5.6 Bivariate distributions

5.6.1 The joint and marginal distributions

We now broaden our previous discussion of the joint properties of two r.v.s (which was restricted to the discrete case). The joint (cumulative) distribution function of two r.v.s \((X, Y)\) is defined as

\[
F(x, y) = P(E \in S : X(E) \leq x, \text{ and } Y(E) \leq y) = P(X \leq x, Y \leq y)
\] (5.40)

The pair of r.v.s is called (jointly) continuous if its joint distribution function can be expressed as

\[
F(x, y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(u, v)\,dudv, \quad \text{for all } x, y
\] (5.41)

where the joint probability density function \(f(x, y), -\infty < x, y < \infty\) has the properties

\[
f(x, y) \geq 0, \quad -\infty < x, y < \infty,
\]

\[
\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y)\,dxdy = 1,
\]

\[
f(x, y) = \begin{cases} \frac{\partial^2 F(x, y)}{\partial x \partial y} & \text{if this derivative exists at } (x, y) \\ 0 & \text{otherwise.} \end{cases}
\]

As in the univariate case, we can give a probability interpretation to \(f(x, y)\) through the approximate relation

\[
P(x < X < x + \delta x, y < Y < y + \delta y) \approx f(x, y)\delta x\delta y.
\]

More generally, if \(A\) is a subset of \(\mathbb{R}^2\), then

\[
P((X, Y) \in A) = \int \int \int_{(x, y) \in A} f_{X,Y}(x, y)\,dxdy.
\] (5.43)

So, for example, at points of differentiability,

\[
f_X(x) = \frac{d}{dx} P(X \leq x) = \frac{d}{dx} \int_{-\infty}^{x} \int_{-\infty}^{\infty} f(u, y)\,dudy = \int_{-\infty}^{x} f(x, y)\,dy, \quad -\infty < x < \infty.
\] (5.44a)

and in this context this is termed the marginal distribution of \(X\). Similarly the marginal distribution of \(Y\) has p.d.f.

\[
f_Y(y) = \int_{-\infty}^{\infty} f(x, y)\,dx, \quad -\infty < y < \infty.
\] (5.44b)

The conditional p.d.f. of \(Y\) given \(X = x\) is defined by

\[
f_{Y|X}(y|x) = \frac{f(x, y)}{f_X(x)}, \quad -\infty < y < \infty.
\] (5.45a)

Similarly, the conditional p.d.f. of \(X\) given \(Y = y\) is defined by

\[
f_{X|Y}(x|y) = \frac{f(x, y)}{f_Y(y)}, \quad -\infty < x < \infty.
\] (5.45b)
5.6.2 Independence

We can no longer define independence of random variables in terms of events like \{X = x\}, since in the continuous case such events have zero probability. A broader definition is as follows:

The random variables \(X\) and \(Y\) are called independent if \(\{X \leq x\}\) and \(\{Y \leq y\}\) are independent events for all \(x, y\). Thus,

\[
X \text{ and } Y \text{ are independent if and only if}
\]

\[
P(X \leq x, Y \leq y) = P(X \leq x)P(Y \leq y) \quad \text{for all } x, y,
\]

i.e.

\[
F(x, y) = F_X(x)F_Y(y) \quad \text{for all } x, y.
\]

(In the discrete case, this can be shown to be equivalent to the definition given previously).

It can be proved that, if \(X\) and \(Y\) are independent, then so are \(g(X)\) and \(h(Y)\) (assuming these functions are also random variables).

It is easily shown from (5.46) that

if \(X\) and \(Y\) are jointly continuous, they are independent if and only if

\[
f(x, y) = f_X(x)f_Y(y) \quad \text{for all } x, y;
\]

or, to state a more general result, if and only if

\[
f(x, y) = (\text{function of } x),(\text{function of } y) = g(x)h(x) \quad \text{say.}
\]

This result is often used in questions of the form ‘...determine the joint p.d.f. of the r.v.s \(U\) and \(V\), then deduce that \(U\) and \(V\) are independent, and find \(f_U(u)\) and \(f_V(v)\).’

Note that if \(X\) and \(Y\) are independent, then

\[
f_Y|X(y|x) = \frac{f(x, y)}{f_X(x)} = \frac{f_X(x)f_Y(y)}{f_X(x)} = f_Y(y)
\]

as expected; i.e. information about \(X\) is irrelevant to the study of \(Y\).

5.6.3 Expectation

One can prove the bivariate form of the ‘law of the unconscious statistician’ for continuous r.v.s \(X, Y\):

\[
E(h(X, Y)) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(x, y)f(x, y)dxdy
\]

whenever this integral converges absolutely. Using this result, it is easily proved that

\[
E(aX + bY) = aE(X) + bE(Y).
\]

Note that this is true whether or not \(X\) and \(Y\) are independent. By a similar proof to that given for discrete r.v.s, it is readily shown that if \(X, Y\) are continuous independent r.v.s, then

\[
E(XY) = E(X)E(Y).
\]

Once again, the converse is false.

The conditional expectation of \(X\) given \(Y = y\) is defined as the mean of the conditional p.d.f. of \(X\) given \(Y = y\): thus

\[
E(X|Y = y) = \int_{-\infty}^{\infty} x f_X|Y(x|y)dx = \int_{-\infty}^{\infty} x \frac{f(x, y)}{f_Y(y)}dx
\]
for any value of \( y \) for which \( f_Y(y) > 0 \). By a proof analogous to that given in §2.4.3 for the discrete case, it is readily shown that

\[
E[E(X|Y)] = \int E(X|Y = y) f_Y(y) dy = E(X),
\]

(5.54)

the integral being over all \( y \) s.t. \( f_Y(y) > 0 \).

The generalisation of the definitions and results for the bivariate case to the multivariate case is generally straightforward and will not be laboured here.

5.7 The bivariate Normal distribution and its generalisation

5.7.1 Bivariate Normal distribution

In its most general form, this distribution has 5 parameters: \( \mu_x, \mu_y, \sigma_x, \sigma_y \) and \( \rho \). The joint p.d.f. of \((X, Y)\) is

\[
f(x, y) = \frac{1}{2\pi \sigma_x \sigma_y \sqrt{1 - \rho^2}} \times \exp \left\{ -\frac{1}{2(1 - \rho^2)} \left[ \left( \frac{x - \mu_x}{\sigma_x} \right)^2 - 2\rho \left( \frac{x - \mu_x}{\sigma_x} \right) \left( \frac{y - \mu_y}{\sigma_y} \right) + \left( \frac{y - \mu_y}{\sigma_y} \right)^2 \right] \right\}
\]

(5.55)

where \(-\infty < x, y < \infty\); the parameters are such that

\[-\infty < \mu_x, \mu_y < \infty; \quad \sigma_x, \sigma_y > 0; \quad -1 < \rho < 1.\]

This is referred to as the \( N(\mu_x, \mu_y; \sigma_x^2, \sigma_y^2, \rho) \) distribution. The joint p.d.f. has an asymmetric bell shape.

The marginal distribution of \( X \) has p.d.f.

\[
f_X(x) = \int_{-\infty}^{\infty} f(x, y) dy, \quad -\infty < x < \infty.
\]
In calculating the integral, we consider $x$ to be fixed: the exponent in $f(x, y)$ can be written (by 'completing the square'):

\[
\left( \frac{x - \mu_x}{\sigma_x} \right)^2 + \left\{ \left( \frac{y - \mu_y}{\sigma_y} \right)^2 - 2\rho \left( \frac{x - \mu_x}{\sigma_x} \right) \left( \frac{y - \mu_y}{\sigma_y} \right) + \rho^2 \left( \frac{x - \mu_x}{\sigma_x} \right)^2 \right\} - \rho^2 \left( \frac{x - \mu_x}{\sigma_x} \right)^2
\]

\[
= (1 - \rho^2) \left( \frac{x - \mu_x}{\sigma_x} \right)^2 + u^2,
\]

where

\[ u = \left( \frac{y - \mu_y}{\sigma_y} \right) - \rho \left( \frac{x - \mu_x}{\sigma_x} \right). \]

Now make a change of variable from $y$ to $u$: \( \frac{du}{dy} = \frac{1}{\sigma_y} \). Then

\[
f_X(x) = \int_{-\infty}^{\infty} \frac{1}{2\pi \sigma_x \sigma_y \sqrt{1 - \rho^2}} \exp \left\{ -\frac{1}{2} \left( \frac{x - \mu_x}{\sigma_x} \right)^2 \right\} \sigma_y du
\]

\[
= \frac{1}{2\pi \sigma_x \sqrt{1 - \rho^2}} \exp \left\{ -\frac{1}{2} \left( \frac{x - \mu_x}{\sigma_x} \right)^2 \right\} \int_{-\infty}^{\infty} \exp \left\{ -\frac{1}{2} \frac{u^2}{1 - \rho^2} \right\} du.
\]

But

\[
\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{1}{2} \frac{u^2}{1 - \rho^2} \right\} du = 1,
\]

since the integrand is the p.d.f. of $N(0, 1 - \rho^2)$. So

\[
f_X(x) = \frac{1}{\sqrt{2\pi \sigma_x}} \exp \left\{ -\frac{1}{2} \left( \frac{x - \mu_x}{\sigma_x} \right)^2 \right\}, \quad -\infty < x < \infty \quad (5.56)
\]

i.e. $X \sim N(\mu_x, \sigma_x^2)$, so that

\[
E(X) = \mu_x, \quad \text{Var}(X) = \sigma_x^2. \quad (5.57a)
\]

Similarly $Y \sim N(\mu_y, \sigma_y^2)$ and

\[
E(Y) = \mu_y, \quad \text{Var}(Y) = \sigma_y^2. \quad (5.57b)
\]

**Note:** It is possible to have a joint p.d.f. which has marginal p.d.f.s which are Normal, yet which is not bivariate normal.

The conditional p.d.f. $f_{Y|X}(y|x)$ is defined as

\[
\frac{f(x, y)}{f_X(x)} = \frac{1}{\sqrt{2\pi \sigma_y \sqrt{1 - \rho^2}}} \exp \left\{ -\frac{1}{2(1 - \rho^2)} \frac{(y - (\mu_y + \rho \frac{\sigma_y}{\sigma_x} (x - \mu_x)))^2}{\sigma_y^2} \right\},
\]

i.e., $f_{Y|X}(y|x)$ is the p.d.f. of

\[
N(\mu_y + \rho \frac{\sigma_y}{\sigma_x} (x - \mu_x), (1 - \rho^2)\sigma_y^2). \quad (5.58a)
\]

Similarly, $f_{X|Y}(x|y)$ is the p.d.f. of

\[
N(\mu_x + \rho \frac{\sigma_x}{\sigma_y} (y - \mu_y), (1 - \rho^2)\sigma_x^2). \quad (5.58b)
\]

Many calculations involving the bivariate Normal distribution can be done in terms of the standard bivariate Normal distribution $N(0, 0; 1, 1; \rho)$. Let

\[
U = \frac{X - \mu_x}{\sigma_x}, \quad V = \frac{Y - \mu_y}{\sigma_y}. \quad (5.59)
\]
Then, as we shall prove later, the joint p.d.f. of \((U, V)\) is the standard bivariate Normal distribution, with

\[
f^{S}(u, v) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left\{ -\frac{1}{2(1-\rho^2)} [u^2 - 2\rho uv + v^2] \right\}, \quad -\infty < u, v < \infty. \tag{5.60}
\]

Then

\[
F(x, y) = \Pr(X \leq x, Y \leq y)
= \Pr\left( \frac{U - \mu_x}{\sigma_x} \leq \frac{y - \mu_y}{\sigma_y} \right)
= F^S\left( \frac{x - \mu_x}{\sigma_x}, \frac{y - \mu_y}{\sigma_y} \right).
\]

Bivariate moments of \(X, Y\) are most easily calculated from the moments of \(U, V\). In particular, we show later that

\[
\rho(X, Y) = \rho(U, V) = \rho. \tag{5.61}
\]

Now when \(\rho = 0, f(x, y) = g(x)h(y)\) for all \(x, y\). So for the bivariate Normal distribution (but not in general)

\[
\rho = 0 \Rightarrow X \text{ and } Y \text{ are independent r.v.s.}
\]

### 5.7.2 Multivariate Normal distribution

(NOTE: Not required for examination purposes.)

The r.v.s \((X_1, X_2, \ldots, X_p)\) have the **multivariate Normal distribution** (or multinormal distribution) if the joint p.d.f. is

\[
f(x_1, x_2, \ldots, x_p) = \frac{1}{(2\pi)^{p/2} |V|^{1/2}} \exp\left\{ -\frac{1}{2} (x - \mu)^t V^{-1} (x - \mu) \right\}, \tag{5.62}
\]

where \(-\infty < x_i < \infty, \quad i = 1, \ldots, p\); here

\[(x - \mu)' = (x_1 - \mu_1, x_2 - \mu_2, \ldots, x_p - \mu_p) \quad (\ldots)' \text{ denotes the transpose},\]

\(V\) is the variance-covariance matrix of \((X_1, \ldots, X_p)\), i.e.

\[
V = \begin{pmatrix}
\text{Var}(X_1) & \text{Cov}(X_1, X_2) & \ldots & \ldots & \text{Cov}(X_1, X_p) \\
\text{Cov}(X_2, X_1) & \text{Var}(X_2) & \ldots & \ldots & \text{Cov}(X_2, X_p) \\
\vdots & \vdots & \ddots & \ddots & \vdots \\
\text{Cov}(X_p, X_1) & \text{Cov}(X_p, X_2) & \ldots & \text{Var}(X_p)
\end{pmatrix},
\]

a symmetric matrix with \((i, j)\)th element

\[
\text{Cov}(X_i, X_j) = \rho(X_i, X_j) \sqrt{\text{Var}(X_i)\text{Var}(X_j)}.
\]

Many marginal distributions can be derived, but in particular

\[
X_i \sim N(\mu_i, \sigma_i^2), \quad \text{where } \sigma_i^2 = \text{Var}(X_i). \tag{5.63}
\]

There is a convenient matrix notation for means, variances, covariances etc. Introduce

\[
X = \begin{bmatrix}
X_1 \\
\vdots \\
X_p
\end{bmatrix},
\]
- a $p \times 1$ vector of random variables, and

$$E(X) = \begin{bmatrix} E(X_1) \\ \vdots \\ E(X_p) \end{bmatrix}$$

a $p \times 1$ vector of means. Then the $p \times p$ covariance matrix (of $X$ or $X_1, \ldots, X_p$) is denoted by $\text{Var}(X)$, and

$$\text{Var}(X) = E[(X - E(X)(X - E(X)))'] = E(XX' - E(X)E(X))'.$$

For the linear combination $\sum_{i=1}^{p} a_i X_i = \mathbf{a}' \mathbf{X}$, we have

$$E(\sum_{i=1}^{p} a_i X_i) = E(\mathbf{a}' \mathbf{X}) = \mathbf{a}'E(\mathbf{X}),$$

- the scalar product of a $1 \times p$ vector and a $p \times 1$ vector, and

$$\text{Var}(\sum_{i=1}^{p} a_i X_i) = \text{Var}(\mathbf{a}' \mathbf{X}) = \mathbf{a}'\text{Var}(\mathbf{X})\mathbf{a}$$

- $(1 \times p) \times (p \times p) \times (p \times 1)$. 
5.8 Functions of several random variables

We first discuss methods of finding the p.d.f. of a function of \((X, Y)\), before generalizing.

5.8.1 Transformation rule

Let the continuous r.v.s \((X, Y)\) have joint p.d.f. \(f_{X,Y}(x, y)\), and let \(A = \{(x, y) : f_{X,Y} > 0\}\). Consider

\[ U = H_1(X, Y), \quad V = H_2(X, Y), \]

where the partial derivatives of \(H_1, H_2\) exist and are continuous at all \((x, y) \in A\). Suppose further that the transformation

\[ u = H_1(x, y), \quad v = H_2(x, y) \quad (5.64a) \]

is one-to-one and maps \(A\) (in the \((x, y)\) plane) onto \(B\) (in the \((u, v)\) plane): then there is an inverse transformation

\[ x = G_1(u, v), \quad y = G_2(u, v). \quad (5.64b) \]

which maps \(B\) onto \(A\). The Jacobian of the original transformation is the determinant

\[
J(u, v; x, y) = \frac{\partial(u, v)}{\partial(x, y)} = \begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{vmatrix}. \quad (5.65a)
\]

The Jacobian of the inverse transformation is

\[
J(x, y; u, v) = \frac{\partial(x, y)}{\partial(u, v)} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix}. \quad (5.65b)
\]

It is the latter that we will require, but since the product of the two Jacobians is 1, we calculate whichever is easier to do. Consider any \(A \subseteq \mathcal{A}\) and suppose that under (5.64a) it is mapped into \(B \subseteq \mathcal{B}\). Then

\[
P((X, Y) \in A) = \int \int_A f_{X,Y}(x, y) dxdy = \int \int_B f_{X,Y}(G_1(u, v), G_2(u, v)) |J(x, y; u, v)| du dv
\]

by a theorem in analysis, while

\[
P((U, V) \in B) = \int \int_B f_{U,V}(u, v) du dv.
\]

But \(P((X, Y) \in A) = P((U, V) \in B)\) for all \(A \subseteq \mathcal{A}\): this implies that

\[
f_{U,V}(u, v) = \begin{cases} f_{X,Y}(G_1(u, v), G_2(u, v)) |J(x, y; u, v)|, & \text{if } (u, v) \in \mathcal{B} \\ 0, & \text{otherwise.} \end{cases} \quad (5.66)
\]

which is the required transformation rule.

Some practical aspects need to be mentioned before we look at examples.

(i) Suppose we wish to find the p.d.f. of the continuous r.v. \(U = H_1(X, Y)\). We introduce a second continuous r.v. \(V = H_2(X, Y)\) such that \(H_1, H_2\) have the above properties, and use the transformation rule (5.66) to obtain the joint p.d.f. of \((U, V)\), \(f_{U,V}(u, v)\). The p.d.f. of \(U\) is then obtained as a marginal p.d.f. of \(f_{U,V}(u, v)\), by integrating \(f_{U,V}(u, v)\) with respect to \(v\) over the appropriate range of \(v\). Naturally, we choose \(V\) so as to make the calculations as easy as possible!
(ii) Suppose we wish to show that the continuous r.v.s \( U = H_1(X,Y) \) and \( V = H_2(X,Y) \) are independent. If we can find the joint p.d.f. of \( (U,V) \), \( f_{U,V}(u,v) \), and it factorises into a function of \( u \) times a function of \( v \), then by (5.48) \( U \) and \( V \) are independent.

(iii) Many-to-one cases can be handled in a manner analogous to univariate transformations (not required in this Module).

(iv) The technique can be extended without difficulty to \( p \)-dimensional r.v.s \( (p \geq 3) \). Now \( p \) r.v.s \( X_1, X_2, ..., X_p \) → \( p \) r.v.s \( U_1, U_2, ..., U_p \) and the Jacobian \( J(x_1, ..., x_p; u_1, ..., u_p) \) is a \( p \times p \) determinant.

5.8.2 Examples

Example 1
Suppose that
\[
(X,Y) \sim N(\mu_x, \mu_y; \sigma_x^2, \sigma_y^2, \rho).
\]
Show that
\[
U = \frac{X - \mu_x}{\sigma_x}, \quad V = \frac{Y - \mu_y}{\sigma_y}
\]
have the standard bivariate Normal distribution (anticipated at end of §5.7.1).

Solution  The transformation
\[
u = \frac{x - \mu_x}{\sigma_x}, \quad v = \frac{y - \mu_y}{\sigma_y}
\]
is one-to-one and has the inverse
\[
x = \mu_x + \sigma_x u, \quad y = \mu_y + \sigma_y v.
\]
Also
\[-\infty < x, y < \infty \quad \rightarrow \quad -\infty < u, v < \infty
\]
and
\[
J(x,y;u,v) = \begin{vmatrix}
\frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\
\frac{\partial y}{\partial u} & \frac{\partial y}{\partial v}
\end{vmatrix} = \begin{vmatrix}
\sigma_x & 0 \\
0 & \sigma_y
\end{vmatrix} = \sigma_x \sigma_y.
\]
Now from (5.55) the joint p.d.f. of \( (X,Y) \) is
\[
f_{X,Y}(x,y) = \frac{1}{2\pi \sigma_x \sigma_y \sqrt{1 - \rho^2}} \exp \left\{-\frac{1}{2(1 - \rho^2)} \left[ \left( \frac{x - \mu_x}{\sigma_x} \right)^2 - 2\rho \left( \frac{x - \mu_x}{\sigma_x} \right) \left( \frac{y - \mu_y}{\sigma_y} \right) + \left( \frac{y - \mu_y}{\sigma_y} \right)^2 \right] \right\}
\]
where \(-\infty < x, y < \infty\). So by (5.66) the joint p.d.f. of \( (U,V) \) is
\[
f_{U,V}(u,v) = f_{X,Y}(\mu_x + \sigma_x u, \mu_y + \sigma_y v) |J(x,y;u,v)|
\]
\[
= \frac{1}{2\pi \rho_x \rho_y \sqrt{1 - \rho^2}} \left\{-\frac{1}{2(1 - \rho^2)} \left[ v^2 - 2\rho uv + u^2 \right] \right\} \rho_x \rho_y, \quad -\infty < u, v < \infty
\]
i.e. \( (U,V) \sim N(0,0;1,1;\rho) \) (see (5.60)). □
Example 2

Let $Z$ and $V$ be independent r.v.s, where $Z \sim N(0, 1)$ and $V \sim \chi_r^2$, i.e.

$$f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}, \quad -\infty < z < \infty;$$

$$f_V(v) = \frac{1}{2^\frac{r}{2} \Gamma\left(\frac{r}{2}\right)} v^{\frac{r}{2} - 1} e^{-\frac{1}{2}v}, \quad 0 \leq v < \infty.$$ 

Find the p.d.f. of the r.v. $T = \frac{Z}{\sqrt{V/r}}$.

**Solution**

Since $Z$ and $V$ are independent, the joint p.d.f. of $(Z, V)$ is

$$f_{Z,V}(z, v) = f_Z(z) f_V(v), \quad -\infty < z < \infty, 0 \leq v < \infty.$$ 

Consider the transformation

$$t = \frac{z}{\sqrt{v/r}}, \quad u = v.$$

It is one-to-one and has inverse

$$z = t\sqrt{u/r}, \quad v = u; \quad -\infty < t < \infty, \quad 0 \leq u < \infty.$$ 

Also

$$J(t, u; z, v) = \begin{vmatrix} \frac{1}{\sqrt{v/r}} & -\frac{z\sqrt{r}}{2v^{3/2}} \\ 0 & \frac{1}{1} \end{vmatrix} = \frac{1}{\sqrt{v/r}} = \frac{1}{\sqrt{u/r}}.$$

So

$$f_{T,U}(t, u) = \frac{f_{Z,V}(t\sqrt{u/r}, u) |J(z, v; t, u)|}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2u/r}, \quad \frac{1}{2^\frac{r}{2} \Gamma\left(\frac{r}{2}\right)} u^{\frac{r}{2} - 1} e^{-u/2}, \quad 0 \leq u < \infty.$$ 

So the p.d.f. of $T$ is

$$f_T(t) = \int_0^\infty f_{T,U}(t, u) du = \frac{1}{\sqrt{2\pi r} 2^\frac{r}{2} \Gamma\left(\frac{r}{2}\right)} \int_0^\infty u^{\frac{r}{2} - 1} e^{-\frac{1}{2}t^2u/r} du.$$ 

We can express this as a Gamma function integral by changing a variable: introduce

$$w = \frac{u}{2}(1 + t^2/r)$$

or

$$u = \frac{2w}{(1 + t^2/r)}.$$ 

Then

$$f_T(t) = \ldots \int_0^\infty \left(\frac{2w}{1 + t^2/r}\right)^{\frac{r}{2} - 1} e^{-w} \left(\frac{2}{1 + t^2/r}\right) dw = \frac{\sqrt{\pi} \Gamma\left(\frac{r}{2}\right)}{\Gamma\left(\frac{r+1}{2}\right)} \frac{\sqrt{\pi} \Gamma\left(\frac{r+1}{2}\right)}{\sqrt{\pi} \Gamma\left(\frac{r}{2}\right)} (1 + t^2/r)^{-\frac{r+1}{2}}, \quad -\infty < t < \infty.$$ 

This is the p.d.f. of Student’s $t$-distribution with $r$ degrees of freedom. \qed
Example 3

Suppose that $X$ and $Y$ are independent r.v.s, where

$$X \sim \chi_m^2, \quad Y \sim \chi_n^2.$$ 

Show that

$$U = X + Y, \quad V = \frac{(X/m)}{(Y/n)}$$

are independent r.v.s and find their distributions.

**Solution** We have

$$f_X(x) = \frac{1}{2^\frac{m}{2}\Gamma(\frac{m}{2})} x^{\frac{m}{2} - 1} e^{-\frac{x}{2}}, \quad 0 \leq x < \infty$$

$$f_Y(y) = C_n y^{\frac{n}{2} - 1} e^{-\frac{y}{2}}, \quad 0 \leq y < \infty.$$ 

The joint p.d.f. of $(U, V)$ is

$$f_{X,Y}(x, y) = f_X(x) f_Y(y) \quad \text{(independence)}$$

$$= C_m C_n \frac{x^{\frac{m}{2} - 1} y^{\frac{n}{2} - 1} e^{-\frac{x+y}{2}}}{(m+n)^{\frac{m+n}{2}}}, \quad 0 \leq x, y < \infty.$$ 

The transformation

$$u = x + y, \quad v = \frac{(x/m)}{(y/n)}$$

is one-to-one with inverse

$$x = \frac{muv}{mv + n}, \quad y = \frac{nu}{mv + n}, \quad 0 \leq u, v < \infty.$$ 

Also

$$J(u, v; x, y) = \begin{vmatrix} \frac{n}{my} & \frac{u}{my} \\ \frac{n}{my} & \frac{v}{my} \end{vmatrix} = -\frac{n(x+y)}{my^2} = \frac{(mv+n)^2}{mnu}.$$ 

So the joint p.d.f. of $(U, V)$ is

$$f_{U,V}(u, v) = f_{X,Y}\left(\frac{mu}{mv+n}, \frac{nu}{mv+n}\right) \left(\frac{(mv+n)^2}{mnu}\right)^{-1}$$

$$= C_m C_n \frac{\left(\frac{nu}{mv+n}\right)^{\frac{n}{2} - 1} e^{-\frac{nu}{2}}}{\left(\frac{nu}{mv+n}\right)^{\frac{n}{2}} (mv+n)^{\frac{m+n}{2}}} \cdot \frac{\left(\frac{mu}{mv+n}\right)^{\frac{m}{2} - 1} e^{-\frac{mu}{2}}}{\left(\frac{mu}{mv+n}\right)^{\frac{m}{2}} (mv+n)^{\frac{m+n}{2}}}, \quad 0 \leq u, v < \infty,$$

i.e. $f_{U,V}(u, v) = \text{(function of u)} \times \text{(function of v)}$ for all $(u, v)$. It follows that $U$ and $V$ are independent r.v.s. Also

$$f_U(u) = Au^{\frac{m+n}{2} - 1} e^{-\frac{u}{2}}, \quad 0 \leq u < \infty;$$

$$f_V(v) = Bv^{\frac{m+n}{2} - 1} \left(\frac{mv+n}{mv+n}\right)^{\frac{m+n}{2}}, \quad 0 \leq v < \infty,$$

where $A$ and $B$ are constants such that

$$\int_0^\infty f_U(u) du = \int_0^\infty f_V(v) dv = 1$$

and also that

$$A \cdot B = C_m C_n m^{\frac{m}{2}} n^{\frac{n}{2}}. \quad (5.67)$$

By inspection, or by integration with respect to $u$ and using the Gamma function, we obtain

$$A = \left[2^{\frac{m+n}{2}} \Gamma\left(\frac{m+n}{2}\right)\right]^{-1} = C_{m+n},$$
and $U \sim \chi^2_{m+n}$; i.e. $U$ has the $\chi^2$-distribution with $(m + n)$ degrees of freedom.

Then from (5.67) we find that

$$B = \frac{m^{\frac{m}{2}} n^{\frac{n}{2}}}{B\left(\frac{m}{2}, \frac{n}{2}\right)}$$

where $B\left(\frac{m}{2}, \frac{n}{2}\right) = \frac{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{m+n}{2}\right)}$,

so that $V \sim F_{m,n}$; i.e. $V$ has the $F$-distribution with $(m, n)$ degrees of freedom.

**Example 4**

Suppose the independent r.v.s $X$ and $Y$ are each uniformly distributed on $[0, 1]$. Find the joint p.d.f. of

$$U = \frac{X}{Y}, \quad V = XY,$$

and hence find the p.d.f. of $U$.

**Solution** We have

$$f_X(x) = \begin{cases} 1, & 0 \leq x \leq 1, \\ 0, & \text{otherwise}; \end{cases} \quad f_Y(y) = \begin{cases} 1, & 0 \leq y \leq 1, \\ 0, & \text{otherwise}. \end{cases}$$

Then

$$f_{X,Y} = f_X(x) f_Y(y) [\text{independence}]$$

$$= \begin{cases} 1, & 0 \leq x, y \leq 1, \\ 0, & \text{otherwise}. \end{cases}$$

The transformation

$$u = \frac{x}{y}, \quad v = xy, \quad 0 \leq x, y \leq 1$$

is one-to-one and has inverse

$$x = \sqrt{uv}, \quad y = \frac{v}{u}.$$ 

Also

$$J(u, v; x, y) = \begin{vmatrix} \frac{1}{y} & -\frac{x}{y^2} \\ \frac{1}{y} & \frac{x}{y} \end{vmatrix} = 2\frac{x}{y} = 2u.$$ 

So

$$f_{U,V}(u, v) = f_{X,Y}(\sqrt{uv}, \sqrt{\frac{v}{u}}) \frac{1}{2u} = \begin{cases} \frac{1}{2u}, & 0 \leq \sqrt{uv} \leq 1, 0 \leq \sqrt{\frac{v}{u}} \leq 1, \\ 0, & \text{otherwise}. \end{cases}$$

Now

$$0 \leq \sqrt{uv} \leq 1, \quad 0 \leq \sqrt{\frac{v}{u}} \leq 1 \quad \implies \quad 0 \leq v \leq \frac{1}{u}, \quad 0 \leq v \leq u.$$ 

Only one or other of these ranges need be retained, depending on whether $u$ is in $[0, 1]$ or $[1, \infty]$. Thus

$$f_{U,V}(u, v) = \begin{cases} \frac{1}{2u}, & 0 \leq u \leq 1, 0 \leq v \leq u; \\ \frac{1}{2u}, & 1 \leq u < \infty, 0 \leq v \leq \frac{1}{u} \\ 0, & \text{otherwise}. \end{cases}$$

Then

$$f_U(u) = \int_{-\infty}^{\infty} f_{U,V}(u, v) dv$$

$$= \begin{cases} \int_{0}^{1} \frac{1}{2u} dv, & 0 \leq u \leq 1 \\ \int_{0}^{\frac{1}{u}} \frac{1}{2u} dv, & 1 \leq u < \infty \end{cases} = \begin{cases} \frac{1}{2}, & 0 \leq u \leq 1 \\ \frac{1}{2u^2}, & 1 \leq u < \infty. \end{cases}$$

$\square$
5.9 Orthogonal transformations

We now proceed to examine a particularly important multivariate situation. Let \( Z_1, Z_2, \ldots, Z_n \) be independent \( N(0, 1) \) r.v.s. Consider the orthogonal transformation

\[
Y_1 = c_{11}Z_1 + \cdots + c_{1n}Z_n \\
Y_2 = c_{21}Z_1 + \cdots + c_{2n}Z_n \\
\vdots \quad \vdots \\
Y_n = c_{n1}Z_1 + \cdots + c_{nn}Z_n
\]

or

\[
Y = CZ
\]

(5.68)

where \( C \) is an orthogonal matrix, i.e.

\[
C' = CC' = I_n
\]

(5.69)

where \( I_n \) is the \( n \times n \) identity matrix

\[
\begin{pmatrix}
1 & 0 & 0 & \cdots & 0 \\
0 & 1 & 0 & \cdots & 0 \\
0 & 0 & 1 & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \cdots & 1
\end{pmatrix}
\]

\( C \) has the following properties:

(i) for any row, the sum of the squared elements is 1, i.e.

\[
\sum_{j=1}^{n} c_{ij}^2 = 1;
\]

(ii) for any two rows, the sum of the products of corresponding elements is 0, i.e.

\[
\sum_{j=1}^{n} c_{ij}c_{kj} = 0, \quad i \neq k;
\]

(iii) \( |C| = \pm 1. \)

Such a transformation is one-to-one with inverse \( Z = C^{-1}Y = C'Y \), and the Jacobian

\[ J(y_1, \ldots, y_n, z_1, \ldots, z_n) = |C| = \pm 1 \]

since \( \frac{\partial y_i}{\partial z_i} = c_{ij} \).

Also

\[
\sum_{i=1}^{n} y_i^2 = Y'Y = Z'C'CZ = Z'I_nZ = Z'Z = \sum_{i=1}^{n} Z_i^2; \quad (5.70)
\]

\[
\sum_{i=1}^{n} y_i^2 = \sum_{i=1}^{n} z_i^2.
\]

Now

\[
f_{Z_i}(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z_i^2}, \quad -\infty < z < \infty \quad (i = 1, \ldots, n).
\]

So the joint p.d.f. of \( Z_1, \ldots, Z_n \) is

\[
f_{Z_1,\ldots,Z_n}(z_1, \ldots, z_n) = f_{Z_1}(z_1)f_{Z_2}(z_2)\cdots f_{Z_n}(z_n) \quad \text{[independence]}
\]

\[
= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}z_i^2\right\}
\]

\[
= \left(\frac{1}{\sqrt{2\pi}}\right)^n \exp\left\{-\frac{1}{2} \sum_{i=1}^{n} z_i^2\right\}, \quad -\infty < z_i < \infty.
\]
Then
\[ f_{Y_1,\ldots,Y_n}(y_1, \ldots, y_n) = f_{Z_1,\ldots,Z_n}(z_1, \ldots, z_n) |J(z_1, \ldots, z_n; y_1, \ldots, y_n)| \]
\[ = \left( \frac{1}{\sqrt{2\pi}} \right)^n \exp\left\{ -\frac{1}{2} \sum_{i=1}^n z_i^2 \right\} \]
\[ = \left( \frac{1}{\sqrt{2\pi}} \right)^n \exp\left\{ -\frac{1}{2} \sum_{i=1}^n y_i^2 \right\} \]
\[ = \prod_{i=1}^n \frac{1}{\sqrt{2\pi}} \exp\left\{ -\frac{1}{2} y_i^2 \right\}, \]
i.e. \( Y_1, \ldots, Y_n \) are independent \( N(0, 1) \) r.v.s.

### 5.10 Some applications to sampling theory

#### 5.10.1 Sampling from \( N(0, 1) \)

Now, in the orthogonal transformation considered in the previous section, let the first row of \( C \) be \( \frac{1}{\sqrt{n}}, \frac{1}{\sqrt{n}}, \ldots, \frac{1}{\sqrt{n}} \), so that
\[ Y_1 = \sum_{i=1}^n \frac{1}{\sqrt{n}} Z_i = \frac{1}{\sqrt{n}} \sum_{i=1}^n Z_i = \sqrt{n} \bar{Z} \]
where \( \bar{Z} = \frac{1}{n} \sum_{i=1}^n Z_i \). The other rows can be chosen in any way which gives an orthogonal \( C \): in any case,
\[ V \equiv \sum_{i=1}^n (Z_i - \bar{Z})^2 = \sum_{i=1}^n Z_i^2 - n \bar{Z}^2 \]
\[ = \sum_{i=1}^n Y_i^2 - Y_1^2 = \sum_{i=2}^n Y_i^2. \]

So, since the \( \{Y_i\} \) are independent r.v.s, \( \sqrt{n} \bar{Z} \) and \( V \) are independent r.v.s, i.e. \( \bar{Z} \) and \( \frac{1}{n-1} \sum_{i=1}^n (Z_i - \bar{Z})^2 \) are independent r.v.s, i.e. the sample mean r.v. and the sample variance r.v. for a random sample from \( N(0, 1) \) are independent r.v.s.

Now
\[ Y_i \sim N(0, 1), \quad i = 1, \ldots, n \]
so
\[ Y_i^2 \sim \chi_1^2 \quad \text{[see HW Ex.6, Qn. 3(ii)].} \]

Hence
\[ \sqrt{n} \bar{Z} = Y_1 \sim N(0, 1) \Rightarrow \bar{Z} \sim N(0, \frac{1}{n}) \quad (5.71) \]
and
\[ V = Y_2^2 + \cdots + Y_n^2 \sim \chi_{n-1}^2 \quad [\text{see Example 3 in §5.8.2}] \quad (5.72) \]
5.10.2 Sampling from $N(\mu, \sigma^2)$

Suppose that $X_1, \ldots, X_n$ are independent $N(\mu, \sigma^2)$ r.v.s corresponding to a random sample $(x_1, \ldots, x_n)$ from $N(\mu, \sigma^2)$. Let

$$Z_i = \frac{X_i - \mu}{\sigma}, \quad i = 1, \ldots, n.$$ 

Then $Z_1, \ldots, Z_n$ are independent $N(0, 1)$ r.v.s.

Consider the sample mean r.v.

$$\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$$

and the sample variance r.v.

$$S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2.$$ 

Since

$$Z = \frac{\bar{X} - \mu}{\sigma}, \quad (n-1)\frac{S^2}{\sigma^2} = \sum_{i=1}^{n} (Z_i - Z)^2,$$

we have that

$$\frac{\sqrt{n}}{\sigma} \frac{\bar{X} - \mu}{\sigma} = \sqrt{n} Z = Y_1,$$

and

$$(n-1)\frac{S^2}{\sigma^2} = V.$$ 

Since $Y_1$ and $V$ are independent, it follows that $\sqrt{n}(\bar{X} - \mu)/\sigma$ and $(n-1)S^2/\sigma^2$ are independent. Hence $\bar{X}$ and $S^2$ are independent r.v.s and

$$\bar{X} = \mu + \frac{\sigma}{\sqrt{n}} Y_1 \sim N(\mu, \frac{\sigma^2}{n}); \quad (n-1)\frac{S^2}{\sigma^2} \sim \chi^2_{n-1}. \quad (5.73)$$

5.10.3 Test statistic for t-test of a mean

Since

$$\frac{\sqrt{n}}{\sigma} \frac{\bar{X} - \mu}{\sigma} \sim N(0, 1)$$

and

$$(n-1)\frac{S^2}{\sigma^2} \sim \chi^2_{n-1}$$

are independent r.v.s,

$$\frac{\sqrt{n}}{\sigma} \frac{\bar{X} - \mu}{\sigma} \div \sqrt{(n-1)\frac{S^2}{\sigma^2} \div (n-1)} \sim t_{n-1}$$

(see Example 2 in §5.8.2), i.e.

$$\frac{\bar{X} - \mu}{\sqrt{S^2/n}} \sim t_{n-1}. \quad (5.74)$$
5.10.4 Test statistic for F-test of two variances

Let $X_{11},...,X_{1m}$ be $m$ independent r.v.s, each distributed $N(\mu_1, \sigma_1^2)$. Let $X_{21},...,X_{2n}$ be $n$ independent r.v.s, each distributed $N(\mu_2, \sigma_2^2)$, independent of the first set. Define the sample variance r.v. for the first distribution:

$$S_1^2 = \frac{1}{m-1} \sum_{j=1}^{m} (X_{1j} - \overline{X}_1)^2, \quad \text{where } \overline{X}_1 = \frac{1}{m} \sum_{j=1}^{m} X_{1j}/m.$$  

Similarly for the second distribution:

$$S_2^2 = \frac{1}{n-1} \sum_{j=1}^{n} (X_{2j} - \overline{X}_2)^2, \quad \text{where } \overline{X}_2 = \frac{1}{n} \sum_{j=1}^{n} X_{2j}/n.$$  

From §5.10.2 above, we have

$$(m-1)\frac{S_1^2}{\sigma_1^2} \sim \chi^2_{m-1},$$  

$$(n-1)\frac{S_2^2}{\sigma_2^2} \sim \chi^2_{n-1},$$

and these r.v.s are also independent. Then (from Example 3, §5.8.2),

$$\frac{(m-1)S_1^2/\sigma_1^2}{(n-1)S_2^2/\sigma_2^2} \sim F_{m-1,n-1}. \quad (5.75a)$$

If $\sigma_1^2 = \sigma_2^2$, this simplifies to

$$\frac{S_1^2}{S_2^2} \sim F_{m-1,n-1}. \quad (5.75b)$$
5.11 Order statistic random variables

Let $x_1, \ldots, x_n$ be a random sample of size $n$ on the continuous r.v. $X$ with p.d.f. $f(x)$, c.d.f. $F(x), -\infty < x < \infty$.

OR

Let $X_1, \ldots, X_n$ be independent r.v.s, each having the same distribution as $X$; and let $x_1$ be a random sample of size 1 on $X_1$, $x_2$ a random sample of size 1 on $X_2$, and so on.

In applications we generally use the first formulation, in theoretical work the second.

Rearrange the sample $x_1, \ldots, x_n$ in ascending order:

$$ x(1) \leq x(2) \leq \cdots \leq x(n). $$

Then $x(i)$ is the $i^{th}$ order statistic of the sample $(x_1, \ldots, x_n)$ and is an observation on the r.v. $X_{(i,n)} = X_{(i)}$ which is described as the $i^{th}$ order statistic r.v. (If we consider repeated samples of size $n$ from the p.d.f. $f(x)$, the $i^{th}$ order statistic values have a distribution with associated r.v. $X_{(i)}$). $\{X_{(1)}, X_{(2)}, \ldots, X_{(n)}\}$ is the set of order statistic r.v.s associated with the set $\{X_1, X_2, \ldots, X_n\}$.

We wish to find, for example, the p.d.f. of $X_{(i)}$, $i = 1, \ldots, n$ and the joint p.d.f. of $X_{(i)}, X_{(j)}$, $i < j$.

We start with two simple cases.

(i) p.d.f. of $X_{(1)}$

($X_{(1)}$ is the r.v. associated with the smallest observation in the sample of size $n$.)

We have that

$$ P(X_{(1)} > x) = P(X_1 > x, X_2 > x, \ldots, X_n > x) = P(X_1 > x)P(X_2 > x)\cdots P(X_n > x) \quad [\text{independence}] $$

$$ = \{1 - F(x)\}^n. $$

So the c.d.f. of $X_{(1)}$ is

$$ F_{(1)}(x) = P(X_{(1)} \leq x) = 1 - \{1 - F(x)\}^n, $$

and the p.d.f. is then

$$ f_{(1)}(x) = \frac{dF_{(1)}(x)}{dx} = n\{1 - F(x)\}^{n-1}f(x), \quad -\infty < x < \infty. \quad (5.76) $$

(ii) p.d.f. of $X_{(n)}$

($X_{(n)}$ is the r.v. associated with the largest observation in the sample of size $n$.)

By a similar argument, we have that the c.d.f. of $X_{(n)}$ is

$$ F_{(n)}(x) = P(X_{(n)} \leq x) = \{F(x)\}^n $$

so the p.d.f. is

$$ f_{(n)}(x) = n\{F(x)\}^{n-1}f(x), \quad -\infty < x < \infty. \quad (5.77) $$
(iii) p.d.f. of $X_{(i)}$

Two derivations of this general p.d.f. can be given.

The first procedure is similar to that used above, i.e. we establish, and then differentiate, the c.d.f. The c.d.f. of $X_{(i)}$ is as follows:

$$F_{(i)}(x) = \frac{P(X_{(i)} \leq x)}{n(i \text{ or more } X_j s \text{ are } \leq x)} = \sum_{k=i}^{n} \binom{n}{k} \{F(x)\}^k (1 - F(x))^{n-k}, \quad 1 \leq i \leq n,$$

where we have invoked the binomial distribution with $p = P(X_j \leq x) = F(x)$.

For $i = n$, differentiation of (the single term of) this expression yields the p.d.f. obtained in (ii). For $i < n$, differentiation gives

$$f_{(i)}(x) = \sum_{k=i}^{n-1} \binom{n}{k} \{F(x)\}^{k-1} (1 - F(x))^{n-k-1} f(x) \{k(1 - F(x)) - (n - k)F(x)\} + n \{F(x)\}^{n-1} f(x).$$

Rearranging terms we get

$$f_{(i)}(x) = i \binom{n}{i} \{F(x)\}^{i-1} (1 - F(x))^{n-i} f(x) - \sum_{k=i}^{n-1} \{(n-k) \binom{n}{k} - (k+1) \binom{n}{k+1}\} \{F(x)\}^k (1 - F(x))^{n-k-1} f(x).$$

In the summation, each term has coefficient 0 (i.e. cancellation occurs), leaving us with

$$f_{(i)}(x) = \frac{n!}{(i-1)!(n-i)!} \{F(x)\}^{i-1} (1 - F(x))^{n-i} f(x). \quad (5.78)$$

- an expression which also holds for $i = n$.

An alternative derivation of (5.78) proceeds as follows. Divide the real axis into 3 parts: $(-\infty, x], (x, x+h], \ (x+h, +\infty)$. Then the probability that $(i-1)$ of the sample values fall in $(-\infty, x]$, one value in $(x, x+h], \text{ and } (n-i)$ values in $(x+h, +\infty)$ is given by the multinomial distribution

$$\frac{n!}{(i-1)!(n-i)!} \{F(x)\}^{i-1} \{1 - F(x)\}^{n-i} \int_x^{x+h} f(t) dt.$$

But this probability can also be written as

$$P(x < X_{(i)} \leq x + h) = \int_x^{x+h} f_{(i)}(t) dt.$$ 

Invoking the mean value theorem for integrals, the integrals can be written

$$\int_x^{x+h} f_{(i)}(t) dt = f_{(i)}(x + h') h', \quad \text{where } 0 \leq h' \leq h;$$

$$\int_x^{x+h} f(t) dt = f(x + h'') h, \quad \text{where } 0 \leq h'' \leq h.$$

So

$$f_{(i)}(x + h') h = \frac{n!}{(i-1)!(n-i)!} \{F(x)\}^{i-1} (1 - F(x))^{n-i} f(x + h'). \ h$$

Then in the limit $h \rightarrow 0, \text{both } h' \text{ and } h'' \rightarrow 0$, giving the result (5.78).
(Note as a check that (5.78) is correct for the cases \( i = 1 \) and \( i = n \) discussed in (i) and (ii) above).

(iv) Joint p.d.f. of \( X_{(i)}, X_{(j)}, i < j \)

Similar arguments can be used here. For example, divide the real axis into

\[ (-\infty, u], \ (u, u+s], \ (u+s, v], \ (v, v+t], \ (v+t, +\infty); \]

then the probability that \( (i-1) \) values are in \( (\infty, u] \), one value in \( (u, u+s] \), \( (j-i-1) \) values in \( (u+s, v] \), one value in \( (v, v+t] \) and \( (n-j) \) values in \( (v+t, +\infty) \) is (again invoking the multinomial distribution)

\[
\int_u^{u+s} \int_v^{v+t} f_{(i)(j)}(x,y) dx dy = \frac{n!}{(i-1)!j!(j-i-1)!(n-j)!} \left\{ F(u) \right\}^{i-1} \left\{ F(v) - F(u) \right\}^{j-i-1} \times \left\{ 1 - F(v) \right\}^{n-j} \int_u^{u+s} f(x) dx \int_v^{v+t} f(y) dy,
\]

\( i < j; -\infty < u \leq v < \infty. \)

Again invoking the mean value theorem for integrals, we have

\[
\begin{align*}
\int_u^{u+s} \int_v^{v+t} f_{(i)(j)}(x,y) dx dy &= f_{(i)(j)}(u+s', v+t'), \quad 0 \leq s' \leq s, 0 \leq t' \leq t; \\
\int_u^{u+s} f(x) dx &= f(u+s''), s, \quad 0 \leq s'' \leq s; \\
\int_v^{v+t} f(y) dy &= f(v+t''), t, \quad 0 \leq t'' \leq t.
\end{align*}
\]

In the limit \( s, t \to 0 \), we obtain

\[
f_{(i)(j)}(u,v) = \frac{n!}{(i-1)!j!(j-i-1)!(n-j)!} \times \left\{ F(u) \right\}^{i-1} \left\{ F(v) - F(u) \right\}^{j-i-1} \left\{ 1 - F(v) \right\}^{n-j} f(u)f(v), \quad i < j; -\infty < u \leq v < \infty. \quad (5.79)
\]

Other joint p.d.f's can be derived by similar arguments.

A useful one-one transformation in the study of order statistic r.v.s is

\[
Y_{(i)} = F(X_{(i)}) \quad \text{or} \quad y = F(x) : \frac{dy}{dx} = f(x). \quad (5.80)
\]

We have

\[
\begin{align*}
f_{Y_{(i)}}(y) &= f_{X_{(i)}}(F^{-1}(y)) \left| \frac{dx}{dy} \right| \\
&= \frac{n!}{(i-1)!} y^{i-1} (1-y)^{n-i} \quad \text{[since } F(F^{-1}(y)) = y] \\
&= \frac{1}{B(i,n-i+1)} y^{i-1} (1-y)^{(n-i+1)-1}, \quad 0 \leq y \leq 1,
\end{align*}
\]

So

\[
Y_{(i)} \sim Beta(i,n-i+1). \quad (5.81)
\]

The following functions of order statistic r.v.s are important in practice:

\[
\begin{align*}
(\text{a}) \ \text{median} &= \begin{cases} X_{(r+1)}, & \text{when } n = 2r+1 \text{ (odd)} \\ \frac{1}{2}(X_{(r)} + X_{(r+1)}), & \text{when } n = 2r \text{ (even)}; \end{cases} \\
(\text{b}) \ \text{range } R &= X_{(n)} - X_{(1)}. \quad (5.82)
\end{align*}
\]