CHAPTER 11

Multivariate distributions

11.1. Joint distribution for discrete and continuous random variables

We are often interested in considering several random variables that might be related to each other. For example, we can be interested in several characteristics of a randomly chosen object, such as gene mutations and a certain disease for a person, different kinds of preventive measures for an infection etc. Each of these characteristics can be thought of as a random variable, and we are interested in their dependence. In this chapter we d study joint distributions of several random variables.

Consider collections of random variables $(X_1, X_2, \ldots, X_n)$, which are known as random vectors. We start by looking at two random variables, though the approach can be easily extended to more variables.

Joint PMF for discrete random variables

The *joint probability mass function* of two discrete random variables X and Y is defined as

$$p_{XY}(x, y) = \mathbb{P}(X = x, Y = y).$$

Recall that here the comma means *and*, or the intersection of two events. If X takes values $\{x_i\}_{i=1}^{\infty}$ and Y takes values $\{y_j\}_{j=1}^{\infty}$, then the range of (X, Y) as a map from the probability space $(S, \mathcal{F}, \mathbb{P})$ to the set $\{(x_i, y_j)\}_{i,j=1}^{\infty}$. Note that p_{XY} is indeed a probability mass function as

$$\sum_{i,j=1}^{\infty} p_{XY}(x_i, y_j) = 1.$$

Definition (Joint PDF for continuous random variables)

Two random variables X and Y are *jointly continuous* if there exists a nonnegative function $f_{XY}: \mathbb{R} \times \mathbb{R} \longrightarrow \mathbb{R}$, such that, for any set $A \subseteq \mathbb{R} \times \mathbb{R}$, we have

$$\mathbb{P}((X, Y) \in A) = \iint_A f_{XY}(x, y)\, dx dy.$$

The function $f_{XY}(x, y)$ is called the *joint probability density function* (PDF) of X and Y.

In particular, for $A = \{(x, y) : a \leqslant x \leqslant b, c \leqslant y \leqslant d\}$ we have

$$\mathbb{P}(a \leqslant X \leqslant b, c \leqslant Y \leqslant d) = \int_a^b \int_c^d f_{XY}(x,y) dy\, dx.$$

Another example is $A = \{(x,y) : x < y\}$, then

$$\mathbb{P}(X < Y) = \iint_{\{x<y\}} f_{X,Y}(x,y) dy\, dx.$$

Example 11.1. If the density $f_{X,Y}(x,y) = ce^{-x}e^{-2y}$ for $0 < x < \infty$ and $x < y < \infty$, what is c?

Solution: We use the fact that a density must integrate to 1. So

$$\int_0^\infty \int_x^\infty ce^{-x}e^{-2y} dy\, dx = 1.$$

Recalling multivariable calculus, this double integral is equal to

$$\int_0^\infty ce^{-x}\tfrac{1}{2}e^{-2x} dx = \frac{c}{6},$$

so $c = 6$.

The multivariate distribution function (CDF) of (X,Y) is defined by $F_{X,Y}(x,y) = \mathbb{P}(X \leqslant x, Y \leqslant y)$. In the continuous case, this is

$$F_{X,Y}(x,y) = \int_{-\infty}^x \int_{-\infty}^y f_{X,Y}(x,y) dy\, dx,$$

and so we have

$$f(x,y) = \frac{\partial^2 F}{\partial x \partial y}(x,y).$$

The extension to n random variables is exactly similar.

Marginal PDFs

Suppose $f_{X,Y}(x,y)$ is a joint PDF of X and Y, then the *marginal densities* of X and of Y are given by

$$f_X(x) = \int_{-\infty}^\infty f_{X,Y}(x,y) dy, \qquad f_Y(y) = \int_{-\infty}^\infty f_{X,Y}(x,y) dx.$$

Example 11.2 (Binomial distribution as a joint distribution)**.** If we have a binomial with parameters n and p, this can be thought of as the number of successes in n trials, and

$$\mathbb{P}(X = k) = \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k}.$$

If we let $k_1 = k$, $k_2 = n - k$, $p_1 = p$, and $p_2 = 1 - p$, this can be rewritten as

$$\frac{n!}{k_1!k_2!}p_1^{k_1}p_2^{k_2},$$

as long as $n = k_1 + k_2$. Thus this is the probability of k_1 successes and k_2 failures, where the probabilities of success and failure are p_1 and p_2, resp.

Example 11.3 (Multivariate binomial distribution)**.** A multivariate random vector is $(X_1, \ldots, X_r)$ with

$$\mathbb{P}(X_1 = n_1, \ldots, X_r = n_r) = \frac{n!}{n_1! \cdots n_r!}p_1^{n_1} \cdots p_r^{n_r},$$

where $n_1 + \cdots + n_r = n$ and $p_1 + \cdots p_r = 1$. Thus this generalizes the binomial to more than 2 categories.

11.2. Independent random variables

Now we describe the concept of *independence* for random variables.

Definition (Independent random variables)

Two discrete random variables X and Y are *independent* if $\mathbb{P}(X = x, Y = y) = \mathbb{P}(X = x)\mathbb{P}(Y = y)$ for all x and y.
Two continuous random variables X and Y are *independent* if

$$\mathbb{P}(X \in A, Y \in B) = \mathbb{P}(X \in A)\mathbb{P}(Y \in B)$$

for all pairs of subsets A, B of the real line $\mathbb{R}$.

Recall that the left-hand side should be understood as

$$\mathbb{P}\left(\{\omega : X(\omega) \text{ is in } A \text{ and } Y(\omega) \text{ is in } B\}\right)$$

and similarly for the right-hand side.

In the discrete case, if we have independence, then in terms of PMFs we have

$$\begin{aligned} p_{X,Y}(x,y) &= \mathbb{P}(X = x, Y = y) = \mathbb{P}(X = x)\mathbb{P}(Y = y) \\ &= p_X(x)p_Y(y). \end{aligned}$$

In other words, the joint probability mass faunction $p_{X,Y}$ factors.

In the continuous case for any a, b, c, d we have

$$\begin{aligned} \int_a^b \int_c^d f_{X,Y}(x,y)dy\,dx &= \mathbb{P}(a \leqslant X \leqslant b, c \leqslant Y \leqslant d) \\ &= \mathbb{P}(a \leqslant X \leqslant b)\mathbb{P}(c \leqslant Y \leqslant d) \\ &= \int_a^b f_X(x)dx \int_c^d f_Y(y)dy \\ &= \int_a^b \int_c^d f_X(x)f_Y(y)\,dy\,dx. \end{aligned}$$

One can conclude from this by taking partial derivatives that the joint density function factors into the product of the density functions. Going the other way, one can also see that if the joint density factors, then one has independence of random variables.

11.1 The joint density function factors for independent random variables

Jointly continuous random variables X and Y are independent if and only if

$$f_{X,Y}(x,y) = f_X(x) f_Y(y).$$

Example 11.4. Suppose one has a floor made out of wood planks and one drops a needle onto it. What is the probability the needle crosses one of the cracks? Suppose the needle is of length L and the wood planks are D across.

Solution: Let X be the distance from the midpoint of the needle to the nearest crack and let Θ be the angle the needle makes with the vertical. Then X and Θ are independent random variables. X is uniform on $[0, D/2]$ and Θ is uniform on $[0, \pi/2]$. A little geometry shows that the needle will cross a crack if $L/2 > X/\cos\Theta$. We have $f_{X,\Theta} = \frac{4}{\pi D}$ and so we have to integrate this constant over the set where $X < L\cos\Theta/2$ and $0 \leqslant \Theta \leqslant \pi/2$ and $0 \leqslant X \leqslant D/2$. The integral is

$$\int_0^{\pi/2} \int_0^{L\cos\theta/2} \frac{4}{\pi D} dx\, d\theta = \frac{2L}{\pi D}.$$

Proposition 11.1 (Sums of independent random variables)

Suppose X and Y are continuous random variables, then the density of $X+Y$ is given by the following *convolution formula*

$$f_{X+Y}(a) = \int f_X(a-y) f_Y(y) dy.$$

Proof. If X and Y are independent, then the joint probability density function factors, and therefore

$$\begin{aligned}
\mathbb{P}(X+Y \leqslant a) &= \iint_{\{x+y \leqslant a\}} f_{X,Y}(x,y) dx\, dy \\
&= \iint_{\{x+y \leqslant a\}} f_X(x) f_Y(y) dx\, dy \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{a-y} f_X(x) f_Y(y) dx\, dy \\
&= \int F_X(a-y) f_Y(y) dy.
\end{aligned}$$

Differentiating with respect to a, we have the convolution formula for the density of $X+Y$ as follows

$$f_{X+Y}(a) = \int f_X(a-y) f_Y(y) dy.$$

$\square$

There are a number of cases where this is interesting.

Example 11.5. If X is a gamma random variable with parameters s and λ and Y is a gamma random variable with parameters t and λ, then a straightforward integration shows that $X+Y$ is a gamma with parameters $s+t$ and λ. In particular, the sum of n independent exponential random variables with parameter λ is a gamma random variable with parameters n and λ.

Example 11.6. If $Z \sim \mathcal{N}(0,1)$, then $F_{Z^2}(y) = \mathbb{P}(Z^2 \leqslant y) = \mathbb{P}(-\sqrt{y} \leqslant Z \leqslant \sqrt{y}) = F_Z(\sqrt{y}) - F_Z(-\sqrt{y})$. Differentiating shows that $f_{Z^2}(y) = ce^{-y/2}(y/2)^{(1/2)-1}$, or Z^2 is a gamma random variable with parameters $\frac{1}{2}$ and $\frac{1}{2}$. So using the previous example, if Z_i are independent $\mathcal{N}(0,1)$ random variables, then $\sum_{i=1}^n Z_i^2$ is a gamma random variable with parameters $n/2$ and $\frac{1}{2}$, i.e., a χ_n^2.

Example 11.7. If X_i is a $\mathcal{N}(\mu_i, \sigma_i^2)$ and the X_i are independent, then some lengthy calculations show that $\sum_{i=1}^n X_i$ is a $\mathcal{N}(\sum \mu_i, \sum \sigma_i^2)$.

Example 11.8. The analogue for discrete random variables is easier. If X and Y take only nonnegative integer values, we have

$$\begin{aligned}
\mathbb{P}(X+Y=r) &= \sum_{k=0}^{r} \mathbb{P}(X=k, Y=r-k) \\
&= \sum_{k=0}^{r} \mathbb{P}(X=k)\mathbb{P}(Y=r-k).
\end{aligned}$$

In the case where X is a Poisson random variable with parameter λ and Y is a Poisson random variable with parameter μ, we see that $X+Y$ is a Poisson random variable with parameter $\lambda+\mu$. To check this, use the above formula to get

$$\begin{aligned}
\mathbb{P}(X+Y=r) &= \sum_{k=0}^{r} \mathbb{P}(X=k)\mathbb{P}(Y=r-k) \\
&= \sum_{k=0}^{r} e^{-\lambda}\frac{\lambda^k}{k!} e^{-\mu}\frac{\mu^{r-k}}{(r-k)!} \\
&= e^{-(\lambda+\mu)}\frac{1}{r!}\sum_{k=0}^{r} \binom{r}{k} \lambda^k \mu^{r-k} \\
&= e^{-(\lambda+\mu)}\frac{(\lambda+\mu)^r}{r!}
\end{aligned}$$

using the binomial theorem.

Note that it is not always the case that the sum of two independent random variables will be a random variable of the same type.

Example 11.9. If X and Y are independent normals, then $-Y$ is also a normal (with $\mathbb{E}(-Y) = -\mathbb{E}Y$ and $\operatorname{Var}(-Y) = (-1)^2 \operatorname{Var} Y = \operatorname{Var} Y$), and so $X - Y$ is also normal.

11.3. Conditioning for random variables

We now extend the notion of conditioning an event on another event to random variables. Suppose we have observed the value of a random variable Y, and we need to update the density function of another random variable X whose value we are still to observe. For this we use the conditional density function of X given Y.

We start by considering the case when X and Y are discrete random variables.

Definition (Conditional discrete random variable)

The *conditional random variable* $X \mid Y = y$ called X *given* $Y = y$, a discrete random variable with the probability mass function

$$\mathbb{P}\left((X \mid Y = y) = x\right) = \mathbb{P}\left(X = x \mid Y = y\right)$$

for values x of X.

In terms of probability mass functions we have

$$\begin{aligned} p_{(X|Y=y)}(x) &= \mathbb{P}_{X|Y}(x \mid y) = \mathbb{P}(X = x \mid Y = y) \\ &= \frac{\mathbb{P}(X = x, Y = y)}{\mathbb{P}(Y = y)} = \frac{p_{XY}(x, y)}{p_Y(y)} \end{aligned}$$

wherever $p_Y(y) \neq 0$.

Analogously, we define conditional random variables in the continuous case.

Definition (Conditional continuous random variable)

Suppose X and Y are jointly continuous, the *conditional probability density function (PDF)* of X given Y is given by

$$f_{X|Y=y}(x)) = \frac{f_{XY}(x, y)}{f_Y(y)}.$$

11.4. Functions of random variables

First we formalize what we saw in the one-dimensional case. First we recall that $g(x)$ is called a *strictly increasing function* if for any $x_1 < x_2$, then $g(x_1) < g(x_2)$. Similarly we can define a *strictly decreasing function*. We also say g is a*strictly monotone function* on $[a, b]$ if it is either strictly increasing or strictly decreasing function on this interval.

Finally, we will use the fact that for a strictly monotone function g we have that for any point y in the range of the function there is a *unique* x such that $g(x) = y$. That is, we have a well-defined g^{-1} on the range of function g.

Theorem 11.1 (PDF for a function of a random variable)

Suppose $g(x)$ is differentiable and $g(x)$ is a strictly monotone function, and X is a continuous random variable. Then $Y = g(X)$ is a continuous random variable with the probability density function (PDF)

$$f_Y(y) = \begin{cases} \frac{f_X(g^{-1}(y))}{|g'(g^{-1}(y))|} = \frac{f_X(x)}{|g'(x)|}, & \text{if } y \text{ is in the range of the function } g \\ 0, & \text{if } y \text{ is } \textbf{not} \text{ in the range of the function } g \end{cases}, \tag{11.4.1}$$

where x is in the range of the random variable X.

Proof. Without loss of generality we can assume that g is strictly increasing. If X has a density f_X and $Y = g(X)$, then

$$\begin{aligned} F_Y(y) &= \mathbb{P}(Y \leqslant y) = \mathbb{P}(g(X) \leqslant y) \\ &= \mathbb{P}(X \leqslant g^{-1}(y)) = F_X(g^{-1}(y)). \end{aligned}$$

Taking the derivative, using the chain rule, and recalling that the derivative of $g^{-1}(y)$ is given by

$$(g^{-1}(y))' = \frac{1}{g'(x)} = \frac{1}{g'(g^{-1}(y))}.$$

Here we use that $y = g(x)$, $x = g^{-1}(y)$, and the assumption that $g(x)$ is an increasing function. $\square$

The higher-dimensional case is very analogous. Note that the function below is assumed to be one-to-one, and therefore it is invertible. Note that in the one-dimensional case strictly monotone functions are one-to-one.

Theorem (PDF for a function of two random variables)

Suppose X and Y are two jointly continuous random variables. Let $(U, V) = g(X, Y) = (g_1(X, Y), g_2(X, Y))$, where $g : \mathbb{R}^2 \longrightarrow \mathbb{R}$ is a continuous one-to-one function with continuous partial derivatives. Denote by $h = g^{-1}$, so $h(U, V) = (h_1(U, V), h_2(U, V)) = (X, Y)$. Then U and $V$are jointly continuous and their joint PDF, $f_{UV}(u, v)$, is defined on the range of (U, V) and is given by

$$f_{UV}(u, v) = f_{XY}(h_1(u, v), h_2(u, v)) |J_h|,$$

where J_h is the Jacobian of the map h, that is,

$$J_h = \det \begin{pmatrix} \frac{\partial h_1}{\partial u} & \frac{\partial h_1}{\partial v} \\ \frac{\partial h_2}{\partial u} & \frac{\partial h_2}{\partial v} \end{pmatrix}.$$

Proof. The proof is based on the change of variables theorem from multivariable calculus, and it is analogous to (11.4.1). Note that to reconcile this formula with some of the applications we might find the following property of the Jacobian for a map and its inverse useful

$$J_{g^{-1}} \circ g = (J_g)^{-1},$$

so

$$J_h \circ g = (J_g)^{-1}.$$

□

Example 11.10. Suppose X_1 is $\mathcal{N}(0, 1)$, X_2 is $\mathcal{N}(0, 4)$, and X_1 and X_2 are independent. Let $Y_1 = 2X_1 + X_2, Y_2 = X_1 - 3X_2$. Then $y_1 = g_1(x_1, x_2) = 2x_1 + x_2, y_2 = g_2(x_1, x_2) = x_1 - 3x_2$, so

$$J = \begin{pmatrix} 2 & 1 \\ 1 & -3 \end{pmatrix} = -7.$$

In general, J might depend on x, and hence on y. Some algebra leads to $x_1 = \frac{3}{7}y_1 + \frac{1}{7}y_2$, $x_2 = \frac{1}{7}y_1 - \frac{2}{7}y_2$. Since X_1 and X_2 are independent,

$$f_{X_1,X_2}(x_1, x_2) = f_{X_1}(x_1) f_{X_2}(x_2) = \frac{1}{\sqrt{2\pi}} e^{-x_1^2/2} \frac{1}{\sqrt{8\pi}} e^{-x_2^2/8}.$$

Therefore

$$f_{Y_1,Y_2}(y_1, y_2) = \frac{1}{\sqrt{2\pi}} e^{-(\frac{3}{7}y_1 + \frac{1}{7}y_2)^2/2} \frac{1}{\sqrt{8\pi}} e^{-(\frac{1}{7}y_1 - \frac{2}{7}y_2)^2/8} \frac{1}{7}.$$

11.5. Further examples and applications

Example 11.11. Suppose we roll two dice with sides $1, 1, 2, 2, 3, 3$. Let X be the largest value obtained on any of the two dice. Let $Y =$ be the sum of the two dice. Find the joint PMF of X and Y.

Solution: First we make a table of all the possible *outcomes*. Note that individually, $X = 1, 2, 3$ and $Y = 2, 3, 4, 5, 6$. The table for possible outcomes of (X, Y) jointly is given by the following table.

outcome	1	2	3
1	$(X = 1, Y = 2) = (1, 2)$	$(2, 3)$	$(3, 4)$
2	$(2, 3)$	$(2, 4)$	$(3, 5)$
3	$(3, 4)$	$(3, 5)$	$(3, 6)$

Using this table we have that the PMF is given by:

$X\backslash Y$	2	3	4	5	6
1	$\mathbb{P}(X = 1, Y = 2) = \frac{1}{9}$	0	0	0	0
2	0	$\frac{2}{9}$	$\frac{1}{9}$	0	0
3	0	0	$\frac{2}{9}$	$\frac{2}{9}$	$\frac{1}{9}$

Example 11.12. Let X, Y have joint PDF

$$f(x, y) = \begin{cases} ce^{-x}e^{-2y} & , 0 < x < \infty, 0 < y < \infty \\ 0 & \text{otherwise} \end{cases}.$$

(a) Find c that makes this a valid PDF;
(b) Find $\mathbb{P}(X < Y)$;
(c) Set up the double integral for $\mathbb{P}(X > 1, Y < 1)$;
(d) Find the marginal $f_X(x)$.

Solution: (a) The region that we integrate over in the first quadrant thus

$$\begin{aligned}1 = &\int_0^\infty \int_0^\infty ce^{-x}e^{-2y}dxdy = c\int_0^\infty e^{-2y}\left[-e^{-x}\right]_0^\infty dy \\ =&c\int_0^\infty e^{-2y}dy = c -\frac{1}{2}e^{-2y}\bigg|_0^\infty = c\frac{1}{2}.\end{aligned}$$

Then $c = 2$.

(b): We need to draw the region Let $D = \{(x,y) \mid 0 < x < y, 0 < y < \infty\}$ and set up the the double integral:

$$\begin{aligned}\mathbb{P}(X < Y) &= \int\int_D f(x,y)dA \\ &= \int_0^\infty \int_0^y 2e^{-x}e^{-2y}dxdy = \frac{1}{3}.\end{aligned}$$

(c):

$$\mathbb{P}(X > 1, Y < 1) = \int_0^1 \int_1^\infty 2e^{-x}e^{-2y}dxdy = (1 - e^{-1})e^{-2}.$$

(d): we have

$$\begin{aligned}f_X(x) &= \int_{-\infty}^\infty f(x,y)dy = \int_0^\infty 2e^{-x}e^{-2y}dy \\ &= 2e^{-x}\left[\frac{-1}{2}e^{-2y}\right]_0^\infty = 2e^{-x}\ 0 + \frac{1}{2}\bigg| \\ &= e^{-x}.\end{aligned}$$

Example 11.13. Let X, Y be a random variable with the joint PDF

$$f_{XY}(x,y) = 6e^{-2x}e^{-3y}\ \ 0 < x < \infty, 0 < y < \infty.$$

Are X, Y independent?

Solution: Find f_X and f_Y and see if f_{XY} factors, that is, $f_{XY} = f_X f_Y$. First

$$\begin{aligned}f_X(x) &= \int_0^\infty 6e^{-2x}e^{-3y}dy = 2e^{-2x}, \\ f_Y(y) &= \int_0^\infty 6e^{-2x}e^{-3y}dx = 3e^{-3y}.\end{aligned}$$

which are both exponential. Thus $f_{XY} = f_X f_Y$, therefore yes, X and Y are independent!

Example 11.14. Let X, Y have the joint PDF

$$f_{X,Y}(x,y) = x + y,\ \ 0 < x < 1, 0 < y < 1$$

Are X, Y independent?

Solution: Note that there is no way to factor $x+y=f_X(x)f_Y(y)$, hence they can not be independent.

11.6. More than two random variables

All the concepts and some of the techniques we introduced for two random variables can be extended to more than two random variables. For example, we can define joint PMF, PDF, CDF, independence for three or more random variables. While many of the explicit expressions can be less tractable, the case of normal variables is tractable. First we comment on independence of several random variables.

Recall that we distinguished between jointly independent events (Definition 3.1) and pairwise independent events.

Definition

A set of n random variables $\{X_1,\ldots,X_n\}$ is *pairwise independent* if every pair of random variables is independent.

As for events, if the set of random variables is pairwise independent, it is not necessarily mutually independent as defined next.

Definition

A set of n random variables $\{X_1,\ldots,X_n\}$ is *mutually independent* if for any sequence of numbers $\{x_1,\ldots,x_n\}$, the events $\{X_1 \leqslant x_1\},\ldots,\{X_n \leqslant x_n\}$ are mutually independent events.

This definition is equivalent to the following condition on the joint cumulative distribution function $F_{X_1,\ldots,X_n}(x_1,\ldots,x_n)$. Namely, a set of n random variables $\{X_1,\ldots,X_n\}$ is mutually independent if

$$F_{X_1,\ldots,X_n}(x_1,\ldots,x_n)=F_{X_1}(x_1)\cdot\ldots\cdot F_{X_n}(x_n) \tag{11.6.1}$$

for all $x_1,\ldots,x_n$. Note that we do not need require that the probability distribution factorizes for all possible subsets as in the case for n events. This is not required because Equation 11.6.1 implies factorization for any subset of $1,...,n$.

The following statement will be easier to prove later once we have the appropriate mathematical tools.

Proposition 11.2

If $X_i \sim \mathcal{N}(\mu_i,\sigma_i^2)$ are *independent* for $1 \leqslant i \leqslant n$ then

$$X_1+\cdots+X_n \sim \mathcal{N}\left(\mu_1+\cdots\mu_n,\sigma_1^2+\cdots+\sigma_n^2\right).$$

In particular if $X \sim \mathcal{N}(\mu_x,\sigma_x^2)$ and $Y \sim \mathcal{N}(\mu_y,\sigma_y^2)$ then $X+Y \sim \mathcal{N}\left(\mu_x+\mu_y,\sigma_x^2+\sigma_y^2\right)$ and $X-Y \sim \mathcal{N}\left(\mu_x-\mu_y,\sigma_x^2+\sigma_y^2\right)$. In general for two independent Gaussian X and Y we have $cX+dY \sim \mathcal{N}\left(c\mu_x+d\mu_y,c\sigma_x^2+d\sigma_y^2\right)$.

Example 11.15. Suppose $T \sim \mathcal{N}(95, 25)$ and $H \sim \mathcal{N}(65, 36)$ represents the grades of T. and H. in their Probability class.

(a) What is the probability that their average grades will be less than 90?
(b) What is the probability that H. will have scored higher than T.?
(c) Answer question (b) if $T \sim \mathcal{N}(90, 64)$ and $H \sim \mathcal{N}(70, 225)$.

Solution: (a) By Proposition 11.2 $T + H \sim \mathcal{N}(160, 61)$. Thus

$$\begin{aligned}\mathbb{P}\left(\frac{T+H}{2} \leqslant 90\right) &= \mathbb{P}(T + H \leqslant 180) \\ &= \mathbb{P}\left(Z \leqslant \frac{180-160}{\sqrt{61}}\right) = \Phi\left(\frac{180-160}{\sqrt{61}}\right) \\ &\approx \Phi(2.56) \approx 0.9961\end{aligned}$$

(b): Using $H - T \sim \mathcal{N}(-30, 61)$ we compute

$$\begin{aligned}\mathbb{P}(H > T) &= \mathbb{P}(H - T > 0) \\ &= 1 - \mathbb{P}(H - T < 0) \\ &= 1 - \mathbb{P}\left(Z \leq \frac{0-(-30)}{\sqrt{61}}\right) \\ &\approx 1 - \Phi(3.84) \approx 0.00006\end{aligned}$$

(c): By Proposition 11.2 $T - H \sim \mathcal{N}(-20, 289)$ and so

$$\begin{aligned}\mathbb{P}(H > T) &= \mathbb{P}(H - T > 0) \\ &= 1 - \mathbb{P}(H - T < 0) \\ &= 1 - \mathbb{P}\left(Z \leqslant \frac{0-(-20)}{17}\right) \\ &\approx 1 - \Phi(1.18) \approx 0.11900\end{aligned}$$

Example 11.16. Suppose X_1, X_2 have joint distribution

$$f_{X_1,X_2}(x_1, x_2) = \begin{cases} x_1 + \frac{3}{2}(x_2)^2 & 0 \leqslant x_1 \leqslant 1, 0 \leqslant x_2 \leqslant 1 \\ 0 & \text{otherwise} \end{cases}.$$

Find the joint PDF of $Y_1 = X_1 + X_2$ and $Y_2 = X_1^2$.

Solution: we use Theorem 11.4 here.

First we find the Jacobian of the following map

$$\begin{aligned}y_1 &= g_1(x_1, x_2) = x_1 + x_2, \\ y_2 &= g_2(x_1, x_2) = x_1^2.\end{aligned}$$

So

$$J(x_1, x_2) = \begin{vmatrix} 1 & 1 \\ 2x_1 & 0 \end{vmatrix} = -2x_1$$

Secondly, we need to invert the map g, that is, solve for x_1, x_2 in terms of y_1, y_2

$$\begin{aligned} x_1 &= \sqrt{y_2}, \\ x_2 &= y_1 - \sqrt{y_2}. \end{aligned}$$

The joint PDF of Y_1, Y_2 then is given by

$$\begin{aligned} f_{Y_1,Y_2}(y_1, y_2) &= f_{X_1,X_2}(x_1, x_2) \left|J(x_1, x_2)\right|^{-1} \\ &= f_{X_1,X_2}\left(\sqrt{y_2}, y_1 - \sqrt{y_2}\right) \frac{1}{2x_1} \\ &= \begin{cases} \frac{1}{2\sqrt{y_2}} \left[\left(\sqrt{y_2} + \frac{3}{2}\left(y_1 - \sqrt{y_2}\right)^2\right), & 0 \leqslant y_2 \leqslant 1,\ 0 \leqslant y_1 - \sqrt{y_2} \leqslant 1 \\ 0, & \text{otherwise} \end{cases} \end{aligned}$$

11.7. Exercises

Exercise 11.1. Suppose that 2 balls are chosen without replacement from an urn consisting of 5 white and 8 red balls. Let X equal 1 if the first ball selected is white and zero otherwise. Let Y equal 1 if the second ball selected is white and zero otherwise.

(A) Find the probability mass function of X, Y.
(B) Find $\mathbb{E}(XY)$.
(C) Is it true that $\mathbb{E}(XY) = (\mathbb{E}X)(\mathbb{E}Y)$?
(D) Are X, Y independent?

Exercise 11.2. Suppose you roll two fair dice. Find the probability mass function of X and Y, where X is the largest value obtained on any die, and Y is the sum of the values.

Exercise 11.3. Suppose the joint density function of X and Y is $f(x, y) = \frac{1}{4}$ for $0 < x < 2$ and $0 < y < 2$.

(A) Find $\mathbb{P}\left(\frac{1}{2} < X < 1, \frac{2}{3} < Y < \frac{4}{3}\right)$.
(B) Find $\mathbb{P}(XY < 2)$.
(C) Find the marginal distributions $f_X(x)$ and $f_Y(y)$.

Exercise 11.4. The joint probability density function of X and Y is given by

$$f(x, y) = e^{-(x+y)}, 0 \leqslant x < \infty, 0 \leqslant y < \infty.$$

Find $\mathbb{P}(X < Y)$.

Exercise 11.5. Suppose X and Y are independent random variables and that X is exponential with $\lambda = \frac{1}{4}$ and Y is uniform on $(2, 5)$. Calculate the probability that $2X + Y < 8$.

Exercise 11.6. Consider X and Y given by the joint density

$$f(x, y) = \begin{cases} 10x^2 y & 0 \leqslant y \leqslant x \leqslant 1 \\ 0 & \text{otherwise .} \end{cases}$$

(A) Find the marginal PDFs, $f_X(x)$ and $f_Y(x)$.
(B) Are X and Y independent random variables?
(C) Find $\mathbb{P}\left(Y \leqslant \frac{X}{2}\right)$.
(D) Find $\mathbb{P}\left(Y \leqslant \frac{X}{4} \mid Y \leqslant \frac{X}{2}\right)$.
(E) Find $\mathbb{E}[X]$.

Exercise 11.7. Consider X and Y given by the joint density

$$f(x, y) = \begin{cases} 4xy & 0 \leqslant x \leqslant 1, 0 \leqslant y \leqslant 1 \\ 0 & \text{otherwise.} \end{cases}$$

(A) Find the marginal PDFs, f_X and f_Y.
(B) Are X and Y independent?
(C) Find $\mathbb{E}Y$.

Exercise 11.8. Consider X, Y given by the joint PDF

$$f(x,y) = \begin{cases} \frac{2}{3}(x+2y) & 0 \leqslant x \leqslant 1, 0 \leqslant y \leqslant 1 \\ 0 & \text{otherwise.} \end{cases}$$

Are X and Y independent random variables?

Exercise 11.9. Suppose that gross weekly ticket sales for UConn basketball games are normally distributed with mean \$2,200,000 and standard deviation \$230,000. What is the probability that the total gross ticket sales over the next two weeks exceeds \$4,600,000?

Exercise 11.10. Suppose the joint density function of the random variable X_1 and X_2 is

$$f(x_1, x_2) = \begin{cases} 4x_1x_2 & 0 < x_1 < 1, 0 < x_2 < 1 \\ 0 & \text{otherwise.} \end{cases}$$

Let $Y_1 = 2X_1 + X_2$ and $Y_2 = X_1 - 3X_2$. What is the joint density function of Y_1 and Y_2?

Exercise 11.11. Suppose the joint density function of the random variable X_1 and X_2 is

$$f(x_1, x_2) = \begin{cases} \frac{3}{2}(x_1^2 + x_2^2) & 0 < x_1 < 1, 0 < x_2 < 1 \\ 0 & \text{otherwise.} \end{cases}$$

Let $Y_1 = X_1 - 2x_2$ and $Y_2 = 2X_1 + 3X_2$. What is the joint density function of Y_1 and Y_2?

Exercise 11.12. We roll two dice. Let X be the minimum of the two numbers that appear, and let Y be the maximum. Find the joint probability mass function of (X, Y), that is, $P(X = i, Y = j)$:

$i \backslash j$	1	2	3	4	5	6
1						
2						
3						
4						
5						
6						

Find the marginal probability mass functions of X and Y. Finally, find the conditional probability mass function of X given that $Y = 5$, that is, $\mathbb{P}(X = i | Y = 5)$, for $i = 1, \ldots, 6$.

11.8. Selected solutions

Solution to Exercise 11.1(A): We have

$$\begin{aligned} p(0,0) &= \mathbb{P}(X=0,Y=0) = \mathbb{P}(RR) = \frac{8\cdot 7}{13\cdot 12} = \frac{14}{39}, \\ p(1.0) &= \mathbb{P}(X=1,Y=0) = \mathbb{P}(WR) = \frac{5\cdot 8}{13\cdot 12} = \frac{10}{39}, \\ p(0,1) &= \mathbb{P}(X=0,Y=1) = \mathbb{P}(RW) = \frac{8\cdot 5}{13\cdot 12} = \frac{10}{39}, \\ p(1,1) &= \mathbb{P}(X=1,Y=1) = \mathbb{P}(WW) = \frac{5\cdot 4}{13\cdot 12} = \frac{5}{39}. \end{aligned}$$

Solution to Exercise 11.1(B):

$$\mathbb{E}(XY) = \mathbb{P}(X=1,Y=1) = \frac{5}{39} \approx 0.1282$$

Solution to Exercise 11.1(C): Not true because

$$(\mathbb{E}X)(\mathbb{E}Y) = \mathbb{P}(X=1)\,\mathbb{P}(Y=1) = \left(\frac{5}{13}\right)^2 = \frac{25}{169} \approx 0.1479$$

Solution to Exercise 11.1(D): X and Y are not independent because

$$\mathbb{P}(X=1,Y=1) = \frac{5}{39} \neq \mathbb{P}(X=1)\,\mathbb{P}(Y=1) = \left(\frac{5}{13}\right)^2.$$

Solution to Exercise 11.2: First we need to figure out what values X, Y can attain. Note that X can be any of $1,2,3,4,5,6$, but Y is the sum and can only be as low as 2 and as high as 12. First we make a table of all possibilities for (X,Y) given the values of the dice. Recall X is the largest of the two, and Y is the sum of them. The possible outcomes are given by the following.

1st die\2nd die	1	2	3	4	5	6
1	$(1,2)$	$(2,3)$	$(3,4)$	$(4,5)$	$(5,6)$	$(6,7)$
2	$(2,3)$	$(2,4)$	$(3,5)$	$(4,6)$	$(5,7)$	$(6,8)$
3	$(3,4)$	$(3,5)$	$(3,6)$	$(4,7)$	$(5,8)$	$(6,9)$
4	$(4,5)$	$(4,6)$	$(4,7)$	$(4,8)$	$(5,9)$	$(6,10)$
5	$(5,6)$	$(5,7)$	$(5,8)$	$(5,9)$	$(5,10)$	$(6,11)$
6	$(6,7)$	$(6,8)$	$(6,9)$	$(6,10)$	$(6,11)$	$(6,12)$

Then we make a table of the PMF $p(x,y)$.

$X\backslash Y$	2	3	4	5	6	7	8	9	10	11	12
1	$\frac{1}{36}$	0	0	0	0	0	0	0	0	0	0
2	0	$\frac{2}{36}$	$\frac{1}{36}$	0	0	0	0	0	0	0	0
3	0	0	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	0	0	0	0	0	0
4	0	0	0	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	0	0	0	0
5	0	0	0	0	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	0	0
6	0	0	0	0	0	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Solution to Exercise 11.3(A): We integrate the PDF over $0 < x < 2$ and $0 < y < 2$ and get

$$\int_{\frac{1}{2}}^{1} \int_{\frac{2}{3}}^{\frac{4}{3}} \frac{1}{4} dydx = \frac{1}{4}\left(1 - \frac{1}{2}\right)\left(\frac{4}{3} - \frac{2}{3}\right) = \frac{1}{12}.$$

Solution to Exercise 11.3(B): We need to find the region that is within $0 < x, y < 2$ and $y < \frac{2}{x}$. (Try to draw the region) We get two regions from this. One with bounds $0 < x < 1, 0 < y < 2$ and the region $1 < x < 2, 0 < y < \frac{2}{x}$. Then

$$\begin{aligned}\mathbb{P}(XY < 2) &= \int_0^1 \int_0^2 \frac{1}{4} dydx + \int_1^2 \int_0^{\frac{2}{x}} \frac{1}{4} dydx \\ &= \frac{1}{2} + \int_1^2 \frac{1}{2x} dx \\ &= \frac{1}{2} + \frac{\ln 2}{2}.\end{aligned}$$

Solution to Exercise 11.3(C): Recall that

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)dy = \int_0^2 \frac{1}{4} dy = \frac{1}{2}$$

for $0 < x < 2$ and 0 otherwise. By symmetry, f_Y is the same.

Solution to Exercise 11.4: Draw a picture of the region and note the integral needs to be set up in the following way:

$$\begin{aligned}\mathbb{P}(X < Y) &= \int_0^{\infty} \int_0^{y} e^{-(x+y)} dxdy = \int_0^{\infty} \left[-e^{-2y} + e^{-y}\right] dy \\ &= \left(\frac{1}{2} e^{-2y} - e^{-y}\right)\Big|_0^{\infty} = 0 - \left(\frac{1}{2} - 1\right) = \frac{1}{2}.\end{aligned}$$

Solution to Exercise 11.5: We know that

$$f_X(x) = \begin{cases} \frac{1}{4}e^{-\frac{x}{4}}, & \text{when } x \geqslant 0 \\ 0, & \text{otherwise} \end{cases}$$

and

$$f_Y(y) = \begin{cases} \frac{1}{3}, & \text{when } 2 < y < 5 \\ 0, & \text{otherwise} \end{cases}$$

Since X, Y are independent then $f_{X,Y} = f_X f_Y$, thus

$$f_{X,Y}(x,y) = \begin{cases} \frac{1}{12}e^{-\frac{x}{4}} & \text{when } x \geqslant 0, 2 < y < 5 \\ 0 & \text{otherwise} \end{cases}$$

Draw the region $(2X + Y < 8)$, which correspond to $0 \leqslant x$, $0 < y < 5$ and $y < 8 - 2x$. Drawing a picture of the region, we get the corresponding bounds of $2 < y < 5$ and $0 < x < 4 - \frac{y}{2}$, so that

$$\begin{aligned} \mathbb{P}(2X + Y < 8) &= \int_2^5 \int_0^{4-\frac{y}{2}} \frac{1}{12}e^{-\frac{x}{4}} dx dy \\ &= \int_2^5 \frac{1}{3}\left(1 - e^{y/8-1}\right) dy \\ &= 1 - \frac{8}{3}\left(e^{-\frac{3}{8}} - e^{-\frac{3}{4}}\right) \end{aligned}$$

Solution to Exercise 11.6(A): We have

$$f_X(x) = \begin{cases} 5x^4, & 0 \leqslant x \leqslant 1 \\ 0, & \text{otherwise} \end{cases}$$

$$f_Y(y) = \begin{cases} \frac{10}{3}y(1 - y^3), & 0 \leqslant y \leqslant 1 \\ 0, & \text{otherwise} \end{cases}.$$

Solution to Exercise 11.6(B): No, since $f_{X,Y} \neq f_X f_Y$.

Solution to Exercise 11.6(C): $\mathbb{P}\left(Y \leqslant \frac{X}{2}\right) = \frac{1}{4}$.

Solution to Exercise 11.6(D): Also $\frac{1}{4}$.

Solution to Exercise 11.6(E): Use f_X and the definition of expected value, which is $5/6$.

Solution to Exercise 11.7(A): $f_X = 2x$ and $f_Y = 2y$.

Solution to Exercise 11.7(B): Yes! Since $f(x, y) = f_X f_Y$.

Solution to Exercise 11.7(C): We have $\mathbb{E}Y = \int_0^1 y \cdot 2y dy = \frac{2}{3}$.

Solution to Exercise 11.8: We get $f_X = \left(\frac{2}{3}x + \frac{2}{3}\right)$ while $f_Y = \frac{1}{3} + \frac{4}{3}y$ and $f \neq f_X f_Y$.

Solution to Exercise 11.9: If $W = X_1 + X_2$ is the sales over the next two weeks, then W is normal with mean $2,200,000 + 2,200,000 = 4,400,00$ and variance $230,000^2 + 230,000^2$.

Thus the variance is $\sqrt{230,000^2 + 230,000^2} = 325,269.1193$. Hence

$$\begin{aligned}\mathbb{P}(W > 5,000,000) &= \mathbb{P}\left(Z > \frac{4,600,000 - 4,400,000}{325,269.1193}\right)\\ &= \mathbb{P}(Z > 0.6149)\\ &\approx 1 - \Phi(0.61) = 0.27.\end{aligned}$$

Solution to Exercise 11.10: First find the Jacobian:

$$\begin{aligned}y_1 &= g_1(x_1, x_2) = 2x_1 + x_2,\\ y_2 &= g_2(x_1, x_2) = x_1 - 3x_2.\end{aligned}$$

Thus

$$J(x_1, x_2) = \begin{vmatrix} 2 & 1 \\ 1 & -3 \end{vmatrix} = -7.$$

Solve for x_1, x_2 and get

$$\begin{aligned}x_1 &= \frac{3}{7}y_1 + \frac{1}{7}y_2\\ x_2 &= \frac{1}{7}y_1 - \frac{2}{7}y_2\end{aligned}$$

The joint PDF for Y_1, Y_2 is given by

$$\begin{aligned}f_{Y_1,Y_2}(y_1, y_2) &= f_{X_1,X_2}(x_1, x_2)\left|J(x_1, x_2)\right|^{-1}\\ &= f_{X_1,X_2}\left(\frac{3}{7}y_1 + \frac{1}{y}y_2, \frac{1}{7}y_1 - \frac{2}{7}y_2\right)\frac{1}{7}.\end{aligned}$$

Since we are given the joint PDF of X_1 and X_2, then plugging them into f_{X_1,X_2}, we have

$$f_{Y_1,Y_2}(y_1, y_2) = \begin{cases} \frac{4}{7^3}(3y_1 + y_2)(y_1 - 2y_2) & 0 < 3y_1 + y_2 < 7, 0 < y_1 - 2y_2 < 2 \\ 0 & \text{otherwise.} \end{cases}$$

Solution to Exercise 11.11:

First find the Jacobian

$$\begin{aligned}y_1 &= g_1(x_1, x_2) = x_1 - 2x_2,\\ y_2 &= g_2(x_1, x_2) = 2x_1 + 3x_2.\end{aligned}$$

This

$$J(x_1, x_2) = \begin{vmatrix} 1 & -2 \\ 2 & 3 \end{vmatrix} = 7.$$

Now solve for x_1, x_2 and get

$$x_1 = \frac{1}{7}\left(3y_1 + 2y_2\right)$$
$$x_2 = \frac{1}{7}\left(-2y_1 + y_2\right)$$

The joint PDF for Y_1, Y_2 is given by

$$\begin{aligned} f_{Y_1,Y_2}\left(y_1, y_2\right) &= f_{X_1,X_2}\left(x_1, x_2\right)\left|J\left(x_1, x_2\right)\right|^{-1} \\ &= f_{X_1,X_2}\left(\frac{1}{7}\left(3y_1 + 2y_2\right), \frac{1}{7}\left(-2y_1 + y_2\right)\right)\frac{1}{7}. \end{aligned}$$

Since we are given the joint PDFs of X_1 and X_2, then plugging them into f_{X_1,X_2}, we have

$$f_{Y_1,Y_2}\left(y_1, y_2\right) = \begin{cases} \frac{3}{2}\cdot\frac{1}{7^3}\left(\left(3y_1 + 2y_2\right)^2 + \left(-2y_1 + y_2\right)^2\right), & 0 < 3y_1 + 2y_2 < 7, 0 < -2y_1 + y_2 < 7 \\ 0, & \text{otherwise.} \end{cases}$$

Solution to Exercise 11.12: The joint probability mass function $\mathbb{P}\left(X = i, Y = j\right)$ of (X, Y) is

$i \backslash j$	1	2	3	4	5	6
1	$1/36$	$1/18$	$1/18$	$1/18$	$1/18$	$1/18$
2	0	$1/36$	$1/18$	$1/18$	$1/18$	$1/18$
3	0	0	$1/36$	$1/18$	$1/18$	$1/18$
4	0	0	0	$1/36$	$1/18$	$1/18$
5	0	0	0	0	$1/36$	$1/18$
6	0	0	0	0	0	$1/36$

The marginal probability mass functions $\mathbb{P}(X = i)$ of X is given by

i	1	2	3	4	5	6
$P(X = i)$	$1/36$	$1/12$	$5/36$	$7/36$	$1/4$	$11/36$

The marginal probability mass functions of Y is given by

i	1	2	3	4	5	6
$P(Y = i)$	$11/36$	$1/4$	$7/36$	$5/36$	$1/12$	$1/36$

The conditional probability mass function of X given that Y is 5, $\mathbb{P}(X = i|Y = 5)$, for $i = 1, \ldots, 6$

i	1	2	3	4	5	6
$P(X = i\|Y = 5)$	2/9	2/9	2/9	2/9	1/9	0