PROBABILITY AND STOCHASTIC PROCESS (MAT 2136)

PROBABILITY

Addition rule: If A and B are two events of an experiment having sample space S, then $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

The conditional probability of an event B, given that the event A already taken place is

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

Baye's Theorem:

Let $B_1, B_2, ...B_k$ are the partitions of S with $P(B_i) \neq 0, i = 1, 2, ...k$ and A be any event of S, then $P(B_i / A) = \dfrac{P(A/B_i)P(B_i)}{\sum_{i=1}^{k} P(A/B_i)P(B_i)}$.

The multiplicative rule of probability : $P(A \cap B) = \begin{cases} P(A)P(B|A), if\ P(A) \neq 0 \\ P(B)P(A|B), if\ P(B) \neq 0 \end{cases}$

If $P(A \cap B) = P(A)P(B)$, then A and B are independent.

Random Variable: Let S be the sample of space of a random experiment. Suppose with each element s of S, a unique real number X is associated according to some rule then X is called random variable. There are two types of random variable, i) Discrete and ii) Continuous.

Discrete Random Variable: A random variable X is said to be discrete, if the number of possible values of X is finite or countably infinite. The probability distribution function (pdf) is named as probability mass function (PMF). The Probability mass function is defined as, let X be a random variable, hence the range space R_X of consists of atmost a countably infinite number of values. The probability mass function is defined as

$p(x_i) = \Pr\{X = x_i\}$, satisfying the conditions i) $p(x_i) \geq 0\ for\ all\ i$

ii) $\sum_{i=1}^{k} p(x_i) = 1$.

Continuous Random Variable: A random variable X is said to be continuous if it can take all possible values between certain limits, here the range space of X is infinite. Therefore the probability distribution function named for such random variable is Probability density function (PDF), which is defined as the pdf of X is a function $f(x)$ satisfying the following properties i) $f(x) \geq 0$

ii) $\int_{-\infty}^{\infty} f(x)dx = 1$

iii) $\Pr\{a \leq X \leq b\} = \int_{a}^{b} f(x)dx$ for any a, b such that $-\infty < a < b < \infty$.

Note: 1. If X is a continuous random variable with pdf f(x), then

$P(a < X < b) = P(a \leq X < b) = P(a < X \leq b) = P(a \leq X \leq b) = \int_{a}^{b} f(x)dx$.

2. $P(X = a) = 0$, if X is a continuous random variable.

Cumulative distribution function: Let X be random variable (discrete or continuous), we define F to be the cumulative distribution function of a random variable X given by $F(x) = \Pr\{X \leq x\}$.

Case i) If X is discrete random variable then
$$F(t) = \Pr\{X \leq t\} = P(x_1) + P(x_2) + \cdots + P(t)$$

Case ii) If x is a continuous random variable then $F(x) = \Pr\{X \leq x\} = \int_{-\infty}^{x} f(x)dx$.

Two dimensional random variable: Let E be an experiment and S be a sample space associated with E. Let X=X(s) and Y=Y(s) be two functions each assigning a real number to each outcome s of S. We call (X, Y) to be two dimensional random variable.

Discrete 2D: If the possible values of (X, Y) are finite or countably infinite then (X, Y) is called discrete and it is defined as $P(x_i, y_j)$ satisfying the following condition,

i) $P(x_i, y_j) \geq 0$ and
ii) $\sum_{j=1}^{\infty} \sum_{i=1}^{\infty} P(x_i, y_j) = 1$. The function $P(x_i, y_j)$ defined is called as Joint probability distribution function (Jpdf).

Continuous 2D: If (X, Y) is a continuous random variable assuming all values in some region R of the Euclidean plane, then the Joint probability density function $f(x, y)$ is a function satisfying the following conditions

i) $f(x, y) \geq 0$ for all (x, y)ϵR
ii) $\iint f(x, y)dx\, dy = 1$ over the region R.

Marginal Probability distribution: The marginal probability distribution is defined as

Case i) In the discrete (X, Y), it is defined as $p(x_i) = P\{X = x_i\} = \sum_{j=1}^{\infty} P(x_i, y_j)$ is the marginal probability distribution of X. Similarly $q(y_j) = P\{Y = y_j\} = \sum_{i=1}^{\infty} P(x_i, y_j)$ is the marginal probability distribution of Y.

Case ii) In the continuous (X, Y), it is defined as the marginal probability function of X is defined as $g(x) = \int_{-\infty}^{\infty} f(x, y)dy$ and the marginal probability function of Y is defined as $h(y) = \int_{-\infty}^{\infty} f(x, y)dx$.

To calculate the conditional probability:

Case i) Discrete: Probability of x_i given y_j is defined as $= \frac{P(x_i, y_j)}{q(y_j)}$, $q(y_j) > 0$

Probability of y_j given x_i is defined as $= \frac{P(x_i, y_j)}{p(x_i)}$, $p(x_i) > 0$

Case ii) Continuous: The pdf of X for given Y=y is $= \frac{f(x,y)}{h(y)}, h(y) > 0$

The pdf off Y for given X=x is $= \frac{f(x,y)}{g(x)}, g(x) > 0.$

Independent Random variable: If X and Y are independent random variable then two-dimensional random variable in case of discrete is defined as $P(x_i, y_j) = p(x_i).q(y_j)$ for all the values of i and j. In case of Continuous it is defined as $f(x, y) = g(x).h(y)$.

Mathematical Expectation: If X is a discrete random variable with pmf p(x), then the expectation of X is given by $E(X) = \sum_x xp(x)$, provided the series is absolutely convergent.

If X is continuous with pdf f(x), then the expectation of X is given by $E(X) = \int xf(x)dx$, provided $\int |x|f(x)dx < \infty$.

Variance of X is given by $V(X) = E(X - E(X))^2 = E(X^2) - \left(E(X)\right)^2$.

Chebyshev's inequality: Let X be random variable with mean μ and variance σ^2 then for any positive real number $k(k > 0)$.

$P\{\, |x - \mu| \geq k\} \leq \frac{\sigma^2}{k^2}$ (upper bound)

$P\{\, |x - \mu| < k\} > 1 - \frac{\sigma^2}{k^2}$ (lower bound)

Note: some other forms

1. $P\{\, |x - \mu| \geq k\sigma\} \leq \frac{1}{k^2}$ and $P\{\, |x - \mu| < k\sigma\} \geq 1$ (upper bound
2. $P\{\, |x - \mu| \geq \epsilon\} \leq \frac{1}{\epsilon^2} E(x - c)^2$ and $P\{\, |x - \mu| < \epsilon\} \geq 1$

DISTRIBUTIONS:

Distribution	PMF/PDF	Mean	Variance
Binomial distribution $X \sim B(n, p)$	$P(x) = {}^{n}C_k\, p^k (1-p)^{n-k} \quad , k = 0,1,2,\ldots,n$	$E(x) = np$	$V(x) = np(1-p)$
Poisson's Distribution $X \sim P(\alpha)$	$P(x) = \dfrac{e^{-\alpha}\alpha^k}{k!}, \ k = 0,1,2,\ldots,\ \alpha > 0$	$E(x) = \alpha = np$	$V(x) = \alpha = np$
Uniform Distribution $X \sim U(a,b)$	$f(x) = \begin{cases} \dfrac{1}{b-a}, & a < x < b \\ 0, & otherwise \end{cases}$	$E(x) = \dfrac{b+a}{2}$	$V(x) = \dfrac{(b-a)^2}{12}$
Normal Distribution $X \sim N(\mu, \sigma^2)$	$f(x) = \dfrac{1}{\sigma\sqrt{2\pi}} e^{\frac{-1}{2}\frac{(x-\mu)^2}{\sigma^2}}, \quad -\infty < x, \mu < \infty,\ \sigma > 0$	$E(x) = \mu$	$V(x) = \sigma^2$
Exponential Distribution $X \sim E(\lambda)$	$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x > 0 \\ 0, & otherwise \end{cases}$	$E(x) = \dfrac{1}{\lambda}$	$V(X) = \dfrac{1}{\lambda^2}$
Gamma Distribution $X \sim G(r, \alpha)$	$f(x) = \begin{cases} \dfrac{x^{r-1}e^{-\alpha x}\alpha^r}{\Gamma(r)}, & x > 0,\ \alpha, r > 0 \\ 0, & elsewhere \end{cases}$	$E(x) = \dfrac{r}{\alpha}$	$V(x) = \dfrac{r}{\alpha^2}$

Chi-square Distribution $X \sim \chi^2(n)$	$f(x) = \begin{cases} \dfrac{x^{\frac{n}{2}-1} e^{-\frac{x}{2}}}{\Gamma(n/2)2^{\frac{n}{2}}}, & x > 0 \\ 0, & elsewhere \end{cases}$	$E(x) = n$	$V(x) = 2n$

Uniform distribution on a two dimensional set: If R is a set in the two-dimensional plane, and R has a finite area, then we may consider the density function equal to the reciprocal of the area of R inside R, and equal to 0 otherwise:

$$f(x,y) = \begin{cases} \frac{1}{area\ R}; & if\ (x,y) \in R \\ 0 & Otherwise \end{cases}.$$

Covariance: $Cov(x,y) = E(xy) - E(x)E(y)$

Correlation coefficient: $\rho_{xy} = \rho = \dfrac{E(xy) - E(x)E(y)}{\sqrt{V(x)V(y)}}$

Properties:

1. $E(c) = c$, where c is a constant.
2. $V(c) = 0$, where c is a constant.
3. If $E(xy) = 0$ then x and y are orthogonal.
4. $V(Ax+b) = A^2V(x)$ when Ax+B is linear function of x.
5. If $\rho = 0$ then X and Y are un correlated.
6. $V(Ax+by) = A^2V(x) + B^2V(y) + 2ABCOV(x,y)$

FUNCTIONS OF ONE DIMENSIONAL RANDOM VARIABLES

Let S be a sample space associated with a random experiment E, then it is known that a random variable X on S is a real valued function, i.e., $X: S \to R$, for each element s $\in$ S, there is a real number associated.

Let X be a random variable defined on S. Let $y = H(x)$ is a real valued function of x. Then $Y = H(X)$ is a random variable on S. i.e., for each element s $\in$ S, there is a real number associated, say $y = H(X(s))$. Here Y is called a function of the random variable X.

Notations:

1. R_X – the set of all possible values of the function X, called the **range space** of the random variable X.
2. R_Y – the set of all possible values of the function $Y = H(X)$, called the **range space** of the random variable Y.

Equivalent Events: Let C be an event associated with the range space R_Y. Let $B \subset R_X$ defined by $B = \{x \in R_X; H(x) \in C\}$, then B and C are called equivalent events.

Distribution function of functions of random variables:

Case 1: Let X be a discrete random variable with p.m.f. $p(x_i) = P(X = x_i)$ for $i = 1,2,3, ...$ Let $Y = H(X)$ then Y is also a discrete random variable. If $Y = H(X)$ is a one to one function then the probability distribution of Y is as follows:

For the possible values of $y_i = H(x_i)$ for $i = 1,2,3, ...$ The p.m.f. of $Y = H(X)$ is $q(y_i) = P(Y = y_i) = P(X = x_i) = p(x_i)$ for $i = 1,2,3, ...$

Case 2: Let X be a discrete random variable with p.m.f. $p(x_i) = P(X = x_i)$ for $i = 1,2,3, ...$ Let $Y = H(X)$ then Y is also a discrete random variable. Suppose tat for one value of $Y = y_i$ there corresponds several values of X say $x_{i_1}, x_{i_2}, ..., x_{i_j}, ...$ then the p.m.f. of of $Y = H(X)$ is

$$q(y_i) = P(Y = y_i) = p\left(x_{i_1}\right) + p\left(x_{i_2}\right) + \cdots + p\left(x_{i_j}\right) + \cdots$$

Case 3: Let X be a continuous random variable with p.d.f. $f(x)$. Let $Y = H(X)$ be a discrete random variable. Then if the set $\{Y = y_i\}$ is equivalent to an event $B \subseteq R_X$ then the p.m.f. of Y is

$$q(y_i) = P(Y = y_i) = \int_B f(x)\, dx$$

Case 4: Let X be a continuous random variable with p.d.f. $f(x)$. Let $Y = H(X)$ be a continuous random variable. Then the p.d.f. of Y, say g is obtained by the following procedure:

Step 1: Obtain the c.d.f. of Y, $\quad G(y) = P(Y \leq y)$, by finding the event $A \subseteq R_X$, which is equivalent to the event $\{Y \leq y_i\}$.

Step 2: Differentiate $G(y)$ with respect to y to get $g(y)$.

Step 3: Determine those values of y in R_Y for which $g(y) > 0$.

Theorem: Let X be a continuous random variable with p.d.f. $f(x)$ where $f(x) > 0$ for $a < x < b$. Suppose that $Y = H(X)$ is strictly monotonic function on $[a, b]$. Then the p.d.f. of the random variable $Y = H(X)$ is given by

$$g(y) = f(x)\left|\frac{dx}{dy}\right|$$

If $Y = H(X)$ is strictly increasing then $g(y) > 0$ for $H(a) < y < H(b)$.

If $Y = H(X)$ is strictly decreasing then $g(y) > 0$ for $H(b) < y < H(a)$.

Theorem: Let X be a continuous random variable with p.d.f. $f(x)$. Let $Y = X^2$ then the p.d.f. of Y is

$$g(y) = \frac{1}{2\sqrt{y}}\left[f(\sqrt{y}) + f(-\sqrt{y})\right]$$

FUNCTIONS OF TWO DIMENSIONAL RANDOM VARIABLES

Let (X, Y) be a two dimensional continuous random variable. Let $Z = H(X, Y)$ be a continuous function of X and Y then $Z = H(X, Y)$ is a continuous one dimensional random variable.

To find the p.d.f. of Z, we introduce another suitable random variable say,

$W = G(X, Y)$ and obtain the joint p.d.f. of the two dimensional random variable (Z, W), say $k(z, w)$. From this distribution, the p.d.f. of Z can be obtained by integrating k with respect to w.

Theorem: Suppose (X, Y) is a two dimensional continuous random variable with joint p.d.f. $f(x, y)$ defined on a region R of the XY-plane. Let $Z = H_1(X, Y)$ and $W = H_2(X, Y)$. Suppose that H_1 and H_2 satisfies the following conditions;

(i) $z = H_1(x, y)$ and $w = H_2(x, y)$ may be uniquely solved for x, y in terms of z & w say, $x = G_1(z, w)$ and $y = G_2(z, w)$.

(ii) The partial derivatives $\frac{\partial x}{\partial z}, \frac{\partial x}{\partial w}, \frac{\partial y}{\partial z}, \frac{\partial y}{\partial w}$ exist and are continuous

Then the joint p.d.f. of (Z, W) say $k(z, w)$ is given by,

$$k(z,w) = f[G_1(z,w), G_2(z,w)]|J(z,w)|$$

where $J(z,w) = \begin{vmatrix} \frac{\partial x}{\partial z} & \frac{\partial x}{\partial w} \\ \frac{\partial y}{\partial z} & \frac{\partial y}{\partial w} \end{vmatrix}$ is called the Jacobian of the transformation

$(x,y) \mapsto (z,w)$. Also, $k(z,w) > 0$ for those values of (z,w) corresponding to the values of (x,y) for which $f(x,y) > 0$.

MOMENT GENERATING FUNCTION (M.G.F.) OF ONE DIMENSIONAL RANDOM VARIABLES

Let X be any one dimensional random variable then the mathematical expectation $E(e^{tX})$ if exists then it is called the moment generating function (m.g.f.) of X. i.e., $M_X(t) = E(e^{tX})$

In particular, if X is discrete then, $M_X(t) = \sum_{i=1}^{i=\infty} e^{tx_i} P(X = x_i)$.

If X is continuous then, $M_X(t) = \int_{-\infty}^{\infty} e^{tx} f(x)\, dx$.

Properties of m.g.f.: Let X be any one dimensional random variable and $M_X(t)$ be the m.g.f. of X then

1. $M_X^n(0) = E(X^n)$ where $M_X^n(0)$ is the n^th derivative of $M_X(t)$ at $t = 0$.
   i.e.; $M_X'(0) = E(X)$
   $M_X''(0) = E(X^2)$
2. $V(X) = M_X''(0) - \left(M_X'(0)\right)^2$
3. Let X be any one dimensional random variable and $M_X(t)$ be the m.g.f. of X. Let $Y = \alpha X + \beta$. Then the m.g.f. of Y is $M_Y(t) = e^{\beta t} M_X(\alpha t)$.
4. Suppose that X and Y are independent random variables. Let $Z = X + Y$. Let $M_X(t), M_Y(t)$ and $M_Z(t)$ be the m.g.f.'s of the random variables X, Y and Z respectively. Then $M_Z(t) = M_X(t) M_Y(t)$
5. Let $X_1, X_2, \dots, X_n$ be n independent random variables which follows a normal distribution $N(\mu_i, \sigma_i^2)$ for $i = 1,2,3,..,n$. Let $Z = X_1 + X_2 + \cdots + X_n$ then $Z \to N(\mu_1 + \mu_2 + \cdots + \mu_n,\ \sigma_1^2 + \sigma_2^2 + \cdots + \sigma_n^2)$.
6. Let $X_1, X_2, \dots, X_n$ be n independent random variables which follows a Poisson distribution with parameter α_i for $i = 1,2,..,n$. Let $Z = X_1 +$

$X_2 + \cdots + X_n$ then Z has a Poisson distribution with parameter $\alpha = \alpha_1 + \alpha_2 + \cdots + \alpha_n$.

7. Let $X_1, X_2, \dots, X_k$ be k independent random variables which follows a Chi-square distribution with degrees of freedom n_i for $i = 1,2,3,..,k$. Let $Z = X_1 + X_2 + \cdots + X_k$ then Z has a Chi-square distribution with degrees of freedom $n = n_1 + n_2 + \cdots + n_k$.
8. Let $X_1, X_2, \dots, X_k$ be k independent random variables, each having distribution $N(0,1)$. Then $S = X_1^2 + X_2^2 + \cdots + X_k^2$ has a Chi-square distribution with degrees of freedom k.
9. Let $X_1, X_2, \dots, X_r$ be r independent random variables, each having exponential distribution with the same parameter α. Let $Z = X_1 + X_2 + \cdots + X_r$ then Z has a Gamma distribution with parameters α and r.
10. Let $X_1, X_2, \dots, X_n, \dots$ be a sequence of random variable with c.d.f.'s $F_1, F_2, \dots, F_n, \dots$ and m.g.f.'s $M_1, M_2, \dots, M_n, \dots$ Suppose that $\lim_{n\to\infty} M_n(t) = M(t)$, where $M(0) = 1$. Then $M(t)$ is the m.g.f. of the random variable X whose c.d.f is $F = \lim_{n\to\infty} F_n(t)$.

MGF of some standard distributions:

1. **Binomial Distributions:** $\boldsymbol{M_X(t)} = M_X(t) = (pe^t + q)^n$
2. **Poisson Distributions:** $\boldsymbol{M_X(t)} = e^{\propto(e^t - 1)}$
3. **Normal Distributions:** $\boldsymbol{M_X(t)} = e^{t\mu + \frac{\sigma^2 t^2}{2}}$
4. **Exponential Distributions:** $\boldsymbol{M_X(t)} = \frac{\propto}{\propto - t}$
5. **Gamma Distributions:** $\boldsymbol{M_X(t)} = \frac{\propto^r}{(\propto - t)^r}$
6. **Chi square Distributions:** $\boldsymbol{M_X(t)} = (1 - 2t)^{-n/2}$

SAMPLING

In statistical investigation, the characteristics of a large group of individuals (called population) is studied. Sampling is a study of the relationship between a population and samples drawn from it.

The population mean and the population variance are denoted by μ and σ^2 respectively.

Sample mean and sample variance: Let X be the random variable which denotes the population with mean μ and variance σ^2. Let $(X_1, X_2, \dots, X_n)$ be a random sample of size n from X. Then,

Sample mean, $\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$ and

Sample variance, $s^2 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n}$

- If $X \to N(\mu, \sigma^2)$ then $\bar{X}$ and s^2 are independent random variables.
- Let X be the random variable with $E(X) = \mu$ and $V(X) = \sigma^2$. Let $(X_1, X_2, \dots, X_n)$ be a random sample of size n from X. Then, $E(\bar{X}) = \mu$ and $V(\bar{X}) = \frac{\sigma^2}{n}$.
- Let $X \to N(\mu, \sigma^2)$ then $\bar{X} \to N(\mu, \frac{\sigma^2}{n})$ and $s^2 \to \chi^2(n-1)$.

Central Limit Theorem: Let $X_1, X_2, \dots, X_n$ be n independent random variables all of which have the same distribution. Let $\mu = E(X_i)$ and $\sigma^2 = V(X_i)$ be the common expectation and variance. Let $S = \sum_{i=1}^{n} X_i$ then $E(S) = n\mu$ and $V(S) = n\sigma^2$ then for large values of n, the random variable $T_n = \frac{S - E(S)}{\sqrt{V(S)}}$ has approximately the distribution $N(0,1)$.

Testing of Hypothesis:

The central limit theorem is used for testing of hypothesis. The purpose of hypothesis testing is to determine whether there is enough statistical evidence in favor of a certain belief, or hypothesis, about a parameter. For the hypothesis tests, the law of probability is assumed to be known, so the sampling distribution is perfectly known and we take a sample to define a decision criterion which help us to accept or reject the hypothesis.

Using the mean, we test a hypothesis H0. This is referred to as the null hypothesis. It is an assumption made on the probability distribution X. The alternate hypothesis is denoted by H1.

The error of the first kind is called Type I error. The error of the second kind is called Type II error. Thus a Type I error is an error in a statistical test which

	Accept H_0	Reject H_0
H_0 is true	Right decision	Wrong decision (error of the first kind α)
H_0 is false	Wrong decision (error of the second kind β)	Right decision

occurs when a false hypothesis is accepted (a false positive in terms of the null hypothesis) and a Type II error is an error in a statistical test which occurs when a true hypothesis is rejected (a false negative in terms of the null hypothesis). Note that the acceptance of H_1 when H_0 is true is called a **Type I error**. The probability of committing a Type I error is called the level of significance and is denoted by α. Also, the failure to reject H_0 when H_1 is true is called a **Type II error**. The probability of committing a Type II error is denoted by β. The probability $1 - \beta$ is called the **power of a test**; it is the probability of taking the correct action of rejecting the null hypothesis when it is false. By increasing n, we can improve the power of a test. For the same α and the same n, the power of test is also used to choose between different tests; a more powerful test is one that yields the correct action with greater frequency.

A statistical hypothesis test may return a value called the p-value. The p-value is the minimum probability of a Type I error with which H_0 can still be rejected. This is a quantity that we can use to interpret or quantify the result of the test and either reject or fail to reject the null hypothesis. This is done by comparing the p-value to a threshold value chosen beforehand called the significance level α.

If p-value > α: Fail to reject the null hypothesis (not significant result).

If p-value <= α: Reject the null hypothesis (significant result).

Some tests do not return a p-value. Instead, they might return a list of critical values and their associated significance levels, as well as a test statistic. The results are interpreted in a similar way. Instead of comparing a single p-value to a pre-specified significance level, the test statistic is compared to the critical value at a chosen significance level.

If test statistic < critical value: Fail to reject the null hypothesis.

If test statistic >= critical value: Reject the null hypothesis.

Moving the critical value provides a trade-off between α and β. A reduction in β is always possible by increasing the size of the critical region, but this increases α. Likewise, reducing α is possible by decreasing the critical region. Note that α and β are related in such a way that decreasing one generally increases the other. This problem is solved with the help of sample size. Both α and β can be reduced simultaneously by increasing the sample size.

Consider H0: $\theta = \theta_0$ vs H_1: $\theta > \theta_0$. This is a one-tailed test with the critical region in the right-tail of the test statistic X. Another one-tailed test could have the form, H_0: $\theta = \theta_0$ vs H_1: $\theta < \theta_0$, in which the critical region is in the left-tail. In a two-tailed test we have : H_0: $\theta = \theta_0$ vs H_1: θ not equal to θ_0.

The first type of test is the most basic: testing the mean of a distribution in which we already know the population variance. Let μ and σ be the mean and standard deviation of X. We take from the population a sample of size n large enough. The sample mean is $\bar{x}$. If $\bar{x}$ is between $\mu - t_\infty . (\sigma/\sqrt{n})$ and $\mu + t_\infty . (\sigma/\sqrt{n})$ then H_0 is accepted. However if $\bar{x}$ is outside of those values, the null hypothesis is rejected (two-tailed test).

Our test statistic is

$$z = \frac{(\bar{x} - \mu)}{(\sigma/\sqrt{n})}$$

where n is the number of observations made when collecting the data for the study, and μ is the true mean when we assume the null hypothesis is true. So to test a hypothesis with given significance level α, we calculate the critical value of z (or critical values, if the test is two-tailed) and then check to see whether or not the value of the test statistic in is in our critical region. This is called a z-test. We are most often concerned with tests involving either $\alpha = .05$ or $\alpha = .01$. When we construct our critical region, we need to decide whether or not our hypotheses in question are one-tailed or two-tailed. If one-tailed, we reject the null hypothesis if $z \geq z\alpha$ (if the hypothesis is right-handed) or if $z \leq z\alpha$ (if the hypothesis is left-handed). If two-tailed, we reject the null hypothesis if $|z| \geq z\alpha/2$. If we do not know the population variance, and if n is large ($n \geq 30$) it suffices for most distributions commonly encountered to replace the unknown population variance with the modified definition of sample variance.

Critical Region: To construct a critical region of size α, we first examine our alternative hypothesis. If our hypothesis is one-tailed, our critical region is either $z \geq z\alpha$ (if the hypothesis is right-handed) or $z \leq -z\alpha$ (if the hypothesis is left-handed). If our hypothesis is two-tailed, then our critical region is $|z| \geq z\alpha/2$.

Student 't' distribution:

If we have a sample of size n from a normal distribution with mean μ and unknown variance, we study $t = (x-\mu)/(s/\sqrt{n})$ and compare this to the Student t-distribution with $(n-1)$ degrees of freedom. If $n \geq 30$ then by the Central Limit Theorem we may instead compare it to the standard normal case.

Thus when we have a small sample size ($n < 30$) taken from a normal distribution of unknown variance, we use the t-test with $(n - 1)$ degrees of freedom.

Critical Region: To construct a critical region of size α, we first examine our alternative hypothesis. If our hypothesis is one-tailed, our critical region is either $t \geq t\alpha, (n-1)$ (if the hypothesis is right-handed) or $t \leq -t\alpha, (n-1)$ (if the hypothesis is left-handed). If our hypothesis is two-tailed, then our critical region is $|t| \geq t\alpha/2, (n-1)$.

Chi-Square Test:

The chi-square test gives a p-value. The p-value tells us whether the test results are significant or not. The chi-square statistic is given by

$$\chi^2 = \sum_{j=1}^{n} \frac{(O_j - E_j)^2}{E_j}$$

where O_j = each Observed (actual) value, E_j= each Expected value.

Calculating the chi-square statistic and comparing it against a critical value from the chi-square distribution allows us to assess whether the observed cell counts in a table are significantly different from the expected cell counts.

Estimation:

Estimator (or estimate) $\hat{\theta}$ for the unknown parameter θ associated with the distribution of a random variable x is called unbiased estimator (or unbiased estimate) if $E(\hat{\theta}) = \theta$ for all θ.

An unbiased estimate of the variance σ^2 of a random variable based on a sample $x_1, \dots, x_n$ is

$$\sigma^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2.$$

Maximum likelihood estimate (M.L.E.):- Based on a random sample $x_1, \dots, x_n$ the M.L.E. $\hat{\theta}$ of θ is that value of θ which maximizes $L(x_1, \dots, x_n; \theta) = f(x_1; \theta)f(x_2; \theta) \cdots f(x_n; \theta)$ when $f(x; \theta)$ is either the p.d.f of x.

Confidence intervals:

Once the sample is observed and the sample mean computed to equal $\bar{x}$, this interval $\left[\bar{x} - z\alpha_{/2}\left(\frac{\sigma}{\sqrt{n}}\right), \bar{x} + z\alpha_{/2}\left(\frac{\sigma}{\sqrt{n}}\right)\right]$ is a known interval for the unknown mean μ.

For example, $\bar{x} \pm 1.96\left(\frac{\sigma}{\sqrt{n}}\right)$ is 95% confidence interval for μ. The number $100(1-\alpha)\%$ or $1-\alpha$ is confidence coefficient.

Let a and b be constants selected for the random sample $x_1, \dots, x_n$. The confidence interval for the variance σ^2 based on the sample variance

$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$ with $100(1-\alpha)\%$

Confidence is given by

$$\left[\sqrt{\frac{(n-1)s^2}{b}}, \sqrt{\frac{(n-1)s^2}{a}}\right] = \left[\sqrt{\frac{(n-1)}{b}}s, \sqrt{\frac{(n-1)}{a}}s\right], \text{ if } P\left(a \le \frac{(n-1)s^2}{\sigma^2} \le b\right) = 1-\alpha.$$

STOCHASTIC PROCESS

1 RANDOM PROCESS

A stochastic process (or Random process) is a probabilistic model used for characterizing the random signals.

The concept of random process is an extension of the idea of a random variable to include time.

Definition 1.1. *Let S be the sample space of a random experiment. A Stochastic process is a mapping $X : S \to \mathbb{R}$ which assigns every outcome $s \in S$ a real valued function of time $x(t, s)$ i.e., $X(s) = x(t, s)$. The collection of all such time functions is denoted by $X(t, s)$, is called a random process.*

Notation 1.1. *We denote a random process by $X(t, s) = \{x(t, s)\} = \{X_t; t \in T\}$*

A one dimensional stochastic process can be classified into one of the following four types;

1. Discrete time discrete state space (ie; both $X(t)$ and t are discrete)
2. Continuous time discrete state space (ie; $X(t)$ is discrete and t is continuous)
3. Continuous time continuous state space (ie; both $X(t)$ and t is continuous)
4. Discrete time continuous state space (ie; t is discrete and $X(t)$ is continuous)

Process with independent increments

If for every $t_1, t_2, ..., t_n$ with $t_1 < t_2 < ... < t_n$ the random variables $X(t_1) - X(t_2), X(t_3) - X(t_2), ..., X(t_n) - X(t_{n-1})$ are independent then $\{X_t; t \in T\}$ is said to be a process with independent increments.

The **mean** of the random process $X(t)$ is the expected value of the random variable $X(t)$ at any time t i.e., $\mu_X(t) = E[X(t)]$ for $-\infty < t < \infty$.

Mean is a function of time t so it is called a mean function.

Correlation coefficient: The correlation coefficient of two random variables

X and Y is given by $\rho = \dfrac{E(XY) - E(X)E(Y)}{\sqrt{V(X)V(Y)}} = \dfrac{COV(X, Y)}{\sqrt{V(X)V(Y)}}$.

Note: Random variables X and Y are uncorrelated if the covariance

$COV(X,Y) = 0$.

Auto Correlation Function(ACF)

The **auto correlation function** of a stochastic process $\{X_t; t \in T\}$ is defined as,

$R_{XX}(t_1, t_2) = E\{X(t_1)X(t_2)\}$

The **average power of** $X(t)$ is defined as $R(t,t) = E\{X(t)X(t)\} = E\{X^2(t)\}$.

Auto covariance is defined as $C(t_1, t_2) = R_{XX}(t_1, t_2) - E\{X(t_1)\}E\{X(t_2)\}$.

So, $C(t,t) = R(t,t) - E\{X(t)\}E\{X(t)\} = (E\{X(t)\})^2$

$= E\{X^2(t)\} - (E\{X(t)\})^2 = V\{X(t)\}$.

Stationary process

A stochastic process $\{X_t; t \in T\}$ is said to be a stationary process of order "n" if for arbitrary $t_1, t_2, ..., t_n$ the joint p.d.f of $(X(t_1), X(t_2), X(t_3), ..., X(t_n))$ and $(X(t_1+h), X(t_2+h), X(t_3+h), ..., X(t_n+h))$ are same for every $h > 0$.

In particular, if we need to say that a random process is said to be first order stationary if its first order probability density function does not change with a shift in time origin. i.e., $f_X(x, t_1) = f_X(x, t_1 + h)$ for any time t_1 and any real number h.

Remark 1.1. *A first order stationary random process has a constant mean.*

Strictly stationary stochastic process

The stationary process is said to be **strictly stationary stochastic process** (SSS) if it is stationary of order "n" for any $n \geq 1$.

The stationarity of a process is the probabilistic structure of the process, is invariant under the translation of the time axis. For SSS the statistics properties of the associated joint distribution of $(X(t_1+h), X(t_2+h), X(t_3+h), ..., X(t_n+h))$ is independent of "h".

Wide Sense Stationary

A stochastic process $\{X_t; t \in T\}$ with finite second order momemts is called wide sense stationary (WSS) [or covariance stationary or weakly stationary] if its mean $E\{X(t)\} = m$, a constant and its auto correlation $R(t_1, t_2) = E\{X(t_1)X(t_2)\}$ depends only on the difference $|t_1 - t_2|$ for every t_1, t_2.

Evolutionary Process

A process which is not stationary in any sense is called an evolutionary process.

2 MARKOV PROCESS

Markov process is a special class of random process, which represent the simplest generalization of independent process which allow the outcome at any instant to depend only on the outcome that precedes it and not on the earlier ones. i.e., in a Markov process the past has no influence on the future if the present is specified.

Definition 2.1. *A random process $X(t)$ is said to be a* ***Markov process*** *if for any $t_1, t_2, ..., t_n$*

$$P\left[X(t_n) \leq x_n | X(t_{n-1}) \leq x_{n-1}, X(t_{n-2}) \leq x_{n-2}, ...X(t_1) \leq x_1\right] = P\left[X(t_n) \leq x_n | X(t_{n-1}) \leq x_{n-1}\right]$$

i.e., the conditional probability distribution of $X(t_n)$ for given values of $X(t_1), X(t_2), ..., X(t_{n-1})$ depends only on $X(t_{n-1})$.

Remark 2.1. *The above definition of Markov Process will hold for continuous Markov process. But we are discussing about discrete state Markov process known as* ***Markov Chains*** *where the system can occupy only a finite or countable number of states.*

Definition 2.2. ***Markov Chains:*** *Let $X(t)$ be a Markov process with states $X(t_i) = X_i$ where $t_0 < t_1 < t_2 < ... < t_n$. If for all n,*

$$P\left[X_n = a_n | X_{n-1} = a_{n-1}, X_{n-2} = a_{n-2}, ...X_0 = a_0\right] = P\left[X_n = a_n | X_{n-1} = a_{n-1}\right]$$

then the sequence of random variables $\{X_n\}$ is called a Markov Chain for $n = 0, 1, 2, 3, ...$ and $a_1, a_2, ..., a_n$ are the states of the Markov Chain.

Definition 2.3. ***Transition probabilities and Transition Matrix:*** *Let $A = \{a_1, a_2, ..., a_n\}$ be the state space of a Markov chain, and for any two states a_i, a_j, let p_{ij} denote the conditional probability that $X_{r+1} = a_j$ given that $X_r = a_i$.*

$$p_{ij} = P\left[X_{r+1} = a_j | X_r = a_i\right]$$

where $X_{r+1} \& X_r = a_i$ are any two successive random variables present in the process.The probabilities p_{ij} are called ***transition probabilities****.*

The square matrix P whose elements are the transition probabilities p_{ij} is called a ***transition probability matrix or transition matrix****.*

$$P = \begin{pmatrix} p_{11} & p_{12} & p_{13} & \cdots & p_{1n} \\ p_{21} & p_{22} & p_{23} & \cdots & p_{2n} \\ \cdots\cdots\cdots\cdots & & & & \\ \cdots\cdots\cdots & & & & \\ & & & & \\ p_{m1} & p_{m2} & p_{m3} & \cdots & p_{mn} \end{pmatrix}$$

Remark 2.2. *In P it is clear that, $p_{ij} \geq 0$ for all $i, j = 1, 2, 3, .., n$ and $\Sigma_{j=1}^{n} p_{ij} = 1$ for $i = 1, 2, 3.., n$.*

Definition 2.4. ***Stochastic Matrix*** *A transition probability matrix $P = [p_{ij}]$ with the above two properties is called a **stochastic matrix.***

Definition 2.5. *A row matrix of the form $Q = \begin{pmatrix} q_1 & q_2 & q_3 & ... & q_n \end{pmatrix}$ where each $q_i \geq 0$ and $\Sigma_{i=1}^{n} q_i = 1$ is called a **probability vector of order n.***

Remark 2.3. *Every row of a stochastic matrix is a probability vector.*

Definition 2.6. *Let $\{X_r\}$ be a Markov chain.*

1. ***Higher transition probabilities:*** *Given that a random variable X_r in the Markov chain takes the value a_i then for some positive integer $k \geq 1$ the probability*

$$p_{ij}^{(k)} = P\left[X_{r+k} = a_j | X_r = a_i\right]$$

   *Here, $p_{ij}^{(k)}$ is called the **k-step transition probabilities.***

2. *It is clear that $p_{ij} = p_{ij}^{(1)}$ are the **1-step transition probabilities.***

3. *The matrix $P^{(k)} = \left[p_{ij}^{(k)}\right]$ is called the **k-step transition matrix.***

4. *$P^{(k)}$ is identically equal to P^k, the k^{th} power of P. i.e., the k-step transition matrix is precisely equal to the k^{th} power of the 1-step transition matrix.*

$$P^{(k)} = \left[p_{ij}^{(k)}\right] = [p_{ij}]^k = P^{(k)}$$

5. *A Stochastic matrix P is said to be **regular** if all the entries in some positive integral power of P are positive.*

   i.e., the transition matrix $P = [p_{ij}]$ is regular if and only if $p_{ij}^{(k)} > 0$ for all possible values of i and j and for some $k \geq 1$.

6. *A Markov chain is said to be irreducible if its transition matrix is a **regular stochastic matrix.***

7. *For a given regular stochastic matrix P of order m, if there exists a probability vector Q of order m such that $QP = Q$ then Q is called a **fixed probability vector** of P. Such vector Q exists and is unique.*

3 Classification of States

Let $A = \{a_1, a_2, ..., a_n\}$ be the state space of a Markov chain $\{X_r\}$. When $\{X_r\} = a_i$ which means that the chain is in the $a_i$state (i.e., visiting state a_i state or in the i^{th} state) at the r^{th} step.

Note 1. *p_{ij} denotes the probability that the system (or Chain) makes a transition (or moves) from the state i to j in one-step.*

Similarly, $p_{ij}^{(k)}$ denotes the probability that the system (or Chain) makes a transition (or moves) from the state i to j in k-steps.

Definition 3.1.

1. ***Absorbing State:*** *A state i is said to be an absorbing state if the transition probabilities p_{ij} are such that*

$$p_{ij} = \begin{cases} 1, & \text{for } i = j \\ 0, & \text{else where} \end{cases}$$

2. ***Transient State:*** *A state i is said to be a transient state if the chain is in this state at some step and there is a chance (i.e., there is a non-zero probability) that it will not return to that state.*

   For example, if we model a Markov chain all the states except the final state are transient states.

3. ***Recurrent State:*** *A state i is said to be a recurrent state if starting from the state i the chain does eventually return to the same state (i.e., the probability of the return is 1).*

4. ***Periodic State:*** *Let i be a recurrent state so that $p_{ii}^{k} > 0$ for some $k \geq 1$. Let $d_i = \gcd\{k \in \mathbb{Z}^{+}; p_{ii}^{k} > 0\}$. Then d_i is called the **period** of te state i.*

   *The recurrent state i is said to be **periodic** if $d_i > 1$ and **aperiodic** if $d_i = 1$*

5. *All the states of an irreducible Markov Chain are of same type.*

6. *If one of the state in a irreducible Markov chain is aperiodic then all the states are aperiodic then such Markov chains are called an **aperiodic Markov chain.***

Appendix D

Tables of Distributions

Prior to the current age of computing, probability tables for certain distributions were part of many text books in probability and statistics. These are not needed any longer. Most statistical computing packages offer easy-to-use calls to determine these probabilities and quantiles. This is certainly true of the language R as we have discussed through out this text. Also, many hand calculators have such functions.

Tables for the following distributions are presented:

Table I Selected quantiles for chi-square distributions.

Table II Cumulative distribution function for the standard normal random variable.

Table III Selected quantiles for t-distributions.

Table IV Selected quantiles for F-distributions.

Table I
Chi-Square Distribution

The following table presents selected quantiles of chi-square distribution, i.e., the values x such that

$$P(X \leq x) = \int_0^x \frac{1}{\Gamma(r/2)2^{r/2}} w^{r/2-1} e^{-w/2}\, dw,$$

for selected degrees of freedom r. The R function `chistable.s` generates this table.

	$P(X \leq x)$							
r	0.010	0.025	0.050	0.100	0.900	0.950	0.975	0.990
1	0.000	0.001	0.004	0.016	2.706	3.841	5.024	6.635
2	0.020	0.051	0.103	0.211	4.605	5.991	7.378	9.210
3	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.345
4	0.297	0.484	0.711	1.064	7.779	9.488	11.143	13.277
5	0.554	0.831	1.145	1.610	9.236	11.070	12.833	15.086
6	0.872	1.237	1.635	2.204	10.645	12.592	14.449	16.812
7	1.239	1.690	2.167	2.833	12.017	14.067	16.013	18.475
8	1.646	2.180	2.733	3.490	13.362	15.507	17.535	20.090
9	2.088	2.700	3.325	4.168	14.684	16.919	19.023	21.666
10	2.558	3.247	3.940	4.865	15.987	18.307	20.483	23.209
11	3.053	3.816	4.575	5.578	17.275	19.675	21.920	24.725
12	3.571	4.404	5.226	6.304	18.549	21.026	23.337	26.217
13	4.107	5.009	5.892	7.042	19.812	22.362	24.736	27.688
14	4.660	5.629	6.571	7.790	21.064	23.685	26.119	29.141
15	5.229	6.262	7.261	8.547	22.307	24.996	27.488	30.578
16	5.812	6.908	7.962	9.312	23.542	26.296	28.845	32.000
17	6.408	7.564	8.672	10.085	24.769	27.587	30.191	33.409
18	7.015	8.231	9.390	10.865	25.989	28.869	31.526	34.805
19	7.633	8.907	10.117	11.651	27.204	30.144	32.852	36.191
20	8.260	9.591	10.851	12.443	28.412	31.410	34.170	37.566
21	8.897	10.283	11.591	13.240	29.615	32.671	35.479	38.932
22	9.542	10.982	12.338	14.041	30.813	33.924	36.781	40.289
23	10.196	11.689	13.091	14.848	32.007	35.172	38.076	41.638
24	10.856	12.401	13.848	15.659	33.196	36.415	39.364	42.980
25	11.524	13.120	14.611	16.473	34.382	37.652	40.646	44.314
26	12.198	13.844	15.379	17.292	35.563	38.885	41.923	45.642
27	12.879	14.573	16.151	18.114	36.741	40.113	43.195	46.963
28	13.565	15.308	16.928	18.939	37.916	41.337	44.461	48.278
29	14.256	16.047	17.708	19.768	39.087	42.557	45.722	49.588
30	14.953	16.791	18.493	20.599	40.256	43.773	46.979	50.892

Table II
Normal Distribution

The following table presents the standard normal distribution. The probabilities tabled are

$$P(Z \leq z) = \Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-w^2/2}\, dw.$$

Note that only the probabilities for $z \geq 0$ are tabled. To obtain the probabilities for $z < 0$, use the identity $\Phi(-z) = 1 - \Phi(z)$. At the bottom of the table, some useful quantiles of the standard normal distribution are displayed. The R function `normaltable.s` generates this table.

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998
3.5	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998

α	0.400	0.300	0.200	0.100	0.050	0.025	0.020	0.010	0.005	0.001
z_α	0.253	0.524	0.842	1.282	1.645	1.960	2.054	2.326	2.576	3.090
$z_{\alpha/2}$	0.842	1.036	1.282	1.645	1.960	2.241	2.326	2.576	2.807	3.291

Table III
t-Distribution

The following table presents selected quantiles of the t-distribution, i.e., the values t such that

$$P(T \leq t) = \int_{-\infty}^{t} \frac{\Gamma[(r+1)/2]}{\sqrt{\pi r}\,\Gamma(r/2)(1+w^2/r)^{(r+1)/2}}\,dw,$$

for selected degrees of freedom r. The last row gives the standard normal quantiles.

	$P(T \leq t)$					
r	0.900	0.950	0.975	0.990	0.995	0.999
1	3.078	6.314	12.706	31.821	63.657	318.309
2	1.886	2.920	4.303	6.965	9.925	22.327
3	1.638	2.353	3.182	4.541	5.841	10.215
4	1.533	2.132	2.776	3.747	4.604	7.173
5	1.476	2.015	2.571	3.365	4.032	5.893
6	1.440	1.943	2.447	3.143	3.707	5.208
7	1.415	1.895	2.365	2.998	3.499	4.785
8	1.397	1.860	2.306	2.896	3.355	4.501
9	1.383	1.833	2.262	2.821	3.250	4.297
10	1.372	1.812	2.228	2.764	3.169	4.144
11	1.363	1.796	2.201	2.718	3.106	4.025
12	1.356	1.782	2.179	2.681	3.055	3.930
13	1.350	1.771	2.160	2.650	3.012	3.852
14	1.345	1.761	2.145	2.624	2.977	3.787
15	1.341	1.753	2.131	2.602	2.947	3.733
16	1.337	1.746	2.120	2.583	2.921	3.686
17	1.333	1.740	2.110	2.567	2.898	3.646
18	1.330	1.734	2.101	2.552	2.878	3.610
19	1.328	1.729	2.093	2.539	2.861	3.579
20	1.325	1.725	2.086	2.528	2.845	3.552
21	1.323	1.721	2.080	2.518	2.831	3.527
22	1.321	1.717	2.074	2.508	2.819	3.505
23	1.319	1.714	2.069	2.500	2.807	3.485
24	1.318	1.711	2.064	2.492	2.797	3.467
25	1.316	1.708	2.060	2.485	2.787	3.450
26	1.315	1.706	2.056	2.479	2.779	3.435
27	1.314	1.703	2.052	2.473	2.771	3.421
28	1.313	1.701	2.048	2.467	2.763	3.408
29	1.311	1.699	2.045	2.462	2.756	3.396
30	1.310	1.697	2.042	2.457	2.750	3.385
∞	1.282	1.645	1.960	2.326	2.576	3.090

Table IV
F-Distribution
Upper 0.05 Critical Points

The following table presents selected 0.95 and 0.99 quantiles of the F-distribution, i.e., for $\alpha = 0.05, 0.01$, the values $F_\alpha(r_1, r_2)$ such that

$$\alpha = P(X \geq F_\alpha(r_1, r_2)) = \int_{F_\alpha(r_1,r_2)}^{\infty} \frac{\Gamma[(r_1+r_2)/2](r_1/r_2)^{r_1/2} w^{r_1/2-1}}{\Gamma(r_1/2)\Gamma(r_2/2)(1+r_1 w/r_2)^{(r_1+r_2)/2}}\, dw,$$

where r_1 and r_2 are the numerator and denominator degrees of freedom, respectively. The R function `fp1.r` generates this table.

$F_{0.05}(r_1, r_2)$									
	r_1								
r_2	1	2	3	4	5	6	7	8	9
1	161.45	199.50	215.71	224.58	230.16	233.99	236.77	238.88	240.54
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04
120	3.92	3.07	2.68	2.45	2.29	2.18	2.09	2.02	1.96
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88

Table IV
F-Distribution, Continued
Upper 0.05 Critical Points

Generated by the R function `fp2.r`.

$F_{0.05}(r_1, r_2)$									
	r_1								
r_2	10	15	20	25	30	40	60	120	∞
1	241.88	245.95	248.01	249.26	250.10	251.14	252.20	253.25	254.31
2	19.40	19.43	19.45	19.46	19.46	19.47	19.48	19.49	19.50
3	8.79	8.70	8.66	8.63	8.62	8.59	8.57	8.55	8.53
4	5.96	5.86	5.80	5.77	5.75	5.72	5.69	5.66	5.63
5	4.74	4.62	4.56	4.52	4.50	4.46	4.43	4.40	4.36
6	4.06	3.94	3.87	3.83	3.81	3.77	3.74	3.70	3.67
7	3.64	3.51	3.44	3.40	3.38	3.34	3.30	3.27	3.23
8	3.35	3.22	3.15	3.11	3.08	3.04	3.01	2.97	2.93
9	3.14	3.01	2.94	2.89	2.86	2.83	2.79	2.75	2.71
10	2.98	2.85	2.77	2.73	2.70	2.66	2.62	2.58	2.54
11	2.85	2.72	2.65	2.60	2.57	2.53	2.49	2.45	2.40
12	2.75	2.62	2.54	2.50	2.47	2.43	2.38	2.34	2.30
13	2.67	2.53	2.46	2.41	2.38	2.34	2.30	2.25	2.21
14	2.60	2.46	2.39	2.34	2.31	2.27	2.22	2.18	2.13
15	2.54	2.40	2.33	2.28	2.25	2.20	2.16	2.11	2.07
16	2.49	2.35	2.28	2.23	2.19	2.15	2.11	2.06	2.01
17	2.45	2.31	2.23	2.18	2.15	2.10	2.06	2.01	1.96
18	2.41	2.27	2.19	2.14	2.11	2.06	2.02	1.97	1.92
19	2.38	2.23	2.16	2.11	2.07	2.03	1.98	1.93	1.88
20	2.35	2.20	2.12	2.07	2.04	1.99	1.95	1.90	1.84
21	2.32	2.18	2.10	2.05	2.01	1.96	1.92	1.87	1.81
22	2.30	2.15	2.07	2.02	1.98	1.94	1.89	1.84	1.78
23	2.27	2.13	2.05	2.00	1.96	1.91	1.86	1.81	1.76
24	2.25	2.11	2.03	1.97	1.94	1.89	1.84	1.79	1.73
25	2.24	2.09	2.01	1.96	1.92	1.87	1.82	1.77	1.71
26	2.22	2.07	1.99	1.94	1.90	1.85	1.80	1.75	1.69
27	2.20	2.06	1.97	1.92	1.88	1.84	1.79	1.73	1.67
28	2.19	2.04	1.96	1.91	1.87	1.82	1.77	1.71	1.65
29	2.18	2.03	1.94	1.89	1.85	1.81	1.75	1.70	1.64
30	2.16	2.01	1.93	1.88	1.84	1.79	1.74	1.68	1.62
40	2.08	1.92	1.84	1.78	1.74	1.69	1.64	1.58	1.51
60	1.99	1.84	1.75	1.69	1.65	1.59	1.53	1.47	1.39
120	1.91	1.75	1.66	1.60	1.55	1.50	1.43	1.35	1.25
∞	1.83	1.67	1.57	1.51	1.46	1.39	1.32	1.22	1.00

Table IV
F-Distribution, Continued
Upper 0.01 Critical Points

The R function `fp3.r` generates this table.

$F_{0.01}(r_1, r_2)$									
	r_1								
r_2	1	2	3	4	5	6	7	8	9
1	4052.2	4999.5	5403.4	5624.6	5763.7	5859.0	5928.4	5981.1	6022.5
2	98.50	99.00	99.17	99.25	99.30	99.33	99.36	99.37	99.39
3	34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.35
4	21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.66
5	16.26	13.27	12.06	11.39	10.97	10.67	10.46	10.29	10.16
6	13.75	10.92	9.78	9.15	8.75	8.47	8.26	8.10	7.98
7	12.25	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72
8	11.26	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91
9	10.56	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35
10	10.04	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94
11	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63
12	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39
13	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	4.19
14	8.86	6.51	5.56	5.04	4.69	4.46	4.28	4.14	4.03
15	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89
16	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78
17	8.40	6.11	5.18	4.67	4.34	4.10	3.93	3.79	3.68
18	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.60
19	8.18	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52
20	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46
21	8.02	5.78	4.87	4.37	4.04	3.81	3.64	3.51	3.40
22	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35
23	7.88	5.66	4.76	4.26	3.94	3.71	3.54	3.41	3.30
24	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26
25	7.77	5.57	4.68	4.18	3.85	3.63	3.46	3.32	3.22
26	7.72	5.53	4.64	4.14	3.82	3.59	3.42	3.29	3.18
27	7.68	5.49	4.60	4.11	3.78	3.56	3.39	3.26	3.15
28	7.64	5.45	4.57	4.07	3.75	3.53	3.36	3.23	3.12
29	7.60	5.42	4.54	4.04	3.73	3.50	3.33	3.20	3.09
30	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07
40	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89
60	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72
120	6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56
∞	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41

Table IV
F-Distribution, Continued
Upper 0.01 Critical Points

The R function `fp4.r` generates this table.

$F_{0.01}(r_1, r_2)$									
	r_1								
r_2	10	15	20	25	30	40	60	120	∞
1	6055.9	6157.3	6208.7	6239.8	6260.7	6286.8	6313.0	6339.4	6365.9
2	99.40	99.43	99.45	99.46	99.47	99.47	99.48	99.49	99.50
3	27.23	26.87	26.69	26.58	26.50	26.41	26.32	26.22	26.13
4	14.55	14.20	14.02	13.91	13.84	13.75	13.65	13.56	13.46
5	10.05	9.72	9.55	9.45	9.38	9.29	9.20	9.11	9.02
6	7.87	7.56	7.40	7.30	7.23	7.14	7.06	6.97	6.88
7	6.62	6.31	6.16	6.06	5.99	5.91	5.82	5.74	5.65
8	5.81	5.52	5.36	5.26	5.20	5.12	5.03	4.95	4.86
9	5.26	4.96	4.81	4.71	4.65	4.57	4.48	4.40	4.31
10	4.85	4.56	4.41	4.31	4.25	4.17	4.08	4.00	3.91
11	4.54	4.25	4.10	4.01	3.94	3.86	3.78	3.69	3.60
12	4.30	4.01	3.86	3.76	3.70	3.62	3.54	3.45	3.36
13	4.10	3.82	3.66	3.57	3.51	3.43	3.34	3.25	3.17
14	3.94	3.66	3.51	3.41	3.35	3.27	3.18	3.09	3.00
15	3.80	3.52	3.37	3.28	3.21	3.13	3.05	2.96	2.87
16	3.69	3.41	3.26	3.16	3.10	3.02	2.93	2.84	2.75
17	3.59	3.31	3.16	3.07	3.00	2.92	2.83	2.75	2.65
18	3.51	3.23	3.08	2.98	2.92	2.84	2.75	2.66	2.57
19	3.43	3.15	3.00	2.91	2.84	2.76	2.67	2.58	2.49
20	3.37	3.09	2.94	2.84	2.78	2.69	2.61	2.52	2.42
21	3.31	3.03	2.88	2.79	2.72	2.64	2.55	2.46	2.36
22	3.26	2.98	2.83	2.73	2.67	2.58	2.50	2.40	2.31
23	3.21	2.93	2.78	2.69	2.62	2.54	2.45	2.35	2.26
24	3.17	2.89	2.74	2.64	2.58	2.49	2.40	2.31	2.21
25	3.13	2.85	2.70	2.60	2.54	2.45	2.36	2.27	2.17
26	3.09	2.81	2.66	2.57	2.50	2.42	2.33	2.23	2.13
27	3.06	2.78	2.63	2.54	2.47	2.38	2.29	2.20	2.10
28	3.03	2.75	2.60	2.51	2.44	2.35	2.26	2.17	2.06
29	3.00	2.73	2.57	2.48	2.41	2.33	2.23	2.14	2.03
30	2.98	2.70	2.55	2.45	2.39	2.30	2.21	2.11	2.01
40	2.80	2.52	2.37	2.27	2.20	2.11	2.02	1.92	1.80
60	2.63	2.35	2.20	2.10	2.03	1.94	1.84	1.73	1.60
120	2.47	2.19	2.03	1.93	1.86	1.76	1.66	1.53	1.38
∞	2.32	2.04	1.88	1.77	1.70	1.59	1.47	1.32	1.00