1

Some basic results in probability and statistics

This chapter contains some basic results in probability and statistics. It is intended as a reference chapter to which you may refer as you read this book. Sometimes, specific references to results in this chapter are made in the text. At other times, you may wish to refer on your own to particular results in this chapter as you feel the need.

You may prefer to scan the results on probability and statistical inference in this chapter before reading Chapter 2, or you may proceed directly to the next chapter.

1.1 SUMMATION AND PRODUCT OPERATORS

Summation operator

The summation operator Σ is defined as follows:

$$\sum_{i=1}^{n} Y_i = Y_1 + Y_2 + \cdots + Y_n \tag{1.1}$$

Some important properties of this operator are:

(1.2a) $$\sum_{i=1}^{n} k = nk \qquad \text{where } k \text{ is a constant}$$

(1.2b) $$\sum_{i=1}^{n} (Y_i + Z_i) = \sum_{i=1}^{n} Y_i + \sum_{i=1}^{n} Z_i$$

(1.2c) $$\sum_{i=1}^{n} (a + cY_i) = na + c\sum_{i=1}^{n} Y_i \qquad \text{where } a \text{ and } c \text{ are constants}$$

The double summation operator $\Sigma\Sigma$ is defined as follows:

(1.3) $$\sum_{i=1}^{n}\sum_{j=1}^{m} Y_{ij} = \sum_{i=1}^{n} (Y_{i1} + \cdots + Y_{im})$$
$$= Y_{11} + \cdots + Y_{1m} + Y_{21} + \cdots + Y_{2m} + \cdots + Y_{nm}$$

An important property of the double summation operator is:

(1.4) $$\sum_{i=1}^{n}\sum_{j=1}^{m} Y_{ij} = \sum_{j=1}^{m}\sum_{i=1}^{n} Y_{ij}$$

Product operator

The product operator Π is defined as follows:

(1.5) $$\prod_{i=1}^{n} Y_i = Y_1 \cdot Y_2 \cdot Y_3 \cdots Y_n$$

1.2 PROBABILITY

Addition theorem

Let A_i and A_j be two events defined on a sample space. Then:

(1.6) $$P(A_i \cup A_j) = P(A_i) + P(A_j) - P(A_i \cap A_j)$$

where $P(A_i \cup A_j)$ denotes the probability of either A_i or A_j or both occurring; $P(A_i)$ and $P(A_j)$ denote, respectively, the probability of A_i and the probability of A_j; and $P(A_i \cap A_j)$ denotes the probability of both A_i and A_j occurring.

Multiplication theorem

Let $P(A_i|A_j)$ denote the conditional probability of A_i occurring, given that A_j has occurred. This conditional probability is defined as follows:

$$(1.7) \qquad P(A_i|A_j) = \frac{P(A_i \cap A_j)}{P(A_j)} \qquad P(A_j) \neq 0$$

The multiplication theorem states:

$$(1.8) \qquad \begin{aligned} P(A_i \cap A_j) &= P(A_i)P(A_j|A_i) \\ &= P(A_j)P(A_i|A_j) \end{aligned}$$

Complementary events

The complementary event of A_i is denoted by $\bar{A}_i$. The following results for complementary events are useful:

$$(1.9) \qquad P(\bar{A}_i) = 1 - P(A_i)$$

$$(1.10) \qquad P(\overline{A_i \cup A_j}) = P(\bar{A}_i \cap \bar{A}_j)$$

1.3 RANDOM VARIABLES

Throughout this section, we assume that the random variable Y assumes a finite number of outcomes. (If Y is a continuous random variable, the summation process is replaced by integration.)

Expected value

Let the random variable Y assume the outcomes $Y_1, \ldots, Y_k$ with probabilities given by the probability function:

$$(1.11) \qquad f(Y_s) = P(Y = Y_s) \qquad s = 1, \ldots, k$$

The expected value of Y is defined:

$$(1.12) \qquad E(Y) = \sum_{s=1}^{k} Y_s f(Y_s)$$

An important property of the expectation operator E is:

$$(1.13) \qquad E(a + cY) = a + cE(Y) \qquad \text{where } a \text{ and } c \text{ are constants}$$

Special cases of this are:

$$(1.13a) \qquad E(a) = a$$

$$(1.13b) \qquad E(cY) = cE(Y)$$

$$(1.13c) \qquad E(a + Y) = a + E(Y)$$

Variance

The variance of the random variable Y is denoted by $\sigma^2(Y)$ and is defined as follows:

(1.14) $$\sigma^2(Y) = E\{[Y - E(Y)]^2\}$$

An equivalent expression is:

(1.14a) $$\sigma^2(Y) = E(Y^2) - [E(Y)]^2$$

The variance of a linear function of Y is frequently encountered. We denote the variance of $a + cY$ by $\sigma^2(a + cY)$ and have:

(1.15) $$\sigma^2(a + cY) = c^2\sigma^2(Y) \qquad \text{where } a \text{ and } c \text{ are constants}$$

Special cases of this result are:

(1.15a) $$\sigma^2(a + Y) = \sigma^2(Y)$$

(1.15b) $$\sigma^2(cY) = c^2\sigma^2(Y)$$

Joint, marginal, and conditional probability distributions

Let the joint probability function for the two random variables Y and Z be denoted by $g(Y, Z)$:

(1.16) $$g(Y_s, Z_t) = P(Y = Y_s \cap Z = Z_t) \qquad s = 1, \ldots, k;\ t = 1, \ldots, m$$

The marginal probability function of Y, denoted by $f(Y)$, is:

(1.17a) $$f(Y_s) = \sum_{t=1}^{m} g(Y_s, Z_t) \qquad s = 1, \ldots, k$$

and the marginal probability function of Z, denoted by $h(Z)$, is:

(1.17b) $$h(Z_t) = \sum_{s=1}^{k} g(Y_s, Z_t) \qquad t = 1, \ldots, m$$

The conditional probability function of Y, given $Z = Z_t$, is:

(1.18a) $$f(Y_s | Z_t) = \frac{g(Y_s, Z_t)}{h(Z_t)} \qquad h(Z_t) \neq 0;\ s = 1, \ldots, k$$

and the conditional probability function of Z, given $Y = Y_s$, is:

(1.18b) $$h(Z_t | Y_s) = \frac{g(Y_s, Z_t)}{f(Y_s)} \qquad f(Y_s) \neq 0;\ t = 1, \ldots, m$$

Covariance

The covariance of Y and Z is denoted by $\sigma(Y, Z)$ and is defined:

(1.19) $$\sigma(Y, Z) = E\{[Y - E(Y)][Z - E(Z)]\}$$

An equivalent expression is:

(1.19a) $$\sigma(Y, Z) = E(YZ) - [E(Y)][E(Z)]$$

The covariance of $a_1 + c_1Y$ and $a_2 + c_2Z$ is denoted by $\sigma(a_1 + c_1Y, a_2 + c_2Z)$, and we have:

(1.20) $$\sigma(a_1 + c_1Y, a_2 + c_2Z) = c_1c_2\sigma(Y, Z)$$ where a_1, a_2, c_1, c_2 are constants

Special cases of this are:

(1.20a) $$\sigma(c_1Y, c_2Z) = c_1c_2\sigma(Y, Z)$$

(1.20b) $$\sigma(a_1 + Y, a_2 + Z) = \sigma(Y, Z)$$

By definition, we have:

(1.21) $$\sigma(Y, Y) = \sigma^2(Y)$$

where $\sigma^2(Y)$ is the variance of Y.

Independent random variables

(1.22) Random variables Y and Z are independent if and only if:

$$g(Y_s, Z_t) = f(Y_s)h(Z_t) \qquad s = 1, \ldots, k;\ t = 1, \ldots, m$$

If Y and Z are independent random variables:

(1.23) $$\sigma(Y, Z) = 0 \qquad \text{when } Y, Z \text{ are independent}$$

(In the special case where Y and Z are jointly normally distributed, $\sigma(Y, Z) = 0$ implies that Y and Z are independent.)

Functions of random variables

Let $Y_1, \ldots, Y_n$ be n random variables. Consider the function Σa_iY_i where the a_i are constants. We then have:

(1.24a) $$E\left(\sum_{i=1}^{n} a_iY_i\right) = \sum_{i=1}^{n} a_iE(Y_i)$$ where the a_i are constants

(1.24b) $$\sigma^2\left(\sum_{i=1}^{n} a_iY_i\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_ia_j\sigma(Y_i, Y_j)$$ where the a_i are constants

Specifically, we have for $n = 2$:

(1.25a) $E(a_1Y_1 + a_2Y_2) = a_1E(Y_1) + a_2E(Y_2)$

(1.25b) $\sigma^2(a_1Y_1 + a_2Y_2) = a_1^2\sigma^2(Y_1) + a_2^2\sigma^2(Y_2) + 2a_1a_2\sigma(Y_1, Y_2)$

If the random variables Y_i are independent, we have:

(1.26) $$\sigma^2\left(\sum_{i=1}^{n} a_iY_i\right) = \sum_{i=1}^{n} a_i^2\sigma^2(Y_i)$$ when the Y_i are independent

Special cases of this are:

(1.26a) $\sigma^2(Y_1 + Y_2) = \sigma^2(Y_1) + \sigma^2(Y_2)$ when Y_1, Y_2 are independent

(1.26b) $\sigma^2(Y_1 - Y_2) = \sigma^2(Y_1) + \sigma^2(Y_2)$ when Y_1, Y_2 are independent

When the Y_i are independent random variables, the covariance of two linear functions Σa_iY_i and Σc_iY_i is:

(1.27) $$\sigma\left(\sum_{i=1}^{n} a_iY_i, \sum_{i=1}^{n} c_iY_i\right) = \sum_{i=1}^{n} a_ic_i\sigma^2(Y_i)$$ when the Y_i are independent

Central limit theorem

(1.28) If $Y_1, \ldots, Y_n$ are independent random observations from a population with probability function $f(Y)$ for which $\sigma^2(Y)$ is finite, the sample mean $\bar{Y}$:

$$\bar{Y} = \frac{\sum_{i=1}^{n} Y_i}{n}$$

is approximately normally distributed when the sample size n is reasonably large, with mean $E(Y)$ and variance $\sigma^2(Y)/n$.

1.4 NORMAL PROBABILITY DISTRIBUTION AND RELATED DISTRIBUTIONS

Normal probability distribution

The density function for a normal random variable Y is:

(1.29) $$f(Y) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left[-\frac{1}{2}\left(\frac{Y-\mu}{\sigma}\right)^2\right] \qquad -\infty < Y < +\infty$$

where μ and σ are the two parameters of the normal distribution and $\exp(a)$ denotes e^a.

The mean and variance of a normal random variable Y are:

(1.30a) $$E(Y) = \mu$$

(1.30b) $$\sigma^2(Y) = \sigma^2$$

Function of normal random variable. A linear function of a normal random variable Y has the following property:

(1.31) If Y is a normal random variable, the transformed variable $Y' = a + cY$ (a and c are constants) is normally distributed, with mean $a + cE(Y)$ and variance $c^2\sigma^2(Y)$.

Standard normal variable. The standard normal variable z:

(1.32) $$z = \frac{Y - \mu}{\sigma} \qquad \text{where } Y \text{ is a normal random variable}$$

is normally distributed, with mean 0 and variance 1. We denote this as follows:

(1.33) $$z \text{ is } N(0, 1)$$

Mean Variance

Table A–1 in the Appendix contains the cumulative probabilities A for percentiles $z(A)$ where:

(1.34) $$P\{z \leq z(A)\} = A$$

For instance, when $z(A) = 2.00$, $A = .9772$. Because the normal distribution is symmetrical about 0, when $z(A) = -2.00$, $A = 1 - .9772 = .0228$.

Function of independent normal random variables. Let $Y_1, \ldots, Y_n$ be independent normal random variables. We then have:

(1.35) When $Y_1, \ldots, Y_n$ are independent normal random variables, the linear combination $a_1Y_1 + a_2Y_2 + \cdots + a_nY_n$ is normally distributed, with mean $\Sigma a_i E(Y_i)$ and variance $\Sigma a_i^2 \sigma^2(Y_i)$.

χ^2 distribution

Let $z_1, \ldots, z_\nu$ be ν independent standard normal variables. We then define:

(1.36) $$\chi^2(\nu) = z_1^2 + z_2^2 + \cdots + z_\nu^2 \qquad \text{where the } z_i \text{ are independent}$$

The χ^2 distribution has one parameter, ν, which is called the *degrees of freedom* (*df*). The mean of the χ^2 distribution with ν degrees of freedom is:

(1.37) $$E[\chi^2(\nu)] = \nu$$

Table A–3 in the Appendix contains percentiles of various χ^2 distributions. We define $\chi^2(A; \nu)$ as follows:

$$P\{\chi^2(\nu) \leq \chi^2(A; \nu)\} = A \tag{1.38}$$

Suppose $\nu = 5$. The 90th percentile of the χ^2 distribution with 5 degrees of freedom is $\chi^2(.90; 5) = 9.24$.

t distribution

Let z and $\chi^2(\nu)$ be independent random variables (standard normal and χ^2, respectively). We then define:

$$t(\nu) = \frac{z}{\left[\dfrac{\chi^2(\nu)}{\nu}\right]^{1/2}} \quad \text{where } z \text{ and } \chi^2(\nu) \text{ are independent} \tag{1.39}$$

The t distribution has one parameter, the *degrees of freedom* ν. The mean of the t distribution with ν degrees of freedom is:

$$E[t(\nu)] = 0 \tag{1.40}$$

Table A–2 in the Appendix contains percentiles of various t distributions. We define $t(A; \nu)$ as follows:

$$P\{t(\nu) \leq t(A; \nu)\} = A \tag{1.41}$$

Suppose $\nu = 10$. The 90th percentile of the t distribution with 10 degrees of freedom is $t(.90; 10) = 1.372$. Because the t distribution is symmetrical about 0, we have $t(.10; 10) = -1.372$.

F distribution

Let $\chi^2(\nu_1)$ and $\chi^2(\nu_2)$ be two independent χ^2 random variables. We then define:

$$F(\nu_1, \nu_2) = \frac{\chi^2(\nu_1)}{\nu_1} \div \frac{\chi^2(\nu_2)}{\nu_2} \quad \text{where } \chi^2(\nu_1) \text{ and } \chi^2(\nu_2) \text{ are independent} \tag{1.42}$$

(ν_1: Numerator *df*; ν_2: Denominator *df*)

The F distribution has two parameters, the *numerator degrees of freedom* and the *denominator degrees of freedom*, here ν_1 and ν_2, respectively.

Table A–4 in the Appendix contains percentiles of various F distributions. We define $F(A; \nu_1, \nu_2)$ as follows:

$$P\{F(\nu_1, \nu_2) \leq F(A; \nu_1, \nu_2)\} = A \tag{1.43}$$

Suppose $\nu_1 = 2$, $\nu_2 = 3$. The 90th percentile of the F distribution with 2 and 3 degrees of freedom, respectively, in the numerator and denominator is $F(.90; 2, 3) = 5.46$.

Percentiles below 50 percent can be obtained by utilizing the relation:

(1.44)
$$F(A; \nu_1, \nu_2) = \frac{1}{F(1 - A; \nu_2, \nu_1)}$$

Thus, $F(.10; 3, 2) = 1/F(.90; 2, 3) = 1/5.46 = .183$.

The following relation exists between the t and F random variables:

(1.45a)
$$[t(\nu)]^2 = F(1, \nu)$$

and the percentiles of the t and F distributions are related as follows:

(1.45b)
$$[t(.5 + A/2; \nu)]^2 = F(A; 1, \nu)$$

1.5 STATISTICAL ESTIMATION

Properties of estimators

(1.46) An estimator $\hat{\theta}$ of the parameter θ is *unbiased* if:

$$E(\hat{\theta}) = \theta$$

(1.47) An estimator $\hat{\theta}$ is a *consistent estimator* of θ if:

$$\lim_{n \to \infty} P(|\hat{\theta} - \theta| \geq \varepsilon) = 0 \qquad \text{for any } \varepsilon > 0$$

(1.48) An estimator $\hat{\theta}$ is a *sufficient estimator* of θ if the conditional joint probability function of the sample observations, given $\hat{\theta}$, does not depend on the parameter θ.

(1.49) An estimator $\hat{\theta}$ is a *minimum variance estimator* of θ if for any other estimator θ^*:

$$\sigma^2(\hat{\theta}) \leq \sigma^2(\theta^*) \qquad \text{for all } \theta^*$$

Maximum likelihood estimators

The method of maximum likelihood is a general method of finding estimators. Suppose we are sampling a population whose probability function $f(Y; \theta)$ involves one parameter, θ. Given independent observations $Y_1, \ldots, Y_n$, the joint probability function of the sample observations is:

(1.50a)
$$g(Y_1, \ldots, Y_n) = \prod_{i=1}^{n} f(Y_i; \theta)$$

When this joint probability function is viewed as a function of θ, with the observations given, it is called the *likelihood function* $L(\theta)$.

(1.50b)
$$L(\theta) = \prod_{i=1}^{n} f(Y_i; \theta)$$

Maximizing $L(\theta)$ with respect to θ yields the maximum likelihood estimator of θ. Under quite general conditions, maximum likelihood estimators are consistent and sufficient.

Least squares estimators

The method of least squares is another general method of finding estimators. The sample observations are assumed to be of the form (for the case of a single parameter θ):

(1.51) $$Y_i = f_i(\theta) + \varepsilon_i \qquad i = 1, \ldots, n$$

where $f_i(\theta)$ is a known function of the parameter θ and the ε_i are random variables, usually assumed to have expectation $E(\varepsilon_i) = 0$.

With the method of least squares, for the given sample observations, the sum of squares:

(1.52) $$Q = \sum_{i=1}^{n} [Y_i - f_i(\theta)]^2$$

is considered as a function of θ. The least squares estimator of θ is obtained by minimizing Q with respect to θ. In many instances, least squares estimators are unbiased and consistent.

1.6 INFERENCES ABOUT POPULATION MEAN—NORMAL POPULATION

We have a random sample of n observations $Y_1, \ldots, Y_n$ from a normal population with mean μ and standard deviation σ. The sample mean and sample standard deviation are:

(1.53a) $$\bar{Y} = \frac{\sum_i Y_i}{n}$$

(1.53b) $$s = \left[\frac{\sum_i (Y_i - \bar{Y})^2}{n - 1}\right]^{1/2}$$

and the estimated standard deviation of the sampling distribution of $\bar{Y}$ is:

(1.53c) $$s(\bar{Y}) = \frac{s}{\sqrt{n}}$$

We then have:

(1.54) $\dfrac{\bar{Y} - \mu}{s(\bar{Y})}$ is distributed as t with $n - 1$ degrees of freedom when the random sample is from a normal population.

Interval estimation

The confidence limits for μ with a confidence coefficient of $1 - \alpha$ are obtained by means of (1.54):

$$\bar{Y} \pm t(1 - \alpha/2; n - 1)s(\bar{Y}) \tag{1.55}$$

Example 1. Obtain a 95 percent confidence interval for μ when:

$$n = 10 \qquad \bar{Y} = 20 \qquad s = 4$$

We require:

$$s(\bar{Y}) = \frac{4}{\sqrt{10}} = 1.265 \qquad t(.975; 9) = 2.262$$

so that the confidence limits are $20 \pm 2.262(1.265)$. Hence, the 95 percent confidence interval for μ is:

$$17.1 \leq \mu \leq 22.9$$

Tests

One-sided and two-sided tests concerning the population mean μ are constructed by means of (1.54), based on the test statistic:

$$t^* = \frac{\bar{Y} - \mu_0}{s(\bar{Y})} \tag{1.56}$$

Table 1.1 contains the decision rules for each of three possible cases, with the risk of making a Type I error controlled at α.

TABLE 1.1 Decision rules for tests concerning mean μ of normal population

Alternatives	*Decision Rule*
	(a)
H_0: $\mu = \mu_0$	If $\lvert t^* \rvert \leq t(1 - \alpha/2; n - 1)$, conclude H_0
H_a: $\mu \neq \mu_0$	If $\lvert t^* \rvert > t(1 - \alpha/2; n - 1)$, conclude H_a
	where: $t^* = \dfrac{\bar{Y} - \mu_0}{s(\bar{Y})}$
	(b)
H_0: $\mu \geq \mu_0$	If $t^* \geq t(\alpha; n - 1)$, conclude H_0
H_a: $\mu < \mu_0$	If $t^* < t(\alpha; n - 1)$, conclude H_a
	(c)
H_0: $\mu \leq \mu_0$	If $t^* \leq t(1 - \alpha; n - 1)$, conclude H_0
H_a: $\mu > \mu_0$	If $t^* > t(1 - \alpha; n - 1)$, conclude H_a

Example 2. Choose between the alternatives:

$$H_0: \mu \leq 20$$
$$H_a: \mu > 20$$

when α is to be controlled at .05 and:

$$n = 15 \qquad \bar{Y} = 24 \qquad s = 6$$

We require:

$$s(\bar{Y}) = \frac{6}{\sqrt{15}} = 1.549$$
$$t(.95; 14) = 1.761$$

so that the decision rule is:

$$\text{If } t^* \leq 1.761, \text{ conclude } H_0$$
$$\text{If } t^* > 1.761, \text{ conclude } H_a$$

Since $t^* = (24 - 20)/1.549 = 2.58 > 1.761$, we conclude H_a.

Example 3. Choose between the alternatives:

$$H_0: \mu = 10$$
$$H_a: \mu \neq 10$$

when α is to be controlled at .02 and:

$$n = 25 \qquad \bar{Y} = 5.7 \qquad s = 8$$

We require:

$$s(\bar{Y}) = \frac{8}{\sqrt{25}} = 1.6$$
$$t(.99; 24) = 2.492$$

so that the decision rule is:

$$\text{If } |t^*| \leq 2.492, \text{ conclude } H_0$$
$$\text{If } |t^*| > 2.492, \text{ conclude } H_a$$

where the symbol $|\,|$ stands for the absolute value. Since $|t^*| = |(5.7 - 10)/1.6| = |-2.69| = 2.69 > 2.492$, we conclude H_a.

***P*-value for sample outcome.** The P-value for a sample outcome is the probability that the sample outcome could have been more extreme than the observed one when $\mu = \mu_0$. Large P-values support H_0 while small P-values support H_a. A test can be carried out by comparing the P-value with the specified α risk. If the P-value equals or is greater than the specified α, H_0 is concluded. If the P-value is less than α, H_a is concluded.

Example 4. In Example 2, $t^* = 2.58$. The P-value for this sample outcome is the probability $P[t(14) > 2.58]$. From Table A–2, we find $t(.985; 14) = 2.415$ and $t(.990; 14) = 2.624$. Hence, the P-value is between .010 and .015. In fact, it can be shown to be .011. Thus, for $\alpha = .05$, H_a is concluded.

Example 5. In Example 3, $t^* = -2.69$. We find from Table A–2 that $P[t(24) < -2.69]$ is between .005 and .0075. In fact, it can be shown to be .0064. Because the test is two-sided and the t distribution is symmetrical, the two-sided P-value is twice the one-sided value, or $2(.0064) = .013$. Hence, for $\alpha = .02$, we conclude H_a.

Relation between tests and confidence intervals. There is a direct relation between tests and confidence intervals. For example, the two-sided confidence limits (1.55) can be used for testing:

$$H_0: \mu = \mu_0$$
$$H_a: \mu \neq \mu_0$$

If μ_0 is contained within the $1 - \alpha$ confidence interval, then the two-sided decision rule in Table 1.1a, with level of significance α, will lead to conclusion H_0, and vice versa. If μ_0 is not contained within the confidence interval, the decision rule will lead to H_a, and vice versa.

There are similar correspondences between one-sided confidence intervals and one-sided decision rules.

1.7 COMPARISONS OF TWO POPULATION MEANS—NORMAL POPULATIONS

Independent samples

There are two normal populations, with means μ_1 and μ_2, respectively, and with the same standard deviation σ. The means μ_1 and μ_2 are to be compared on the basis of independent samples for each of the two populations:

$$\text{Sample 1: } Y_1, \ldots, Y_{n_1}$$
$$\text{Sample 2: } Z_1, \ldots, Z_{n_2}$$

Estimators of the two population means are the sample means:

(1.57a)
$$\bar{Y} = \frac{\sum_i Y_i}{n_1}$$

(1.57b)
$$\bar{Z} = \frac{\sum_i Z_i}{n_2}$$

and an estimator of $\mu_1 - \mu_2$ is $\bar{Y} - \bar{Z}$.

An estimator of the common variance σ^2 is:

(1.58) $$s^2 = \frac{\sum_i (Y_i - \bar{Y})^2 + \sum_i (Z_i - \bar{Z})^2}{n_1 + n_2 - 2}$$

and an estimator of $\sigma^2(\bar{Y} - \bar{Z})$, the variance of the sampling distribution of $\bar{Y} - \bar{Z}$, is:

(1.59) $$s^2(\bar{Y} - \bar{Z}) = s^2\left[\frac{1}{n_1} + \frac{1}{n_2}\right]$$

We have:

(1.60) $\dfrac{(\bar{Y} - \bar{Z}) - (\mu_1 - \mu_2)}{s(\bar{Y} - \bar{Z})}$ is distributed as t with $n_1 + n_2 - 2$ degrees of freedom when the two independent samples come from normal populations with the same standard deviation.

Interval estimation. The confidence limits for $\mu_1 - \mu_2$ with confidence coefficient $1 - \alpha$ are obtained by means of (1.60):

(1.61) $$(\bar{Y} - \bar{Z}) \pm t(1 - \alpha/2; n_1 + n_2 - 2)s(\bar{Y} - \bar{Z})$$

Example 6. Obtain a 95 percent confidence interval for $\mu_1 - \mu_2$ when:

$$\begin{array}{lll} n_1 = 10 & \bar{Y} = 14 & \Sigma(Y_i - \bar{Y})^2 = 105 \\ n_2 = 20 & \bar{Z} = 8 & \Sigma(Z_i - \bar{Z})^2 = 224 \end{array}$$

We require:

$$s^2 = \frac{105 + 224}{10 + 20 - 2} = 11.75$$

$$s^2(\bar{Y} - \bar{Z}) = 11.75\left(\frac{1}{10} + \frac{1}{20}\right) = 1.7625$$

$$s(\bar{Y} - \bar{Z}) = 1.328$$

$$t(.975; 28) = 2.048$$

$$3.3 = (14 - 8) - 2.048(1.328) \leq \mu_1 - \mu_2 \leq (14 - 8) + 2.048(1.328) = 8.7$$

Tests. One-sided and two-sided tests concerning $\mu_1 - \mu_2$ are constructed by means of (1.60). Table 1.2 contains the decision rules for each of three possible cases, based on the test statistic:

(1.62) $$t^* = \frac{\bar{Y} - \bar{Z}}{s(\bar{Y} - \bar{Z})}$$

with the risk of making a Type I error controlled at α.

TABLE 1.2 Decision rules for tests concerning means μ_1 and μ_2 of two normal populations ($\sigma_1 = \sigma_2 = \sigma$)

Alternatives	*Decision Rule*
	(a)
H_0: $\mu_1 = \mu_2$	If $\lvert t^*\rvert \leq t(1 - \alpha/2; n_1 + n_2 - 2)$, conclude H_0
H_a: $\mu_1 \neq \mu_2$	If $\lvert t^*\rvert > t(1 - \alpha/2; n_1 + n_2 - 2)$, conclude H_a
	where: $t^* = \dfrac{\bar{Y} - \bar{Z}}{s(\bar{Y} - \bar{Z})}$
	(b)
H_0: $\mu_1 \geq \mu_2$	If $t^* \geq t(\alpha; n_1 + n_2 - 2)$, conclude H_0
H_a: $\mu_1 < \mu_2$	If $t^* < t(\alpha; n_1 + n_2 - 2)$, conclude H_a
	(c)
H_0: $\mu_1 \leq \mu_2$	If $t^* \leq t(1 - \alpha; n_1 + n_2 - 2)$, conclude H_0
H_a: $\mu_1 > \mu_2$	If $t^* > t(1 - \alpha; n_1 + n_2 - 2)$, conclude H_a

Example 7. Choose between the alternatives:

$$H_0\colon \mu_1 = \mu_2$$
$$H_a\colon \mu_1 \neq \mu_2$$

when α is to be controlled at .10 and the data are those of Example 6. We require $t(.95; 28) = 1.701$, so that the decision rule is:

$$\text{If } \lvert t^*\rvert \leq 1.701, \text{ conclude } H_0$$
$$\text{If } \lvert t^*\rvert > 1.701, \text{ conclude } H_a$$

Since $\lvert t^*\rvert = \lvert (14 - 8)/1.328\rvert = \lvert 4.52\rvert = 4.52 > 1.701$, we conclude H_a.

The one-sided P-value here is the probability $P[t(28) > 4.52]$. We see from Table A–2 that this P-value is less than .0005. In fact, it can be shown to be .00005. Hence, the two-sided P-value is .0001. For $\alpha = .10$, the appropriate conclusion therefore is H_a.

Paired observations

When the observations in the two samples are paired (e.g., attitude scores Y_i and Z_i for the ith sample employee before and after a year's experience on the job), we use the differences:

$$W_i = Y_i - Z_i \qquad i = 1, \ldots, n \tag{1.63}$$

in the fashion of a sample from a single population. Thus, when the W_i can be treated as observations from a normal population, we have:

(1.64) $\dfrac{\bar{W} - (\mu_1 - \mu_2)}{s(\bar{W})}$ is distributed as t with $n - 1$ degrees of freedom when the differences W_i can be considered to be observations from a normal population and:

$$\bar{W} = \frac{\sum_i W_i}{n}$$

$$s^2(\bar{W}) = \frac{\sum_i (W_i - \bar{W})^2}{n - 1} \div n$$

1.8 INFERENCES ABOUT POPULATION VARIANCE—NORMAL POPULATION

When sampling from a normal population, the following holds for the sample variance s^2 where s is defined in (1.53b):

(1.65) $\dfrac{(n-1)s^2}{\sigma^2}$ is distributed as χ^2 with $n - 1$ degrees of freedom when the random sample is from a normal population.

Interval estimation

The lower confidence limit L and the upper confidence limit U in a confidence interval for the population variance σ^2 with confidence coefficient $1 - \alpha$ are obtained by means of (1.65):

(1.66) $$L = \frac{(n-1)s^2}{\chi^2(1 - \alpha/2; n - 1)} \qquad U = \frac{(n-1)s^2}{\chi^2(\alpha/2; n - 1)}$$

Example 8. Obtain a 98 percent confidence interval for σ^2, using the data of Example 1 ($n = 10$, $s = 4$).

We require:

$$s^2 = 16 \qquad \chi^2(.01; 9) = 2.09 \qquad \chi^2(.99; 9) = 21.67$$

$$6.6 = \frac{9(16)}{21.67} \leq \sigma^2 \leq \frac{9(16)}{2.09} = 68.9$$

Tests

One-sided and two-sided tests concerning the population variance σ^2 are constructed by means of (1.65). Table 1.3 contains the decision rule for each of three possible cases, with the risk of making a Type I error controlled at α.

TABLE 1.3 Decision rules for tests concerning variance σ^2 of normal population

Alternatives	*Decision Rule*
	(a)
H_0: $\sigma^2 = \sigma_0^2$	If $\chi^2(\alpha/2; n-1) \leq \frac{(n-1)s^2}{\sigma_0^2} \leq \chi^2(1-\alpha/2; n-1)$, conclude H_0
H_a: $\sigma^2 \neq \sigma_0^2$	Otherwise conclude H_a
	(b)
H_0: $\sigma^2 \geq \sigma_0^2$	If $\frac{(n-1)s^2}{\sigma_0^2} \geq \chi^2(\alpha; n-1)$, conclude H_0
H_a: $\sigma^2 < \sigma_0^2$	If $\frac{(n-1)s^2}{\sigma_0^2} < \chi^2(\alpha; n-1)$, conclude H_a
	(c)
H_0: $\sigma^2 \leq \sigma_0^2$	If $\frac{(n-1)s^2}{\sigma_0^2} \leq \chi^2(1-\alpha; n-1)$, conclude H_0
H_a: $\sigma^2 > \sigma_0^2$	If $\frac{(n-1)s^2}{\sigma_0^2} > \chi^2(1-\alpha; n-1)$, conclude H_a

1.9 COMPARISONS OF TWO POPULATION VARIANCES—NORMAL POPULATIONS

Independent samples are selected from two normal populations, with means and variances of μ_1 and σ_1^2 and μ_2 and σ_2^2, respectively. Using the notation of Section 1.7, the two sample variances are:

(1.67a)
$$s_1^2 = \frac{\sum_i (Y_i - \bar{Y})^2}{n_1 - 1}$$

(1.67b)
$$s_2^2 = \frac{\sum_i (Z_i - \bar{Z})^2}{n_2 - 1}$$

We have:

(1.68) $\frac{s_1^2}{\sigma_1^2} \div \frac{s_2^2}{\sigma_2^2}$ is distributed as $F(n_1 - 1, n_2 - 1)$ when the two independent samples come from normal populations.

Interval estimation

The lower and upper confidence limits L and U for σ_1^2/σ_2^2 with confidence coefficient $1 - \alpha$ are obtained by means of (1.68):

$$L = \frac{s_1^2}{s_2^2} \frac{1}{F(1 - \alpha/2; n_1 - 1, n_2 - 1)} \tag{1.69}$$
$$U = \frac{s_1^2}{s_2^2} \frac{1}{F(\alpha/2; n_1 - 1, n_2 - 1)}$$

Example 9. Obtain a 90 percent confidence interval for σ_1^2/σ_2^2 when the data are:

$$n_1 = 16 \qquad n_2 = 21$$
$$s_1^2 = 54.2 \qquad s_2^2 = 17.8$$

We require:

$$F(.05; 15, 20) = 1/F(.95; 20, 15) = 1/2.33 = .429$$
$$F(.95; 15, 20) = 2.20$$
$$1.4 = \frac{54.2}{17.8} \frac{1}{2.20} \leq \frac{\sigma_1^2}{\sigma_2^2} \leq \frac{54.2}{17.8} \frac{1}{.429} = 7.1$$

Tests

One-sided and two-sided tests concerning σ_1^2/σ_2^2 are constructed by means of (1.68). Table 1.4 contains the decision rules for each of three possible cases, with the risk of making a Type I error controlled at α.

TABLE 1.4 Decision rules for tests concerning variances σ_1^2 and σ_2^2 of two normal populations

Alternatives	*Decision Rule*
	(a)
H_0: $\sigma_1^2 = \sigma_2^2$	If $F(\alpha/2; n_1 - 1, n_2 - 1) \leq \frac{s_1^2}{s_2^2}$
	$\leq F(1 - \alpha/2; n_1 - 1, n_2 - 1)$, conclude H_0
H_a: $\sigma_1^2 \neq \sigma_2^2$	Otherwise conclude H_a
	(b)
H_0: $\sigma_1^2 \geq \sigma_2^2$	If $\frac{s_1^2}{s_2^2} \geq F(\alpha; n_1 - 1, n_2 - 1)$, conclude H_0
H_a: $\sigma_1^2 < \sigma_2^2$	If $\frac{s_1^2}{s_2^2} < F(\alpha; n_1 - 1, n_2 - 1)$, conclude H_a
	(c)
H_0: $\sigma_1^2 \leq \sigma_2^2$	If $\frac{s_1^2}{s_2^2} \leq F(1 - \alpha; n_1 - 1, n_2 - 1)$, conclude H_0
H_a: $\sigma_1^2 > \sigma_2^2$	If $\frac{s_1^2}{s_2^2} > F(1 - \alpha; n_1 - 1, n_2 - 1)$, conclude H_a

Example 10. Choose between the alternatives:

$$H_0\colon \sigma_1^2 = \sigma_2^2$$
$$H_a\colon \sigma_1^2 \neq \sigma_2^2$$

when α is to be controlled at .02 and the data are those of Example 9.

We require:

$$F(.01; 15, 20) = 1/F(.99; 20, 15) = 1/3.37 = .297$$
$$F(.99; 15, 20) = 3.09$$

so that the decision rule is:

$$\text{If } .297 \leq \frac{s_1^2}{s_2^2} \leq 3.09, \text{ conclude } H_0$$
$$\text{Otherwise conclude } H_a$$

Since $s_1^2/s_2^2 = 54.2/17.8 = 3.04$, we conclude H_0.