and 1 white beads are drawn?

2.2 Continuous Distribution

2.2.1 Uniform (Rectangular) Distribution

A **uniform density function** is a density function that is constant, (*Ie all the values are* equally likely outcomes over the domain). Often referred as the *Rectangular distribution* because the graph of the pdf has the form of a rectangle, making it the simplest kind of density function. The uniform distribution lies between two values on the x-axis. The total area is equal to 1.0 or 100% within the rectangle

Definition: A random variable X has a uniform distribution over the range $[a, b]$ If

$$\mathrm{f(x)} = \begin{cases} \frac{1}{b-a}, & a \le x \le b \\ 0, & elsewhere \end{cases}$$ We denote this distribution by $X \sim U(a, b)$

where: a = smallest value the variable can assume and b = largest value.

The expected Value and the Variance of X are given by $\mu = \dfrac{a+b}{2}$ and $\sigma^2 = \dfrac{(b-a)^2}{12}$

respectively. The cdf F(x) is given by

$$\mathrm{F(x)} = \frac{1}{b-a}\int_a^x dt = \frac{x-a}{b-a} \qquad \Rightarrow \quad \mathrm{F(x)} = \begin{cases} 0 & x < a \\ \frac{x-a}{b-a}, & a \le x \le b \\ 1 & x > b \end{cases}$$

Example Prof Hinga travels always by plane. From past experience he feels that take off time is uniformly distributed between 80 and 120 minutes after check in. determine the probability that: a) he waits for more than 15 minutes for take off after check in. b) the waiting time will be between 1.5 standard deviation from the mean,

Solution

$$X \sim U(80,120) \quad \Rightarrow \quad \mathrm{f(x)} = \begin{cases} \frac{1}{40}, & 80 \le x \le 120 \\ 0, & elsewhere \end{cases}$$

$$P(X > 105) = 1 - P(X \le 105) = 1 - \tfrac{105-80}{40} = \tfrac{3}{8}$$

$$P(\mu - 1.5\sigma \le x \le \mu + 1.5\sigma) = \int_{\mu-1.5\sigma}^{\mu+1.5\sigma} \tfrac{1}{40}\,dx = \tfrac{1}{40}\left[x\right]_{\frac{1}{4\mu-1.5\sigma 0}}^{\mu-1.5\sigma} = \frac{3\sigma}{40} \qquad \text{But } \sigma = \frac{b-a}{\sqrt{12}} = \frac{40}{\sqrt{12}}$$

$$P(\mu - 1.5\sigma \le x \le \mu + 1.5\sigma) = \frac{3\sigma}{40} = \frac{3}{\sqrt{12}}$$

Exercise

1. Uniform: The amount of time, in minutes, that a person must wait for a bus is uniformly distributed between 0 and 15 minutes, inclusive. What is the probability that a person waits fewer than 12.5 minutes? What is the probability that will be between 0.5 standard deviation from the mean,
2. Slater customers are charged for the amount of salad they take. Sampling suggests that the amount of salad taken is uniformly distributed between 5 ounces and 15 ounces. Let x = salad plate filling weight, find the expected Value and the Variance of x. What is the probability hat a customer will take between 12 and 15 ounces of salad?
3. The average number of donuts a nine-year old child eats per month is uniformly distributed from 0.5 to 4 donuts, inclusive. Determine the probability that a randomly selected nine-year old child eats an average of;
    a) more than two donuts
    b) more than two donuts given that his or her amount is more than 1.5 donuts.
4. Starting at 5 pm every half hour there is a flight from Nairobi to Mombasa. Suppose that none of these plane tickets are completely sold out and they always have room for

passagers. A person who wants to fly to Mombasa arrives at the airport at a random time between 8.45 AM and (.45 AM. Determine the probability that he waits for

a) At most 10 minutes b) At least 15 minutes

2.2.2 Exponential Distribution

The exponential distribution is often concerned with the amount of time until some specific event occurs. For example, the amount of time (beginning now) until an earthquake occurs has an exponential distribution. Other examples include the length, in minutes, of long distance business telephone calls, and the amount of time, in months, a car battery lasts. It can be shown, too, that the amount of change that you have in your pocket or purse follows an exponential distribution. Values for an exponential random variable occur in the following way. There are fewer large values and more small values. For example, the amount of money customers spend in one trip to the supermarket follows an exponential distribution. There are more people that spend less money and fewer people that spend large amounts of money. The exponential distribution is widely used in the field of reliability. Reliability deals with the amount of time a product lasts

In brief this distribution is commonly used to model waiting times betweenoccurrences of rare events, lifetimes of electrical or mechanical devices

Definition: A RV X is said to have an exponential distribution with parameter $\lambda > 0$ if the pdf of X is: $\text{f(x)} = \begin{cases} \lambda e^{-\lambda x} \text{for x} \geq 0 \text{ and } \lambda > 0 \\ 0 \quad otherwise \end{cases}$ we abbreviate this as $\text{X} \sim \exp(\lambda)$

λ is called the *rate parameter*

The mean and variance of this distribution are $\mu = \dfrac{1}{\lambda}$ and $\sigma^2 = \dfrac{1}{\lambda^2}$ respectively

The cumulative distribution function is F(x) is given by $\text{F(x)} = \begin{cases} 1 - e^{-\lambda x} \text{for x} \geq 0 \\ 0 \quad otherwise \end{cases}$

Example Torch batteries have a lifespan T years with pdf $\text{f(t)} = \begin{cases} 0.01 e^{-0.01t}, \text{T} \geq 0 \\ 0 \quad otherwise \end{cases}$. Determine the probability that the battery; a) Falls before 25 hours. b) life is between 35 and 50 hours. *c)* life exceeds 120 hours. d) life exceeds the mean lifespan.

S*olution*

a) $\text{P(T} < 25) = \text{F}(25) = \int_0^{25} e^{-0.01t} dt = 1 - e^{-0.01(25)} \approx 0.2212$

b) $\text{P}(35 \leq \text{T} \leq 50) = \int_{35}^{50} e^{-0.01t} dt = e^{-0.35} - e^{-0.50} \approx 0.0982$

c) $\text{P(T} > 120) = \int_{120}^{\infty} e^{-0.01t} dt = e^{-1.2} - 0 \approx 0.3012$

d) $\mu = \dfrac{1}{0.01} = 100 \Rightarrow \text{P(T} > 100) = \int_{100}^{\infty} e^{-0.01t} dt = e^{-1} \approx 0.3679$

Exercise:

1. Jobs are sent to a printer at an average of 3 jobs per hour.
   a) What is the expected time between jobs?
   b) What is the probability thatthe next job is sent within 5 minutes?
2. The time required to repair a machine is an exponential random variable with rate λ= 0.5 downs/hour
   a) what is the probability that a repair time exceeds 2 hours?

b) what is the probability that the repair time will take at least 4 hours given that the repair man has been working on the machine for 3 hours?

3. Buses arrive to a bus stop according to an exponential distribution with rate λ= 4 busses/hour. If you arrived at 8:00 am to the bus stop,
   a) what is the expected time of the next bus?
   b) Assume you asked one of the people waiting for the bus about the arrival time of the last bus and he told you that the last bus left at 7:40 am. What is the expected time of the next bus?
4. Break downs occur on an old car with rate λ= 5 break-downs/month. The owner of the car is planning to have a trip on his car for 4 days.
   a) What is the probability that he will return home safely on his car.
   b) If the car broke down the second day of the trip and the car was fixed, what is the probability that he doesn't return home safely on his car.
5. Suppose that the amount of time one spends in a bank is exponentially distributed with mean 10 minutes. What is the probability that a customer will spend more than 15 minutes in the bank? What is the probability that a customer will spend more than 15 minutes in the bank given that he is still in the bank after 10 minutes?
6. Suppose the lifespan in hundreds of hours, T, of a light bulb of a home lamp is exponentially distributed with lambda = 0.2. compute the probability that the light bulb will last more than 700 hours Also, the probability that the light bulb will last more than 900 hours
7. Let X = amount of time (in minutes) a postal clerk spends with his/her customer. The time is known to have an exponential distribution with the average amount of time equal to 4 minutes.
   a) Find the probability that a clerk spends four to five minutes with a randomly selected customer.
   b) Half of all customers are finished within how long? (Find median)
   c) Which is larger, the mean or the median?
8. On the average, a certain computer part lasts 10 years. The length of time the computer part lasts is exponentially distributed.
   a) What is the probability that a computer part lasts more than 7 years?
   b) On the average, how long would 5 computer parts last if they are used one after another?
   c) Eighty percent of computer parts last at most how long?
   d) What is the probability that a computer part lasts between 9 and 11 years?
9. Suppose that the length of a phone call, in minutes, is an exponential random variable with decay parameter = 1/12 . If another person arrives at a public telephone just before you, find the probability that you will have to wait more than 5 minutes. Let X = the length of a phone call, in minutes. What is median mean and standard deviation of X?

2.2.3 Gamma Distribution

Gamma Function

Let $\alpha > 0$ we define $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx$ called the gamma function with parameter

Theorem

i) $\Gamma(\alpha) = (\alpha-1)\Gamma(\alpha-1)$ if $\alpha > 1$ ii) $\Gamma(\alpha) = (\alpha-1)!$ for $\alpha \in Z^+$

iii) $\Gamma(\alpha) = 2\int_0^\infty u^{2\alpha-1} e^{-u^2} du$ and iv) $\Gamma(\frac{1}{2}) = \sqrt{\pi}$

Proof

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx = -\left[x^{\alpha-1} e^{-x}\right]_0^\infty + (\alpha-1)\int_0^\infty x^{\alpha-2} e^{-x} dx = (\alpha-1)\Gamma(\alpha-1)$$

For $\alpha = 1$, $\Gamma(1) = \int_0^\infty e^{-x} dx = -\left[e^{-x}\right]_0^\infty = 1$ Now suppose it holds for $\alpha = k$ ie

$\Gamma(k) = (k-1)!$ for $k \in Z^+$ then $\Gamma(k+1) = k(k-1)! = k!$ for $k \in Z^+$. Thus if the results holds for $\alpha = k$ then they must also hold for $\alpha = k+1$. But the results are true for $\alpha = 1$

Therefore the results are true for any $\alpha \in Z^+$

from $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx$ put $x = u^2 \Rightarrow dx = 2udu \Rightarrow \Gamma(\alpha) = \int_0^\infty u^{2(\alpha-1)} e^{-u^2} 2udu = 2\int_0^\infty u^{2\alpha-1} e^{-u^2} du$

Gamma Distribution

The $\text{Gamma}(\alpha, \beta)$ distribution models the time required for *a* events to occur, given that the events occur randomly in a Poisson process with a mean time between events of β. For example, an insurance company observes that large commercial fire claims occur randomly in time with a mean of 0.7 years between claims. Not only in real life, the Gamma distribution is also wildly used in many scientific areas, like Reliability Assessment, Queuing Theory, Computer Evaluations, or biological studies. In a nut shell, this distribution is used to model total waiting time of a procedure that consists of α independent stages, each stage with a waiting time having a distribution $\text{Exp}(\beta)$. Then the total time has a Gamma distribution with parameters α and β.

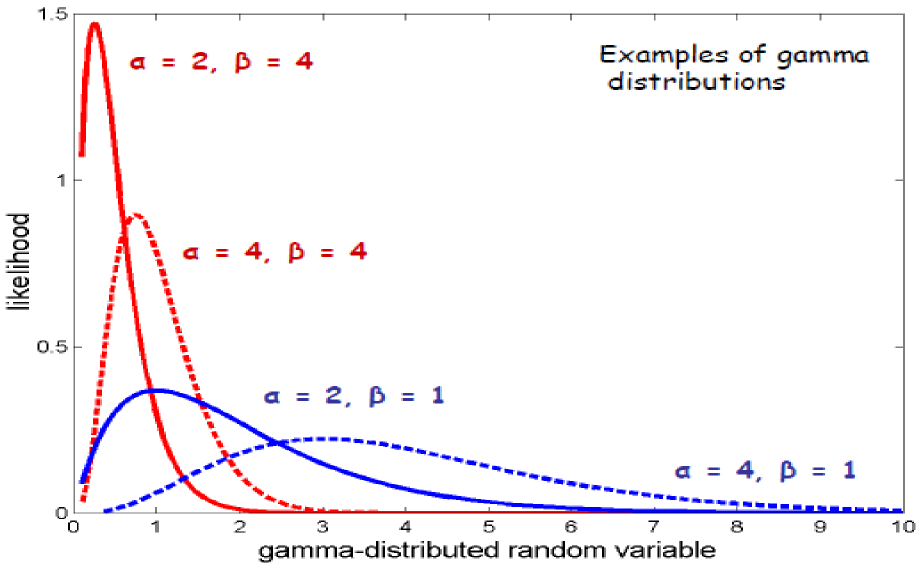
Definition A random variable X has *gamma density* if it pdf is given by

$$f(x) = \begin{cases} \dfrac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}, & x \geq 0 \\ 0 & otherwise \end{cases}$$

We say that the random variable $X \sim \text{Gamma}(\alpha, \beta)$

The parameter α is called the shape parameter while β is called the rate or scale parameter $\Gamma(\alpha)$ is the Gamma function.

Some examples of gamma distributions are plotted below. Notice that the modes shift to the right as the ratio of $\frac{\alpha}{\beta}$ increases.



Remarks

a) If $\beta = 1$ then we have the standard gamma distribution.

b) If $\alpha = 1$ then we have exponential density function.

The cdf, F(x) is of the form $F(x) = \frac{\beta^{\alpha}}{\Gamma(\alpha)}\int_0^x t^{\alpha-1}e^{-\beta t}dt$ and it's computation is not trivial.

Theorem: If X has a gamma distribution with parameters α and β, then

$$E(X) = \mu = \tfrac{\alpha}{\beta} \quad \text{and.} \quad Var(X) = \sigma^2 = \tfrac{\alpha}{\beta^2}$$

Definition: Let ν be a positive integer. A random variable X is said to have a **chi-square** distribution with ν degree of freedom if and only if is a **gamma-**distributed random variable with parameters $\alpha = \frac{\nu}{2}$ and $\beta = 2$

Theorem: If X is a chi-square random variable with degrees of freedom ν, then $E(X) = \mu = \nu$ and. $Var(X) = \sigma^2 = 2\nu$

Example 1 Suppose the reaction time of a randomly selected individual to a certain stimulus has a standard gamma distribution with $\alpha = 2$ sec. Find the probability that reaction time will be (a) between 3 and 5 seconds (b) greater than 4 seconds

Solution

$\Gamma(2) = 1!$. Therefore $f(x) = xe^{-x}, x > 0$ and $f(x) = 0$ *elsewhere*

$$P(3 \le X \le 5) = \int_3^5 xe^{-x}dx = -(x+1)e^{-x}\Big|_3^5 = 4e^{-3} - 6e^{-5} \approx 0.1587$$

$$P(X > 4) = 1 - P(X \le 4) = 1 - \int_0^4 xe^{-x}dx = 1 - \left[-(x+1)e^{-x}\Big|_0^4\right] = 1 - \left[1 - 5e^{-4}\right] \approx 0.09158$$

Example 2 Suppose the survival time X in weeks of a randomly selected male mouse exposed to 240 rads of gamma radiation has a gamma distribution with $\alpha = 8$ and $\beta = 15$.

a) Find the expected value and the standard deviation of the survival time.
b) What is the probability that a mouse survives (i) between 60 and 120 weeks.(ii) at least 30 weeks

Solution

$$X \sim \Gamma(8,15) \quad \Rightarrow \quad f(x) = \begin{cases} \frac{1}{15^8\Gamma(8)}x^7e^{-\frac{x}{15}}, & x \ge 0 \\ 0 & otherwise \end{cases}$$

a) $E(X) = \mu) = (8)(15) = 120$ weeks $\qquad \sigma = \sqrt{8(15)^2} \approx 42.4264$ weeks

$b)$ $$P(60 \le X \le 120) = \int_{60}^{120}\frac{1}{15^8\Gamma(8)}x^7e^{-\frac{x}{15}}dx = \int_4^8\frac{1}{\Gamma(8)}y^7e^{-y}dy$$

$$= -\left(y^7 + 7y^6 + 42y^5 + 210y^4 + 840y^3 + 2520y^2 + 5040y + 5040\right)\frac{e^{-y}}{7!}\Bigg|_4^8$$

$$= \frac{261104e^{-4} - 6805296e^{-8}}{7!} \approx 0.4959$$

$c)$ $$PX \ge 30) = \int_{30}^{\infty}\frac{1}{15^8\Gamma(8)}x^7e^{-\frac{x}{15}}dx = \int_2^{\infty}\frac{1}{7!}y^7e^{-y}dy$$

$$= -\left(y^7 + 7y^6 + 42y^5 + 210y^4 + 840y^3 + 2520y^2 + 5040y + 5040\right)\frac{e^{-y}}{7!}\Bigg|_2^{\infty}$$

$$= \frac{37200e^{-2} - 0}{7!} \approx 0.9989$$

Questionn The time between failures of a laser machine is exponentially distributed with a mean of 25,000 hours. What is the expected time until the second failure? What is the probability that the time until the third failure exceeds 50,000 hours?

2.2.4 Beta Distribution

Beta Function

We define a beat function as $\mathrm{B}(\alpha,\beta)$ as $\mathrm{B}(\alpha,\beta)=\int_0^1 x^{\alpha-1}(1-x)^{\beta-1}dx$ with parameters $\alpha>0$ and $\beta>0$

Theorem:

i) $\mathrm{B}(\alpha,\beta)=\mathrm{B}(\beta,\alpha)$ ie by putting $y=1-x$ in the function

ii) $\mathrm{B}(\alpha,\beta)=\int_0^\infty \frac{u^{\alpha-1}}{(1+u)^{\beta-1}}du$ ie by putting $u=\frac{x}{1-x}$ or $x=\frac{u}{1+u}$ in the function

iii) $\mathrm{B}(\alpha,\beta)=2\int_0^{\pi/2}\cos^{2\alpha-1}t\sin^{2\beta-1}tdt$ ie by putting $x=\cos^2 t$ in the function

iv) $\mathrm{B}(\alpha,\beta)=\frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$

v) $\Gamma(\frac{1}{2})=\sqrt{\pi}$ use $\mathrm{B}(\frac{1}{2},\frac{1}{2})=\frac{\Gamma(\frac{1}{2})\Gamma(\frac{1}{2})}{\Gamma(1)}=[\Gamma(\frac{1}{2})]^2$ but

$\mathrm{B}(\frac{1}{2},\frac{1}{2})=2\int_0^{\pi/2}\cos^{2\times\frac{1}{2}-1}t\sin^{2\times\frac{1}{2}-1}tdt=2\int_0^{\pi/2}dt=\pi \quad \Rightarrow \quad \Gamma(\frac{1}{2})=\sqrt{\pi}$

Note $\Gamma(\frac{1}{2})=\int_0^\infty x^{-1/2}e^{-x}dx=2\int_0^\infty e^{-u^2}du=\sqrt{\pi}$. Simply $\int_0^\infty e^{-x^2}dx=\frac{1}{2}\sqrt{\pi}$

Beta Distribution

Definition: A random variable X is said to have a standard beta distribution with parameters α and β if it's probability density function is given by

$$\text{f(x)} = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha,\beta)}, 0\le x\le 1 \text{ and f(x) = 0 } elsewhere \text{ we denote this as } X \sim Beta(\alpha,\beta)$$

Theorem: If X has a standard beta distribution with parameters α and β, then

$$\mathrm{E(X)}=\mu=\frac{\alpha}{\alpha+\beta} \quad \text{and.} \quad \mathrm{Var\ (X)}=\sigma^2=\frac{\alpha\beta}{(\alpha+\beta+1)(\alpha+\beta)^2}$$

3. MOMENTS AND MOMENT-GENERATING FUNCTIONS

Definition: The k[th] **moment** of a r.v. X taken about zero, or about origin is defined to be $E[X^k]$ and denoted by μ_k' .

Definition: The k[th] **moment** of a r.v. X taken about its mean, or the k[th] **central moment** of X, is defined to be $E(X-\mu)^k$ and denoted by μ_k .

Definition: The **moment-generating function** (mgf), $m(t)$, for a r.v. X is defined to be $M_x(t)$ or simply $M(t)=E[e^{tx}]$

We say that an mgf for X exists if there is $b>0$ such that $M(t)<\infty$ for $|t|\le b$.

$$e^{tx} = 1 + \frac{tx}{1!} + \frac{(tx)^2}{2!} + \frac{(tx)^3}{3!} + \frac{(tx)^4}{4!} + ...$$

$$M(t) = E\left[e^{tx}\right] = \sum_{all\ x}\left(1 + \frac{tx}{1!} + \frac{(tx)^2}{2!} + \frac{(tx)^3}{3!} + \frac{(tx)^4}{4!} + ...\right)P(X = x)$$

$$= \sum_{all\ x} P(X = x) + t\sum_{all\ x} xP(X = x) + \frac{t^2}{2!}\sum_{all\ x} x^2 P(X = x) + \frac{t^3}{3!}\sum_{all\ x} x^3 P(X = x) + ...$$

$$= 1 + tE(X) + \frac{t^2}{2!}E(X^2) + \frac{t^3}{3!}E(X^3) + \ldots \text{ie a function of all moments about the origin}$$

Theorem: If $M(t)$ exists, then for any $k \in N$ $\left.\frac{d^k M(t)}{dt^k}\right|_{t-0} = \mu_k^/ = E(X^k)$

Proof

$$M(t) = 1 + t\mu_1^/ + \frac{t^2}{2!}\mu_2^/ + \frac{t^3}{3!}\mu_3^/ + \ldots$$

$$M^{|}(t) = \mu_1^/ + \frac{2t}{2!}\mu_2^/ + \frac{3t^2}{3!}\mu_3^/ + \ldots \quad \Rightarrow \quad M^{|}(0) = \mu_1^/ = E(X)$$

$$M^{||}(t) = \mu_2^/ + \frac{2t}{2!}\mu_3^/ + \frac{3t^2}{3!}\mu_4^/ + \ldots \quad \Rightarrow \quad M^{||}(0) = \mu_2^/ = E(X^2)$$

Remark: The mgf of a particular distribution is unique and we can recognize the pdf if we are given the mgf.

Example 1

The mgf of a r.v Y is given by $M(t) = \frac{1}{6}e^t + \frac{1}{3}e^{2t} + \frac{1}{2}e^{3t}$ Find the mean and variance of Y

Solution

$$E(Y) = M^{|}(0) = \left.\left(\tfrac{1}{6}e^t + \tfrac{2}{3}e^{2t} + \tfrac{3}{2}e^{3t}\right)\right|_{t-0} = \tfrac{1}{6} + \tfrac{2}{3} + \tfrac{3}{2} = \tfrac{7}{3}$$

$$E(Y^2) = M^{||}(0) = \left.\left(\tfrac{1}{6}e^t + \tfrac{4}{3}e^{2t} + \tfrac{9}{2}e^{3t}\right)\right|_{t-0} = \tfrac{1}{6} + \tfrac{4}{3} + \tfrac{9}{2} = 6$$

$$Var(Y) = E(Y^2) - \mu^2 = 6 - \left(\tfrac{7}{3}\right)^2 = \tfrac{5}{9}$$

Example 2 Find the mgf of a r.v $X \sim Po(\lambda)$

Solution

$$M(t) = E\left[e^{tx}\right] = \sum_{x=0}^{\infty}\left[e^{tx}\right]\frac{\lambda^x e^{-\lambda}}{x!} = e^{-\lambda}\sum_{x=0}^{\infty}\frac{\left(\lambda e^t\right)^x}{x!} = e^{-\lambda}e^{\lambda e^t} = e^{\lambda\left(e^t - 1\right)}$$

Now $E(X) = M^{|}(0) = \left.\lambda e^t e^{\lambda\left(e^t - 1\right)}\right|_{t-0} = \lambda$

Example 3 Find the mgf of a r.v X whose pmf is given by $f(x) = \begin{cases} \frac{1}{6}\left(\frac{5}{6}\right)^x \text{ for } x = 0,1,2,3..... \\ 0 \quad elsewhere \end{cases}$

hence obtain the mean and variance of X

Solution

$$.\,M(t) = E\left[e^{tx}\right] = \sum_{x=0}^{\infty} e^{tx}\,\tfrac{1}{6}\left(\tfrac{5}{6}\right)^x = \sum_{x=0}^{\infty}\tfrac{1}{6}\left(\tfrac{5}{6}e^t\right)^x = \frac{\frac{1}{6}}{1 - \frac{5}{6}e^t} = \frac{1}{6 - 5e^t} = \left(6 - 5e^t\right)^{-1}$$

$M^{|}(t) = 5e^t\left(6-5e^t\right)^{-2}$ and $M^{||}(t) = 5e^t\left(6-5e^t\right)^{-2} + 50e^{2t}\left(6-5e^t\right)^{-3}$

$\Rightarrow E(X) = 5e^t\left(6-5e^t\right)^{-2}\Big|_{t-0} = 5$ and $E(Y^2) = \left[5e^t\left(6-5e^t\right)^{-2} + 50e^{2t}\left(6-5e^t\right)^{-3}\right]_{=0} = 55$

$Var(X) = E(X^2) - \mu^2 = 55 - 5^2 = 30$

Exercise

1) The mgf of a r.v Y is given by; a) $M(t) = e^{2t^2+3t}$ b) $M(t) = \exp\left\{\frac{1}{2}\sigma^2 t^2 + t\mu\right\}$ Find the mean and variance of Y
2) A r.v X has a gamma distribution with parameters , Find the mgf of X hence obtain the mean and variance of X

3.1 The Mgf of a Sum of Independent Random Variables

The mgf of the sum of n independent random variable is the product of their individual mgf's
The mean (variance) of the sum of n independent random variable is the sum of their individual means (variances).
The mgf about $X = a$ is given by $M_{x,a}(t) = E\left[e^{t(x-a)}\right] = e^{-at}E\left[e^{tx}\right] = e^{-at}M_x(t)$

4 NORMAL DISTRIBUTION

4.1 Introduction

The normal, or Gaussian, distribution is one of the most important distributions in probability theory. It is widely used in statistical inference. One reason for this is that sums of random variables often approximately follow a normal distribution.

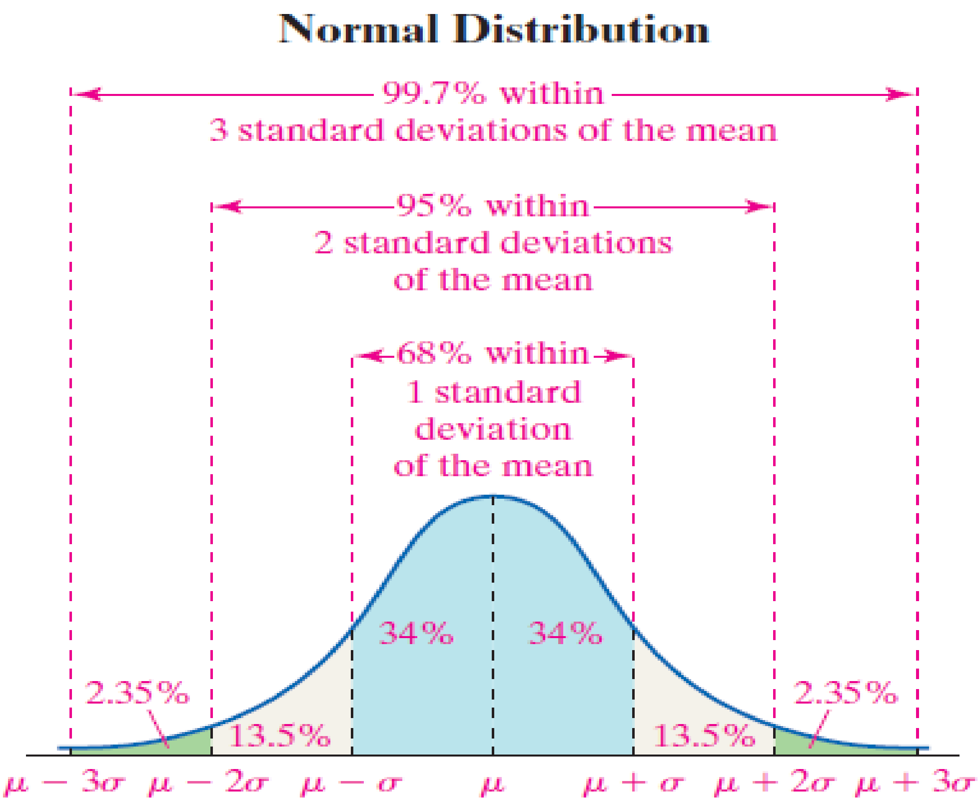
Definition A r.v X has a normal distribution with parameters μ and σ^2, abbreviated $X \sim N\left(\mu, \sigma^2\right)$ if it has probability density function

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}\exp\left\{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right\} \text{ for } -\infty < x < \infty \text{ and } \sigma > 0$$

Where μ is the mean and σ is the standard deviation.

4.1.1 Properties of normal distribution

1) The normal distribution curve is bell-shaped and symmetric, about the mean
2) The curve is asymptotic to the horizontal axis at the extremes.
3) The highest point on the normal curve is at the mean, which is also the median and mode.
4) The mean can be any numerical value: negative, zero, or positive
5) The standard deviation determines the width of the curve: larger values result in wider, flatter curves
6) Probabilities for the normal random variable are given by areas under the curve. The total area under the curve is 1 (0.5 to the left of the mean and 0.5 to the right).
7) It has inflection points at $\mu - \sigma$ and $\mu + \sigma$.
8) Empirical Rule:
   a) 68.26% of values of a normal random variable are within ± 1 standard deviation of its mean. ie $P\left(\mu - \sigma \le X \le \mu + \sigma\right) = 0.6826$
   b) 95.44% of values of a normal random variable are within ± 2 standard deviation of its mean. ie $P\left(\mu - 2\sigma \le X \le \mu + 2\sigma\right) = 0.9544$
   c) 99.72% of values of a normal random variable are within ± 3 standard deviation of its mean. ie $P\left(\mu - 3\sigma \le X \le \mu + 3\sigma\right) = 0.9972$



4.2 Standard Normal Probability Distribution

A random variable having a normal distribution with a mean of 0 and a varuance of 1 is said to have a **standard normal** probability distribution

Definition The random variable Z is said to have the standard normal distribution if $Z \sim N(0,1)$. Therefore, the density of Z, which is usually denoted $\phi(z)$ is given by;

$$\phi(z) = \frac{1}{\sqrt{2\pi}} \exp\left\{-\tfrac{1}{2} z^2\right\} \text{ for } -\infty < z < \infty$$

The cumulative distribution function of a standard normal random variable is denoted $\Phi(z)$, and is given by

$$\Phi(z) = \int_{-\infty}^{z} \phi(t)dt$$

4.2.1 Computing Normal Probabilities

It is very important to understand how the standardized normal distribution works, so we will spend some time here going over it. There is no simple analytic expression for $\Phi(z)$ in terms of elementary functions. but the values of $\Phi(z)$ has been exhaustively tabulated. This greatly simplifies the task of computing normal probabilities..

Table 1 below reports the cumulative normal probabilities for normally distributed variables in standardized form (i.e. Z-scores). That is, this table reports $P(Z \le z) = \Phi(z)$). For a given value of Z, the table reports what proportion of the distribution lies below that value. For example, $P(Z \le 0) = \Phi(0) = 0.5$; half the area of the standardized normal curve lies to the left of $Z = 0$.

Theorem: It may be useful to keep in mind that

i) $P(Z > z) = 1 - \Phi(z)$ complementary law

ii) $P(Z \le -z) = P(Z \ge z) = 1 - \Phi(z)$ ie due to symmetry

$$\Rightarrow \quad \Phi(z) + \Phi(-z) = 1 \text{ Since } P(Z \le z) + P(Z \ge z) = 1$$

iii) $P(a \le z \le b) = \Phi(b) - \Phi(a)$

iv) $P(-a \le z \le a) = 2\Phi(a) - 1$ since

$$P(-a \le z \le a) = \Phi(a) - \Phi(-a) = \Phi(a) - [1 - \Phi(a)] = 2\Phi(a) - 1$$

v) If we now make $\Phi(a)$ the subject, then $\Phi(a) = \frac{1}{2}[1 + P(-a \le z \le a)]$

Table 1

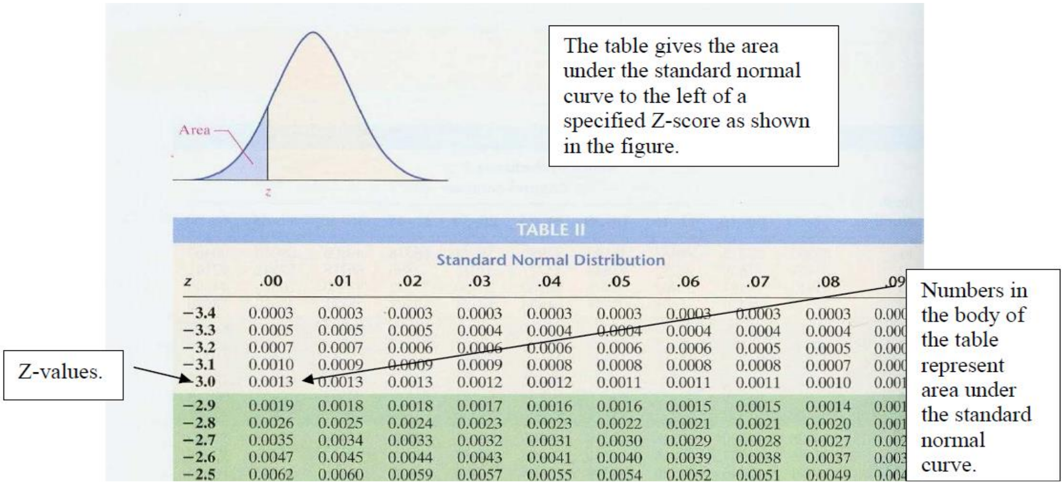


TABLE II

Standard Normal Distribution

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	0.000
−3.3	0.0005	0.0005	0.0005	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.000
−3.2	0.0007	0.0007	0.0006	0.0006	0.0006	0.0006	0.0006	0.0005	0.0005	0.000
−3.1	0.0010	0.0009	0.0009	0.0009	0.0008	0.0008	0.0008	0.0008	0.0007	0.000
−3.0	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.001
−2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.001
−2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.001
−2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.002
−2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.003
−2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.004

Example 1

Given $Z \sim N(0,1)$, find;

a) $P(Z \le z)$ if $z = 1.65, -1.65, 1.0, -1.0$

b) $P(Z > z)$ for $z = 1.02, -1.65$

c) $P(0.365 \le z \le 1.75)$

d) $P(-0.696 \le z \le 1.865)$

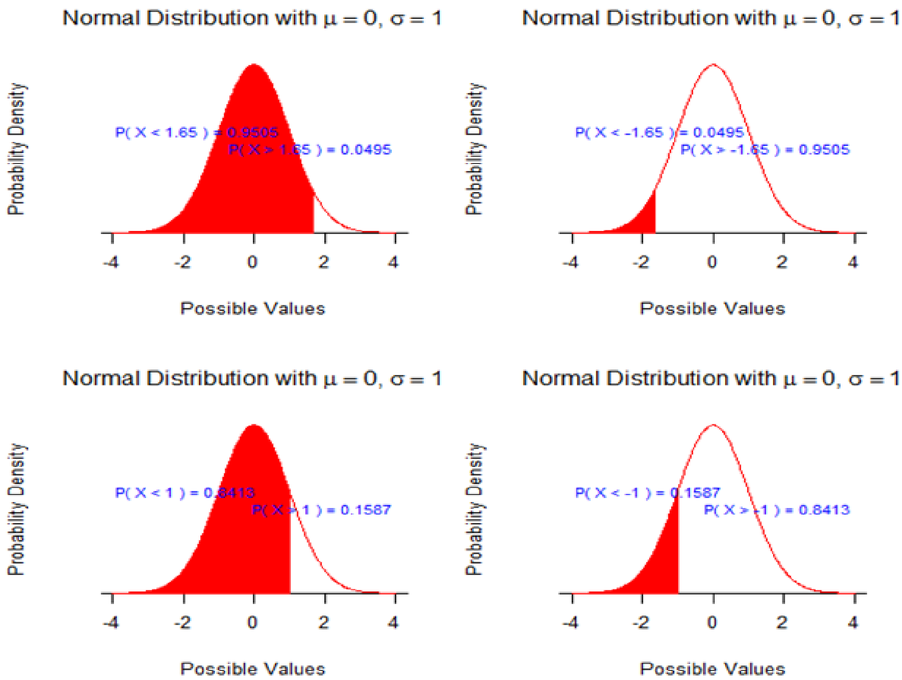
e) $P(-2.345 \le z \le -1.65)$

f) $P(|z| \le 1.43)$

Solution

a) Look up and report the value for $\Phi(z)$ from the standard normal probabilities table

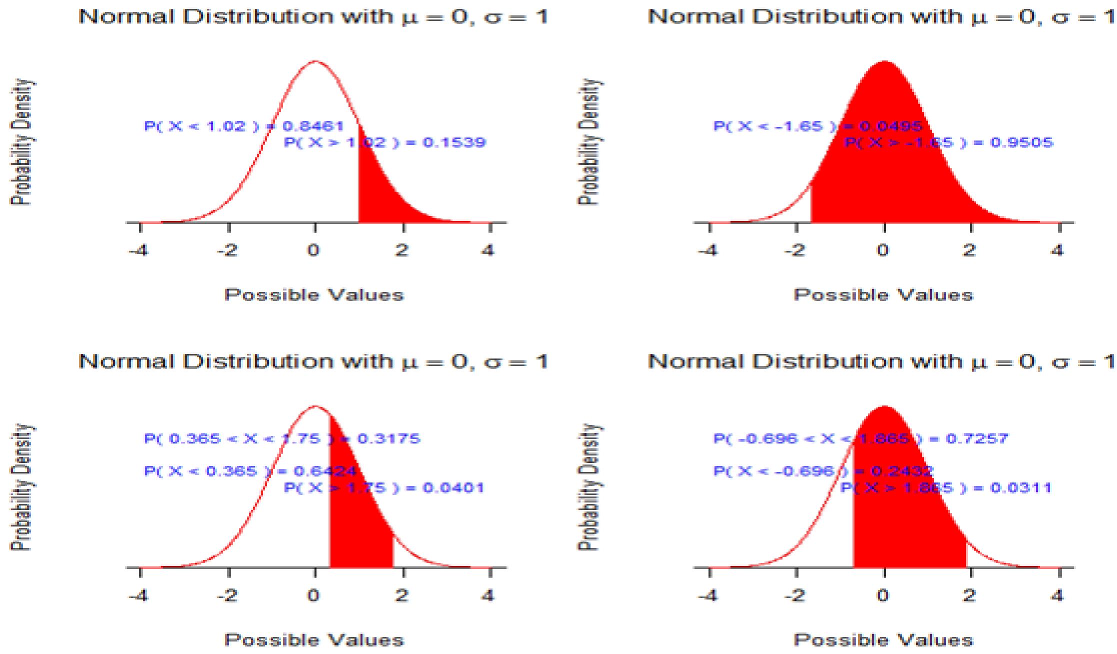
$$P(Z \le 1.65) = \Phi(1.65) = 0.9505 \quad \Phi(-1.65) = 0.0495 \quad \Phi(1.0) = 0.8413 \quad \Phi(-1.0) = 0.1587$$



b) $P(Z > z) = \Phi(-z)$ Thus $P(Z > 1.02) = \Phi(-1.02) = 0.1515 \quad P(Z > -1.65) = \Phi(1.65) = 0.9505$

c) $P(0.365 \le z \le 1.75) = \Phi(1.75) - \Phi(0.365) = 0.9599 - 0.6350 = 0.3249$

d) $P(-0.696 \le z \le 1.865) = \Phi(1.865) - \Phi(-0.696) = 0.9689 - 0.2432 = 0.3249 = 0.7257$



e) $P(-2.345 \le z \le -1.65) = \Phi(-1.65) - \Phi(-2.345) = 0.0505 - 0.0095 = 0.0410$

f) $P(|z| \le 1.43) = P(-1.43 \le z \le 1.43) = 2\Phi(1.43) - 1 = 2(0.9236) - 1 = 0.8472$

Example 2

If $Z \sim N(0,1)$, find the value of t for which;

a) $P(Z \le t) = = 0.6026, \quad 0.9750, \quad 0.3446$

b) $P(Z > t) == 0.4026, 0.7265, \ 0.5446$

c) $P(-0.28 \le z \le t) = 0.2665$

d) $P(-t \le z \le t) = 0.9972, \ 0.9505, \ 0.9750$

Solution

Here we find the probability value in Table I, and report the corresponding value for Z.

a) $\Phi(t) = 0.6026 \Rightarrow t = 0.26 \quad \Phi(t) = 0.950 \Rightarrow t = 1.96 \quad \Phi(t) = 0.3446 \Rightarrow t = -0.40$

b) $P(Z > t) = 0.4026 \Rightarrow \Phi(t) = 0.5974 \Rightarrow t = 0.25$

$P(Z > t) = 0.7265 \Rightarrow \Phi(t) = 0.2735 \Rightarrow t = -0.60$

$P(Z > t) = 0.5446 \Rightarrow \Phi(t) = 0.4554 \Rightarrow t = -0.11$

c) $P(-0.28 \le z \le t) = \Phi(t) - \Phi(-0.28) = 0.2665 \Rightarrow \Phi(t) = 0.3897 + 0.2665 \Rightarrow t = 0.40$

d) $P(-t \le z \le t) = 2\,\Phi(t) - 1 = 0.9972 \Rightarrow \Phi(t) = 0.9986 \Rightarrow t = 2.99$

$P(-t \le z \le t) = 2\,\Phi(t) - 1 = 0.9505 \Rightarrow \Phi(t) = 0.9753 \Rightarrow t = 1.96$

$P(-t \le z \le t) = 2\,\Phi(t) - 1 = 0.9750 \Rightarrow \Phi(t) = 0.9875 \Rightarrow t = 2.24$

Exercise

1..Given $Z \sim N(0,1)$, find;

a) $P(Z \le z)$ if
$z = 1.95, -1.89, 1.074, \ -1.53$

b) $P(Z > z)$ for $z = 1.72, -1.15$

c) $P(0 \le z \le 1.05)$

d) $P(-1.396 \le z \le 1.125)$

e) $P(-1.96 \le z \le -1.65)$

f) $P(|z| \le 2.33)$

2..If $Z \sim N(0,1)$, find the value of z for which;

a) $P(Z \le a) = = 0.973, \quad 0.6693, \quad 0.4634$

b) $P(Z > a) == 0.3719,\ 0.9545,\ 0.7546$

c) $P(-1.21 \le z \le t) = 0.6965$

d) $P(|z| \le t) = 0.9544,\ 0.9905,\ 0.3750$

4.3 The General Normal Density

Consider $Z \sim N(0,1)$ and let $X = \mu + \sigma Z$ for $\sigma > 0$. Then $X \sim N(\mu, \sigma^2)$ But we know that

$f(x) = \frac{1}{\sigma}\phi\left(\frac{X-\mu}{\sigma}\right)$ from which the claim follows. Conversely, if $X \sim N(\mu, \sigma^2)$, then $Z = \frac{X-\mu}{\sigma} \sim N(0,1)$. It is also easily shown that the cumulative distribution function satisfies

$$F(x) = \Phi\left(\frac{X-\mu}{\sigma}\right)$$

and so the cumulative probabilities for any normal random variable can be calculated using the tables for the standard normal distribution..

Definition A variable X is said to be standardized if it has been adjusted (or transformed) such that its mean equals 0 and its standard deviation equals 1. Standardization can be accomplished using the formula for a z-score: $Z = \frac{X-\mu}{\sigma} \sim N(0,1)$. The z-score represents the number of standard deviations that a data value is away fromthe m ean.

Let $X \sim N(\mu, \sigma^2)$ then $P(a \le X \le b) = P\left(\frac{a-\mu}{\sigma} \le Z \le \frac{b-\mu}{\sigma}\right) = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$ where $Z = \frac{X-\mu}{\sigma} \sim N(0,1)$

Example 1 A r.v $X \sim N(50, 25)$ compute $P(45 \le z \le 60)$

Solution

$\mu = 50$ and $\sigma = 5 \Rightarrow Z = \frac{x-50}{5} \sim N(0,1)$

$P(45 \le X \le 60) = P\left(\frac{45-50}{5} \le Z \le \frac{60-50}{5}\right) = \Phi(2) - \Phi(-1) = 0.9772 - 0.1587 = 0.8185$

Example 2 Suppose $X \sim N(30, 16)$. Find; a) $P(X < 40)$ b) $P(X > 21)$ c) $P(30 < X < 35)$

Solution

$X \sim N(30,16) \Rightarrow Z = \frac{X-30}{4} \sim N(0,1)$

a) $P(X \le 40) = P\left(Z \le \frac{40-30}{4}\right) = \Phi(2.5) = 0.9938$

b) $P(X > 21) = P\left(Z > \frac{21-30}{4}\right) = P(Z > -2.25) = P(Z \le 2.25) = \Phi(2.25) = 0.9878$

c) $P(30 < X < 35) = P\left(\frac{30-30}{4} \le Z \le \frac{35-30}{4}\right) = P(0 < Z < 1.25) = 0.8944 - 0.5 = 0.3944$

Example 3

The top 5% of applicants (as measured by GRE scores) will receive scholarships. If $GRE \sim N(500, 100^2)$, how high does your GRE score have to be to qualify for a scholarship?

Solution

Let $X = GRE$. We want to find x such that $P(X \ge x) = 0.05$ this is too hard to solve as it stands - so instead, compute $Z = \frac{X-500}{100} \sim N(0,1)$ and find z for the problem,

$$P(Z \ge z) = 1 - \Phi(z) = 0.05 \quad \Rightarrow \quad \Phi(z) = 0.95 \quad \Rightarrow \quad z = 1.645$$

To find the equivalent x, compute $X = \mu + \sigma Z \Rightarrow x = 500 + 100(1.645) = 66.5$

Thus, your GRE score needs to be 665 or higher to qualify for a scholarship.

Example 4

Family income is believed to be normally distributed with a mean of $25000 and a standard deviation on $10000. If the poverty level is $10,000, what percentage of the population lives in poverty? A new tax law is expected to benefit "middle income" families, those with incomes between $20,000 and $30,000. What percentage of the population will benefit from the law?

Solution

Let X = Family income. We want to find P(X ≤ $10,000)., so

$X \sim \mathrm{N}\left(25000, 10000^2\right) \Rightarrow \; Z = \frac{X-25000}{10000} \sim \mathrm{N}\left(0,1\right)$

$\mathrm{P}(X \le 10{,}000) = \mathrm{P}(Z \le -1.5) = \Phi(-1.5) = 0.0668$.

Hence, a slightly below 7% of the population lives in poverty.

$\mathrm{P}(20{,}000 \le X \le 30{,}000) = \mathrm{P}(-0.5 \le Z \le 0.5) = 2\Phi(0.5) - 1 = 2 \times 0.6915 - 1 - 0.383$

Thus, about 38% of the taxpayers will benefit from the new law.

Exercise

1) Suppose $X \sim \mathrm{N}(130, 25)$. Find; a) $P(X < 140)$ b) $P(X > 120)$ c) $P(130 < X < 135)$
2) The random variable X is normally distributed with mean 500 and standard deviation 100. Find; (i) $P(X < 400)$, (ii) $P(X > 620)$ (iii) the 90th percentile (iv) the lower and upper quartiles. Use graphs with labels to illustrate your answers.
3) A radar unit is used to measure speeds of cars on a motorway. The speeds are normally distributed with a mean of 90 km/hr and a standard deviation of 10 km/hr. What is the probability that a car picked at random is travelling at more than 100 km/hr?
4) For a certain type of computers, the length of time bewteen charges of the battery is normally distributed with a mean of 50 hours and a standard deviation of 15 hours. John owns one of these computers and wants to know the probability that the length of time will be between 50 and 70 hours
5) Entry to a certain University is determined by a national test. The scores on this test are normally distributed with a mean of 500 and a standard deviation of 100. Tom wants to be admitted to this university and he knows that he must score better than at least 70% of the students who took the test. Tom takes the test and scores 585. Will he be admitted to this university?
6) A large group of students took a test in Physics and the final grades have a mean of 70 and a standard deviation of 10. If we can approximate the distribution of these grades by a normal distribution, what percent of the student; (a) scored higher than 80? (b) should pass the test (grades≥60)? (c) should fail the test (grades<60)?
7) A machine produces bolts which are N(4 0.09) where measurements are in cm. Bolts are measured accurately and any bolt smaller than 3.5 cm or larger than 4.4 cm is rejected. Out of 500 bolts how many would be accepted? Ans 430
8) Suppose IQ ~ N(100,22.5).a woman wants to form an Egghead society which only admits people with the top 1% IQ score. What should she have to set the cut-off in the test to allow this to happen? Ans 134.9
9) A manufacturer does not know the mean and standard deviation of ball bearing he is producing. However a sieving system rejects all the bearings larger than 2.4 cm and those under 1.8 cm in diameter. Out of 1,000 ball bearings, 8% are rejected as too small and 5.5% as too big. What is the mean and standard deviation of the ball bearings produced? Ans mean=2.08 sigma=0.2

4.4 Normal Approximation to the Binomial Distribution.

4.4.1 Introduction

Suppose a fair coin is tossed 10 times, whar is the probability of observing: a) exactly 4 heads b) at most 4 heads?

Solution

Let X be the r.v the number of heads observed then $X \sim \text{Bin}(10, 0.5)$

$\Rightarrow \; P(X = x) = {}_{10}C_x \times (0.5)^x (0.5)^{10-x} = {}_{10}C_x \times \left(\tfrac{1}{2}\right)^{10}$ for x = 0,1,2,….,10

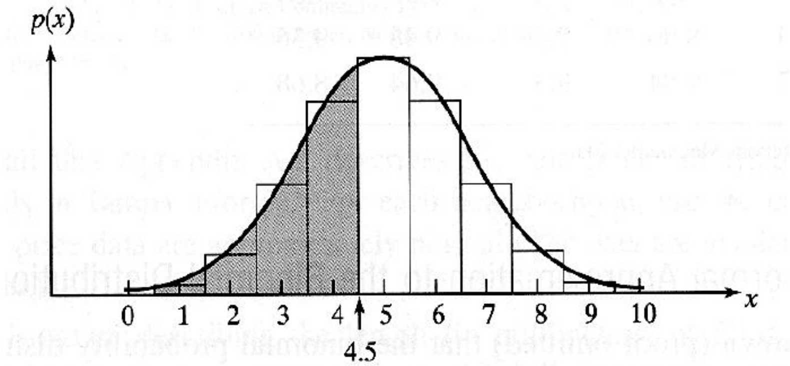
a) $\Rightarrow \; P(X = 4) = {}_{10}C_4 \times \left(\tfrac{1}{2}\right)^{10} = \dfrac{105}{512} \approx 0.2051$

b) $P(X \le 4) = \left[{}_{10}C_0 + {}_{10}C_1 + {}_{10}C_2 + {}_{10}C_3 + {}_{10}C_4\right](0.5)^{10} = \dfrac{193}{512} \approx 0.3770$

4.4.2 Normal approximation:

Many interesting problems can be addressed via the binomial distribution. However, for large n, it is sometimes difficult to directly compute probabilities for a binomial (*n, p*) random variable, X. **Eg**: Compute $P(X \le 12)$ for 25 tosses of a fair coin. Direct calculations can get cumbersome very quickly. . Fortunately, as n becomes large, the binomial distribution becomes more and more symmetric, and begins to converge to a normal distribution. That is, for a large enough n, a binomial variable X is approximately ~ N(*np, npq*). Hence, the normal distribution can be used to approximate the binomial distribution.

To get a feel for why this might work, let us draw the probability histogram for 10 tosses of a fair coin. The histogram looks bell-shaped, as long as the number of trials is not too small



In general, the distribution of a binomial random variable may be accurately approximated by that of a normal random variable, as long as $np \ge 5$ and $nq \ge 5$, and assuming that a .continuity correction. is made to account for the fact that we are using a continuous distribution (the normal) to approximate a discrete one (the binomial).

In approximating the distribution of a binomial random variable X, we will use the normal distribution with mean $\mu = np$ and variance $\sigma^2 = npq$ where $q = 1 - p$. Why are these reasonable choices of μ, σ2?

4.4.3 Continuity Correction

In the binomial, $P(X \le a) + P(X \ge a+1) = 1$ whenever a is an integer. But if we sum the area under the normal curve corresponding to $P(X \le a) + P(X \ge a+1)$, this area does not sum to 1 because the area from a to (a + 1) is missing.

The usual way to solve this problem is to associate 1/2 of the interval from a to a + 1 with each adjacent integer. The continuous approximation to the probability $P(X \le a)$ would thus be $P(X \le a + \frac{1}{2})$, while the continuous approximation to $P(X \ge a+1)$ would be $P(X \ge a + \frac{1}{2})$. This adjustment is called a continuity correction. More specifically,

$$\underbrace{P(X \le x) = P(X < x+1)}_{\text{Binomialdistribution}} \to \underbrace{P(X \le x+0.5)}_{\text{Normal approximation}} = \underbrace{P\left(z \le \tfrac{x+0.5-np}{\sqrt{npq}}\right)}_{\text{Standardd Normal approx}}$$

$$\underbrace{P(X \ge x) = P(X > x-1)}_{\text{Binomialdistribution}} \to \underbrace{P(X \ge x-0.5)}_{\text{Normal approximation}} = \underbrace{P\left(z \le \tfrac{x-0.5-np}{\sqrt{npq}}\right)}_{\text{Standardd Normal approx}}$$

$$\underbrace{P(a \le X \le b) = P(a-1 < X < b+1)}_{\text{Binomialdistribution}} \to = \underbrace{P(a-0.5 \le X \le b+0.5)}_{\text{Normal approximation}} = P\left(\tfrac{a-0.5-np}{\sqrt{npq}} \le z \le \tfrac{b+0.5-np}{\sqrt{npq}}\right)$$

$$\underbrace{P(X = x) = P(x-1 < X < x+1)}_{\text{Binomialdistribution}} \to \underbrace{P(x-0.5 \le X \le x+0.5)}_{\text{Normal approximation}} = \underbrace{P\left(\tfrac{x-0.5-np}{\sqrt{npq}} \le z \le \tfrac{x+0.5-np}{\sqrt{npq}}\right)}_{\text{Standardd Normal approx}}$$

NB: For the binomial distribution, the values to the right of each = sign are primarily included for illustrative purposes. The equalities which hold in the binomial distribution do not hold in the normal distribution, because there is a gap between consecutive values of a. The normal approximation deals with this by “splitting” the difference.
For example, in the binomial, $P(X \le 6) = P(X < 7)$ since 6 is the next possible value of X that is less than 7. In the normal, we approximate this by finding $P(X \le 6.5)$. And, in the binomial, $P(X \ge 6) = P(X > 5)$, because 6 is the next value of X that is greater than 5. In the normal, we approximate this by finding $P(X \ge 5.5)$

Returning to the case of coin tossing Suppose we wish to find $P(X \le 4)$, the probability that the binomial r.v is less than or equal to 4. In the diagram above, the bars represent the binomial distribution with n = 10, p = 0.5. The superimposed curve is a normal density f(x). The mean of the normal is $\mu = np = 5$ and the standard deviation is $\sigma = \sqrt{10(0.5)(0.5)} \approx 1.58$ Using the normal approximation, we need to calculate the probability that our normal r.v is less than or equals to 4.5. ie

$$\underbrace{P(X \le 4)}_{Binomial} = \underbrace{P(X \le 4.5)}_{Normal} = \underbrace{P\left(Z \le \tfrac{4.5-5}{1.58}\right)}_{std\ normal} = \Phi(-0.3162) = 0.3759$$ which is very close to the actual answer of 0.377

Example 1 Suppose 50% of the population approves of the job the governor is doing, and that 20 individuals are drawn at random from the population. Solve the following, using the normal approximation to the binomial. What is the probability that;.

a) exactly 7 people will support the governor?
b) at least 7 people will support the governor?
c) more than 11 people will support the governor?
d) 11 or fewer will support the governor?

Solution

Note that $n = 20,\ p = 0.5 \Rightarrow \mu = np = 10$ and $\sigma = \sqrt{npq} = \sqrt{5}$ Since $np \ge 5$ and $nq \ge 5$, it is probably safe to assume that $X \sim N(10,5)$

a) $$\underbrace{P(X = 7)}_{Binomial} = \underbrace{P(6.5 \le X \le 7.5)}_{Normal} = \underbrace{P\left(\tfrac{6.5-10}{\sqrt{5}} \le Z \le \tfrac{7.5-10}{\sqrt{5}}\right)}_{std\ normal} = \underbrace{P(-1.565 \le Z \le -1.118)}_{std\ normal}$$

$$= \Phi(-1.118) - \Phi(-1.565) = 0.1318 - 0.0588 = 0.0730$$

b) $\underbrace{P(X \le 7)}_{Binomial} = \underbrace{P(X \le 7.5)}_{Normal} = \underbrace{P(Z \le -1.118)}_{std\ normal} = 0.1318$

c) $\underbrace{P(X > 11)}_{Binomial} = \underbrace{P(X \ge 11.5)}_{Normal} = \underbrace{P(Z \ge 0.6708)}_{std\ normal} = 1 - \Phi(0.6708) = 1 - 0.7488 = 0.2512$

d) $\underbrace{P(X \le 11)}_{Binomial} = \underbrace{P(X \le 11.5)}_{Normal} = \underbrace{P(Z \le 0.6708)}_{std\ normal} = \Phi(0.6708) = 0.7488$

Example 2
In each of 25 races, the Democrats have a 60% chance of winning. What are the odds that the Democrats will win 19 or more races? Use the normal approximation to the binomial
Solution
Note that $n = 25,\ p = 0.6 \Rightarrow \mu = np = 15$ and $\sigma = \sqrt{npq} = \sqrt{6}$ Since $np > 5$ and $nq > 5$, it is probably safe to assume that $X \sim N(15, 6)$. .
Using the normal approximation to the binomial,
$\underbrace{P(X \ge 19)}_{Binomial} = \underbrace{P(X \ge 18.5)}_{Normal} = \underbrace{P(Z \ge 1.4289)}_{std\ normal} = 1 - \Phi(1.4289) = 1 - 0.9235 = 0.0765$
Hence, Democrats have a little less than an 8% chance of winning 19 or more races.

Example 3 Tomorrow morning Iberia flight to Madrid can seat 370 passengers. From past experience, Iberia knows that the probability is 0.90 that a given ticket-holder will show up for the flight. They have sold 400 tickets, deliberately overbooking the flight. How confident can Iberia be that no passenger will need to be .bumped. (denied boarding)?
Solution:
We will assume that the number (X) of passengers showing up for the flight has a binomial distribution with mean $\mu = 400 \times 0.9 = 360$ and standard deviation $\sigma = \sqrt{400 \times 0.9 \times 0.1} = 6$
We want $\underbrace{P(X \le 370)}_{Binomial} = \underbrace{P(X \le 370.5)}_{Normal} = \underbrace{P(Z \le 1.75)}_{std\ normal} = 0.9599$ So the probability that nobody gets bumped is approximately 0.9599. (Almost 96%).

Exercise (Use the normal approximation to the binomial in the entire exercise)
1. A coin is loaded such that heads is thrice as likely as the tails. Find the probability of observing between 4 and 7 heads inclusive with 12 tosses of the coin.
2. Based upon past experience, 40% of all customers at Miller's Automotive Service Station pay for their purchases with a credit card. If a random sample of 200 customers is selected, what is the *approximate* probability that;
   a) at least 75 pay with a credit card?
   b) not more than 70 pay with a credit card?
   c) between 70 and 75 customers, inclusive, pay with a credit card?
3. The probability that Ronado scores a goal in any game against a tough opponent in soccer is 0.3. What is the probability that he scores 30 goals in the next 100 games in which he plays?
4. Crafty Computers limited produces PCs. The probability that one of their computers has a virus is 0.25. JKUAT ICSIT buys 300 computers from the company. What is the probability that between 70 and 80 PCs inclusive have a virus? Would you advice the director JKUAT ICSIT to buy Computers from this company in future?
5. For overseas flights, an airline has three different choices on its dessert menu—ice cream, apple pie, and chocolate cake.Based on past experience the airline feels that each dessert is equally likely to be chosen.

a) If a random sample of four passengers is selected, what is the probability that at least two will choose ice cream for dessert
b) If a random sample of 21 passengers is selected, what is the *approximate* probability that at least eight will choose ice cream for dessert?

6. In a family of 11 children, what is the probability that there will be more boys than girls?.
7. A baseball player has a long term batting average of 0.300. What is the chance he gets an average of 0.330 or higher in his next 100 bats?
8. Suppose we draw a Simple Random Sample of 1,500 Americans and want to assess whether the representation of blacks in the sample is accurate. We know that about 12% of Americans are black, what is the probability that the sample contains 170 or fewer blacks?
9. Let T be the lifetime in years of new bus engines. Suppose that T is continuous with probability density function $f(t)=\begin{cases} 0 \text{ for } t<1 \\ {}^{d}\!/\!_{x^3} \text{ for } t>1 \end{cases}$ for some constant d.
   a) Find the value of d and the mean and median of T.
   b) Suppose that 240 new bus engines are installed at the same time, and that their lifetimes are independent. By using a normal approximation to the bonomial, find the probability that at most 10 of the engines last for 4 years or more.

4.5 Sums of Independent Random Variables

Theorem 1 If $X_1, X_2, ..., X_n$ are n independent continuous r.v each with mean μ and variance $\sigma^2 < \infty$ then r.v $W = \frac{1}{n}\sum_{i=1}^{n} x_i$ has mean μ and variance $\frac{\sigma^2}{n}$

Theorem 2 If $X_1 \sim N(\mu_1, \sigma_1^2)$ and $X_2 \sim N(\mu_2, \sigma_2^2)$ are 2 independent r.v, then $X_1 \pm X_2 \sim N(\mu_1 \pm \mu_2\ ,\ \sigma_1^2 + \sigma_2^2)$

Proof

Let the mgf of X_1 be $M_1(t) = e^{t\mu_1 + \frac{1}{2}t^2\sigma_1^2}$ and for X_2 be $M_2(t) = e^{t\mu_2 + \frac{1}{2}t^2\sigma_2^2}$. Let $Y_1 = X_1 + X_2$ so Y_1 has mgf $M_1(t) = M_1(t) \times M_2(t) = e^{t\mu_1 + \frac{1}{2}t^2\sigma_1^2} \times e^{t\mu_2 + \frac{1}{2}t^2\sigma_2^2} = e^{t(\mu_1+\mu_2) + \frac{1}{2}t^2(\sigma_1^2+\sigma_2^2)}$ which is the mgf of a normal r.v with mean $\mu_1 + \mu_2$ and variance $\sigma_1^2 + \sigma_2^2$

Note $Y_2 = -X_2 \sim N(-\mu_2, +\sigma_2^2) \Rightarrow X_1 - X_2 \sim N(\mu_1 - \mu_2\ ,\ \sigma_1^2 + \sigma_2^2)$

Theorem 3 If $X_1, X_2, ...X_n$ are n independent r.vs and each $X_i \sim N(\mu_i, \sigma_i^2)$, then the r.v $X_1 \pm X_2 \pm ..... \pm X_n \sim N(\mu_1 \pm \mu_2 \pm ..... \pm \mu_n\ ,\ \sigma_1^2 + \sigma_2^2 + ... + \sigma_n^2)$

Proof is by induction (left as exercise to the learner)

Example If $X \sim N(60,16)$ and $Y \sim N(70,9)$ are 2 independent r.v, Find (a) $P(X+Y \le 140)$ (b) $P(120 \le X+Y \le 135)$ (c) $P(Y-X > 7)$ (d) $P(2 \le Y-X \le 12)$

Solution

$X+Y \sim N(130,25)$ and $Y-X \sim N(10,25)$ therefore

a) $P(X+Y \le 140) = P(Z \le \frac{140-130}{5}) = \Phi(2) = 0.9772$

b) $P(120 \le X+Y \le 135) = P(\frac{120-130}{5} \le Z \le \frac{135-130}{5}) = \Phi(1) - \Phi(-2) = 0.8413 - 0.0228 = 0.8185$

c) $P(Y-X > 7) = P(Z > \frac{7-10}{5}) = P(Z > -0.6) = P(Z \le 0.6) = \Phi(0.6) =\ 0.7257$

d) $P(2 \le Y-X \le 12) = P(\frac{2-10}{5} \le Z \le \frac{12-10}{5}) = \Phi(0.4) - \Phi(-1.6) = 0.6554 - 0.0548 = 0.6006$

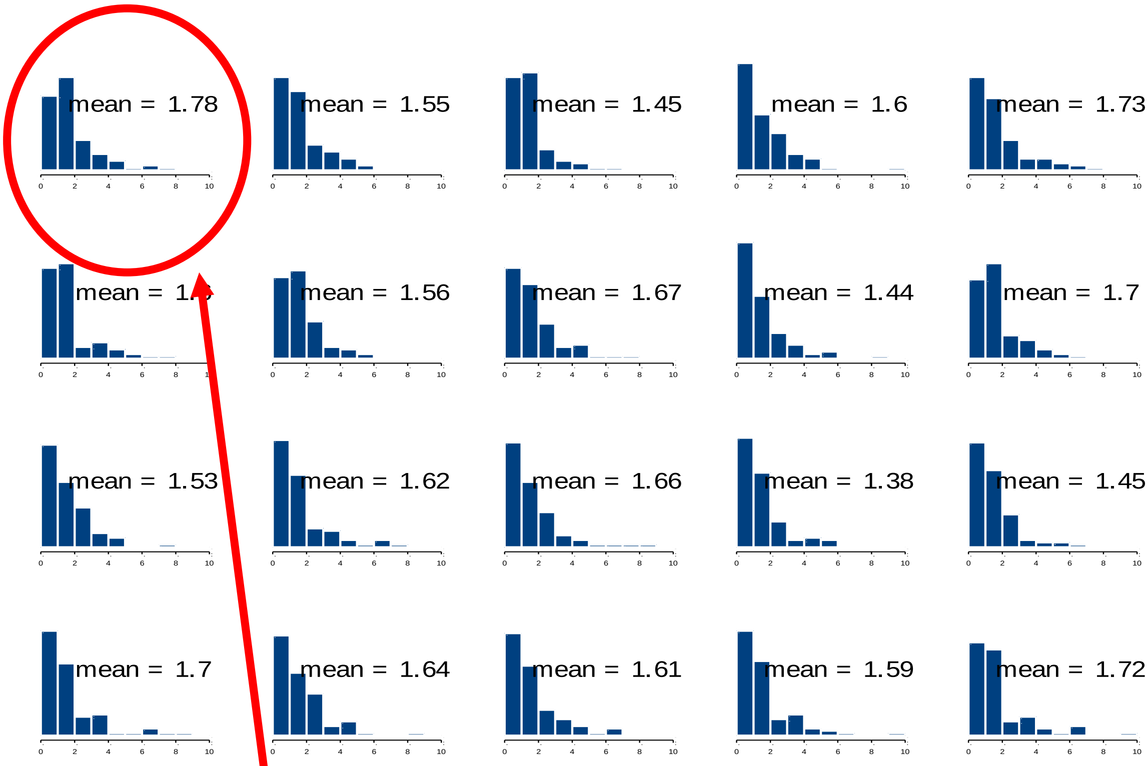
Exercise

1. If $X \sim \mathrm{N}(65, 28)$ and $Y \sim \mathrm{N}(85, 36)$ are 2 independent r.v, Find (a) $P(X+Y \le 142)$ (b) $P(134 \le X+Y \le 166)$ (c) $P(Y-X>4)$ (d) $P(12 \le Y-X \le 24)$
2. Each day Mr. Njoroge walks to the library bto read a newspaper. Total time spent walking is normally distributed with mean 15 minutes and standard deviation 2 minutes. Total time spent in the library is also normally distributed with mean 25 minutes and standard deviation $\sqrt{12}$ minutes. Find the probability that on one day;
   a) he is away from his home for more than 45 minutes.
   b) he spends more time walking than in the library

4.6 Sampling Distributions

In many investigations the data of interest can take on many possible values and it is often of interest to estimate the population mean, μ. A common estimator for μ is the sample mean $\bar{x}$. Consider the following set up: We observe a sample of size *n* from some population and compute the mean $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ .Since the particular individuals included in our sample are *random*, we would observe a different value of $\bar{x}$ if we repeated the procedure. That is, $\bar{x}$ *is also a random quantity*. Its value is determined partly by which people are randomly chosen to be in the sample. If we repeatedly drew samples of size *n* and calculated $\bar{x}$, we could ascertain the sampling distribution of $\bar{x}$.

Many possible samples, many possible $\bar{x}$'s



We only see one!

We will have a better idea of how good our one estimate is if we have good knowledge of how $\bar{x}$ behaves; that is, if we know the probability distribution of $\bar{x}$.

4.6.1 Properties of the Sampling Distribution of the Sample Means (Summary)

When all of the possible sample means are computed, then the following properties are true:

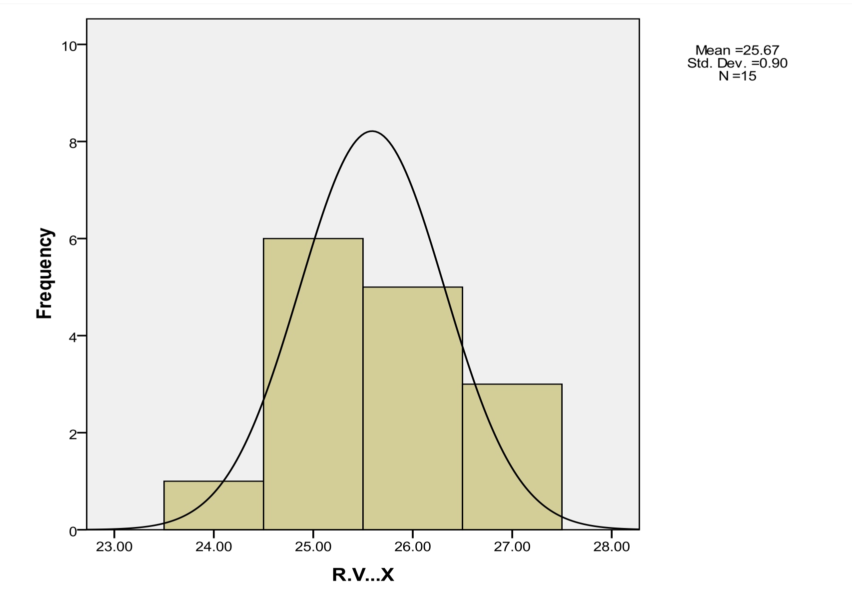
1. The mean of the sample means will be the mean of the population
2. The variance of the sample means will be the variance of the population divided by the sample size.
3. The standard deviation of the sample means (known as the standard error of the mean) will be smaller than the population mean and will be equal to the standard deviation of the population divided by the square root of the sample size.
4. If the population has a normal distribution, then the sample means will have a normal distribution.
5. If the population is not normally distributed, but the sample size is sufficiently large, then the sample means will have an approximately normal distribution. Some books define sufficiently large as at least 30 and others as at leas t25.

Example

The law firm of Hoya and Associates has six partners (A, B, C, D, E, F). At their weekly partners meeting each reported the number of hours they charged clients for their services last week. A 24, B 26, C 28, D 26, E 24. F26 (eg, Mr. E charged 24 hrs)
If n=2, (ie two partners) are selected randomly, how many different samples are possible?
This is the combination of 6 objects taken 2 at a time. Ie there are ${}^{6}c_2=15$ possible samples.
15 sample means are given below: (e.g. if the sample has A and B, sample mean is 24)
AB 25, AC 26, AD 25, AE 24, AF 25, BC 27, BD 26, BE 25, BF 26, CD 27, CE 26, CF 27, DE 25, DF 26, EF 25 putting this in a frequency table we get

$\bar{x}$	24	25	26	27
f	1	6	5	3

This is almost the sampling distribution of means. If we divide individual frequencies by total frequency (ie 15) we get "relative frequency" or probability. These probabilities add up to one, so we have a prob. distribution. The above information says that the probability that sample mean is 24 is 1 out of 15 or 0.066667. Now draw a histogram for sampling distribution of means and fit a normal curve on it.



Note the shape is similar to Normal distribution

The sampling distribution is simply this probability distribution defined over all possible samples of size n from the population of size N. In the real world problems N will be large (e.g. 200 million US population) and n will be also be large (e.g., 1000 people surveyed) and ${}^{N}c_{n}$ will be astronomical number. Then the sampling distribution can only be imagined. We have chosen a simple example of N=6, n=2 so that the entire sampling distribution can be explicitly computed and visualized. Now the random variable is $\bar{x}$, it is no longer just X.

Definitions
Central Limit Theorem:- Stats that as the sample size increases, the sampling distribution of the sample means will become approximately normally distributed.
Sampling Distribution of the Sample Means:- Distribution obtained by using the means computed from random samples of a specific size.
Sampling Error :- Difference which occurs between the sample statistic and the population parameter due to the fact that the sample isn't a perfect representation of the population.
Standard Error or the Mean:- The standard deviation of the sampling distribution of the sample means. It is equal to the standard deviation of the population divided by the square root of the sample size.

4.6.2 The Mean and Standard Deviation of $\bar{x}$

What are the mean and standard deviation of $\bar{x}$?
Let's be more specific about what we mean by a sample of size *n*. We consider the sample to be a collection of *n independent and identically distributed (*or *iid)*random variables $X_1, X_2, ..., X_n$ with common mean μ and common standard deviation σ.

Thus, $E(\overline{X}) = E\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right) = \frac{1}{n}\sum_{i=1}^{n} E(x_i) = \frac{1}{n}\sum_{i=1}^{n}\mu = \frac{1}{n}(n\mu) = \mu$

$$Var(\overline{X}) = Var\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right) = \frac{1}{n^2}\sum_{i=1}^{n} Var(x_i) = \frac{1}{n^2}\sum_{i=1}^{n}\sigma^2 = \frac{1}{n^2}(n\sigma^2) = \frac{\sigma^2}{n} \quad \Rightarrow SD(\overline{X}) = \frac{\sigma}{\sqrt{n}}$$

4.6.3 The Central Limit Theorem

Now we know that $\bar{x}$ has mean μ and standard deviation $\sigma/\sqrt{n}$, but what is its distribution?
If $X_1, X_2, ..., X_n$ are *normally distributed*, then $\bar{x}$ is also normally distributed. Thus, $X_i \sim N(\mu, \sigma^2) \Rightarrow \overline{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$. If $X_1, X_2, ..., X_n$ are *not* normally distributed, then the Central Limit Theorem tells us that $\bar{x}$ is *approximately N*ormal.
In brief if $X_1, X_2, ..., X_n$ are *iid* random variables with mean μ and finite standard deviation σ. Then for a sufficiently large n, the sampling distribution of $\overline{X}$ is approximately Normal with mean μ and variance $\frac{\sigma^2}{n}$.

Remarks

- Central limit theorem involves two different distributions: the distribution of the original population and the distribution of the sample means
- The formula for a z-score when working with the sample means is: $z = \dfrac{\bar{x} - \mu}{\sigma/\sqrt{n}} \sim$ N(0,1)

Example

Intelligence Quotient (IQ) is normally distributed with mean 110 and standard deviation of 10. A moron is a person with IQ less than 80. Find the probability that a randomly chosen person is a moron. Let idiot be defined as one with an IQ less than 90. Find the probability that a randomly chosen person is an idiot. (Hint this random variable is for a single person X) If a sample of 25 students is available, what is the probability that the **average** IQ exceeds 105? What is the probability that the average IQ exceeds 115 (Hint this random variable is for an average over 25 persons or $\overline{X}$)

Solution

$IQ = X \sim \mathrm{N}(110_1, 10^2)$, and therefore for a sample of 25 people average $IQ = \overline{X} \sim \mathrm{N}(110_1, 4)$

The probability that a randomly chosen person is a moron is given by

$P(X < 80) = P\left(Z < \frac{80-110}{10}\right) = \Phi(-3) = 0.0013$

The probability that a randomly chosen person is an idiot is given by

$P(X < 90) = P\left(Z < \frac{90-110}{10}\right) = \Phi(-2) = 0.0228$

The probability that the **average** IQ exceeds 105 is $P(\overline{X} > 105)$

The random variable under consideration here is the average. Hence, a sampling distribution is relevant when we consider average IQ as the variable of interest, not he IQ of an individual student, but the average over 25 students. Standard deviation of the sampling distribution = Standard Error SE = $\sigma/\sqrt{n} = 10/5 = 2$. Now

$P(\overline{X} > 105) = P(Z > \frac{105-110}{2}) = \underbrace{P(Z \le 2.5)}_{due\,to\,symmetry} = \Phi(2.5) = 0.9938$

We now find probability that the average IQ exceeds 115 ie

$P(\overline{X} > 115) = P(Z > \frac{115-110}{2}) = \underbrace{P(Z \le -2.5)}_{due\,to\,symmetry} = \Phi(-2.5) = 0.0062$

Exercise

1) The annual salaries of employees in a large company are approximately normally distributed with a mean of \$50,000 and a standard deviation of \$10,000. If a random sample of 50 employees is taken, what is the probability that their average salary is;
   a) less than \$45,000?
   b) between \$45,000 and \$65,000?
   c) more than \$70,000
2) Library usually has 13% of its books checked out. Find the probability that in a sample of 588 books greater than 14% are checked out. ANS= 0.2358
3) The length of similar components produced by a company are approximated by a normal distribution model with a mean of 5 cm and a standard deviation of 0.02 cm.
   a) If a component is chosen at random what is the probability that the length of this component is between 4.98 and 5.02 cm?
   b) what is the probability that the average length of of a sample of 25 component is between 4.96 and 5.04 cm?
4) The length of life of an instrument produced by a machine has a normal distribution with a mean of 12 months and standard deviation of 2 months. Find the probability that in a random sample of 4 instrument produced by this machine, the average length of life

   a) less than 10.5 months. b) between 11and 13 months.

5) The time taken to assemble a car in a certain plant is a random variable having a normal distribution of 20 hours and a standard deviation of 2 hours. What is the probability that a car can be assembled at this plant in a period of time
   a) less than 19.5 hours? b) between 20 and 22 hours?

5 STATISTICAL INFERENCES

5.1 Introduction

In research, one always has some fixed ideas about cetain population parameters based on say, prior experiments, surveys or experience. However, these are only ideas. There is therefore a need to ascertain whether these ideas /claims are correct or not.
The ascertaining of claims is done by first collecting information in the form of sample data. We then decide whether our sample observations (statistic) have come from a postulated population or not.

Definitions
A **hypothesis** is a claim (assumption) about a population parameters such as the population mean, the population proportion or the population standard deviation is a postulated or a stipulated value of a parameter
Example: The mean monthly cell phone bill in this city is μ = \$42
The proportion of adults in this city with cell phones is π = 0.68
On the basis of observation data, one then performs a test to decide whether the postulated hypothesis should be accepted or not. However, we note that the decision aspect is prone to error/risk.

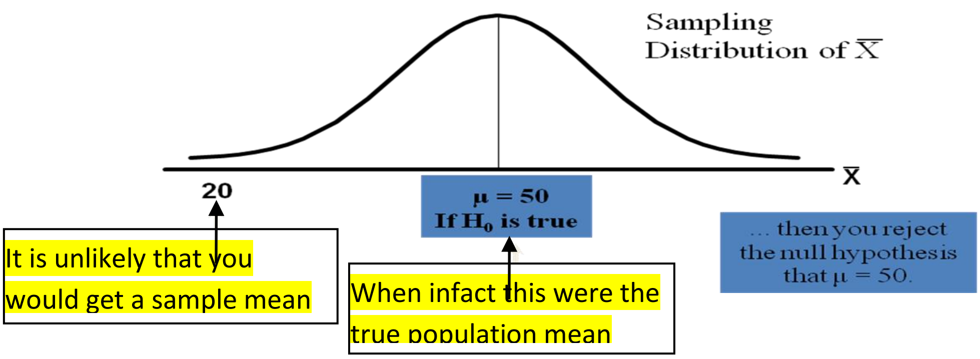
Null Hypothesis (denoted H_0): Statement of zero or no change and is the hypothesis which is to be actually tested for acceptance or rejection. If the original claim includes equality (<=, =, or >=), it is the null hypothesis. If the original claim does not include equality (<, not equal, >) then the null hypothesis is the complement of the original claim. The null hypothesis *always* includes the equal sign. The decision is based on the null hypothesis.
Eg: The average number of TV sets in U.S. Homes is equal to three ($H_0 : \mu = 3$)
It's always about a population parameter, and not about a sample statistic
Ie $H_0 : \mu = 3$ but **NOT** $H_0 : \bar{x} = 3$
We begin with the assumption that the null hypothesis is true
- Similar to the notion of innocent until proven guilty

Alternative Hypothesis (denoted H_1 or H_a): Statement which is true if the null hypothesis is false. it Challenges the status quo. It Is generally the hypothesis that the researcher is trying to prove and it is accepted when H0 is rejected and vice versa. The type of test (left, right, or two-tail) is based on the alternative hypothesis.

5.2 The Hypothesis Testing Process

Claim: The population mean age is 50. Ie Hypothesis H0: μ = 50, vs H_1: $\mu \neq 50$
Sample the population and find sample mea. Suppose the sample mean age was $\bar{x} = 20$. This is significantly lower than the claimed mean population age of 50. If the null hypothesis were

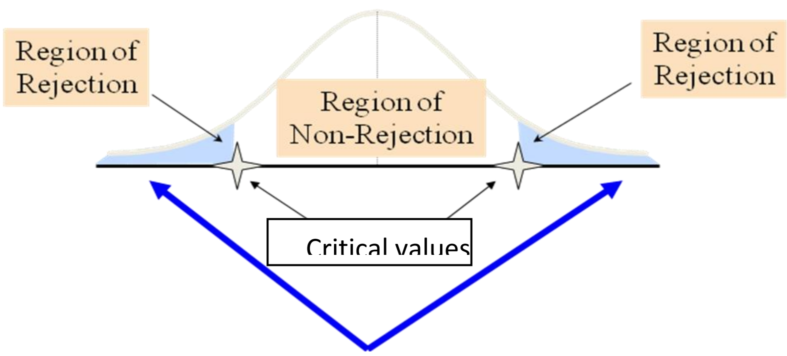
true, the probability of getting such a different sample mean would be very small, so you reject the null hypothesis .In other words, getting a sample mean of 20 is so unlikely if the population mean was 50, you conclude that the population mean must not be 50.



- If the sample mean is close to the assumed population mean, the null hypothesis is not rejected.
- If the sample mean is far from the assumed population mean, the null hypothesis is rejected.

How far is "far enough" to reject H_0? The critical value of a test statistic creates a "line in the sand" for decision making -- it answers the question of how far is far enough.

Sampling Distribution of the test statistic



"Too Far Away" From Mean of Sampling Distribution

5.2.1 Possible Errors in Hypothesis Test Decision Making

When taking a decision about the acceptance or rejection of a null hypothesis/ alternative hypothesis, there is a risk of committing an error. These errors are of two types:

Type I error; Mistake of rejecting the null hypothesis when it is true (saying false when true). It is usually the more serious error.

The probability of a Type I Error is (denoted α) is Called the level of significance of the test and it is Set by researcher in advance. α= 0.05 and α = 0.01 are common. If no level of significance is given, use α = 0.05. The level of significance is the complement of the level of confidence in estimation.

Type II error: Mistake of failing to reject the null hypothesis when it is false (saying true when false). The probability of a Type II Error is denoted by $\boldsymbol{\beta}$

Remarks

1) The confidence coefficient ($1-\alpha$) is the probability of not rejecting H_0 when it is true.
2) The confidence level of a hypothesis test is 100 ($1-\alpha$)%.
3) The power of a statistical test ($1-\beta$) is the probability of rejecting H_0 when it is false

Possible Hypothesis Test Outcomes		
	Actual Situation	
Decision	H_0 True	H_0 False
Do Not Reject H_0	No Error Probability $1 - \alpha$	Type II Error Probability β
Reject H_0	Type I Error Probability α	No Error Probability $1 - \beta$

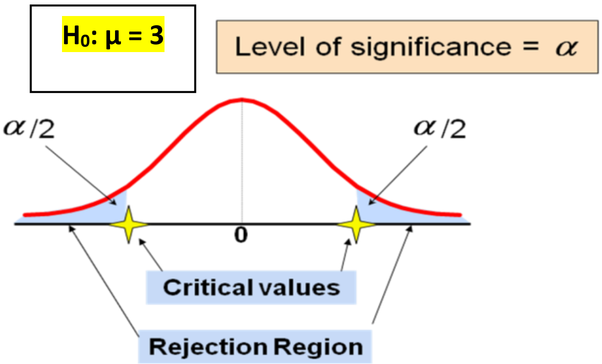
5.2.1 Relationship between Type I & Type II Error

Type I and Type II errors cannot happen at the same time
- A Type I error can only occur if H_0 is true
- A Type II error can only occur if H_0 is false

If Type I error probability (α) increases, then Type II error probability (β) decreases

5.2.3 Level of Significance and the Rejection Region

Critical region: Set of all values which would cause us to reject H_0
Critical value(s): The value(s) which separate the critical region from the non-critical region. The critical values are determined independently of the sample statistics.



This is a two-tail test because there is a rejection region in both tails

Test statistic: Sample statistic used to decide whether to reject or fail to reject the null hypothesis
Probability Value (P-value): The probability of getting the results obtained if the null hypothesis is true. If this probability is too small (smaller than the level of significance), then we reject the null hypothesis. If the level of significance is the area beyond the critical values, then the probability value is the area beyond the test statistic.
Decision: A statement based upon the null hypothesis. It is either "reject the null hypothesis" or "fail to reject the null hypothesis". We will never accept the null hypothesis.

Conclusion: A statement which indicates the level of evidence (sufficient or insufficient), at what level of significance, and whether the original claim is rejected (null) or supported (alternative).

5.2.4 Steps in Hypothesis Testing

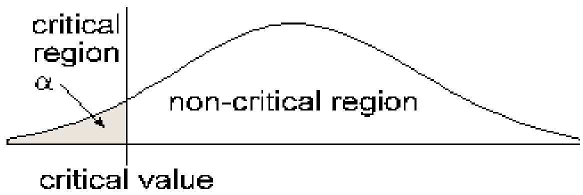
Any hypothesis testing is done under the assumption that the null hypothesis is true.
Here are the steps to performing hypothesis testing

a) Write the null and alternative hypothesis.
b) Use the alternative hypothesis to identify the type of test.
c) specify the level of significance, α and find the critical value using the tables
d) Compute the test statistic
e) Make a decision to reject or fail to reject the null hypothesis.
f) Write the conclusion

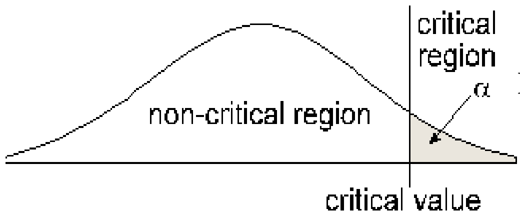
Remarks

The first thing to do when given a claim is to write the claim mathematically (if possible), and decide whether the given claim is the null or alternative hypothesis. If the given claim contains equality, or a statement of no change from the given or accepted condition, then it is the null hypothesis, otherwise, if it represents change, it is the alternative hypothesis.
The type of test is determined by the *Alternative Hypothesis* (H_1)

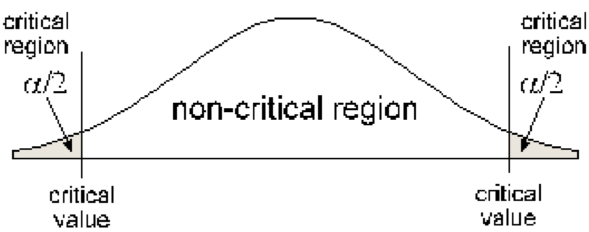
Left Tailed Test
H_1: parameter < value
Notice the inequality points to the left
Decision Rule: Reject H_0 if t.s. < c.v.



Right Tailed Test
H_1: parameter > value
Notice the inequality points to the right
Decision Rule: Reject H_0 if t.s. > c.v.



Two Tailed Test
H_1: parameter **not equal** to a value
Notice the inequality points to both sides
Decision Rule: Reject H_0 if t.s. < c.v. (left) or t.s. > c.v. (right)



If the test statistic falls into the non rejection region, do not reject the null hypothesis H_0. If the test statistic falls into the rejection region, reject the null hypothesis. Express the managerial conclusion in the context of the problem
Conclusions are sentence answers which include whether there is enough evidence or not (based on the decision) and whether the original claim is supported or rejected. Conclusions are based on the original claim, which may be the null or alternative hypotheses.

5.3 Approaches to Hypothesis Testing

There are three approaches to hypothesis testing namely Classical Approach, p vale approach and the confidence interval approach

5.3.1 The Classical Approach

The Classical Approach to hypothesis testing is to compare a test statistic and a critical value. It is best used for distributions which give areas and require you to look up the critical value (like the Student's t distribution) rather than distributions which have you look up a test statistic to find an area (like the normal distribution).

The Classical Approach also has three different decision rules, depending on whether it is a left tail, right tail, or two tail test.

One problem with the Classical Approach is that if a different level of significance is desired, a different critical value must be read from the table.

5.3.2 P-Value Approach

The P-Value Approach, short for Probability Value, approaches hypothesis testing from a different manner. Instead of comparing z-scores or t-scores as in the classical approach, you're comparing probabilities, or areas.

The level of significance (alpha) is the area in the critical region. That is, the area in the tails to the right or left of the critical values.

The p-value is the area to the right or left of the test statistic. If it is a two tail test, then look up the probability in one tail and double it.

If the test statistic is in the critical region, then the p-value will be less than the level of significance. It does not matter whether it is a left tail, right tail, or two tail test. This rule always holds.

Reject the null hypothesis if the p-value is less than the level of significance.

You will fail to reject the null hypothesis if the p-value is greater than or equal to the level of significance.

The p-value approach is best suited for the normal distribution when doing calculations by hand. However, many statistical packages will give the p-value but not the critical value. This is because it is easier for a computer or calculator to find the probability than it is to find the critical value.

Another benefit of the p-value is that the statistician immediately knows at what level the testing becomes significant. That is, a p-value of 0.06 would be rejected at an 0.10 level of significance, but it would fail to reject at an 0.05 level of significance. Warning: Do not decide on the level of significance after calculating the test statistic and finding the p-value.

Here are a couple of statements to help you keep the level of significance the probability value straight.

The Level of Significance is pre-determined before taking the sample. It **does not** depend on the sample at all. It is the area in the critical region,that is the area beyond the **c**ritical values. It is the probability at which we consider something unusua**l**.

The Probability-Value can only be found after taking the sample. It depends on the sample. It is the area beyond the test statistic**.** It is the probability of getting the results we obtained if the null hypothesis is true.

5.3.3 Confidence Intervals as Hypothesis Tests

Using the confidence interval to perform a hypothesis test only works with a two-tailed test.

a) If the hypothesized value of the parameter lies within the confidence interval with a 1-alpha level of confidence, then the decision at an alpha level of significance is to fail to reject the null hypothesis.

b) If the hypothesized value of the parameter lies outside the confidence interval with a 1-alpha level of confidence, then the decision at an alpha level of significance is to reject the null hypothesis.

However, it has a couple of problems.

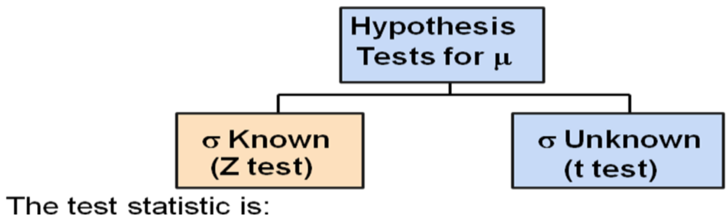
- It only works with two-tail hypothesis tests.
- It requires that you compute the confidence interval first. This involves taking a z-score or t-score and converting it into an x-score, which is more difficult than standardizing an x-score.

5.4 Testing a Single Mean

The value for all population parameters in the test statistics come from the null hypothesis. This is true not only for means, but all of the testing we're going to be doing.

The following hypotheses are to be tested: $H_0 : \mu = \mu_0$ vs $H_1 : \mu \neq \mu_0$ 0r $H_0 : \mu > \mu_0$ 0r $H_0 : \mu < \mu_0$ Where μ_0 is some hypothesised value.

The statistic and the critical values depends on whether **σ**, is known or unknown.



5.4.1 Population Standard Deviation Known

If the population standard deviation **σ**, is known, then the population mean has a normal distribution, and you will be using the z-score formula for sample means. The test statistic is the standard formula you've seen before. $z = \dfrac{\bar{x} - \mu}{\sigma / \sqrt{n}}$

The critical value is obtained from the normal table.

Example Test at 5% level the claim that the true mean # of TV sets in US homes is equal to 3. Suppose the sample results are n = 100, $\bar{x} = 2.84$ (σ = 0.8 is assumed known**)**

Solution

State the appropriate null and alternative hypotheses

- H_0: μ = 3 H_1: μ ≠ 3 (This is a two-tail test)

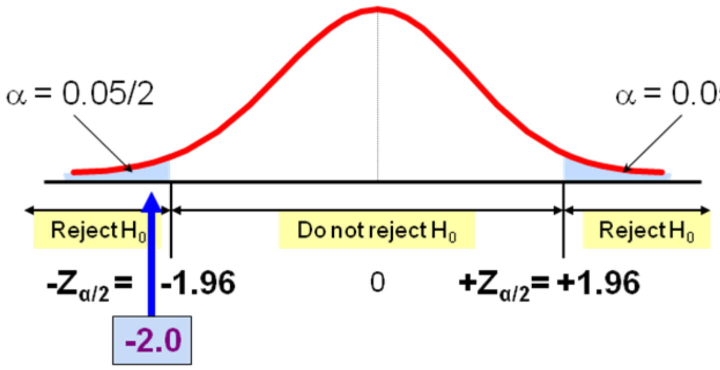
Determine the appropriate technique

- σ is assumed known so this is a Z test.

Determine the critical values

- For α = 0.05 the critical Z values are ±1.96

Compute the test statistic Z_{STAT} so the test statistic is: $Z_{STAT} = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} = \frac{2.84 - 3}{0.8/\sqrt{100}} = -2.0$



Since $Z_{STAT} = -2.0 < -1.96$, reject the null hypothesis and conclude there is sufficient evidence that the mean number of TVs in US homes is not equal to 3

Exercise

1. A simple random sample of 10 people from a certain population has a mean age of 27. Can we conclude that the mean age of the population is less than 30? The variance is known to be 20. Let α = .05.
2. Bon Air Elementary School has 300 students. The principal of the school thinks that the average IQ of students at Bon Air is at least 110. To prove her point, she administers an IQ test to 20 randomly selected students. Among the sampled students, the average IQ is 108. Assuming variance is known to be 100, should the principal accept or reject her original hypothesis? at 5% level of significance
3. Central bank believes that if consumer confidence is too high, the economy risks over heating. Low confidence is a warning that rcession might be on the way. In either case, the bank may choose to intervene by altering interest rates. The ideal value for the bank's chosen measure is 50. We may assume the measure is normally distributed with standard deviation 10. The bank takes a survey of 25 people. Which returned a sample mean of 54 for the index. What would you advice the bank to do? Use α = .05.
4. A manager will switch to a new technology if the production process exceeds 80 units per hour. The manager asks the company statistician to test the null hypothesis: H_0: $\mu = 80$ against the alternative hypothesis: H_1: $\mu > 80$ If there is strong evidence to reject the null hypothesis then the new technology will be adopted. Past experience has shown that the standard deviation is 8. A data set with n = 25 for the new technology has a sample mean of 83 Does this justify adoption of the new technology?

5.4.2 Population Standard Deviation Unknown

If the population standard deviation **σ**, is unknown, then the population mean has a student's t distribution, and you will be using the t-score formula for sample means. The test statistic is very similar to that for the z-score, except that sigma has been replaced by s and z has been replaced by t. ie $t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$

The critical value is obtained from the t-table. The degrees of freedom is n-1.

Example 1 A fertilizer mixing machine is set to give 12 kg of nitrate for every 100kg bag of fertilizer. Ten 100kg bags are examined. The percentages of nitrate are as follows: 11, 14, 13, 12, 13, 12, 13, 14, 11, 12. Is there reason to believe that the machine is defective at 5% level of significance?

Solution

Hypothesis H_0: $\mu = 12$ H_1: $\mu \neq 12$ (This is a two-tail test)

σ is unknown so this is a t test. Use the unbiased estimator ie $s = \sqrt{\frac{\sum(x-\bar{x})^2}{n-1}}$

Critical Region based on $\alpha = 0.05$ and 9 degrees freedom

$t_{9,0.025} = 2.262$ ie reject H_0: $\mu = 12$ if $|t_c| \geq 2.262$

From calculator $\bar{x} = 12.5$ and $s = 1.0801$

Test statistic $t_c = \frac{\bar{x}-\mu}{s/\sqrt{n}} = \frac{12.5-12}{1.0801/\sqrt{10}} = 1.4639$

Decision since $|t_c| = 1.2639 < 2.262$, we fail to reject H_0 and conclude that the machine is not defective.

Example 2 The following figures give the end of year profits of ten randomly selected Chemists in Nairobi county.

Profit(in million shillings)	21.8	24.8	27.3	29.3	30.8	31.8	32.8	32.5	32.1	31.3

On the basis of this data, test whether the average profit is greater than 30M KSH at 1% level of significance

Solution

Hypothesis H_0: $\mu = 30$ H_1: $\mu > 30$ (This is a 1-tail test)

σ is unknown so this is a t test. Use the unbiased estimator ie $s = \sqrt{\frac{\sum(x-\bar{x})^2}{n-1}}$

Critical Region based on $\alpha = 0.01$ and 9 degrees freedom

$t_{9,0.01} = 2.82$ ie reject H_0: $\mu = 12$ if $|t_c| \geq 2.82$

From calculator $\bar{x} = 29.415$ and $s = 3.6601$

Test statistic $t_c = \frac{\bar{x}-\mu}{s/\sqrt{n}} = \frac{29.415-30}{3.6601/\sqrt{10}} = -0.51$

Decision since $|t_c| = 0.51 < 2.82$, we don't reject H_0 and conclude that the average profit is not greater than 30M KSH.

Exercise

1. Identify the critical *t* value for each of the following tests:
   a. A two-tailed test with α=0.05 and 11 degrees of freedom
   b. A one-tailed test with α=0.01 and n=17
2. Consider a sample with $n = 20$ $\bar{x} = 8.0$ and $s = 2$ Do the following hypothesis tests.
   a) . H_0: μ =8.7 H_1: μ > 8.7 at α=0.01
   b) H_0: μ =8.7 H_1: $\mu \neq$ 8.7 at α=0.05
3. It is widely believed that the average body temperature for healthy adults is 98.6 degrees Fahrenheit. A study was conducted a few years go to examine this belief. The body temperatures of *n* = 130 healthy adults were measured (half male and half female). The average temperature from the sample was found to be $\bar{x} = 98.249$ with a standard

deviation s = 0.7332. Do these statistics contradict the belief that the average body temperature is 98.6? test at 1% level of significance

4. A study is to be done to determine if the cognitive ability of children living near a lead smelter is negatively impacted by increased exposure to lead. Suppose the average IQ for children in the United States is 100. From a pilot study, the mean and standard deviation were estimated to be $\bar{x} = 89$ and s = 14.4 respectively. Test at 5% level whether there is a negative impact.
5. The average cost of a hotel room in New York is said to be \$168 per night. To determine if this is true, a random sample of 25 hotels is taken and resulted in $\bar{x} = \$172.5$ and $s = \$15.40$. Test the appropriate hypotheses at α = 0.05.
6. A sample of eleven plants gave the following shoot lengths
   Shoot length (cm) 10.1 21.5 11.7 12.9 14.8 11.0 19.2 11.4 22.6 10.8 10.2
   An earlier study reported that the mean shoot length is 15cm. Test whether the experimental data confirms the old view at 5% level of significance.
7. A simple random sample of 14 people from a certain population gives body mass indices as shown in Table 7.2.1. Can we conclude that the BMI is not 35? Let α = .05.

subject	1	2	3	4	5	6	7	8	9	10	11	12	13	14
BM	23	25	21	37	39	21	23	24	32	57	23	26	31	45

8. A company selling licenses for a franchise operation claims that, in the first year, the yield on an initial investment is 10%. How should a hypothesis test be stated? If there is strong evidence that the mean return on the investment is below 10% this will give a cautionary warning to a potential investor. Therefore, test the null hypothesis: H0 : μ = 10 against the alternative hypothesis: H1 : μ < 10 From a sample of n = 10 observations, the sample statistics are: $\bar{x} = 8.82$ x and s = 2.40
9. We know the distance that an athlete can jump is normally distributed but we do not know the standard deviation. We record 15 jumps: 7.48 7.34 7.97 5.88 7.48 7.67 7.49 7.48 8.51 5.79 7.13 6.80 6.19 6.95 5.93 Test whether these values are consistent with a mean jump length of 7m. Do you have any reservations about this test?
10. The manufacturing process should give a weight of 20 ounces. Does the data show evidence that the process is operating correctly? Test the null hypothesis: H0 : μ = 20 the process is operating correctly against the alternative: H1 : μ ¹ 20 the process is not operating correctly From the data set, the sample statistics are: n = 9, $\bar{x} = 20.356$ (ounces) and s = 0.6126