Section 2.7. Transformations for Several Random Variables

Note. We now extend the ideas of Section 2.2, "Transformations: Bivariate Random Variables," from two random variables to several random variables. As a consequence, we have no new theorems or definitions in this section, we only have some new computational techniques. This requires us to consider the change of variables result for for integrals of several variables, which we now state.

Theorem 2.7.A. Consider an integral of the form

$$\int \cdots \iint_A f(x_1, x_2, \ldots, x_n)\, dx_1\, dx_2 \cdots dx_n$$

where $A \subset \mathcal{S} \subset \mathbb{R}^n$ is a "nice" (i.e., measurable) set and $f : \mathbb{R}^n \to \mathbb{R}^n$ is an integrable function on $\mathcal{S}$. Let

$$y_1 = u_1(x_1, x_2, \ldots, x_m),\ y_2 = u_2(x_1, x_2, \ldots, x_n), \ldots,\ y_n = u_n(x_1, x_2, \ldots, x_n)$$

define a one to one transformation that maps $\mathcal{S} \subset \mathbb{R}^n$ onto $\mathcal{T} \subset \mathbb{R}^n$. Let the first partial derivatives of the inverse functions be continuous and let the $n \times n$ determinant, called the *Jacobian* of the transformation,

$$J = \begin{vmatrix} \partial x_1/\partial y_1 & \partial x_1/\partial y_2 & \cdots & \partial x_1/\partial y_n \\ \partial x_2/\partial y_1 & \partial x_2/\partial y_2 & \cdots & \partial x_2/\partial y_n \\ \vdots & \vdots & \ddots & \vdots \\ \partial x_n/\partial y_1 & \partial x_n/\partial y_2 & \cdots & \partial x_n/\partial y_n \end{vmatrix}$$

not be identically zero in $\mathcal{T}$. Then

$$\int \cdots \iint_A f(x_1, x_2, \ldots, x_n)\, dx_1\, dx_2 \cdots dx_n$$

$$= \int \cdots \iint_B f(w_1(y_1, y_2, \ldots, y_n), w_2(y_1, y_2, \ldots, y_n),$$
$$\ldots, w_n(y_1, y_2, \ldots, y_n))|J|\, dy_1\, dy_2 \cdots dy_n.$$

Note. For a more general statement of the above result, see Peter D. Lax's "Change of Variables in Multiple Integrals," *The American Mathematical Monthly*, Vol. 106 #6 (June-July) 1999, 497–501.

Note 2.7.A. With the notation of Theorem 2.7.A we have, as in the case of two random variables, that if the probability density function of $X_1, X_2, \ldots, X_n$ is $f(x_1, x_2, \ldots, x_n)$ then the probability density function of $Y_1, Y_2, \ldots, Y_n$ is

$$g(y_1, y_2, \ldots, y_n) = f(w_1(y_1, y_2, \ldots, y_n), w_2(y_1, y_2, \ldots, y_n), \ldots, w_n(y_1, y_2, \ldots, y_n))|J|$$

for $(y_1, y_2, \ldots, y_n) \in \mathcal{T}$ and g is 0 elsewhere.

Example 2.7.1. Let random variables X_1, X_2, X_3 have the joint probability density function

$$f(x_1, x_2, x_3) = \begin{cases} 48x_1x_2x_3 & \text{for } 0 < x_1 < x_2 < x_3 < 1 \\ 0 & \text{elsewhere.} \end{cases}$$

If $Y_1 = X_1/X_2$, $Y_2 = X_2/X_3$, and $Y_2 = X_3$ (so that $u_1(x_1, x_2, x_3) = x_1/x_2$, $u_2(x_1, x_2, x_3) = x_2/x_3$, and $u_3(x_1, x_2, x_3) = x_3$) then the inverse transformation is given by $x_1 = w_1(y_1, y_2, y_3) = y_1y_2y_3$, $x_2 = w_2(y_1, y_2, y_3) = y_2y_3$, and $x_3 =$

$u_2(y_1, y_2, y_3) = y_3$ (solve in order for X_3, X_2, and then X_1). So the Jacobian is

$$J = \begin{vmatrix} \partial x_1/\partial y_1 & \partial x_1/\partial y_2 & \partial x_1/\partial y_3 \\ \partial x_2/\partial y_1 & \partial x_2/\partial y_2 & \partial x_2/\partial y_3 \\ \partial x_3/\partial y_1 & \partial x_n/\partial y_2 & \partial x_3/\partial y_3 \end{vmatrix} = \begin{vmatrix} y_2y_3 & y_1y_3 & y_1y_2 \\ 0 & y_3 & y_2 \\ 0 & 0 & 1 \end{vmatrix} = y_2y_3^2.$$

The four inequalities defining the support of f, $0 < x_1 < x_2 < x_3 < 1$, correspond to the inequalities $0 < y_1y_2y_3 < y_2y_3 < y_3 < 1$ (notice that this implies that each y_i is positive) which yield $0 < y_3 < 1$, $0 < y_2 < 1$, and $0 < y_1 < 1$ so that the support $\mathcal{T}$ of the joint probability density function g of Y_1, Y_2, Y_3 is $\mathcal{T} = \{(y_1, y_2, y_3) \mid 0 < y_i < 1 \text{ for } i = 1, 2, 3\}$. This implies

$$g(y_1, y_2, y_3) = 48(y_1y_2y_3)(y_2y_3)(y_3)|y_2y_3^2| = 48y_1y_2^3y_3^5$$

where $0 < y_i < 1$ for $i = 1, 2, 3$ and 0 elsewhere. By Note 2.1.C, the marginal probability density functions are

$$\begin{aligned} g_1(y_1) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(y_1, y_2, y_3)\, dy_2\, dy_3 = \int_0^1\int_0^1 48y_1y_2^3y_3^5\, dy_2\, dy_3 \\ &= \int_0^1 48y_1\left(\frac{1}{4}y_2^4\right)y_3^5\Big|_{y_2=0}^{y_2=1} dy_3 = \int_0^1 12y_1y_3^5 \\ &= 12y_1\left(\frac{1}{6}y_3^6\right)\Big|_{y_3=0}^{y_3=1} = 2y_1 \text{ for } 0 < y_1 < 1, \end{aligned}$$

$$\begin{aligned} g_2(y_2) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(y_1, y_2, y_3)\, dy_1\, dy_3 = \int_0^1\int_0^1 48y_1y_2^3y_3^5\, dy_1\, dy_3 \\ &= \int_0^1 48\left(\frac{1}{2}y_1^2\right)y_2^3y_3^5\Big|_{y_1=0}^{y_1=1} dy_3 = \int_0^1 24y_2^3y_3^5\, dy_3 \\ &= 24y_2^3\left(\frac{1}{6}y_3^6\right)\Big|_{y_3=0}^{y_3=1} = 4y_2^3 \text{ for } 0 < y_2 < 1, \text{ and} \end{aligned}$$

$$g_3(y_3) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(y_1, y_2, y_3)\, dy_1\, dy_3 = \int_0^1\int_0^1 48y_1y_2^3y_3^5\, dy_1\, dy_2$$

$$= \int_0^1 48\left(\frac{1}{2}y_1^2\right)y_2^3y_3^5\Big|_{y_1=0}^{y_1=1}\,dy_2 = \int_0^1 24y_2^3y_3^5\,dy_2$$
$$= 24\left(\frac{1}{4}y_2^4\right)y_3^5\Big|_{y_2=0}^{y_2=1} = 6y_3^5 \text{ for } 0 < y_3 < 1.$$

Notice that $g(y_1, y_2, y_3) = g_1(y_1)g_2(y_2)g_3(y_3)$ so that Y_1, Y_2, Y_3 are mutually independent (by definition; see Section 2.6).

Example 2.7.2. Let random variables X_1, X_2, X_3 be independent and identical in distribution ("idd," see Section 2.6) with common probability density function

$$f(x) = \begin{cases} e^{-x} & \text{for } 0 < x < \infty \\ 0 & \text{elsewhere.} \end{cases}$$

Since X_1, X_2, X_3 are mutually independent then the joint probability density function of X_1, X_2, X_3 is

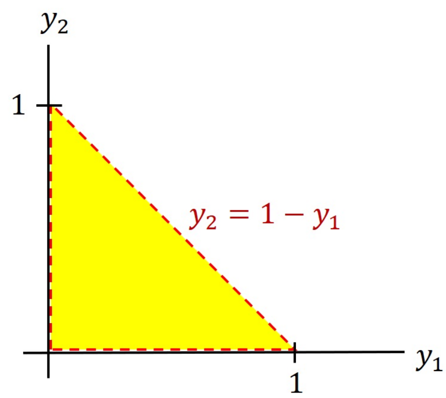
$$f_{X_1,X_2,X_3}(x_1,x_2,x_3) = f_{X_1}(x_1)f_{X_2}(x_2)f_{X_3}(x_3) = \begin{cases} e^{-x_1}e^{-x_2}e^{-x_3} & \text{for } 0 < x_i < \infty, i = 1,2,3 \\ 0 & \text{elsewhere.} \end{cases}$$

Define random variables Y_1, Y_2, Y_3 as $Y_1 = \dfrac{X_1}{X_1 + X_2 + X_3}$, $Y_2 = \dfrac{X_2}{X_1 + X_2 + X_3}$, $Y_3 = X_1 + X_2 + X_3$. So the inverse transformation is (as is easily verified) $x_1 = y_1y_3$, $x_2 = y_2y_3$, and $x_3 = y_3 - y_1y_3 - y_2y_3$. So the Jacobian is

$$J = \begin{vmatrix} \partial x_1/\partial y_1 & \partial x_1/\partial y_2 & \partial x_1/\partial y_3 \\ \partial x_2/\partial y_1 & \partial x_2/\partial y_2 & \partial x_2/\partial y_3 \\ \partial x_3/\partial y_1 & \partial x_3/\partial y_2 & \partial x_3/\partial y_3 \end{vmatrix} = \begin{vmatrix} y_3 & 0 & y_1 \\ 0 & y_3 & y_2 \\ -y_3 & -y_3 & 1 - y_1 - y_2 \end{vmatrix}$$

$$= (y_3)((y_3)(1-y_1-y_2) - (y_2)(-y_3)) + (y_1)(0 - (y_3)(-y_3)) = y_3^2(1-y_1) + y_1y_3^2 = y_3^2.$$

The support $\mathcal{S}$ of X_1, X_2, X_3, $\mathcal{S} = \{(x_1, x_2, x_3) \mid 0 < x_i < \infty, i = 1, 2, 3\}$, maps onto the set $\mathcal{T}$ where $0 < y_1y_3 < \infty$, $0 < y_2y_3 < \infty$, $0 < y_3(1 - y_1 - y_2) < \infty$, so the support of the joint probability g of Y_1, Y_2, Y_3 is $\mathcal{T} = \{(y_1, y_2, y_3) \mid 0 < y_1, 0 < y_2, 0 < y_3, 0 < 1 - y_1 - y_2\}$. In the y_1y_2-plane we have the region:



So $\mathcal{T}$ is then a right triangular cylinder with this as its base and with $0 < y_3$. The joint probability density function is then

$$g(y_1, y_2, y_3) = e^{-(y_1y_3)}e^{-(y_2y_3)}e^{-(y_3-y_1y_3-y_2y_3)}|y_3^2| = y_3^2e^{-y_3}$$

where $(y_1, y_2, y_3) \in \mathcal{T}$, and 0 elsewhere. By Note 2.1.C, the marginal probability density functions are

$$\begin{aligned}
g_1(y_1) &= \int_0^{1-y_1}\int_0^{\infty} y_3^2e^{-y_3}\,dy_3\,dy_2 \\
&= \int_0^{1-y_1}\left(-y_3^2e^{-y_3} + 2\int y_3e^{-y_3}\,dy_3\right)\Bigg|_0^{\infty} dy_2 \text{ by integration by parts} \\
&= \int_0^{1-y_1}(-y_3^2e^{-y_3} + 2(-y_3e^{-y_3} - e^{-y_3})\Big|_0^{\infty} dy_2 \text{ by integration by parts} \\
&= = \int_0^{1-y_1}(2)\,dy_2 = 2y_2\Big|_{y_2=0}^{y_2=1-y_1} = 2(1-y_1) \text{ where } 0 < y_1 < 1, \\
g_2(y_2) &= \int_0^{1-y_2}\int_0^{\infty} y_3^2e^{-y_3}\,dy_3\,dy_1 = 2(1-y_2) \text{ where } 0 < y_2 < 1 \text{ (as with } g_1(y_1)) \\
g_3(y_3) &= \int_0^1\int_0^{1-y_1} y_3^2e^{-y_3}\,dy_2\,dy_1 = \int_0^1 y_3^2e^{-y_3}(1-y_1)\,dy_1 \\
&= y_3^2e^{-y_3}(y_1 - y_1^2/2)\Big|_{y_1=0}^{y_1=1} = \frac{1}{2}y_3^2e^{-y_3} \text{ where } 0 < y_3.
\end{aligned}$$

Notice that $g(y_1y_2, y_3) = y_3^2e^{-y_3} \neq g(y_1)g_2(y_2)g_3(y_3) = 2(1-y_1)(1-y_2)y_3^2e^{-y_3}$ so that Y_1, Y_2, Y_3 are not mutually independent (by definition; see Section 2.6). Hence, the joint probability density function of Y_1 and Y_3 (obtained from g by integrating over y_2) is

$$g_{13}(y_1, y_3) = \int_0^{1-y_1} y_3^2e^{-y_3}\,dy_2 = (1-y_1)y_3^2e^{-y_3} \text{ where } 0 < y_1 < 1 \text{ and } 0 < y_3.$$

So $g_{13}(y_1, y_3) = g_1(y_1)g_3(y_3) = (2(1-y_1))(y_3^2e^{-y_3}/2)$ so that Y_1 and Y_3 are independent. Similarly, $g_{12}(y_1, y_2) = \int_0^\infty y_3^2e^{-y_3}\,dy_3 = 2$ (as shown above) where $0 < y_1$, $0 < y_2$, and $y_1 + y_2 < 1$. So $g_{12}(y_1, y_2) \neq g_1(y_1)g_2(y_2) = (2(1-y_1))(2(1-y_2))$ and Y_1 and Y_2 are dependent. □

Note. We now describe a transformation which is not one-to-one. Let X be a random variable with the Cauchy probability density function $f(x) = \dfrac{1}{\pi(1+x^2)}$ where $-\infty < x < \infty$, and let $Y = X^2$. We want the probability density function of Y. Now the support of X is $\mathcal{S} = \mathbb{R}$ and the transformation $y = x^2$ maps $\mathcal{S}$ onto $\mathcal{T} = \{y \mid 0 \leq y < \infty\}$, but the transformation is two-to-one except at $x = 0$. Since X is a continuous random variable, no probabilities change if we redefine $(f(0)$ as 0 (instead of $1/\pi$). This then modifies the support of f so that it becomes $\mathcal{S} = \mathbb{R} \setminus \{0\}$ and modifies $\mathcal{T}$ to $\mathcal{T} = \{y \mid 0 < y < \infty\}$. Next we partition $\mathcal{S}$ as $A_1 \mathbin{\dot\cup} A_2$ where $A_1 = \{x \mid -\infty < x < 0\}$ and $A_2 = \{x \mid 0 < x < \infty\}$. Then the transformation $y = x^2$ maps A_1 onto $\mathcal{T}$ in a one-to-one way and has inverse $x = -\sqrt{y}$, and $y = x^2$ maps A_2 onto $\mathcal{T}$ in a one-to-one way and has inverse $x = \sqrt{y}$. Consider the probability $P(Y \in B)$ where $B \subset \mathcal{T}$ (and B is measurable). Let $A_3 = \{x \mid x = -\sqrt{y}, y \in B\} \subset A_1$ and $A_4 = \{x \mid x = \sqrt{y}, y \in B\} \subset A_2$. Then

$Y \in B$ if and only if either $X \in A_3$ or $X \in A_4$. So

$$P(Y \in B) = P(X \in A_3) + P(X \in A_4) = \int_{A_3} f(x)\,dx + \int_{A_4} f(x)\,dx.$$

In the first integral we have $x = -\sqrt{y}$ and the Jacobian is $J_1 = \frac{d}{dx}0 - \sqrt{y}] = \frac{-1}{2\sqrt{y}}$; in the second integral we have $x = \sqrt{y}$ and the Jacobian is $J_2 = \frac{d}{dx}[\sqrt{y}] = \frac{1}{2\sqrt{y}}$ (of course the Jacobian is just the coefficient of "dy" in this change of variables). Also, $x = -\sqrt{y}$ maps A_3 onto B and $x = \sqrt{y}$ maps A_4 onto B. So we have

$$\begin{aligned} P(Y \in B) &= \int_B f(-\sqrt{y}) \left| \frac{-1}{2\sqrt{y}} \right| dy + \int_B f(\sqrt{y}) \left| \frac{1}{2\sqrt{y}} \right| dy \\ &= \int_B (f(-\sqrt{y}) + f(\sqrt{y})) \frac{1}{2\sqrt{y}}\,dy. \end{aligned}$$

Hence the probability density function of Y is $g(y) = \frac{1}{2\sqrt{y}} f(-\sqrt{y}) + f(\sqrt{y}))$ for $y \in \mathcal{T}$. Since f is the Cauchy probability density function (well, except at the single point $x = 0$) we have

$$g(y) = \frac{1}{2\sqrt{y}} \left(\frac{1}{\pi(1 + (-\sqrt{y})^2)} + \frac{1}{\pi(1 + (\sqrt{y})^2)} \right) = \frac{1}{\pi\sqrt{y}(1+y)} \text{ for } 0 < y < \infty.$$

Note. Inspired by the previous example, we now consider continuous random variables $X_1, X_2, \ldots, X_n$ with joint probability density function $f(x_1, x_2, \ldots, x_n)$. Let $\mathcal{S}$ be the support of f and consider the transformation $y_1 = u_1(x_1, x_2, \ldots, x_n), y_2 = u_2(x_1, x_2, \ldots, x_n), \ldots, y_n = u_n(x_1, x_2, \ldots, x_n)$ which maps $\mathcal{S}$ onto $\mathcal{T}$ where $\mathcal{T}$ is in the $y_1 y_2 \cdots y_n$-space. But the transformation may not be one-to-one. *Suppose* that $\mathcal{S}$ can be written as the union of a finite number, say k, of mutually disjoint sets $A_1, A_2, \ldots, A_k$, so that $\mathcal{S} = \cup_{i=1}^k A_i$, and the transformation is one-to-one on each

A_i and maps each A_i onto $\mathcal{T}$. Thus to each point in $\mathcal{T}$ there corresponds exactly one point in each of $A_1, A_2, \ldots, A_k$ (so the transformation is a k-to-one mapping). Since the transformation is one-to-one on each A_i and onto $\mathcal{T}$ then there is an inverse transformation mapping $\mathcal{T}$ onto A_i. Say the inverse transformation is

$$x_1 = w_{1i}(y_1, y_2, \ldots, y_n), x_2 = w_{2i}(y_1, y_2, \ldots, y_n), \ldots, x_n = w_{ni}(y_1, y_2, \ldots, y_n)$$

for $i = 1, 2, \ldots, k$. *Suppose* the first partial derivatives are continuous and define for $i - 1, 2, \ldots, k$

$$J_i = \begin{vmatrix} \partial w_{1i}/\partial y_1 & \partial w_{1i}/\partial y_2 & \cdots & \partial w_{1i}/\partial y_n \\ \partial w_{2i}/\partial y_1 & \partial w_{2i}/\partial y_2 & \cdots & \partial w_{2i}/\partial y_n \\ \vdots & \vdots & \ddots & \vdots \\ \partial w_{ni}/\partial y_1 & \partial w_{ni}/\partial y_2 & \cdots & \partial w_{ni}/\partial y_n \end{vmatrix}$$

and *suppose* each J_i is not identically equal to 0 in $\mathcal{T}$. As in the previous example,

$$\begin{aligned} P(X \in A) &= \sum_{i=1}^k P(X \in A_i) = \sum_{i=1}^k \left(\int_{A_i} f(\vec{x})\, d\vec{x} \right) \\ &= \sum_{i=1}^k \left(\int_B f(x_1, x_2, \ldots, x_n)\, dx_1\, dx_2 \cdots dx_n \right) \\ &= \sum_{i=1}^k \int_B f(w_{1i}(y_1, y_2, \ldots, y_n), w_{2i}(y_1, y_2, \ldots, y_n), \\ &\quad \ldots, w_{ni}(y_1, y_2, \ldots, y_n))|J_i|\, dy_1\, dy_2 \cdots dy_n \\ &= \sum_{i=1}^k \left(\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(w_{1i}(y_1, y_2, \ldots, y_n), w_{2i}(y_1, y_2, \ldots, y_n), \right. \\ &\quad \left. \cdots, w_{ni}(y_1, y_2, \ldots, y_n))|J_i|\, dy_1\, dy_2 \cdots dy_n \right) \\ &\quad \text{since the support of } f \text{ is } \mathcal{S} \\ &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \infty_{-\infty}^{\infty} \left(\sum_{i=1}^k f(w_{1i}(y_1, y_2, \ldots, y_n), w_{2i}(y_1, y_2, \ldots, y_n), \right. \end{aligned}$$

$$\left.\cdots, w_{ni}(y_1, y_2, \ldots, y_n))|J_i|\, dy_1\, dy_2 \cdots dy_n\right)$$
$$= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(y_1, y_2, \ldots, y_n)\, dy_1\, dy_2 \cdots dy_n.$$

So the joint probability density function of $Y_1 = u_1(X_1, X_2, \ldots, X_n)$, $Y_2 = u_2(X_1, X_2, \ldots, X_n)$, $\ldots$, $Y_n = u_n(X_1, X_2, \ldots, X_n)$ is

$$g(y_1, y_2, \ldots, y_n) = \sum_{i=1}^{k} f(w_{1i}(y_1, y_2, \ldots, y_n), w_{2i}(y_1, y_2, \ldots, y_n), \ldots, w_{ni}(y_1, y_2, \ldots, y_n))|J_i|$$

for $(y_1, y_2, \ldots, y_n) \in \mathcal{T}$ and is 0 elsewhere.

Example 2.7.3. Let X_1 and X_2 be continuous random variables with joint probability density function over the unit circle by

$$f(x_1, x_2) - \begin{cases} 1/\pi & \text{for } 0 < x_1^2 + x_2^2 < 1 \\ 0 & \text{elsewhere.} \end{cases}$$

Define random variables $Y_1 = X_1^2 + X_2^2$ and $Y_2 = \dfrac{X_1^2}{X_1^2 + X_2^2}$. Then $y_1 y_2 = x_1^2$ and $x_2^2 = y_1 - x_1^2 = y_1 - y_1 y_2 = y_1(1 - y_2)$. Notice $y_2 = \dfrac{x_1^2}{x_1^2 + x_2^2}$ satisfies $0 < y_1 < 1$ so that the support $\mathcal{S} = \{(x_1, x_2) \mid 0 < x_1^2 + x_2^2 < 1\}$ is mapped onto $\mathcal{T} = \{(y_1, y_2) \mid 0 < y_i < 1, i = 1, 2\}$. But each $(y_1, y_2) \in \mathcal{T}$ is the image of four points in $\mathcal{S}$, namely

$$\begin{aligned}
(x_1, x_2) &= (\sqrt{y_1 y_2}, \sqrt{y_1(1 - y_2)}), \\
(x_1, x_2) &= (\sqrt{y_1 y_2}, -\sqrt{y_1(1 - y_2)}), \\
(x_1, x_2) &= (-\sqrt{y_1 y_2}, \sqrt{y_1(1 - y_2)}), \text{ and} \\
(x_1, x_2) &= (-\sqrt{y_1 y_2}, -\sqrt{y_1(1 - y_2)}).
\end{aligned}$$

We find that $|J_i| = \dfrac{1}{4\sqrt{y_2(1 - y_2)}}$ for $i = 1, 2, 3, 4$; for example, based on the one-to-one and onto mapping $(x_1, x_2) = (\sqrt{y_1 y_2}, -\sqrt{y_1(1-y_2)})$ we have

$$J_2 = \begin{vmatrix} \frac{1}{2}\sqrt{\frac{y_2}{y_1}} & \frac{1}{2}\sqrt{\frac{y_1}{y_2}} \\ \frac{-1}{2}\sqrt{\frac{1-y_2}{y_1}} & \frac{1}{2}\sqrt{\frac{y_1}{1-y_2}} \end{vmatrix} = \frac{1}{4}\frac{\sqrt{y_2}1-y_2}{+}\frac{1}{4}\sqrt{\frac{1-y_2}{y_2}}$$

$$= \frac{1}{4}\sqrt{\frac{y_2^2}{y_2(1-y_2)}} + \frac{1}{4}\sqrt{\frac{(1-y_2)^2}{y_2(1-y_2)}} = \frac{1}{4\sqrt{y_2(1-y_2)}}.$$

So the joint probability density function of Y_1 and Y_2 is

$$g(y_1, y_2) = f(\sqrt{y_1 y_2}, \sqrt{y_1(1-y_2)})|J_1| + f(\sqrt{y_1 y_2}, -\sqrt{y_1(1-y_2)})|J_2|$$

$$+f(-\sqrt{y_1 y_2}, \sqrt{y_1(1-y_2)})|J_3| + f(-\sqrt{y_1 y_2}, -\sqrt{y_1(1-y_2)})|J_4|$$

$$= 4\left(\frac{1}{\pi}\right)\frac{1}{4\sqrt{y_2(1-y_2)}} = \frac{1}{\pi\sqrt{y_2(1-y_2)}}$$

for $(y_1, y_2) \in \mathcal{T}$ and 0 elsewhere. Notice that Theorem 2.4.1 implies that Y_1 and Y_2 are independent.

Note. We can extend the moment generating function concept from the bivariate case (see Definition 2.1.2) to the multivariate case. With $Y = g(X_1, X_2, \ldots, X_n)$ an d$f(x_1, x_2, \ldots, x_n)$ the joint probability density function of $X_1, X_2, \ldots, X_n$, the moment generating function of Y is

$$E(e^{tY}) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} r^{tg(x_1,x_2,\ldots,x_n)} f(x_1, x_2, \ldots, x_n)\, dx_1\, dx_2 \cdots dx_n$$

where the random variables are continuous. Hogg and Craig claim (see page 149): "This procedure is particularly useful in cases in which we are dealing with linear functions of independent random variables." This idea is illustrated in Example 2.7.4 and the following example (the linear combination allows us to expand the exponential).

Example 2.7.5. Let X_1, X_2, X_3, X_4 be independent continuous random variables with common probability density function

$$f(x) = \begin{cases} e^{-x} & \text{for } x > 0 \\ 0 & \text{elsewhere.} \end{cases}$$

Define $Y = X_1 + X_2 + X_3 + X_4$. The moment generating function of Y is

$$E(e^{tY}) = E(e^{t(X_1+X_2+X_3+X_4)}) = E(e^{yX_1}e^{tX_2}e^{tX_3}e^{yX_4}) = E(e^{tX_1})E(e^{tX_2})E(e^{tX_3})E(e^{tX_4})$$

by Theorem 2.4.4, since X_1, X_2, X_3, X_4 are independent. In Example 1.9.A we saw that $E(e^{tX}) = \dfrac{1}{1-t}$, so that $E(e^{tX_i}) = \frac{1}{1-t}$ for $i = 1, 2, 3, 4$ and hence $E(e^{tY}) = \frac{1}{(1-t)^4}$. It will be shown in Section 3.3, "The Γ, χ^2, and β Distributions," that this is the moment generating function of a distribution with probability density function

$$f_Y(y) = \begin{cases} \frac{1}{3!}y^3e^{-y} & \text{where } 0 < y < \infty \\ 0 & \text{elsewhere.} \end{cases}$$

So this is the probability density function of Y (this is the Γ-distribution $\Gamma(3, 0)$).

Revised: 10/10/2022