Continuous joint distributions (continued)

Example 1 (Uniform distribution on the triangle). Consider the random vector \((X, Y)\) whose joint distribution is

\[
f(x, y) = \begin{cases} 
2 & \text{if } 0 \leq x < y \leq 1, \\
0 & \text{otherwise.}
\end{cases}
\]

This is a density function [on a triangle].

(1) What is the distribution of \(X\)? How about \(Y\)?

We have

\[
f_X(a) = \int_{-\infty}^{\infty} f(a, y) \, dy.
\]

If \(a \notin (0, 1)\), then \(f(a, y) = 0\) regardless of the value of \(y\) [draw a picture!]. Therefore, for \(a \notin (0, 1)\), \(f_X(a) = 0\). If on the other hand \(0 < a < 1\), then [draw a picture!],

\[
f_X(a) = \int_{a}^{1} 2 \, dy = 2(1 - a).
\]

That is,

\[
f_X(a) = \begin{cases} 
2(1 - a) & \text{if } 0 < a < 1, \\
0 & \text{otherwise.}
\end{cases}
\]

Similarly,

\[
f_Y(b) = \int_{0}^{b} 2 \, dx = 2b \quad \text{if } 0 < b < 1,
\]

and \(f_Y(b) = 0\), otherwise.
(2) Are X and Y independent?
No, there exist [many] choices of (x, y) such that f(x, y) = 2 ≠ f_X(x)f_Y(y). In fact, P{X < Y} = \int f = 1 [check!].

(3) Find EX and EY. Also compute the SDs of X and Y.
Let us start with the means:

\[ EX = \int_0^1 x \frac{f_X(x)}{2} dx = 2 \int_0^1 x dx - 2 \int_0^1 x^2 dx = \frac{1}{3}, \]

similarly,

\[ EY = \int_0^1 y \frac{f_Y(y)}{2} dy = \frac{2}{3}. \]

Also:

\[ E(X^2) = \int_0^1 x^2 2(1 - x) dx = \frac{1}{6} \Rightarrow \text{Var}(X) = \frac{1}{6} - \frac{1}{9} = \frac{1}{18}. \]

Similarly,

\[ E(Y^2) = \int_0^1 y^2 2y dy = \frac{1}{2} \Rightarrow \text{Var}(Y) = \frac{1}{2} - \frac{4}{9} = \frac{1}{18}. \]

Consequently, SD(X) = SD(Y) = 1/\sqrt{18}.

(4) Compute E(XY).
After we draw a picture [of the region of integration], we find that

\[ E(XY) = \int_0^1 \int_x^1 2xy dy dx = 2 \int_0^1 y \left( \int_0^y x dx \right) dy = 2 \int_0^1 1/2 y^2 dy = \frac{1}{4}. \]

(5) Define correlation as in the discrete. Then what is the correlation between X and Y?
The correlation is

\[ \rho := \frac{E(XY) - EXEY}{SD(X)SD(Y)} = \frac{1}{4} - \left( \frac{1}{2} \times \frac{2}{3} \right) = \frac{1}{2}. \]

The distribution of a sum
Suppose (X, Y) has joint density f(x, y). Question: What is the distribution of X + Y in terms of the function f?
The distribution of a sum

\[ F_{X+Y}(a) = P\{X + Y \leq a\} = \int_{-\infty}^{\infty} \int_{-\infty}^{-x+a} f(x,y)\,dy\,dx \]

\[ = \int_{-\infty}^{a} \int_{-\infty}^{a} f(x, z-x)\,dz\,dx. \]

Differentiate \([d/da]\) to obtain the density of \(X + Y\), using the fundamental theorem of calculus:

\[ f_{X+Y}(a) = \int_{-\infty}^{\infty} f(x, a-x)\,dx. \]

An important special case: \(X\) and \(Y\) are independent if \(f(x, y) = f_X(x)f_Y(y)\) for all pairs \((x, y)\). If \(X\) and \(Y\) are independent, then

\[ f_{X+Y}(a) = \int_{-\infty}^{\infty} f_X(x)f_Y(a-x)\,dx. \]

This is called the convolution of the functions \(f_X\) and \(f_Y\).

**Example 2.** Suppose \(X\) and \(Y\) are independent exponentially-distributed random variables with common parameter \(\lambda\). What is the distribution of \(X + Y\)?

We know that \(f_X(x) = \lambda e^{-\lambda x}\) for \(x > 0\) and \(f_X(x) = 0\) otherwise. And \(f_Y\) is the same function as \(f_X\). Therefore,

\[ f_{X+Y}(a) = \int_{-\infty}^{\infty} f_X(x)f_Y(a-x)\,dx \]

\[ = \int_{0}^{\infty} \lambda e^{-\lambda x}f_Y(a-x)\,dx = \int_{0}^{a} \lambda e^{-\lambda x}\lambda e^{-\lambda(a-x)}\,dx \]

\[ = \lambda^2 a e^{-\lambda a}, \]

provided that \(a > 0\). And \(f_{X+Y}(a) = 0\) if \(a \leq 0\). In other words, the sum of two independent exponential (\(\lambda\)) random variables has a gamma density with parameters \((2, \lambda)\). We can generalize this (how?) as follows: If \(X_1, \ldots, X_n\) are independent exponential random variables with common parameter \(\lambda > 0\), then \(X_1 + \cdots + X_n\) has a gamma distribution with parameters \(r = n\) and \(\lambda\). A special case, in applications, is when \(\lambda = \frac{1}{2}\). A gamma distribution with parameters \(r = n\) and \(\lambda = \frac{1}{2}\) is also known as a \(\chi^2\) distribution [pronounced “chi squared”] with \(n\) “degrees of freedom.” This distribution arises in many different settings, chief among them in multivariable statistics and the theory of continuous-time stochastic processes.
The distribution of a sum (discrete case)

It is important to understand that the preceding "convolution formula" is a procedure that we ought to understand easily when \(X\) and \(Y\) are discrete instead.

**Example 3** (Two draws at random, Pitman, p. 144). We make two draws at random, without replacement, from a box that contains tickets numbered 1, 2, and 3. Let \(X\) denote the value of the first draw and \(Y\) the value of the second draw. The following tabulates the function \(f(x,y) = P\{X = x, Y = y\}\) for all possible values of \(x\) and \(y\):

<table>
<thead>
<tr>
<th>possible value for (x)</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>possible values for (Y)</td>
<td>3/6</td>
<td>1/6</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>2/6</td>
<td>1/6</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>1/6</td>
<td>1/6</td>
<td>1/6</td>
</tr>
</tbody>
</table>

We want to know the distribution of \(X + Y = \) the total number of dots rolled. Here is a way to compute that: First of all, the possible values of \(X + Y\) are 3, 4, 5. Next, we note that

\[
P\{X + Y = 3\} = P\{X = 2, Y = 1\} + P\{X = 1, Y = 1\} = \frac{1}{3},
\]

\[
P\{X + Y = 4\} = P\{X = 1, Y = 3\} + P\{X = 3, Y = 1\} = \frac{1}{3},
\]

\[
P\{X + Y = 5\} = P\{X = 2, Y = 3\} + P\{X = 3, Y = 2\} = \frac{1}{3}.
\]

The preceding example can be generalized: If \((X, Y)\) are distributed as a discrete random vector, then

\[
P\{X + Y = a\} = \sum_x P\{X = x, Y = a - x\};
\]

When \(X\) and \(Y\) are independent, the preceding simplifies to

\[
P\{X + Y = a\} = \sum_x P\{X = x\} \cdot P\{Y = a - x\};
\]

This is a "discrete convolution" formula.

The distribution of a ratio

The preceding ideas can be used to answer other questions as well. For instance, suppose \((X, Y)\) is jointly distributed with joint density \(f(x,y)\). Then what is the density of \(Y/X\)?
We proceed as we did for sums:

\[
F_{Y/X}(a) = P\left[ \frac{Y}{X} \leq a \right]
\]

\[
= P\left[ \frac{Y}{X} \leq a \quad Y > 0 \right] + P\left[ \frac{Y}{X} \leq a \quad Y < 0 \right]
\]

\[
= P\{Y \leq aX, \quad X > 0\} + P\{Y \geq aX, \quad X < 0\}
\]

\[
= \int_0^\infty \int_{-\infty}^{ax} f(x, y) \, dy \, dx + \int_{-\infty}^{0} \int_{ax}^{\infty} f(x, y) \, dy \, dx
\]

\[
= \int_0^\infty \int_{-\infty}^{a} f(x, ax) \, dx \, dx + \int_{-\infty}^{0} \int_{ax}^{\infty} f(x, x) \, dx \, dx.
\]

Differentiate, using the fundamental theorem of calculus, to arrive at

\[
f_{Y/X}(a) = \int_0^\infty f(x, ax) \, dx - \int_{-\infty}^{0} f(x, ax) \, dx
\]

\[
= \int_{-\infty}^{\infty} f(x, ax) |x| \, dx.
\]

In the important special case that \(X\) and \(Y\) are independent, this yields the following formula:

\[
f_{Y/X}(a) = \int_{-\infty}^{\infty} f_X(x) f_Y(ax) |x| \, dx.
\]

**Example 4.** Suppose \(X\) and \(Y\) are independent exponentially-distributed random variables with respective parameters \(\alpha\) and \(\beta\). Then what is the density of \(Y/X\)? The answer is

\[
f_{Y/X}(a) = \int_0^\infty ae^{-ax} f_Y(ax) \, dx
\]

\[
= \int_0^\infty ae^{-ax} \beta e^{-\beta x} x \, dx \quad \text{[if \(a > 0\); else, \(f_{Y/X}(a) = 0\]}
\]

\[
= a\beta \int_0^\infty xe^{-(\alpha + \beta a)x} \, dx
\]

\[
= \frac{a\beta}{(\alpha + \beta a)^2} \cdot \int_0^\infty ye^{-y} \, dy \quad [y := (\alpha + \beta a)x]
\]

\[
= \frac{a\beta}{(\alpha + \beta a)^2} \cdot \Gamma(2) = \frac{a\beta}{(\alpha + \beta a)^2}.
\]
for \( a > 0 \) and \( f_{Y/X}(a) = 0 \) for \( a \leq 0 \). In the important case that \( \alpha = \beta \), we have
\[
f_{Y/X}(a) = \begin{cases} 
\frac{1}{(1 + a)^2} & \text{if } a > 0, \\
0 & \text{otherwise}.
\end{cases}
\]

Note, in particular, that
\[
E \left( \frac{Y}{X} \right) = \int_0^\infty \frac{a}{(1 + a)^2} da = \infty.
\]

**Example 5.** Suppose \( X \) and \( Y \) are independent standard normal random variables. Then a similar computation shows that
\[
f_{Y/X}(a) = \frac{1}{\pi(1 + a^2)} \quad \text{for all real } a.
\]

[See Example 5, p. 383 of your text] This is called the *standard Cauchy density*. Note that the Cauchy density does not have a well-defined expectation, although
\[
E \left( \left| \frac{Y}{X} \right| \right) = \frac{1}{\pi} \cdot \int_{-\infty}^{\infty} \frac{|a|}{1 + a^2} da = \frac{2}{\pi} \cdot \int_0^\infty \frac{a}{1 + a^2} da = \infty.
\]

**Exercise.** One might wish to know about the distribution of \( V/X \) when \( Y \) and \( X \) are discrete random variables. Check that if \( X \) and \( Y \) are discrete and \( P\{X = 0\} = 0 \), then
\[
P \left( \frac{Y}{X} = a \right) = \sum_{x \neq 0} P\{X = x\} \cdot P\{Y = ax\}.
\]

Note that if we replace the sum by an integral and probabilities with densities we do *not* obtain the correct formula for continuous random variables [\([|x| \text{ is missing}]\).

**Functions of a random vector**

Basic problem: If \( (X, Y) \) has joint density \( f \), then what, if any, is the joint density of \( (U, V) \), where \( U = u(X, Y) \) and \( V = v(X, Y) \)? Or equivalently, \( (U, V) = T(X, Y) \), where
\[
T(x, y) = \begin{pmatrix} u(x, y) \\ v(x, y) \end{pmatrix}.
\]
Example 6. Let \((X, Y)\) be distributed uniformly in the circle of radius \(R > 0\) about the origin in the plane. Thus,
\[
f_{X,Y}(x,y) = \begin{cases} 
\frac{1}{\pi R^2} & \text{if } x^2 + y^2 \leq R^2, \\
0 & \text{otherwise.}
\end{cases}
\]

We wish to write \((X, Y)\), in polar coordinates, as \((R, \Theta)\), where
\[
R = \sqrt{x^2 + y^2} \quad \text{and} \quad \Theta = \arctan(Y/X).
\]

Then, we compute first the joint distribution function \(F_{R,\Theta}\) of \((R, \Theta)\) as follows:
\[
F_{R,\Theta}(a, b) = P\{R \leq a, \Theta \leq b\} \\
= P\{(X, Y) \in A\},
\]
where \(A\) is the “partial cone” \(\{(x, y) : x^2 + y^2 \leq a^2, \arctan(y/x) \leq b\}\). If \(a\) is not between 0 and \(R\), or \(b \notin (-\pi, \pi)\), then \(F_{R,\Theta}(a, b) = 0\). Else,
\[
F_{R,\Theta}(a, b) = \int_A f_{X,Y}(x,y) \, dx \, dy \\
= \int_0^b \int_0^a \frac{1}{\pi R^2} r \, dr \, d\theta,
\]
after the change of variables \(r = \sqrt{x^2 + y^2}\) and \(\theta = \arctan(y/x)\). Therefore, for all \(a \in (0, R)\) and \(b \in (-\pi, \pi)\),
\[
F_{R,\Theta}(a, b) = \begin{cases} 
\frac{a^2 b}{\pi R^2} & \text{if } 0 < a < R \text{ and } -\pi < b < \pi, \\
0 & \text{otherwise.}
\end{cases}
\]

It is easy to see that
\[
f_{R,\Theta}(a, b) = \frac{\partial^2 F_{R,\Theta}}{\partial a \partial b}(a, b).
\]

Therefore,
\[
f_{R,\Theta}(a, b) = \begin{cases} 
\frac{a}{\pi R^2} & \text{if } 0 < a < R \text{ and } -\pi < b < \pi, \\
0 & \text{otherwise.}
\end{cases}
\]

The previous example can be generalized.

Suppose \(T\) is invertible with inverse function
\[
T^{-1}(u, v) = \begin{pmatrix} x(u, v) \\ y(u, v) \end{pmatrix}.
\]

The Jacobian of this transformation is
\[
J(u, v) = \frac{\partial x}{\partial u} \frac{\partial y}{\partial v} - \frac{\partial x}{\partial v} \frac{\partial y}{\partial u}.
\]
Theorem 1. If $T$ is “nice,” then
$$f_{U,V}(u, v) = f_{X,Y}(x(u, v), y(u, v))|J(u, v)|.$$ 

Example 7. In the polar coordinates example ($r = u, \theta = v$),
$$r(x, y) = \sqrt{x^2 + y^2},$$
$$\theta(x, y) = \arctan(y/x) = \theta,$$
$$x(r, \theta) = r \cos \theta,$$
$$y(r, \theta) = r \sin \theta.$$ 
Therefore, for all $r > 0$ and $\theta \in (-\pi, \pi),$
$$f(r, \theta) = (\cos(\theta) \times r \cos(\theta)) - (-r \sin(\theta) \times \sin(\theta))$$
$$= r \cos^2(\theta) + u \sin^2(\theta) = r.$$ 
Hence,
$$f_{R,\Theta}(r, \theta) = \begin{cases} r f_{X,Y}(r \cos \theta, r \sin \theta) & \text{if } r > 0 \text{ and } \pi < \theta < \pi, \\ 0 & \text{otherwise.} \end{cases}$$ 
You should check that this yields Example 6, for instance.

Example 8. Let us compute the joint density of $U = X$ and $V = X + Y.$ Here,
$$u(x, y) = x$$
$$v(x, y) = x + y$$
$$x(u, v) = u$$
$$y(u, v) = v - u.$$ 
Therefore,
$$f(u, v) = (1 \times 1) - (0 \times -1) = 1.$$ 
Consequently,
$$f_{U,V}(u, v) = f_{X,Y}(u, v - u).$$ 
This has an interesting by-product: The density function of $V = X + Y$ is
$$f_V(v) = \int_{-\infty}^{\infty} f_{U,V}(u, v) \, du$$
$$= \int_{-\infty}^{\infty} f_{X,Y}(u, v - u) \, du.$$