Distribution of the Function of Random Variable (M.G.F. Technique) & Order Statistics
The Moment Generating Function Technique

**Definition 2.** Let $X_1, X_2, \ldots, X_n$ denote $n$ mutually stochastically independent random variables, each of which has the same but possibly unknown p.d.f. $f(x)$; that is, the probability density functions of $X_1, X_2, \ldots, X_n$ are, respectively, $f_1(x_1) = f(x_1), f_2(x_2) = f(x_2), \ldots, f_n(x_n) = f(x_n)$, so that the joint p.d.f. is $f(x_1)f(x_2)\cdots f(x_n)$. The random variables $X_1, X_2, \ldots, X_n$ are then said to constitute a random sample from a distribution that has p.d.f. $f(x)$.

**Definition 1.** A function of one or more random variables that does not depend upon any unknown parameter is called a statistic.

In accordance with this definition, the random variable $Y = \sum_{i=1}^{n} X_i$ discussed above is a statistic. But the random variable $Y = (X_1 - \mu)/\sigma$ is not a statistic unless $\mu$ and $\sigma$ are known numbers. It should be noted that, although a statistic does not depend upon any unknown parameter, the distribution of that statistic may very well depend upon unknown parameters.

We should recall that a moment-generating function, when it exists, is unique and that it uniquely determines the distribution of a probability.

Let $\varphi(x_1, x_2, \ldots, x_n)$ denote the joint p.d.f. of the $n$ random variables $X_1, X_2, \ldots, X_n$. These random variables may or may not be the items of a random sample from some distribution that has a given p.d.f. $f(x)$. Let $Y_1 = u_1(X_1, X_2, \ldots, X_n)$. We seek $g(y_1)$, the p.d.f. of the random variable $Y_1$. Consider the moment-generating function of $Y_1$. If it exists, it is given by
\[ M(t) = E(e^{tY_1}) = \int_{-\infty}^{\infty} e^{tY_1} g(y_1) \, dy_1 \]

in the continuous case. It would seem that we need to know \( g(y_1) \) before we can compute \( M(t) \). That this is not the case is a fundamental fact. To see this consider

\[ \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \exp \left[ tu_1(x_1, \ldots, x_n) \right] \varphi(x_1, \ldots, x_n) \, dx_1 \cdots dx_n, \]

which we assume to exist for \(-h < t < h\). We shall introduce \( n \) new variables of integration. They are \( y_1 = u_1(x_1, x_2, \ldots, x_n), \ldots, y_n = u_n(x_1, x_2, \ldots, x_n) \). Momentarily, we assume that these functions define a one-to-one transformation. Let \( x_i = w_i(y_1, y_2, \ldots, y_n), i = 1, 2, \ldots, n \), denote the inverse functions and let \( J \) denote the Jacobian. Under this transformation, display (1) becomes

\[ \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{tu_1} |J| \varphi(w_1, \ldots, w_n) \, dw_1 \cdots dw_n \, dy_1. \]

In accordance with Section 4.5,

\[ |J| \varphi[w_1(y_1, y_2, \ldots, y_n), \ldots, w_n(y_1, y_2, \ldots, y_n)] \]

is the joint p.d.f. of \( Y_1, Y_2, \ldots, Y_n \). The marginal p.d.f. \( g(y_1) \) of \( Y_1 \) is obtained by integrating this joint p.d.f. on \( y_2, \ldots, y_n \). Since the factor \( e^{ty_1} \) does not involve the variables \( y_2, \ldots, y_n \), display (2) may be written as

\[ \int_{-\infty}^{\infty} e^{ty_1} g(y_1) \, dy_1. \]

The reader will observe that we have assumed the transformation to be one-to-one. We did this for simplicity of presentation. If the transformation is not one-to-one, let

\[ x_j = w_{j_1}(y_1, \ldots, y_n), \quad j = 1, 2, \ldots, n, \quad i = 1, 2, \ldots, k, \]
denote the $k$ groups of $n$ inverse functions each. Let $J_i, i = 1, 2, \ldots, k,$
dequate the $k$ Jacobians. Then

It should be noted that the expectation, subject to its existence, of any function of $Y_1$ can be computed in like manner. That is, if $w(y_1)$ is a function of $y_1$, then

$$E[w(Y_1)] = \int_{-\infty}^{\infty} w(y_1)g(y_1) \, dy_1$$

$$= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} w[u_1(x_1, \ldots, x_n)]\varphi(x_1, \ldots, x_n) \, dx_1 \cdots dx_n.$$

**Example 1.** Let the stochastically independent random variables $X_1$ and $X_2$ have the same p.d.f.

$$f(x) = \frac{x}{6}, \quad x = 1, 2, 3,$$

$$= 0 \text{ elsewhere;}$$

that is, the p.d.f. of $X_1$ is $f(x_1)$ and that of $X_2$ is $f(x_2)$; and so the joint p.d.f. of $X_1$ and $X_2$ is

$$f(x_1)f(x_2) = \frac{x_1x_2}{36}, \quad x_1 = 1, 2, 3, x_2 = 1, 2, 3,$$

$$= 0 \text{ elsewhere.}$$

A probability, such as $\Pr (X_1 = 2, X_2 = 3)$, can be seen immediately to be $(2)(3)/36 = \frac{1}{6}$. However, consider a probability such as $\Pr (X_1 + X_2 = 3)$. The computation can be made by first observing that the event $X_1 + X_2 = 3$ is the union, exclusive of the events with probability zero, of the two mutually exclusive events $(X_1 = 1, X_2 = 2)$ and $(X_1 = 2, X_2 = 1).$ Thus

$$\Pr (X_1 + X_2 = 3) = \Pr (X_1 = 1, X_2 = 2) + \Pr (X_1 = 2, X_2 = 1)$$

$$= \frac{(1)(2)}{36} + \frac{(2)(1)}{36} = \frac{4}{36}.$$
More generally, let $y$ represent any of the numbers $2, 3, 4, 5, 6$. The probability of each of the events $X_1 + X_2 = y$, $y = 2, 3, 4, 5, 6$, can be computed as in the case $y = 3$. Let $g(y) = \Pr(X_1 + X_2 = y)$. Then the table

<table>
<thead>
<tr>
<th>$y$</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>$g(y)$</td>
<td>$\frac{1}{36}$</td>
<td>$\frac{4}{36}$</td>
<td>$\frac{10}{36}$</td>
<td>$\frac{12}{36}$</td>
<td>$\frac{9}{36}$</td>
</tr>
</tbody>
</table>

gives the values of $g(y)$ for $y = 2, 3, 4, 5, 6$. For all other values of $y$, $g(y) = 0$.

What we have actually done is to define a new random variable $Y$ by $Y = X_1 + X_2$, and we have found the p.d.f. $g(y)$ of this random variable $Y$. We shall now solve the same problem, and by the moment-generating-function technique.

Now the moment-generating function of $Y$ is

$$M(t) = \mathbb{E}(e^{t(X_1 + X_2)})$$
$$= \mathbb{E}(e^{tx_1}e^{tx_2})$$
$$= \mathbb{E}(e^{tx_1})\mathbb{E}(e^{tx_2}),$$

$$E(e^{tx_1}) = E(e^{tx_2}) = \frac{1}{6}e^t + \frac{2}{6}e^{2t} + \frac{3}{6}e^{3t}.$$

$$M(t) = (\frac{1}{6}e^t + \frac{2}{6}e^{2t} + \frac{3}{6}e^{3t})^2$$
$$= \frac{1}{36}e^{2t} + \frac{4}{36}e^{3t} + \frac{10}{36}e^{4t} + \frac{12}{36}e^{5t} + \frac{9}{36}e^{6t}.$$

This form of $M(t)$ tells us immediately that the p.d.f. $g(y)$ of $Y$ is zero except at $y = 2, 3, 4, 5, 6$, and that $g(y)$ assumes the values $\frac{1}{36}, \frac{4}{36}, \frac{10}{36}, \frac{12}{36}, \frac{9}{36}$, respectively, at these points where $g(y) > 0$. This is, of course, the same result that was obtained in the first solution.
Example 2. Let $X_1$ and $X_2$ be stochastically independent with normal distributions $n(\mu_1, \sigma_1^2)$ and $n(\mu_2, \sigma_2^2)$, respectively. Define the random variable $Y$ by $Y = X_1 - X_2$. The problem is to find $g(y)$, the p.d.f. of $Y$. This will be done by first finding the moment-generating function of $Y$. It is

$$M(t) = E(e^{t(X_1 - X_2)})$$
$$= E(e^{tX_1}e^{-tX_2})$$
$$= E(e^{tX_1})E(e^{-tX_2}),$$

since $X_1$ and $X_2$ are stochastically independent. It is known that

$$E(e^{tX_1}) = \exp \left( \mu_1 t + \frac{\sigma_1^2 t^2}{2} \right)$$

and that

$$E(e^{tX_2}) = \exp \left( \mu_2 t + \frac{\sigma_2^2 t^2}{2} \right)$$

for all real $t$. Then $E(e^{-tX_2})$ can be obtained from $E(e^{tX_2})$ by replacing $t$ by $-t$. That is,

$$E(e^{-tX_2}) = \exp \left( -\mu_2 t + \frac{\sigma_2^2 t^2}{2} \right).$$

Finally, then,

$$M(t) = \exp \left( \mu_1 t + \frac{\sigma_1^2 t^2}{2} \right) \exp \left( -\mu_2 t + \frac{\sigma_2^2 t^2}{2} \right)$$
$$= \exp \left( (\mu_1 - \mu_2)t + \frac{(\sigma_1^2 + \sigma_2^2)t^2}{2} \right).$$

it is seen that $Y$ has the p.d.f. $g(y)$, which is

$$n(\mu_1 - \mu_2, \sigma_1^2 + \sigma_2^2).$$
Theorem 1. Let $X_1, X_2, \ldots, X_n$ be mutually stochastically independent random variables having, respectively, the normal distributions $n(\mu_1, \sigma_1^2), n(\mu_2, \sigma_2^2), \ldots, \text{and } n(\mu_n, \sigma_n^2).$ The random variable $Y = k_1X_1 + k_2X_2 + \cdots + k_nX_n$, where $k_1, k_2, \ldots, k_n$ are real constants, is normally distributed with mean $k_1\mu_1 + \cdots + k_n\mu_n$ and variance $k_1^2\sigma_1^2 + \cdots + k_n^2\sigma_n^2$. That is, $Y$ is $n\left(\sum_{1}^{n} k_i\mu_i, \sum_{1}^{n} k_i^2\sigma_i^2 \right)$.

Theorem 2. Let $X_1, X_2, \ldots, X_n$ be mutually stochastically independent variables that have, respectively, the chi-square distributions $\chi^2(r_1), \chi^2(r_2), \ldots, \text{and } \chi^2(r_n).$ Then the random variable $Y = X_1 + X_2 + \cdots + X_n$ has a chi-square distribution with $r_1 + \cdots + r_n$ degrees of freedom; that is, $Y$ is $\chi^2(r_1 + \cdots + r_n)$.

Theorem 3. Let $X_1, X_2, \ldots, X_n$ denote a random sample of size $n$ from a distribution that is $n(\mu, \sigma^2)$. The random variable

$$Y = \sum_{1}^{n} \left(\frac{X_i - \mu}{\sigma}\right)^2$$

has a chi-square distribution with $n$ degrees of freedom.

EXERCISES

4.68. Let the stochastically independent random variables $X_1$ and $X_2$ have the same p.d.f. $f(x) = \frac{1}{6}, x = 1, 2, 3, 4, 5, 6$, zero elsewhere. Find the p.d.f. of $Y = X_1 + X_2$. Note, under appropriate assumptions, that $Y$ may be interpreted as the sum of the spots that appear when two dice are cast.
4.69. Let \(X_1\) and \(X_2\) be stochastically independent with normal distributions \(n(6, 1)\) and \(n(7, 1)\), respectively. Find \(\Pr(X_1 > X_2)\). \textit{Hint.} Write \(\Pr(X_1 > X_2) = \Pr(X_1 - X_2 > 0)\) and determine the distribution of \(X_1 - X_2\).

4.70. Let \(X_1\) and \(X_2\) be stochastically independent random variables. Let \(X_1\) and \(Y = X_1 + X_2\) have chi-square distributions with \(r_1\) and \(r\) degrees of freedom, respectively. Here \(r_1 < r\). Show that \(X_2\) has a chi-square distribution with \(r - r_1\) degrees of freedom. \textit{Hint.} Write \(M(t) = E(e^{tX_1 + X_2})\) and make use of the stochastic independence of \(X_1\) and \(X_2\).

4.71. Let the stochastically independent random variables \(X_1\) and \(X_2\) have binomial distributions with parameters \(n_1, \rho_1 = \frac{1}{3}\) and \(n_2, \rho_2 = \frac{1}{3}\), respectively. Show that \(Y = X_1 - X_2 + n_2\) has a binomial distribution with parameters \(n = n_1 + n_2, \rho = \frac{1}{3}\).

4.72. Let \(X\) be \(n(0, 1)\). Use the moment-generating-function technique to show that \(Y = X^2\) is \(\chi^2(1)\). \textit{Hint.} Evaluate the integral that represents \(E(e^{tX^2})\) by writing \(w = x\sqrt{1 - 2t}, t < \frac{1}{2}\).

4.8 The Distributions of \(\bar{X}\) and \(nS^2/\sigma^2\)

Let \(X_1, X_2, \ldots, X_n\) denote a random sample of size \(n \geq 2\) from a distribution that is \(n(\mu, \sigma^2)\). In this section we shall investigate the distributions of the mean and the variance of this random sample, that is, the distributions of the two statistics \(\bar{X} = \frac{1}{n} \sum X_i\) and \(S^2 = \frac{1}{n-1} \sum (X_i - \bar{X})^2\).

The problem of the distribution of \(\bar{X}\), the mean of the sample, is solved by the use of Theorem 1 of Section 4.7. We have here, in the notation of the statement of that theorem, \(\mu_1 = \mu_2 = \cdots = \mu_n = \mu\),
\( \sigma_1^2 = \sigma_2^2 = \cdots = \sigma_n^2 = \sigma^2 \), and \( k_1 = k_2 = \cdots = k_n = 1/n \). Accordingly, \( Y = X \) has a normal distribution with mean and variance given by

\[
\sum_{1}^{n} \left( \frac{1}{n} \mu \right) = \mu, \quad \sum_{1}^{n} \left[ \frac{1}{n} \sigma^2 \right] = \frac{\sigma^2}{n},
\]

respectively. That is, \( X \) is \( n(\mu, \sigma^2/n) \).

We now take up the problem of the distribution of \( S^2 \), the variance of a random sample \( X_1, \ldots, X_n \) from a distribution that is \( n(\mu, \sigma^2) \). To do this, let us first consider the joint distribution of \( Y_1 = X_1 \), \( Y_2 = X_2, \ldots, Y_n = X_n \). The corresponding transformation

\[
x_1 = n y_1 - y_2 - \cdots - y_n \\
x_2 = y_2 \\
\vdots \quad \vdots \\
x_n = y_n
\]

has Jacobian \( n \). Since

\[
\sum_{1}^{n} (x_i - \mu)^2 = \sum_{1}^{n} (x_i - \bar{x} + \bar{x} - \mu)^2 \\
= \sum_{1}^{n} (x_i - \bar{x})^2 + n(\bar{x} - \mu)^2
\]

because \( 2(\bar{x} - \mu) \sum_{1}^{n} (x_i - \bar{x}) = 0 \), the joint p.d.f. of \( X_1, X_2, \ldots, X_n \) can be written

\[
\left( \frac{1}{\sqrt{2\pi\sigma}} \right)^n \exp \left[ - \frac{\sum (x_i - \bar{x})^2}{2\sigma^2} - \frac{n(\bar{x} - \mu)^2}{2\sigma^2} \right],
\]
where $\bar{x}$ represents $(x_1 + x_2 + \cdots + x_n)/n$ and $-\infty < x_i < \infty$, $i = 1, 2, \ldots, n$. Accordingly, with $y_1 = \bar{x}$, we find that the joint p.d.f. of $Y_1, Y_2, \ldots, Y_n$ is

$$n \left( \frac{1}{\sqrt{2\pi}\sigma} \right)^n \exp \left\{ -\frac{n(y_1 - y_2 - \cdots - y_n - y_1)^2}{2\sigma^2} \right. \right.$$ 

$$\left. - \frac{\sum_{i=2}^{n} (y_i - y_1)^2}{2\sigma^2} - \frac{n(y_1 - \mu)^2}{2\sigma^2} \right\},$$

$-\infty < y_i < \infty$, $i = 1, 2, \ldots, n$. The quotient of this joint p.d.f. and the p.d.f.

$$\sqrt{n} \left( \frac{1}{\sqrt{2\pi}\sigma} \right)^{n-1} \exp \left( -\frac{q}{2\sigma^2} \right),$$

of $Y_1 = X$ is the conditional p.d.f. of $Y_2, Y_3, \ldots, Y_n$, given $Y_1 = y_1$,

$$\sqrt{n} \left( \frac{1}{\sqrt{2\pi}\sigma} \right)^{n-1} \exp \left( -\frac{q}{2\sigma^2} \right),$$

where $q = (ny_1 - y_2 - \cdots - y_n - y_1)^2 + \sum_{i=2}^{n} (y_i - y_1)^2$. Since this is a joint conditional p.d.f., it must be, for all $\sigma > 0$, that

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \sqrt{n} \left( \frac{1}{\sqrt{2\pi}\sigma} \right)^{n-1} \exp \left( -\frac{q}{2\sigma^2} \right) \, dy_2 \cdots dy_n = 1.$$

Now consider

$$nS^2 = \sum_{1}^{n} (X_i - \bar{X})^2$$

$$= (nY_1 - Y_2 - \cdots - Y_n - Y_1)^2 + \sum_{2}^{n} (Y_i - Y_1)^2 = Q.$$ 

The conditional moment-generating function of $nS^2/\sigma^2 = Q/\sigma^2$, given $Y_1 = y_1$, is
\[
E(e^{tQ_1\sigma^2} | y_1) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \sqrt{n} \left( \frac{1}{\sqrt{2\pi\sigma}} \right)^{n-1} \exp \left[ -\frac{(1 - 2t)q}{2\sigma^2} \right] dy_2 \cdots dy_n
\]

where \(0 < 1 - 2t\), or \(t < \frac{1}{2}\). However, this latter integral is exactly the same as that of the conditional p.d.f. of \(Y_2, Y_3, \ldots, Y_n\), given \(Y_1 = y_1\), with \(\sigma^2\) replaced by \(\sigma^2/(1 - 2t) > 0\), and thus must equal 1. Hence the conditional moment-generating function of \(nS^2/\sigma^2\), given \(Y_1 = y_1\) or equivalently \(\tilde{X} = \tilde{x}\), is

\[
E(e^{tnS^2/\sigma^2} | \tilde{x}) = (1 - 2t)^{-(n-1)/2}, \quad t < \frac{1}{2}.
\]

That is, the conditional distribution of \(nS^2/\sigma^2\), given \(\tilde{X} = \tilde{x}\), is \(\chi^2(n - 1)\). Moreover, since it is clear that this conditional distribution does not depend upon \(\tilde{x}\), \(\tilde{X}\) and \(nS^2/\sigma^2\) must be stochastically independent or, equivalently, \(\tilde{X}\) and \(S^2\) are stochastically independent.

To summarize, we have established, in this section, three important properties of \(\tilde{X}\) and \(S^2\) when the sample arises from a distribution which is \(n(\mu, \sigma^2)\):

(a) \(\tilde{X}\) is \(n(\mu, \sigma^2/n)\).
(b) \(nS^2/\sigma^2\) is \(\chi^2(n - 1)\).
(c) \(\tilde{X}\) and \(S^2\) are stochastically independent.
EXERCISES

4.83. Let \( \bar{X} \) be the mean of a random sample of size 5 from a normal distribution with \( \mu = 0 \) and \( \sigma^2 = 125 \). Determine \( c \) so that \( \Pr (\bar{X} < c) = 0.90 \).

4.84. If \( \bar{X} \) is the mean of a random sample of size \( n \) from a normal distribution with mean \( \mu \) and variance 100, find \( n \) so that \( \Pr (\mu - 5 < \bar{X} < \mu + 5) = 0.954 \).

4.86. Find the mean and variance of \( S^2 = \frac{1}{n} \sum (X_i - \bar{X})^2 \), where \( X_1, X_2, \ldots, X_n \) is a random sample from \( n(\mu, \sigma^2) \). *Hint.* Find the mean and variance of \( nS^2/\sigma^2 \).

Expectations of Functions of Random Variables

*Example 2.* Let \( X_i \) denote a random variable with mean \( \mu_i \) and variance \( \sigma_i^2 \), \( i = 1, 2, \ldots, n \). Let \( X_1, X_2, \ldots, X_n \) be mutually stochastically independent and let \( k_1, k_2, \ldots, k_n \) denote real constants. We shall compute the mean and variance of the linear function \( Y = k_1X_1 + k_2X_2 + \cdots + k_nX_n \). Because \( E \) is a linear operator, the mean of \( Y \) is given by

\[
\mu_Y = E(k_1X_1 + k_2X_2 + \cdots + k_nX_n) = k_1E(X_1) + k_2E(X_2) + \cdots + k_nE(X_n) = k_1\mu_1 + k_2\mu_2 + \cdots + k_n\mu_n = \sum_{i=1}^{n} k_i\mu_i.
\]
The variance of $Y$ is given by

$$
\sigma_Y^2 = E\left\{ \left[ (k_1 X_1 + \cdots + k_n X_n) - (k_1 \mu_1 + \cdots + k_n \mu_n) \right]^2 \right\}
= E\left\{ \left[ k_1(X_1 - \mu_1) + \cdots + k_n(X_n - \mu_n) \right]^2 \right\}
= E\left\{ \sum_{i=1}^{n} k_i^2(X_i - \mu_i)^2 + 2 \sum_{i<j} k_i k_j(X_i - \mu_i)(X_j - \mu_j) \right\}
= \sum_{i=1}^{n} k_i^2 E[(X_i - \mu_i)^2] + 2 \sum_{i<j} k_i k_j E[(X_i - \mu_i)(X_j - \mu_j)].
$$

Consider $E[(X_i - \mu_i)(X_j - \mu_j)]$, $i < j$. Because $X_i$ and $X_j$ are stochastically independent, we have

$$
E[(X_i - \mu_i)(X_j - \mu_j)] = E(X_i - \mu_i)E(X_j - \mu_j) = 0.
$$

Finally, then,

$$
\sigma_Y^2 = \sum_{i=1}^{n} k_i^2 E[(X_i - \mu_i)^2] = \sum_{i=1}^{n} k_i^2 \sigma_i^2.
$$

$$
E[(X_i - \mu_i)(X_j - \mu_j)] = \rho_{ij} \sigma_i \sigma_j, \quad i < j.
$$

If we refer to Example 2, we see that again $\mu_Y = \sum_{i=1}^{n} k_i \mu_i$. But now

$$
\sigma_Y^2 = \sum_{i=1}^{n} k_i^2 \sigma_i^2 + 2 \sum_{i<j} k_i k_j \rho_{ij} \sigma_i \sigma_j.
$$

Thus we have the following theorem.

**Theorem 4.** Let $X_1, \ldots, X_n$ denote random variables that have means $\mu_1, \ldots, \mu_n$ and variances $\sigma_1^2, \ldots, \sigma_n^2$. Let $\rho_{ij}$, $i \neq j$, denote the correlation coefficient of $X_i$ and $X_j$ and let $k_1, \ldots, k_n$ denote real constants. The mean and the variance of the linear function

$$
Y = \sum_{i=1}^{n} k_i X_i
$$

are, respectively,

$$
\mu_Y = \sum_{i=1}^{n} k_i \mu_i
$$

and

$$
\sigma_Y^2 = \sum_{i=1}^{n} k_i^2 \sigma_i^2 + 2 \sum_{i<j} k_i k_j \rho_{ij} \sigma_i \sigma_j.
$$
\[ \sigma^2_Y = \sum_{i=1}^{n} k_i^2 \sigma^2_i + 2 \sum_{i<j} k_i k_j \rho_{ij} \sigma_i \sigma_j. \]

**Corollary.** Let \( X_1, \ldots, X_n \) denote the items of a random sample of size \( n \) from a distribution that has mean \( \mu \) and variance \( \sigma^2 \). The mean and the variance of \( Y = \sum_{i=1}^{n} k_i X_i \) are, respectively, \( \mu_Y = \left( \sum_{i=1}^{n} k_i \right) \mu \) and \( \sigma^2_Y = \left( \sum_{i=1}^{n} k_i^2 \right) \sigma^2. \)

**Example 3.** Let \( \bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i \) denote the mean of a random sample of size \( n \) from a distribution that has mean \( \mu \) and variance \( \sigma^2 \). In accordance with the Corollary, we have \( \mu_{\bar{X}} = \mu \sum_{i=1}^{n} (1/n) = \mu \) and \( \sigma^2_{\bar{X}} = \sigma^2 \sum_{i=1}^{n} (1/n)^2 = \sigma^2/n. \) We have seen, in Section 4.8, that if our sample is from a distribution that is \( n(\mu, \sigma^2) \), then \( \bar{X} \) is \( n(\mu, \sigma^2/n) \). It is interesting that \( \mu_{\bar{X}} = \mu \) and \( \sigma^2_{\bar{X}} = \sigma^2/n \) whether the sample is or is not from a normal distribution.

**EXERCISES**

4.90. Let \( X_1, X_2, X_3, X_4 \) be four mutually stochastically independent random variables having the same p.d.f. \( f(x) = 2x, \ 0 < x < 1, \) zero elsewhere. Find the mean and variance of the sum \( Y \) of these four random variables.

4.91. Let \( X_1 \) and \( X_2 \) be two stochastically independent random variables so that the variances of \( X_1 \) and \( X_2 \) are \( \sigma_1^2 = k \) and \( \sigma_2^2 = 2 \), respectively. Given that the variance of \( Y = 3X_2 - X_1 \) is 25, find \( k \).

4.92. If the stochastically independent variables \( X_1 \) and \( X_2 \) have means \( \mu_1, \mu_2 \) and variances \( \sigma_1^2, \sigma_2^2 \), respectively, show that the mean and variance of the product \( Y = X_1X_2 \) are \( \mu_1 \mu_2 \) and \( \sigma_1^2 \sigma_2^2 + \mu_1^2 \sigma_2^2 + \mu_2^2 \sigma_1^2 \), respectively.

4.93. Find the mean and variance of the sum \( Y \) of the items of a random sample of size 5 from the distribution having p.d.f. \( f(x) = 6x(1 - x), \ 0 < x < 1, \) zero elsewhere.
Distributions of Order Statistics

Let $X_1, X_2, \ldots, X_n$ denote a random sample from a distribution of the *continuous type* having a p.d.f. $f(x)$ that is positive, provided that $a < x < b$. Let $Y_1$ be the smallest of these $X_i$, $Y_2$ the next $X_i$ in order of magnitude, $\ldots$, and $Y_n$ the largest $X_i$. That is, $Y_1 < Y_2 < \cdots < Y_n$ represent $X_1, X_2, \ldots, X_n$ when the latter are arranged in ascending order of magnitude. Then $Y_i$, $i = 1, 2, \ldots, n$, is called the $i$th order statistic of the random sample $X_1, X_2, \ldots, X_n$. It will be shown that the joint p.d.f. of $Y_1, Y_2, \ldots, Y_n$ is given by

$$g(y_1, y_2, \ldots, y_n) = (n!) f(y_1)f(y_2) \cdots f(y_n),$$

$$a < y_1 < y_2 < \cdots < y_n < b,$$

$$= 0 \text{ elsewhere.}$$

It will first be shown how the marginal p.d.f. of $Y_n$ may be expressed in terms of the distribution function $F(x)$ and the p.d.f. $f(x)$ of the random variable $X$. If $a < y_n < b$, the marginal p.d.f. of $y_n$ is given by

$$g_n(y_n) = n! \frac{[F(y_n)]^{n-1}}{(n-1)!} f(y_n)$$

$$= n[F(y_n)]^{n-1} f(y_n), \quad a < y_n < b,$$

$$= 0 \text{ elsewhere.}$$
\[ g_1(y_1) = n[1 - F(y_1)]^{n-1}f(y_1), \quad a < y_1 < b, \]
\[ = 0 \text{ elsewhere.} \]

\[ (2) \quad g_k(y_k) = \frac{n!}{(k - 1)! (n - k)!} [F(y_k)]^{k-1}[1 - F(y_k)]^{n-k}f(y_k), \]
\[ a < y_k < b, \]
\[ = 0 \text{ elsewhere.} \]

**Example 2.** Let \( Y_1 < Y_2 < Y_3 < Y_4 \) denote the order statistics of a random sample of size 4 from a distribution having p.d.f.

\[ f(x) = 2x, \quad 0 < x < 1, \]
\[ = 0 \text{ elsewhere.} \]

We shall express the p.d.f. of \( Y_3 \) in terms of \( f(x) \) and \( F(x) \) and then compute \( \Pr \left( \frac{1}{2} < Y_3 \right) \). Here \( F(x) = x^2 \), provided that \( 0 < x < 1 \), so that

\[ g_3(y_3) = \frac{4!}{2! 1!} (y_3^2)(1 - y_3^2)(2y_3), \quad 0 < y_3 < 1, \]
\[ = 0 \text{ elsewhere.} \]

\[ \Pr \left( \frac{1}{2} < Y_3 \right) = \int_{1/2}^{\infty} g_3(y_3) \, dy_3 \]
\[ = \int_{1/2}^{1} 24(y_3^5 - y_3^3) \, dy_3 = \frac{24}{255}. \]

Finally, the joint p.d.f. of any two order statistics, say \( Y_i < Y_j \), is as easily expressed in terms of \( F(x) \) and \( f(x) \). We have
Example 2. Let $Y_1 < Y_2 < Y_3 < Y_4$ denote the order statistics of a random sample of size 4 from a distribution having p.d.f.

$$f(x) = 2x, \quad 0 < x < 1,$$

$$= 0 \text{ elsewhere.}$$

We shall express the p.d.f. of $Y_3$ in terms of $f(x)$ and $F(x)$ and then compute $\Pr \left( \frac{1}{2} < Y_3 \right)$. Here $F(x) = x^2$, provided that $0 < x < 1$, so that

$$g_3(y_3) = \frac{4!}{2! \cdot 1!} (y_3^2)(1-y_3)(2y_3), \quad 0 < y_3 < 1,$$

$$= 0 \text{ elsewhere.}$$

$$\Pr \left( \frac{1}{2} < Y_3 \right) = \int_{1/2}^{\infty} g_3(y_3) \, dy_3$$

$$= \int_{1/2}^{1/3} 24(y_3^3 - y_3^2) \, dy_3 = \frac{2}{15}.$$  

Example 3. Let $Y_1, Y_2, Y_3$ be the order statistics of a random sample of size 3 from a distribution having p.d.f.

$$f(x) = 1, \quad 0 < x < 1,$$

$$= 0 \text{ elsewhere.}$$

We seek the p.d.f. of the sample range $Z_1 = Y_3 - Y_1$.

Since $F(x) = x$, $0 < x < 1$, the joint p.d.f. of $Y_1$ and $Y_3$ is

$$g_{13}(y_1, y_3) = 6(y_3 - y_1), \quad 0 < y_1 < y_3 < 1,$$

$$= 0 \text{ elsewhere.}$$

In addition to $Z_1 = Y_3 - Y_1$, let $Z_2 = Y_3$. Consider the functions $z_1 = y_3 - y_1$, $z_2 = y_3$, and their inverses $y_1 = z_2 - z_1$, $y_3 = z_2$, so that the corresponding Jacobian of the one-to-one-transformation is
\[ J = \begin{vmatrix} \frac{\partial y_1}{\partial z_1} & \frac{\partial y_1}{\partial z_2} \\ \frac{\partial y_3}{\partial z_1} & \frac{\partial y_3}{\partial z_2} \end{vmatrix} = \begin{vmatrix} -1 & 1 \\ 0 & 1 \end{vmatrix} = -1. \]

Thus the joint p.d.f. of \( Z_1 \) and \( Z_2 \) is
\[ h(z_1, z_2) = \begin{vmatrix} -1 \end{vmatrix} \cdot 6z_1 = 6z_1, \quad 0 < z_1 < z_2 < 1. \]
\[ = 0 \text{ elsewhere.} \]

Accordingly, the p.d.f. of the range \( Z_1 = Y_3 - Y_1 \) of the random sample of size 3 is
\[ h_1(z_1) = \int_{z_1}^{1} 6z_1 \, dz_2 = 6z_1(1 - z_1), \quad 0 < z_1 < 1, \]
\[ = 0 \text{ elsewhere.} \]

**EXERCISES**

4.50. Let \( Y_1 < Y_2 < Y_3 < Y_4 \) be the order statistics of a random sample of size 4 from the distribution having p.d.f. \( f(x) = e^{-x}, \) \( 0 < x < \infty, \) zero elsewhere. Find \( \Pr(3 \leq Y_4). \)

4.51. Let \( X_1, X_2, X_3 \) be a random sample from a distribution of the continuous type having p.d.f. \( f(x) = 2x, \) \( 0 < x < 1, \) zero elsewhere. Compute the probability that the smallest of these \( X_i \) exceeds the median of the distribution.

4.53. Let \( Y_1 < Y_2 < Y_3 < Y_4 < Y_5 \) denote the order statistics of a random sample of size 5 from a distribution having p.d.f. \( f(x) = e^{-x}, \) \( 0 < x < \infty, \) zero elsewhere. Show that \( Z_1 = Y_2 \) and \( Z_2 = Y_4 - Y_2 \) are stochastically independent. *Hint.* First find the joint p.d.f. of \( Y_2 \) and \( Y_4. \)

4.54. Let \( Y_1 < Y_2 < \cdots < Y_n \) be the order statistics of a random sample of size \( n \) from a distribution with p.d.f. \( f(x) = 1, \) \( 0 < x < 1, \) zero elsewhere. Show that the \( k \)th order statistic \( Y_k \) has a beta p.d.f. with parameters \( \alpha = k \) and \( \beta = n - k + 1. \)
number of degrees of freedom of the random variable that has the chi-square distribution. Some approximate values of

\[ \Pr (T \leq t) = \int_{-\infty}^{t} g_1(w) \, dw \]

for selected values of \( r \) and \( t \), can be found in Table IV in Appendix B.

Next consider two stochastically independent chi-square random variables \( U \) and \( V \) having \( r_1 \) and \( r_2 \) degrees of freedom, respectively. The joint p.d.f. \( \varphi(u, v) \) of \( U \) and \( V \) is then

\[
\varphi(u, v) = \frac{1}{\Gamma(r_1/2) \Gamma(r_2/2) 2^{(r_1+r_2)/2}} \frac{u^{r_1/2-1}v^{r_2/2-1} e^{-(u+v)/2}}{u^{r_1/2}v^{r_2/2}},
\]

\[ 0 < u < \infty, \; 0 < v < \infty, \]

\[ = 0 \text{ elsewhere}. \]

### 4.4 The \( t \) and \( F \) Distributions

It is the purpose of this section to define two additional distributions quite useful in certain problems of statistical inference. These are called, respectively, the (Student's) \( t \) distribution and the \( F \) distribution.

#### \( T \) Distribution

Let \( W \) denote a random variable that is \( n(0, 1) \); let \( V \) denote a random variable that is \( \chi^2(r) \); and let \( W \) and \( V \) be stochastically independent. Then the joint p.d.f. of \( W \) and \( V \), say \( \varphi(w, v) \), is the product of the p.d.f. of \( W \) and that of \( V \) or

\[
\varphi(w, v) = \frac{1}{\sqrt{2\pi}} e^{-w^2/2} \frac{1}{\Gamma(r/2) 2^{r/2}} \frac{v^{r/2-1} e^{-v/2}}{u^{r/2}},
\]

\[ -\infty < w < \infty, \; 0 < v < \infty, \]

\[ = 0 \text{ elsewhere}. \]
\[ g(t, u) = \varphi \left( \frac{t\sqrt{u}}{\sqrt{r}}, \frac{u}{\sqrt{r}} \right) |J| \]

\[ = \frac{1}{\sqrt{2\pi \Gamma(r/2)2^{r/2}}} u^{r/2-1} \exp \left[ -\frac{u}{2} \left( 1 + \frac{t^2}{r} \right) \right] \frac{\sqrt{u}}{\sqrt{r}}, \]

\[-\infty < t < \infty, 0 < u < \infty, \]

\[ = 0 \text{ elsewhere.} \]

The marginal p.d.f. of \( T \) is then

\[ g_1(t) = \int_{-\infty}^{\infty} g(t, u) \, du \]

\[ = \int_{0}^{\infty} \frac{1}{\sqrt{2\pi \Gamma(r/2)2^{r/2}}} u^{(r+1)/2-1} \exp \left[ -\frac{u}{2} \left( 1 + \frac{t^2}{r} \right) \right] \, du. \]

In this integral let \( z = u[1 + (t^2/r)]/2 \), and it is seen that

\[ g_1(t) = \int_{0}^{\infty} \frac{1}{\sqrt{2\pi \Gamma(r/2)2^{r/2}}} \left( \frac{2z}{1 + t^2/r} \right)^{(r+1)/2-1} e^{-z} \left( \frac{2}{1 + t^2/r} \right) \, dz \]

\[ = \frac{\Gamma[(r + 1)/2]}{\sqrt{\pi r \Gamma(r/2)}} \frac{1}{(1 + t^2/r)^{(r+1)/2}}, \quad -\infty < t < \infty. \]

Thus, if \( W \) is \( n(0, 1) \), if \( V \) is \( \chi^2(r) \), and if \( W \) and \( V \) are stochastically independent, then

\[ T = \frac{W}{\sqrt{V/r}} \]

has the immediately preceding p.d.f. \( g_1(t) \). The distribution of the random variable \( T \) is usually called a \emph{t distribution}. It should be observed that a \( t \) distribution is completely determined by the parameter \( r \), the
number of degrees of freedom of the random variable that has the chi-square distribution. Some approximate values of
\[
\Pr (T \leq t) = \int_{-\infty}^{t} g_1(w) \, dw
\]
for selected values of \( \tau \) and \( t \), can be found in Table IV in Appendix B.

Next consider two stochastically independent chi-square random variables \( U \) and \( V \) having \( r_1 \) and \( r_2 \) degrees of freedom, respectively. The joint p.d.f. \( \varphi(u, v) \) of \( U \) and \( V \) is then
\[
\varphi(u, v) = \frac{1}{\Gamma(r_1/2)\Gamma(r_2/2)2^{(r_1+r_2)/2}} u^{r_1/2-1}v^{r_2/2-1}e^{-(u+v)/2},
\]
\[0 < u < \infty, \, 0 < v < \infty,\]
\[= 0 \text{ elsewhere.}\]

**F Distribution**

We define the new random variable
\[
F = \frac{U}{r_1} / \frac{V}{r_2}
\]
and we propose finding the p.d.f. \( g_1(f) \) of \( F \). The equations
\[
f = \frac{u/r_1}{v/r_2}, \quad z = v,
\]
define a one-to-one transformation that maps the set \( \mathscr{A} = \{(u, v); \, 0 < u < \infty, \, 0 < v < \infty\} \) onto the set \( \mathscr{B} = \{(f, z); \, 0 < f < \infty, \, 0 < z < \infty\} \). Since \( u = (r_1/r_2)f \), \( v = z \), the absolute value of the Jacobian of the transformation is \( |J| = (r_1/r_2)z \). The joint p.d.f. \( g(f, z) \) of the random variables \( F \) and \( Z = V \) is then
\[ g(f, z) = \frac{1}{\Gamma(r_1/2) \Gamma(r_2/2) 2^{(r_1 + r_2)/2}} \left( \frac{r_1 f}{r_2} \right)^{r_1/2 - 1} z^{r_2/2 - 1} \]

\[ \times \exp \left[ -\frac{z}{2} \left( \frac{r_1 f}{r_2} + 1 \right) \right] \frac{r_1 z}{r_2} \]

provided that \((f, z) \in \mathcal{B}\), and zero elsewhere. The marginal p.d.f. \(g_1(f)\) of \(F\) is then

\[ g_1(f) = \int_{-\infty}^{\infty} g(f, z) \, dz \]

\[ = \int_{0}^{\infty} \frac{(r_1/r_2)^{r_1/2}(f)^{r_1/2 - 1}}{\Gamma(r_1/2) \Gamma(r_2/2) 2^{(r_1 + r_2)/2}} z^{(r_1 + r_2)/2 - 1} \exp \left[ -\frac{z}{2} \left( \frac{r_1 f}{r_2} + 1 \right) \right] \, dz. \]

If we change the variable of integration by writing

\[ y = \frac{z}{2} \left( \frac{r_1 f}{r_2} + 1 \right), \]

it can be seen that

\[ g_1(f) = \int_{0}^{\infty} \frac{(r_1/r_2)^{r_1/2}(f)^{r_1/2 - 1}}{\Gamma(r_1/2) \Gamma(r_2/2) 2^{(r_1 + r_2)/2}} \left( \frac{2y}{r_1 f/r_2 + 1} \right)^{(r_1 + r_2)/2 - 1} e^{-y} \]

\[ \times \left( \frac{2}{r_1 f/r_2 + 1} \right) \, dy \]

\[ = \frac{\Gamma[(r_1 + r_2)/2] (r_1/r_2)^{r_1/2}}{\Gamma(r_1/2) \Gamma(r_2/2)} \frac{(f)^{r_1/2 - 1}}{(1 + r_1 f/r_2)^{(r_1 + r_2)/2}}, \quad 0 < f < \infty, \]

\[ = 0 \text{ elsewhere.} \]
Accordingly, if $U$ and $V$ are stochastically independent chi-square variables with $r_1$ and $r_2$ degrees of freedom, respectively, then

$$F = \frac{U/r_1}{V/r_2}$$

has the immediately preceding p.d.f. $g_1(f)$. The distribution of this random variable is usually called an $F$ distribution. It should be observed that an $F$ distribution is completely determined by the two parameters $r_1$ and $r_2$. Table V in Appendix B gives some approximate values of

$$\Pr (F \leq f) = \int_0^f g_1(w) \, dw$$

for selected values of $r_1$, $r_2$, and $f$.

**EXERCISES**

4.34. Let $T$ have a $t$ distribution with 10 degrees of freedom. Find $\Pr (|T| > 2.228)$ from Table IV.

4.35. Let $T$ have a $t$ distribution with 14 degrees of freedom. Determine $b$ so that $\Pr (-b < T < b) = 0.90$.

4.36. Let $F$ have an $F$ distribution with parameters $r_1$ and $r_2$. Prove that $1/F$ has an $F$ distribution with parameters $r_2$ and $r_1$. 