King Saud University

College of Science

Department of Statistics and Operations Research

STAT 223

Theory of Statistics 1

Lectures' Notes

1438/1439

Dr. Samah Alghamdi

Outline of the Course:

Course code and number: STAT 223.

Course name: Theory of Statistics 1.

Credits: 3(2+1).

Pre-requisite: STAT 215.

Instructor:

Dr. Samah Abdullah Alghamdi.

Office: 346 – Building 5.

E-mail: samalghamdi@ksu.edu.sa

Web Site: http://fac.ksu.edu.sa/samalghamdi

References:

1. Probability and Statistics for Engineers and Scientists, 2012, Ninth Edition by R. Walpole, R. Myers, S. Myers and K. Ye, Person.
2. Introduction to Mathematical Statistics, 2005, Sixth Edition by R. Hogg, J. McKean, and A. Craig, Prentice Hall.
3. Introduction to the Theory of Statistics, 2007, Third Edition by A. Mood, F. Graybill and D. Boes, McGrow-Hill.
4. Statistical Inference, 2002, Second Edition by G. Casella and R. Berger, Duxbury.

Main Topics:

1. Introduction: Basic Definitions and Concepts.
2. Sampling Distribution.
3. Central Limit Theorem.
4. Point Estimation.
5. Estimator Properties: Unbiasedness – Mean Squared Error – Consistency – Sufficiency – Minimal Sufficiency.
6. Exponential Family Theorem.
7. Completeness.
8. Fisher Information.
9. Minimum Variance Unbiased Estimator.
10. Cramer-Rao Inequality.
11. Rao-Blackwell Theorem.
12. Lehmann-Scheffé Theorem.

13. Estimation Methods: Method of Moments – Maximum Likelihood Method and its Properties.
14. Interval Estimation: Confidence Interval – Pivotal Quantity – Large Sample Confidence Interval.
15. Bayesian Estimate: Prior and Posterior Distributions – Loss Function Approach – Bayesian Confidence Intervals.

Marking Scheme:

25 Marks: First Mid-Term Exam.

25 Marks: Second Mid-Term Exam.

10 Marks: Assailments and Quizzes.

40 Marks: Final Exam.

Chapter 1: Introduction

This chapter introduce a brief review of some basic definitions and statistical distributions.

1.1 Definition and Basic Concept

In this chapter, we give some basic definitions and concepts.

Data:

Data is facts or numbers about something collected to be examined and considered and used to help decision-making.

Experiment:

Experiment is a process that generates a set of data.

Sample Space:

In an experiment, the sample space, Ω, is the set of all possible outcomes of the experiment.

Event:

A subset of sample space is known as event.

Equally Likely Outcomes:

Equally likely outcomes always have the same chance of occurrence.

Probability:

If A be an event, the probability of event A is denoted by

$$P(A) = \frac{\#(\text{elements in } A)}{\#(\text{elements in } \Omega)}.$$

where $0 \leq P(A) \leq 1$ for each event A.

Sure and impossible Events:

Any subset of Ω, including Ω itself is called sure event and the empty set ϕ is called impossible event, where $P(\Omega) = 1$ and $P(\phi) = 0$.

Some Set-Theoretic Notions in Terms of Events:

Concept	Notation	Meaning
Union	$A \cup B$	Either A occurs or B occurs or both occur
Intersection	$A \cap B$	A and B both occur
Complement	A^c	A does not occur
Difference	$A - B$ or $A \cap B^c$	A occurs but B does not

Additive Rule:

$P(A \cup B) = P(A) + (B) - P(A \cap B)$.

Disjoint Event:

Two events called disjoint or mutually exclusive events if they have empty intersection i.e. $A \cap B = \phi$ or $P(A \cap B) = 0$.

Independent Events:

Two events A and B are said to be independent events if one of the following satisfied:

$$P(A \cap B) = P(A)P(B) \text{ or } P(A|B) = P(A) \text{ or } P(B|A) = P(B).$$

Conditional Probability:

Conditional probability of A, given B, denoted by $P(A|B)$, is defined by

$$P(A|B) = \frac{P(A \cap B)}{P(B)},\ P(B) > 0.$$

Example 1.1:

Suppose the experiment is the toss of a dice. The six numbers that can appear face up, from 1 to 6, are the 6 possible outcomes of the experiment. Hence, the sample space is:

$$\Omega = \{1,2,3,4,5,6\}.$$

Define events A, B and C as following:

A=even numbers, B=odd numbers, C=numbers less than 5.

1. Find $A, B, C, A \cap B, B \cup C, C^c, A - C$.
2. $P(A), P(C^c), P(A - C), P(C|A)$.
3. Are A and B disjoint events?

Solution:

1. A= {2,4,6}, B= {1,3,5}, C= {1,2,3,4}.

$A \cap B = \emptyset$, $B \cup C = \{1,2,3,4,5\}$, $C^c = \{5,6\}$, $A - C = A \cap C^c = \{6\}$.

2. $P(A) = \frac{3}{6}$, $P(C^c) = \frac{2}{6}$, $P(A - C) = \frac{1}{6}$.
   $C \cap A = \{2,4\}$, $P(C \cap A) = \frac{2}{6}$, $P(C|A) = \frac{P(C \cap A)}{P(A)} = \frac{2}{3}$.
3. Yes, they are disjoint, since $A \cap B = \phi$.

Population:

- A population is the largest collection of elements or individuals in which we are interested in a particular time and about which we want to make some statement or conclusion.
- The population values usually denoted by $\underline{X} = (X_1, X_2, \dots, X_N)$, where N is the number of elements in the population, called the population size.

Sample:

- A sample is a subset of a population on which we collect data.
- The sample values usually denoted by $\underline{x} = (x_1, x_2, \dots, x_n)$, where n is the number of elements in the sample, called the sample size.

Parameter:

- A parameter is a measure (or number) obtained from the population values.
- Values of the parameters are unknown in general.

Statistic:

- A statistic is a measure (or number) obtained from the sample values.
- Values of the statistic are known in general.

Random Variable:

- A random variable X is a function that associates a real number with each element in the sample space.
- Most of the time, statisticians deal with two special kinds of random variables, that are discrete and continuous random variables.

Discrete Random Variable:

A random variable X is discrete if:

1. It can take on values from finite or countable values.

2. It has a discrete distribution, called the **probability mass function** (**pmf**) of X if, for each possible outcome x

$$f_X(x) \geq 0,$$
$$\sum_x f_X(x) = 1,$$
$$\text{and} \qquad f_X(x) = P(X = x).$$

Continuous Random Variable:

A random variable X is continuous if:

1. It can take on values from an interval or not countable values.
2. It has a continuous distribution, called the **probability density function** (**pdf**) for X, defined over the set of real numbers, if

$$f_X(x) \geq 0 \quad \text{for all } x \in R,$$
$$\int_{-\infty}^{\infty} f_X(x)dx = 1,$$
$$\text{and} \qquad P(a \leq X \leq b) = \int_a^b f_X(x)dx \ .$$

Cumulative Distribution Function:

Let X be a random variable. The cumulative distribution function (**distribution function** or **cdf**) of X is a function such that

$$F_X(x) = P(X \leq x), \ \text{for} \ -\infty < x < \infty.$$

Random Sample:

A random sample is a sample that is chosen randomly. Random sample are used to avoid bias and other unwanted effects.

Joint Probability distribution:

The function $f(x, y)$ is a joint probability distribution of the random variables X and Y if:

1. $f(x, y) \geq 0$, for all (x, y).
2. $\sum_x \sum_y f(x, y) = 1$ if X and Y are discrete

   $\int_x \int_y f(x, y)dydx = 1$ if X and Y are continuous.

Independent Random Variables:

Let $X_1, X_2, \dots, X_n$ be a n random variables, discrete or continuous, with joint probability distribution $f(x_1, x_2, \dots, x_n)$. The random variables $X_1, X_2, \dots, X_n$ are said to be **mutually statistically independent** if and only if

$$f(x_1, x_2, \dots, x_n) = f_1(x_1) f_2(x_2) \dots f_n(x_n).$$

For all $(x_1, x_2, \dots, x_n)$ within their range.

Expectations and Moments:

The ***r*th moment about the origin** of the random variable X is given by

$$\mu_r' = E(X^r) = \begin{cases} \sum x^r f_X(x), & \text{If } X \text{ is discrete,} \\ \int x^r f_X(x) dx, & \text{If } X \text{ is continuous.} \end{cases}$$

The first moment (**mean** or **expected value)** and the **second moment** are given by $\acute{\mu}_1 = \mu = E(X)$ and $\acute{\mu}_2 = E(X^2)$, respectively.

The **variance** is defined as

$$Var(X) = \sigma^2 = \acute{\mu}_2 - \acute{\mu}_1^2 = E[(X - \mu)^2] = E(X^2) - \left(E(X)\right)^2.$$

The standard deviation is the square root of the variance denoted as

$$\sigma = \sqrt{\sigma^2} = \sqrt{E(X^2) - \left(E(X)\right)^2}.$$

The ***r*th central moment** of X is defined as

$$E[(X - \mu)^r] = \begin{cases} \sum (x - \mu)^r f_X(x), & \text{If } X \text{ is discrete,} \\ \int (x - \mu)^r f_X(x) dx, & \text{If } X \text{ is continuous.} \end{cases}$$

Remark:

If $Y = aX \pm b$, then the mean and the variance of Y are given by

$$E(Y) = aE(X) \pm b \quad \text{and} \quad Var(Y) = a^2 Var(X)$$

Example1.2:

Consider the following distribution:

X	-2	0	1	2	Total
$f(x)$	0.33	0.24	0.15	0.28	1

Find: 1. $P(X = 0), P(X < 1), P(X \geq 2)$.

2. The cdf of X.

3. The mean and the variance.

Solution:

1. $P(X=0)=0.24,\ P(X<1)=0.57,\ P(X\geq 2)=0.28.$
2. The cdf of X is

X	-2	0	1	2
$F(x)$	0.33	0.57	0.72	1

3. The mean and the variance are given as

X	-2	0	1	2	Total
$f(x)$	0.33	0.24	0.15	0.28	1
$E(X)$	-0.66	0	0.15	0.56	0.05
$E(X^2)$	1.32	0	0.15	1.12	2.59

Then,
$E(X)=0.05.$
$Var(X)=E(X^2)-\left(E(X)\right)^2=2.59-0.05^2=2.5875.$

Example 1.3:

Let X be a continuous random variable whose probability density function is

$$f(x)=3x^2, \quad \text{for } 0<x<1.$$

Find:

1. Prove $f(x)$ is a pdf.
2. $P(0.5<X<1)$.
3. The cdf of X.
4. $E(X)$ and $Var(X)$.

Solution:

1. Since $f(x)\geq 0$ for all $x\in(0,1)$ and

$\int_0^1 f(x)dx=\int_0^1 3x^2dx= \left. x^3\right]_0^1=1$. Thus, $f(x)$ is a pdf.

2. $P(0.5<X<1)=\int_{0.5}^1 3x^2dx= \left. x^3\right]_{0.5}^1=1-0.5^3=0.875.$
3. $F(x)=\int_0^x 3x^2dx= \left. x^3\right]_0^x=x^3.$
4. $E(X)=\int_0^1 3x^3dx=\frac{3}{4}\left. x^4\right]_0^1=\frac{3}{4}.$

$$E(X^2)=\int_0^1 3x^4dx=\frac{3}{5}\left. x^5\right]_0^1=\frac{3}{5}.$$

$$Var(X) = E(X^2) - \big(E(X)\big)^2 = \frac{3}{5} - \left(\frac{3}{4}\right)^2 = 0.0375.$$

Moment-Generation Function:

The moment-generation function (**mgf**) of a random variable X is given by $E(e^{tX})$ and is denoted by, $M_X(t)$. Hence,

$$M_X(t) = E(e^{tX}) = \begin{cases} \sum_x e^{tx} f(x), & \text{if } X \text{ is discrete,} \\ \int_x e^{tx} f(x)dx, & \text{if } X \; is \text{ continuous} \end{cases}.$$

Some properties of the mgf:

1. $M_{X+a}(t) = e^{at} M_X(t)$.
2. $M_{aX}(t) = M_X(at)$.

1.2 Discrete Probability Distributions

In this section, we present some commonly used distributions for the discrete random variable.

1.2.1 Bernoulli and Binomial Distribution

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

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

where $\binom{n}{x} = \frac{n!}{x!(n-x)!}$.

The mean, variance and mgf of the binomial distribution, $Binomial(n, p)$, are

$$\mu = np, \;\; \sigma^2 = npq \;\; \text{and } M(t) = (pe^t + q)^n.$$

Example 1.4:

The probability that a certain kind of component will survive a shock test is 0.75. Find:

1. The probability that exactly 2 of the next 4 components tested survive.
2. The probability that more than 2 of the next 4 components tested survive.
3. The mean and the standard deviation.

Solution:

Assuming that the tests are independent and $p = 0.75$ for each of the $n = 4$ tests, we obtain:

$$f(x; 4,0.75) = \binom{4}{x}(0.75)^x(0.25)^{4-x}, \quad x = 0, 1, 2, 3, 4.$$

1. $f(2; 4,0.75) = \binom{4}{2}0.75^2 0.25^2 = 0.2109$
2. $P(X > 2) = f(3; 4,0.75) + f(4; 4,0.75)$

$$= \binom{4}{3}0.75^3 0.25^1 + \binom{4}{4}0.75^4 0.25^0 = 0.7383$$

3. $\mu = np = (4)(0.75) = 3$ and $\sigma = \sqrt{npq} = \sqrt{4(0.75)(0.25)} = 0.866.$

1.2.2 Poisson Distribution

The probability distribution of the Poisson random variable X with parameter λ, $Poisson(\lambda)$, representing the number of outcomes occurring in a given time interval or specified region denoted by t, is

$$f(x; \lambda t) = \frac{e^{-\lambda t}(\lambda t)^x}{x!}, \quad x = 0, 1, 2, \ldots$$

where $\lambda > 0$ is the average number of outcomes per unit time, distance or area.

The mean and the variance of the Poisson distribution are

$$\mu = \sigma^2 = \lambda t.$$

Example 1.5:

Births in a hospital occur randomly at an average rate of 1.6 births per hour. Calculate:

1. The probability of observing 4 births in a given hour.
2. The probability of observing more than or equal to 2 births in a given hour.
3. The mean of births per hour.
4. The probability of observing 1 birth per 2 hours.
5. The variance of births per 30 minutes.

Solution:

Let X be the number of births in a given hour and $\lambda t = 1.6$ per hour. The pdf of X is given as

$$f(x;1.6)=\frac{e^{-1.6}(1.6)^x}{x!},\quad x=0,1,2,\ldots$$

1. $f(4;1.6)=\frac{e^{-1.6}(1.6)^4}{4!}=0.0551$
2. $P(X\geq 2)=1-P(X<2)=1-[f(1;1.6)+f(0;1.6)]=0.4751$
3. $\mu=\lambda t=1.6$
4. $\lambda t=(1.6)(2)=3.2$
$$f(x;3.2)=\frac{e^{-3.2}(3.2)^x}{x!},\quad x=0,1,2,\ldots$$
$$f(1;3.2)=\frac{e^{-3.2}(3.2)^1}{1!}=0.1304$$
5. $\sigma^2=\lambda t=(1.6)(0.5)=0.8.$

1.3 Continuous Probability Distributions

1.3.1 Uniform Distribution

The density function of the continuous uniform random variable X on the interval $[a,b]$ is

$$f(x;a;b)=\frac{1}{b-a},\ \ a\leq x\leq b.$$

The mean and the variance of the uniform distribution, $Uniform(a,b)$, are

$$\mu=\frac{a+b}{2}\ \text{ and }\ \sigma^2=\frac{(b-a)^2}{12}.$$

Example 1.6:

Suppose that a large conference room at a certain company can be reserved for no more than 4 hours. In fact, it can be assumed that the length X of a conference has a uniform distribution on interval $[0,4]$.

(a) What is the probability density function?
(b) What is the probability that any given conference lasts at least 3 hours?

Solution:

(a) The appropriate density function for the uniformly distributed random variable X in the situation is
$$f(x)=\frac{1}{4},\quad 0\leq x\leq 4$$
(b) $P[X\geq 3]=\int_3^4\frac{1}{4}\,dx=\frac{1}{4}.$

1.3.2 Exponential Distribution

The pdf of the exponential distribution for a continuous random variable X with parameter $\theta > 0$, denoted as $Expnential\left(\frac{1}{\theta}\right)$, is given as

$$f(x;\theta) = \theta e^{-\theta x}, \quad x \geq 0$$

The mean and the variance of this distribution are

$$E(X) = \frac{1}{\theta} \quad \text{and} \quad V(x) = \frac{1}{\theta^2}.$$

The cdf and mgf obtained as

$$F(x) = 1 - e^{-\theta x} \ \text{ and } \ M(t) = \frac{\theta}{\theta - t} = \left(1 - \frac{t}{\theta}\right)^{-1}, t < \theta.$$

1.3.3 Gamma Distribution

The continuous random variable X has a gamma distribution with parameters α and $\frac{1}{\beta}$, $Gamma\left(\alpha, \frac{1}{\beta}\right)$ if its density function is given by

$$f\left(x;\alpha,\frac{1}{\beta}\right) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}, \quad x \geq 0$$

where $\alpha > 0, \beta > 0$ and $\Gamma(\alpha)$ is a gamma function defined as

$$\Gamma(\alpha) = (\alpha - 1)! = \int_0^\infty y^{\alpha-1} e^{-y}\, dy$$

The mean, the variance and the mgf are

$$E(X) = \frac{\alpha}{\beta}, \quad V(x) = \frac{\alpha}{\beta^2} \ \text{ and } \ M(t) = \left(\frac{\beta}{\beta - t}\right)^\alpha = \left(1 - \frac{t}{\beta}\right)^{-\alpha}, t < \beta.$$

Note:

1. The exponential distribution is a special case of gamma distribution with $\frac{1}{\beta}$ parameter when $\alpha = 1$.
2. $\int_0^\infty x^{\alpha-1} e^{-\beta x} dx = \frac{\Gamma(\alpha)}{\beta^\alpha}$.

1.3.4 Weibull Distribution

The continuous random variable X has a Weibull distribution, with parameters α and $\frac{1}{\beta}$, if its pdf is given by

$$f\left(x;\alpha,\frac{1}{\beta}\right)=\alpha\beta x^{\beta-1}e^{-\alpha x^{\beta}},\quad x\geq 0$$

where $\alpha>0$ and $\beta>0$.

The cumulative distribution function for the Weibull distribution is given by

$$F(x)=1-e^{-\alpha x^{\beta}}.$$

Note: For $\beta=1$, the Weibull density reduces to the exponential density function.

1.3.5 Chi-Squared Distribution

The random variable X has a chi-squared distribution with $v>0$ degrees of freedom, denoted as, $X\sim\chi^2(v)$, if its pdf is given by

$$f(x,v)=\frac{1}{2^{\left(\frac{v}{2}\right)}\Gamma\left(\frac{v}{2}\right)}x^{\frac{v}{2}-1}e^{-\frac{x}{2}},\quad x>0$$

The mean and the variance are

$$E(X)=v \quad \text{and} \quad V(x)=2v.$$

The mgf of this distribution is $M(t)=(1-2t)^{-\frac{v}{2}}, t>\frac{1}{2}$.

Note: It is a special case of gamma distribution in which $\alpha=\frac{v}{2}$ and $\beta=\frac{1}{2}$.

Example 1.7:

Let X be a $\chi^2(10)$. Find:

1. Find $P(X>20.5)$.
2. a, if $P(X>a)=0.05$.

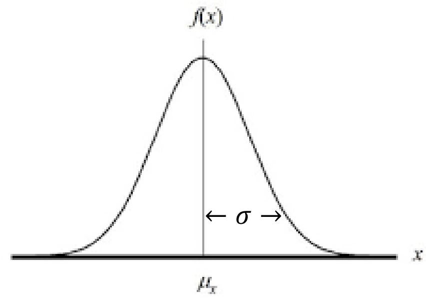
Solution:

By χ^2 Table (Table I) and $v=10$, we get

1. $P(X>20.5)=0.025$
2. $P(X>a)=0.05$, thus $a=18.31$.

1.3.6 Normal Distribution

The most important continuous probability distribution in the entire field of statistics is the **normal distribution**. Its graph, called the **normal curve**, is the bell-shaped curve of following figure, which approximately describes many phenomena that occur in nature, industry, and research.



Definition:

The density of the normal random variable X, with mean μ and variance σ^2, $X \sim N(\mu, \sigma^2)$, is

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

where $-\infty < \mu < \infty$ and $\sigma > 0$.

The properties of the normal curves:

1. The mode = median = mean = μ.
2. The curve is symmetric about the mean μ.
3. The normal curve depends on the parameters μ and σ, its mean and standard deviation, respectively.
4. The mean μ and the variance σ^2 determine the location and the shape of the normal curve, respectively.
5. The total area under the curve and above the horizontal axis is equal to 1.
6. The mgf is given by $M(t) = e^{\mu t + \frac{1}{2}\sigma^2 t^2}$.

1.3.7 Standard Normal Distribution

The distribution of a normal random variable with mean 0 and variance 1 is called a **standard normal distribution** and defined as

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}, -\infty < z < \infty.$$

The properties of the standard normal curves:

1. The mode = median = mean = 0.
2. The curve is symmetric about the mean 0.
3. The total area under the curve and above the horizontal axis is equal to 1.
4. The mgf is given by $M(t) = e^{\frac{1}{2}t^2}$.

Application: we are able to transform all the observations of any normal random variable X into a new set of observations of a normal random variable Z with mean 0 and variance 1. This can be done by mean of the transformation i.e.

$$\text{If } X \sim N(\mu, \sigma^2), \text{ then } Z = \frac{X-\mu}{\sigma} \sim N(0,1).$$

Example 1.8:

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

1. to the left of $z = 1.84$.
2. to the right of $z = 1.84$.

Solution: From Table II,

1. the area to the left of $z = 1.84$ is equal to,
$$P(Z < 1.84) = 0.9671.$$
2. the area to the right of $z = 1.84$ is equal to,
$$P(Z > 1.84) = 1 - P(Z < 1.84) = 1 - 0.9671 = 0.0329.$$

Normal Approximation to the Binomial:

Theorem 1.1:

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

$$Z = \frac{X - np}{\sqrt{npq}} \sim N(0,1)$$

as $n \to \infty$.

1.3.8 *T*-Distribution

A continuous random variable *T* is said to have a t-distribution with parameter $v > 0$ if its pdf defined as

$$f(t;v) = \frac{\Gamma\left(\frac{v+1}{2}\right)}{\sqrt{v\pi}\,\Gamma\left(\frac{v}{2}\right)} \left(1+\frac{t^2}{v}\right)^{-\frac{v+1}{2}}; \quad -\infty < t < \infty.$$

The properties of the standard normal curves:

1. The mode = median = mean = 0.
2. The curve is symmetric about the mean 0.
3. Compared to the standard normal distribution, the *t*-distribution is less peaked in the center and has higher tails.
4. It depends on the degrees of freedom *v*.
5. T-distribution approaches the standard normal distribution as $v \to \infty$.
6. The total area under the curve and above the horizontal axis is equal to 1.

Example 1.9: Find:

1. $P(T < 2.145)$ when $v = 14$.
2. $t_{0.995}$ when $v = 7$.

Solution: From Table III,

1. $P(T < 2.145) = 0.975$ when $v = 14$.
2. $t_{0.995} = 3.499$ when $v = 7$.

1.3.9 *F*-Distribution

If a random variable *X* has a F-distribution with parameters r and *v*, we write $X \sim F(r, v)$. Then the probability density function for *X* is given by

$$f(x;r,v) = \frac{1}{B\left(\frac{r}{2},\frac{v}{2}\right)} \left(\frac{r}{v}\right)^{\frac{r}{2}} x^{\frac{r}{2}-1} \left(1+\frac{r}{v}x\right)^{-\left(\frac{r+v}{2}\right)}$$

For real $x \geq 0$. Here is $B(a,b) = \int_0^1 y^{a-1}(1-y)^{b-1}dy$ is the beta function and $r, v > 0$.

Theorem 1.2:

If $F_\alpha(r, v)$ has F-distribution with r and v degrees of freedom, then

$$F_{1-\alpha}(v, r) = \frac{1}{F_\alpha(r, v)}$$

has F-distribution with v and r degrees of freedom.

Example 1.10: Find:

1. $F_{0.01}$ with $r = 10$ and $v = 6$.
2. $F_{0.95}(9,15)$.

Solution: Table IV:

1. $F_{0.01}(10,6) = 7.87$.
2. $F_{0.95}(9,15) = \frac{1}{F_{0.05}(15,9)} = \frac{1}{3.01} = 0.3322$.

1.4 Transformation of Variables

In standard statistical methods, the result of statistical hypotheses testing, estimation, or even statistical graphics does not involve a single random variable but, rather, **functions of one or more random variables**. As a result, statistical inference requires the distribution of these functions. In this section, we represent methods to find the distribution of these functions.

1.4.1 Discrete Random Variable

1.4.1.1 One-to-One Transformation:

Theorem 1.3:

Suppose that X is a discrete random variable with probability distribution $f(x)$. Let $Y = u(X)$ define a one-to-one transformation between the values of X and Y so that the equation $y = u(x)$ can be uniquely solved for x in terms of y, say $x = w(y)$. Then the probability distribution of Y is

$$g(y) = f[w(y)].$$

Example 1.11:

Let X be a discrete random variable with pmf as

$$f(x) = \frac{x}{4}\ ,\ \ x = 0, 1, 3.$$

Find the pmf of the random variable $Y = X^2$.

Solution:

Since the value of X are all positive, the transformation defines a one-to-one correspondence between the x and y values.

Hence,

Since $x = 0, 1, 3 \Longrightarrow y = 0, 1, 9$ and $y = x^2 \Longrightarrow x = \sqrt{y}$.

Then, the pmf of Y is given by

$$g(y) = f\left(\sqrt{y}\right) = \frac{\sqrt{y}}{4},\ \ y = 0, 1, 9.$$

Similarly, for a two-dimension transformation.

Theorem 1.4:

Suppose that X_1 and X_2 are discrete random variables with joint probability distribution $f(x_1, x_2)$. Let $Y_1 = u_1(X_1, X_2)$ and $Y_2 = u_2(X_1, X_2)$ define a one-to-one transformation between the points (x_1, x_2) and (y_1, y_2) so that the equations

$$y_1 = u_1(x_1, x_2) \text{ and } y_2 = u_2(x_1, x_2)$$

may be uniquely solved for x_1 and x_2 in terms of y_1 and y_2, say $x_1 = w_1(y_1, y_2)$ and $x_2 = w_2(y_1, y_2)$. Then the joint probability distribution of Y_1 and Y_2 is

$$g(y_1, y_2) = f[w_1(y_1, y_2), w_2\,(y_1, y_2)]$$

1.4.2 Continuous Random Variable

This section introduced three methods of transformation to find the distribution of continuous random variable.

1.4.2.1 One-to-One Transformation

Theorem 1.5:

Suppose that X is a continuous random variable with probability distribution $f(x)$. Let $Y = u(X)$ define a one-to-one correspondence between the values of X and Y so that the equation $y = u(x)$ can be uniquely solved for x in terms of y, say $x = w(y)$. Then the probability distribution of Y is

$$g(y) = f[w(y)].\, |J|$$

where $|J| = |w'(y)| = \left|\frac{\partial x}{\partial y}\right|$ and is called the **Jacobian** of the transformation.

Example 1.12:

Let X be a continuous random variable with probability distribution

$$f(x) = \begin{cases} \dfrac{x}{12}, & 1 < x < 5, \\ 0, & \text{elsewhere.} \end{cases}$$

Find the probability distribution of the random variable $Y = 2X - 3$.

Solution:

The inverter solution of $y = 2x - 3$ yields $x = (y+3)/2$, from which we obtain $J = w'(y) \ = \frac{dx}{dy} = \frac{1}{2}$.

Therefore,

$$1 < x < 5 \ \Rightarrow \ 1 < \frac{y+3}{2} < 5 \Rightarrow \ 2 < y+3 < 10 \Rightarrow \ -1 < y < 7$$

Using Theorem 1.5, we find the density function of Y to be

$$g(y) = \begin{cases} \dfrac{(y+3)/2}{12}\left(\frac{1}{2}\right) = \frac{y+3}{48}, & -1 < y < 7 \\ 0, & \text{elswhere} \end{cases}.$$

Theorem 1.6:

Suppose that X_1 and X_2 are continuous random variable with joint probability distribution $f(x_1, x_2)$. Let $Y_1 = u_1(X_1, X_2)$ and $Y_2 = u_2(X_1, X_2)$ define a one-to-one transformation between the points (x_1, x_2) and (y_1, y_2) so that the equations $y_1 = u_1(x_1, x_2)$ and $y_2 = u_2(x_1, x_2)$ may be uniquely solved for

x_1 and x_2 in terms of y_1 and y_2, say $x_1 = w_1(y_1, y_2)$ and $x_2 = w_2(y_1, y_2)$. Then the joint probability distribution of Y_1 and Y_2 is

$$g(y_1, y_2) = f[w_1(y_1, y_2), w_2(y_1, y_2)].|J|$$

where the Jacobian is 2×2 determinant as

$$|J| = \begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} \end{vmatrix}.$$

1.4.2.2 Distribution Function Method (cdf Method):

The general method works as follows:

If X be an independent random variable with pdf $f_X(x)$ and $Y = u(X)$ be a function of X. Then, find

1. $F_X(x)$, cdf of X.
2. The region of Y.
3. $F_Y(y) = P(Y \le y) = P(u(X) \le y) = P(X \le w(Y)) = F_X(w(Y))$.
4. The density function $f_Y(y)$ by differentiating $F_Y(y)$.

Example 1.13:

Suppose the random variable X has a pdf

$$f_X(x) = 3x^2, \quad 0 < x < 1.$$

Find the pdf of $Y = 2X + 3$.

Solution:

From Example 1.3, we get $F_X(x) = x^3$.

Since $0 < x < 1 \Longrightarrow 0 < 2x < 2 \Rightarrow 3 < y < 5$.

$F_Y(y) = P(Y \le y) = P(2X + 3 \le y) = P(2X \le y - 3)$

$$= P\left(X \le \frac{y-3}{2}\right) = F_X\left(\frac{y-3}{2}\right) = \left(\frac{y-3}{2}\right)^3.$$

Then, the pdf of Y is

$$f_Y(y) = \frac{dF_Y(y)}{dy} = \frac{3}{8}(y-3)^2, \qquad 3 < y < 5.$$

1.4.2.3 Moment-Generating Method:

Theorem 1.7: (uniqueness Theorem)

Let X and Y be two random variables with moment-generating functions $M_X(t)$ and $M_Y(t)$, respectively, if $M_X(t) = M_Y(t)$ for all values of t, then X and Y have the same probability distribution.

Theorem 1.8:

If $X_1, X_2, \dots\dots, X_n$ are independent random variable with moment-generating functions $M_{X_1}(t), M_{X_2}(t), \dots\dots, M_{X_n}(t)$, respectively, and $Y = X_1 + X_2 + \cdots. + X_n$, then

$$M_Y(t) = M_{X_1}(t). M_{X_2}(t) \dots M_{X_n}(t).$$

Moreover, if $M_{X_1}(t), M_{X_2}(t), \dots\dots, M_{X_n}(t)$are equals. Then,

$$M_Y(t) = \left(M_{X_1}(t)\right)^n.$$

Example 1.14:

If $X_1, X_2, \dots\dots, X_n$ are independent, each with an exponential distribution with parameter $\frac{1}{\theta}$. Show that $Y = \sum_{i=1}^{n} X_i$ has a gamma distribution with parameters *n* and $\frac{1}{\theta}$.

Solution:

Since that the mgf of $expnential\left(\frac{1}{\theta}\right)$ is $M_X(t) = \frac{\theta}{\theta - t}$. Thus, the mgf of Y is given by

$$M_Y(t) = M_{\sum_{i=1}^{n} X_i}(t) = M_{X_1+X_2+\cdots.+X_n}(t)$$

$$= M_{X_1}(t). M_{X_2}(t) \dots M_{X_n}(t) = \left(\frac{\theta}{\theta - t}\right)^n.$$

which is the mgf of $Gamma\left(n, \frac{1}{\theta}\right)$.