Section 2.2. Transformations: Bivariate Random Variables

Note. We now consider transformations of random vectors, say $Y = g(X_1, X_2)$. We desire to find the cumulative distribution function of Y. We give several examples, but state no new theorems.

Note. In the discrete case, let $p_{X_1,X_2}(x_x, x_2)$ be the joint probability mass function of discrete random vector (X_1, X_2) with support $\mathcal{S}$. Let $y_1 = u_1(x_1, x_2)$ and $y_2 = u_2(x_1, x_2)$ define a one-to-one transformation that map $\mathcal{S}$ onto some set $\mathcal{T} \subset \mathbb{R}^2$. Since $U : \mathcal{S} \to \mathcal{T}$, where $U(x_1, x_2) = (y_1, y_2) = (u_1(x_1, x_2), u_2(x_1, x_2))$, is one-to-one then there is an inverse $U^{-1} : \mathcal{T} \to \mathcal{S}$ where $U^{-1}(y_1, y_2) = (x_1, x_2) = (w_1(y_1, y_2), w_2(y_1, y_2))$. The joint probability mass function of (Y_1, Y_2) is

$$p_{Y_1,Y_2}(y_1, y_2) = \begin{cases} p_{X_1,X_2}(w_1(y_1, y_2), w_2(y_1, y_2)) & \text{for } (y_1, y_2) \in \mathcal{T} \\ 0 & \text{elsewhere.} \end{cases}$$

Example 2.2.1. In a large city, there are two strains of flu, strain A and strain B. For a given week, let X_1 and X_2 be the respective number of reported cases of strains A and B with the joint probability mass function

$$p_{X_1,X_2}(x_1, x_2) = \frac{\mu_1^{x_1}\mu_2^{x_2}e^{-\mu_1}e^{-\mu_2}}{x_1!x_2!} \text{ for } x_1 = 0, 1, 2, \ldots, x_2 = 0, 1, 2, , \ldots$$

and 0 elsewhere (here, μ_1 and μ_2 are some positive real numbers). Now

$$\begin{aligned} E[X_1] &= \sum_{x_1=0}^{\infty}\sum_{x_2=0}^{\infty} x_1 p_{X_1,X_2}(x_1, x_2) = \sum_{x_1=0}^{\infty}\sum_{x_2=0}^{\infty} x_1 \frac{\mu_1^{x_2}\mu_2^{x_2}e^{-\mu_1}e^{-\mu_2}}{x_1!x_2!} \\ &= e^{-\mu_1}e^{-\mu_2}\left(\sum_{x_1=0}^{\infty} x_1 \frac{\mu_2^{x_1}}{x_1!}\right)\left(\sum_{x_2=0}^{\infty} \frac{\mu_2^{x_2}}{x_2!}\right) \end{aligned}$$

$$
\begin{aligned}
&= e^{-\mu_1} e^{-\mu_2} \left(\sum_{x_1=1}^{\infty} x_1 \frac{\mu_1^{x_1}}{x_1!} \right) e_2^{\mu} \\
&= e^{-\mu_1} \left(\sum_{x_1=0}^{\infty} x_1 \frac{\mu_1^{x_1+1}}{(x_1+1)!} \right) = e^{-\mu_1} \left(\sum_{x_1=0}^{\infty} \mu_1 \frac{\mu_1^{x_1}}{x_1!} \right) \\
&= e^{-\mu_1} \mu_1 e^{\mu_1} = \mu_1.
\end{aligned}
$$

Similarly $E[X_2] = \mu_2$. We are interested in the random variable $Y_1 = X_1 + X_2$ (the total number of cases of strain A and strain B flue combined in a week). We have by Theorem 2.1.1 that $E[Y_1] = E[X_1 + X_2] = E[X_1] + E[X_2] = \mu_1 + \mu_2$. To illustrate the above note, we need a second random variable Y_2. We take $Y_2 = X_2$. Then we have $y_1 = u_1(x_1, x_2) = x_1 + x_2$ and $y_2 = u_2(x_1, x_2) = x_2$ and this is a one-to-one mapping of $\mathcal{S}$ to $\mathcal{T} = \{(y_1, y_2) \mid y_1 = 0, 1, 2, \ldots, y_2 = 0, 1, 2 \ldots\}$ with inverse $x_1 = w_1(y_1, y_2) = y_1 - y_2$ and $x_2 = w_2(y_1, y_2) = y_1$. So the joint probability mass function of Y_1 and Y_2 is

$$
p_{Y_1,Y_2}(y_1, y_2) = p_{X_1,X_2}(w_1(y_1, y_2), w_2(y_1, y_2)) = \frac{\mu_1^{y_1 - y_2} \mu_2^{y_2} e^{-\mu_1} e^{\mu_2}}{(y_1 - y_2)! y_2!}
$$

for $y_1 = 0, 1, 2, \ldots$, $y_2 = 0, 1, 2, \ldots$ and 0 elsewhere. The marginal probability mass function of Y_1 is then

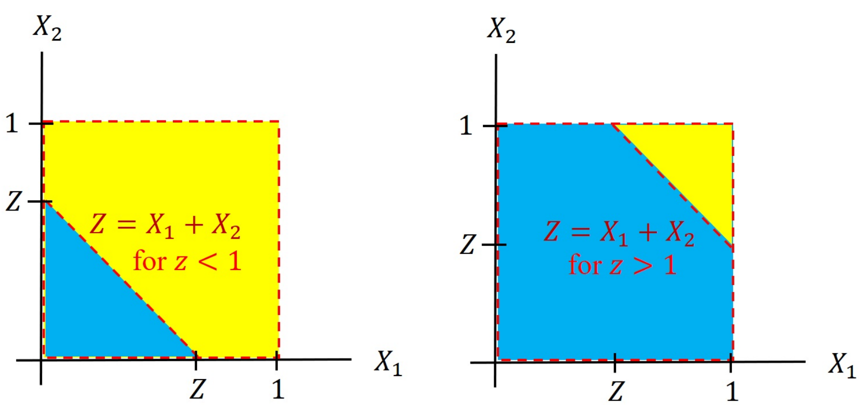
$$
\begin{aligned}
p_{Y_1}(y_1) &= \sum_{y_2=0}^{y_1} p_{Y_1,Y_2}(y_1, y_2) \text{ (notice for given } y_1 \text{ that } y_2 \text{ can be} \\
&\quad\ 0, 1, 2, \ldots, y_1 \text{ since } x_1 = y_1 - y_2) \\
&= \sum_{y_2=0}^{y_1} \frac{\mu_1^{y_1 - y_2} \mu_2^{y_2} e^{-\mu_1} e^{-\mu_2}}{(y_1 - y_2)! y_2!} \\
&= \frac{e^{-\mu_1 - \mu_2}}{y_1!} \sum_{y_2=0}^{y_1} \frac{y_1!}{(y_1 - y_2)! y_2!} \mu_1^{y_1 - y_2} \mu_2^{y_2} \\
&= \frac{e^{-\mu_1 - \mu_2}}{y_1!} (\mu_1 + \mu_2)^{y_1} \text{ by the Binomial Theorem.}
\end{aligned}
$$

Example 2.2.2. Consider an experiment where a point (X_1, X_2) is chosen at random from the unit square $\mathcal{S} = \{(x_1, x_2) \mid 0 < x_1 < 1, 0 < x_2 < 1\}$ according to the uniform probability density function

$$f_{X_1,X_2}(x_1, x_2) = \begin{cases} 1 & \text{if } 0 < x_1 < 1, 0 < x_2 < 1 \\ 0 & \text{elsewhere.} \end{cases}$$

With $Z = X_1 + X_2$ we have the cumulative distribution function of Z of (see the figures below)

$$\begin{aligned} F_Z(z) = P(X_1 + X_2 \le z) &= \begin{cases} 0 & \text{for } z < 0 \\ \int_0^z \int_{x_2=0}^{x_2=z-x_1} dx_2\, dx_1 & \text{for } 1 \le z < 2 \\ 1 - \int_{z-1}^1 \int_{x_2=z-x_1}^{x_2=1} dx_2\, dx_1 & \text{for } 1 \le z < 2 \\ 1 & \text{for } 2 \le z \end{cases} \\ &= \begin{cases} 0 & \text{for } z < 0 \\ z^2/2 & \text{for } 0 \le z < 1 \\ 1 - (2-z)^2/2 & \text{for } 1 \le z < 2 \\ 1 & \text{for } 2 \le z. \end{cases} \end{aligned}$$



Since F'_Z exists for all z, the probability density function for Z is

$$f_Z(z) = \begin{cases} z & \text{for } 0 \le z < 1 \\ 2 - z & \text{for } 1 \le z < 2 \\ 0 & \text{elsewhere.} \end{cases}$$

Note. In using the probability density function f_{X_1,X_2} to calculate probabilities of events such as $\{(x_1, x)2 \mid (x_1, x_2) \in A\}$ for $A \subset \mathcal{S}$, we need to explore substitution in double integrals. Just as there is a "du" term when we change variables in integrals of a single variable function, there is a corresponding term in double integrals. The following discussion is based on my Calculus 3 (MATH 2110) notes on 15.8. Substitution in Multiple Integrals. Suppose G is a region in the uv-plane that is transformed onto a region R in the xy-plane that is transformed onto a region R in the xy-plane by the equations $x = g(u, v)$ and $y = h(u, v)$. For function $f(x, y)$ defined on R, we can interpret f as a function of u and v as $f(g(u, v), h(u, v))$ on G. So we want to relate the double integrals

$$\iint_R f(x, y)\, dxdy \text{ and } \iint_G f(g(u, v), h(u, v))\, du\, dv.$$

But the second integral does not contain the "du" term in it entirety. We see in Calculus 3 that

$$\iint_R f(x, y)\, dx\, dy = \iint_G f(g(u, v), h(u, v))|J(u, v)|\, du\, dv$$

(the term "$|J(u, v)|$" represents the absolute value of $J(u, v)$) where $J(u, v)$ is the *Jacobian* defined as

$$J(u, v) = \frac{\partial(x, y)}{\partial(u, v)} = \begin{vmatrix} \partial x/\partial u & \partial x/\partial v \\ \partial y/\partial u & \partial y/\partial v \end{vmatrix} = \frac{\partial u}{\partial u}\frac{\partial y}{\partial v} - \frac{\partial x}{\partial v}\frac{\partial y}{\partial u}.$$

For this to hold, we need g, h, and f to have continuous partial derivatives and $J(u,v)$ to be 0 only at isolated points. I we relate rectangular coordinates (x,y) to polar coordinates $(u,v)=(r,\theta)$ as $x=r\cos\theta$ and $y=r\sin\theta$ then we have

$$J(u,v)=J(r,\theta)=\begin{vmatrix} \frac{\partial[r\cos\theta]}{\partial r} & \frac{\partial[r\cos\theta]}{\partial\theta} \\ \frac{\partial[r\sin\theta]}{\partial r} & \frac{\partial[r\sin\theta]}{\partial\theta} \end{vmatrix} = (\cos\theta)(r\cos\theta)-(-r\sin\theta)(\sin\theta)=r$$

and so

$$\iint_R f(x,y)\,dx\,dy=\iint_G f(r\cos\theta,r\sin\theta)r\,dr\,d\theta$$

where G is a representation of R in polar coordinates. Notice that this is consistent with the double integral of a function in polar coordinates, as seen in Calculus 3 (see 15.4. Double Integrals in Polar Form).

Note. Let (X_1,X_2) have continuous joint probability density function $f_{X_1,X_2}(x_1,x_2)$ and support $\mathcal{S}$. Let $(Y_1,Y_2)=T(X_1,X_2)=(u_1(X_1,X_2),u_2(X_1,X_2))$ where T is a one-to-one transformation with continuous partial derivatives of components. Since T is one-to-one, the inverse T^{-1} exists, say $T^{-1}(y_1,y_2)=(w_1(y_1,y_2),w_2(y_1,y_2))=(x_1,x_2)$. Then the Jacobian of T is

$$J=\begin{vmatrix} \partial x_1/\partial y_1 & \partial x_1/\partial y_2 \\ \partial x_2/\partial y_1 & \partial x_2/\partial y_2 \end{vmatrix} = \frac{\partial x_1}{\partial y_1}\frac{\partial x_2}{\partial y_2}-\frac{\partial x_1}{\partial y_2}\frac{\partial x_2}{\partial y_1}.$$

Let B be any ("nice") subset of $\mathcal{T}$ and let $A=T^{-1}(B)$. Then, since T is one-to-one, $P((X_1,X_2)\in A)=P(T(X_1,X_2)\in T(A))=P((Y_1,Y_2)\in B)$. So

$$P((X_1,X_2)\in A)=\iint_A f_{X_1,X_2}(x_1,x_2)\,dx_1\,dx_2$$

$$=\iint_{T(A)} f_{X_1,X_2}(x_1,x_2)(T^{-1}(y_1,y_2))|J|\,dy_1\,dy_2$$

$$= \iint_B f_{X_1,X_2}(w_1(y_1,y_2), w_2(y_1,y_2))|J|\,dy_1\,dy_2.$$

Since we can take $B = \mathcal{T}$, then the integrand here must be the probability density function of (Y_1, Y_2):

$$f_{Y_1,Y_2}(y_1,y_2) = \begin{cases} f_{X_1,X_2}(w_1(y_1,y_2), w_2(y_1,y_2))|J| & \text{for } (y_1,y_2) \in \mathcal{T} \\ 0 & \text{elsewhere.} \end{cases}$$

Example 2.2.5. Let continuous random vector (X_1, X_2) have the joint probability density function

$$f_{X_1,X_2}(x_1,x_2) = \begin{cases} 10x_1x_2^2 & \text{for } 0 < x_1 < x_2 < 1 \\ 0 & \text{elsewhere.} \end{cases}$$

Let $Y_1 = X_1/X_2$ and $Y_2 = X_2$. So $y_1 = u_1(x_1,x_2) = x_1/x_2$ and $y_2 = u_2(x_1,x_2) = x_2$. Notice that the transformation mapping $(x_1,x_2) \mapsto (y_1,y_2)$ is one-to-one and the inverse transformation has components $x_1 = w_1(y_1,y_2) = y_1y_2$ and $x + 2 = w_2(y_1,y_2) = y_2$. So the Jacobian is

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

The support $\mathcal{S}$ of (x_1,x_2) is $0 < x_1 < x_2 < 1$ and this corresponds to the support $\mathcal{T}$ of (Y_1,Y_2) of $0 < y_1y_2 < y_2 < 1$, or $0 < y_1 < 1$ and $0 < y_2 < 1$. So the joint probability density function of (Y_1, Y_2) is

$$f_{Y_1,Y_2}(y_1,y_2) = f_{X_1,X_2}(w_1(y_1,y_2), w_2(y_1,y_2))|J| \text{ for } (y_1,y_2) \in \mathcal{T}$$

$$= 10(y_1y_2)(y_2)^2|y_2| = 10y_1y_2^4 \text{ for } (y_1,y_2) \in \mathcal{T}.$$

Note. We can also use moment generating functions to find distributions. We illustrate this technique by finding a moment generating function and then recognizing it as the moment generating function of some known distribution—this works because the moment generating function uniquely determines the cumulative distribution function (see Theorem 1.9.2, the note in these class notes following Definition 2.1.2, and the comment on page 96 following Definition 2.1.2).

Example 2.2.7. Let continuous random vector (X_1, X_2) have joint probability density function

$$f_{X_1,X_2}(x_1,x_2) = \begin{cases} \frac{1}{4}\exp\left(-\frac{x_1+x_2}{2}\right) & \text{for } 0 < x_1 < \infty, 0 < x_2 < \infty \\ 0 & \text{elsewhere.} \end{cases}$$

Define $Y = \frac{1}{2}(X_1 - X_2)$. The moment generating function of Y is

$$\begin{aligned} E[e^{tY}] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x_1,x_2) f_{X_1,X_2}(x_1,x_2)\,dx_1\,dx_2 \\ &= \int_0^{\infty}\int_0^{\infty} e^{t(x_1-x_2)/2}\frac{1}{4}e^{-(x_1+x_2)/2}\,dx_1\,dx_2 \\ &= \int_0^{\infty}\frac{1}{2}e^{-x_1(1-t)/2}\,dx_1 \int_0^{\infty}\frac{1}{2}e^{-x_2(1+t)/2}\,dx_2 \\ &= \left(\frac{1}{2}\frac{-2}{1-t}e^{-x_1(1-t)/2}\right)\Bigg|_0^{\infty}\left(\frac{1}{2}\frac{-2}{1+t}e^{-x_2(t-1)/2}\right)\Bigg|_0^{\infty} \\ &= \frac{1}{1-t}\frac{1}{1+t} = \frac{1}{1-t^2} \end{aligned}$$

if $1-t>0$ and $1+t>0$, or if $-1<t<1$. But by Exercise 1.9.20, we have that $1/(1-t^2)$ is the moment generating function of the probability density function $f(x) = \frac{1}{2}e^{-|x|}$, $-\infty < x < \infty$ (called the *Laplace distribution*). So the probability density function of Y is also $f_Y(y) = \frac{1}{2}e^{-|y|}$, $-\infty < y < \infty$.