6 Jointly continuous random variables

Again, we deviate from the order in the book for this chapter, so the subsections in this chapter do not correspond to those in the text.

6.1 Joint density functions

Recall that X is continuous if there is a function $f(x)$ (the density) such that

$$\mathbf{P}(X \le t) = \int_{-\infty}^{t} f_X(x)\, dx$$

We generalize this to two random variables.

Definition 1. *Two random variables X and Y are jointly continuous if there is a function $f_{X,Y}(x,y)$ on $\mathbb{R}^2$, called the joint probability density function, such that*

$$\mathbf{P}(X \le s, Y \le t) = \int\int_{x \le s, y \le t} f_{X,Y}(x,y)\, dxdy$$

The integral is over $\{(x,y) : x \le s, y \le t\}$. We can also write the integral as

$$\begin{aligned} \mathbf{P}(X \le s, Y \le t) &= \int_{-\infty}^{s} \left(\int_{-\infty}^{t} f_{X,Y}(x,y)\, dy \right) dx \\ &= \int_{-\infty}^{t} \left(\int_{-\infty}^{s} f_{X,Y}(x,y)\, dx \right) dy \end{aligned}$$

In order for a function $f(x,y)$ to be a joint density it must satisfy

$$\begin{aligned} f(x,y) &\ge 0 \\ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y) dxdy &= 1 \end{aligned}$$

Just as with one random variable, the joint density function contains all the information about the underlying probability measure if we only look at the random variables X and Y. In particular, we can compute the probability of any event defined in terms of X and Y just using $f(x,y)$.

Here are some events defined in terms of X and Y :
$\{X \le Y\}$, $\{X^2 + Y^2 \le 1\}$, and $\{1 \le X \le 4, Y \ge 0\}$. They can all be written in the form $\{(X,Y) \in A\}$ for some subset A of $\mathbb{R}^2$.

Proposition 1. *For $A \subset \mathbb{R}^2$,*

$$\mathbf{P}((X,Y) \in A) = \int\int_A f(x,y)\,dxdy$$

The two-dimensional integral is over the subset A of $\mathbb{R}^2$. Typically, when we want to actually compute this integral we have to write it as an iterated integral. It is a good idea to draw a picture of A to help do this.

A rigorous proof of this theorem is beyond the scope of this course. In particular we should note that there are issues involving σ-fields and constraints on A. Nonetheless, it is worth looking at how the proof might start to get some practice manipulating integrals of joint densities.

If $A = (-\infty, s] \times (-\infty, t]$, then the equation is the definition of jointly continuous. Now suppose $A = (-\infty, s] \times (a, b]$. The we can write it as $A = [(-\infty, s] \times (-\infty, b]] \setminus [(-\infty, s] \times (-\infty, a]]$ So we can write the event

$$\{(X,Y) \in A\} = \{(X,Y) \in (-\infty, s] \times (-\infty, b]\} \setminus \{(X,Y) \in (-\infty, s] \times (-\infty, a]\}$$

MORE !!!!!!!!!

Example: Let $A \subset \mathbb{R}^2$. We can X and Y are uniformly distributed on A if

$$f(x) = \begin{cases} \frac{1}{c}, & \text{if } (x,y) \in A \\ 0, & \text{otherwise} \end{cases}$$

where c is the area of A.

Example: Let X, Y be uniform on $[0,1] \times [0,2]$. Find $\mathbf{P}(X + Y \leq 1)$.

Example: Let X, Y have density

$$f(x,y) = \frac{1}{2\pi} \exp(-\frac{1}{2}(x^2 + y^2))$$

Compute $\mathbf{P}(X \leq Y)$ and $\mathbf{P}(X^2 + Y^2 \leq 1)$.

Example: Now suppose X, Y have density

$$f(x,y) = \begin{cases} e^{-x-y} & \text{if } x, y \geq 0 \\ 0, & \text{otherwise} \end{cases}$$

Compute $\mathbf{P}(X + Y \leq t)$.

End of October 19 lecture

What does the pdf mean? In the case of a single discrete RV, the pmf has a very concrete meaning. $f(x)$ is the probability that $X = x$. If X is a single continuous random variable, then

$$\mathbf{P}(x \le X \le x+\delta) = \int_x^{x+\delta} f(u)\,du \approx \delta f(x)$$

If X, Y are jointly continuous, than

$$\mathbf{P}(x \le X \le x+\delta, y \le Y \le y+\delta) \approx \delta^2 f(x,y)$$

6.2 Independence and marginal distributions

Suppose we know the joint density $f_{X,Y}(x,y)$ of X and Y. How do we find their individual densities $f_X(x)$, $f_Y(y)$. These are called *marginal densities.* The cdf of X is

$$\begin{aligned} F_X(x) &= \mathbf{P}(X \le x) = \mathbf{P}(-\infty < X \le x, -\infty < Y < \infty) \\ &= \int_{-\infty}^{x} \left[\int_{-\infty}^{\infty} f_{X,Y}(u,y)\,dy \right] du \end{aligned}$$

Differentiate this with respect to x and we get

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

In words, we get the marginal density of X by integrating y from $-\infty$ to ∞ in the joint density.

Proposition 2. *If X and Y are jointly continuous with joint density $f_{X,Y}(x,y)$, then the marginal densities are given by*

$$\begin{aligned} f_X(x) &= \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy \\ f_Y(y) &= \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dx \end{aligned}$$

We will define independence of two contiunous random variables differently than the book. The two definitions are equivalent.

Definition 2. *Let X, Y be jointly continuous random variables with joint density $f_{X,Y}(x,y)$ and marginal densities $f_X(x)$, $f_Y(y)$. We say they are independent if*

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

If we know the joint density of X and Y, then we can use the definition to see if they are independent. But the definition is often used in a different way. If we know the marginal densities of X and Y **and** we know that they are independent, then we can use the definition to find their joint density.

Example: If X and Y are independent random varialbes and each has the standard normal distribution, what is their joint density?

$$f(x,y) = \frac{1}{2\pi} \exp(-\frac{1}{2}(x^2 + y^2))$$

Example: Suppose that X and Y have a joint density that is uniform on the disc centered at the origin with radius 1. Are they independent?

Example: In the homework you will show that if X and Y have a joint density that is uniform on the square $[a,b] \times [c,d]$, then they are independent.

Example: Suppose that X and Y have joint density

$$f(x,y) = \begin{cases} e^{-x-y} & \text{if } x, y \geq 0 \\ 0, & \text{otherwise} \end{cases}$$

Are X and Y independent?

Example: Suppose that X and Y are independent. X is uniform on $[0,1]$ and Y has the Cauchy density.
(a) Find their joint density.
(b) Compute $\mathbf{P}(0 \leq X \leq 1/2, 0 \leq Y \leq 1)$
(c) Compute $\mathbf{P}(Y \geq X)$.

6.3 Expected value

If X and Y are jointly continuously random variables, then the mean of X is still defined by

$$\mathbf{E}[X] = \int_{-\infty}^{\infty} x\, f_X(x)\, dx$$

If we write the marginal $f_X(x)$ in terms of the joint density, then this becomes

$$\mathbf{E}[X] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x\, f_{X,Y}(x,y)\, dxdy$$

Now suppose we have a function $g(x,y)$ from $\mathbb{R}^2$ to $\mathbb{R}$. Then we can define a new random variable by $Z = g(X,Y)$. In a later section we will see how to compute the density of Z from the joint density of X and Y. We could then compute the mean of Z using the density of Z. Just as in the discrete case there is a shortcut.

Theorem 1. *Let X, Y be jointly continuous random variables with joint density $f(x,y)$. Let $g(x,y) : \mathbb{R}^2 \to \mathbb{R}$. Define a new random variable by $Z = g(X,Y)$. Then*

$$\mathbf{E}[Z] = \int_{\infty}^{\infty} \int_{\infty}^{\infty} g(x,y)\, f(x,y)\, dxdy$$

provided

$$\int_{\infty}^{\infty} \int_{\infty}^{\infty} |g(x,y)|\, f(x,y)\, dxdy < \infty$$

An important special case is the following

Corollary 1. *If X and Y are jointly continuous random variables and a, b are real numbers, then*

$$\mathbf{E}[aX + bY] = a\mathbf{E}[X] + b\mathbf{E}[Y]$$

Example: X and Y have joint density

$$f(x,y) = \begin{cases} x+y & \text{if } 0 \leq x \leq 1, 0 \leq y \leq 1 \\ 0, & \text{otherwise} \end{cases}$$

Let $Z = X + Y$. Find the mean and variance of Z.

We now consider independence and expectation.

Theorem 2. *If X and Y are independent and jointly continuous, then*

$$\mathbf{E}[XY] = \mathbf{E}[X]\,\mathbf{E}[Y]$$

Proof. Since they are independent, $f_{X,Y}(x,y) = f_X(x) f_Y(y)$. So

$$\begin{aligned}
\mathbf{E}[XY] &= \int\int xy\, f_X(x)\, f_Y(y)\, dxdy \\
&= \left[\int x\, f_X(x)\, dx\right]\left[\int y\, f_Y(y)\, dy\right] = \mathbf{E}[X]\mathbf{E}[Y]
\end{aligned}$$

□

6.4 Function of two random variables

Suppose X and Y are jointly continuous random variables. Let $g(x,y)$ be a function from $\mathbb{R}^2$ to $\mathbb{R}$. We define a new random variable by $Z = g(X,Y)$. Recall that we have already seen how to compute the expected value of Z. In this section we will see how to compute the density of Z. The general strategy is the same as when we considered functions of one random variable: we first compute the cumulative distribution function.

Example: Let X and Y be independent random variables, each of which is uniformly distributed on $[0,1]$. Let $Z = XY$. First note that the range of Z is $[0,1]$.

$$F_Z(z) = \mathbf{P}(Z \le z) = \int\int_A 1\, dxdy$$

Where A is the region

$$A = \{(x,y) : 0 \le x \le 1, 0 \le y \le 1, xy \le z\}$$

PICTURE

$$\begin{aligned}
F_Z(z) &= z + \int_z^1 \left[\int_0^{z/x} 1\, dy\right] dx \\
&= z + \int_z^1 \left[\int_0^{z/x} 1\, dy\right] dx
\end{aligned}$$

$$
\begin{aligned}
&= z + \int_z^1 \frac{z}{x}\, dx \\
&= z + z \ln x|_z^1 = z - z \ln z
\end{aligned}
$$

This is the cdf of Z. So we differentiate to get the density.

$$
\frac{d}{dz} F_Z(z) = \frac{d}{dz} z - z \ln z = 1 - \ln z - z \frac{1}{z} = -\ln z
$$

$$
f_Z(z) = \begin{cases} -\ln z, & \text{if } 0 \leq z \leq 1 \\ 0, & \text{otherwise} \end{cases}
$$

Example: Let X and Y be independent random variables, each of which is exponential with parameter λ. Let $Z = X + Y$. Find the density of Z.

Should get gamma with same λ and $w = 2$.

This is special case of a much more general result. The sum of gamma(λ, w_1) and gamma(λ, w_2) is gamma($\lambda, w_1 + w_2$). We could try to show this as we did the previous example. But it is much easier to use moment generating functions which we will introduce in the next section.

End of October 21 lecture

One of the most important examples of a function of two random variables is $Z = X + Y$. In this case

$$
\begin{aligned}
F_Z(z) &= \mathbf{P}(Z \leq z) = \mathbf{P}(X + Y \leq z) \\
&= \int_{-\infty}^{\infty} \left[\int_{-\infty}^{z-x} f(x,y)\, dy \right] dx
\end{aligned}
$$

To get the density of Z we need to differentiate this with respect to Z. The only z dependence is in the upper limit of the inside integral.

$$
\begin{aligned}
f_Z(z) = \frac{d}{dz} F_Z(z) &= \int_{-\infty}^{\infty} \left[\frac{d}{dz} \int_{-\infty}^{z-x} f(x,y)\, dy \right] dx \\
&= \int_{-\infty}^{\infty} f(x, z-x) dx
\end{aligned}
$$

If X and Y are independent, then this becomes

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z-x) dx$$

This is known as a convolution. We can use this formula to find the density of the sum of two independent random variables. But in some cases it is easier to do this using generating functions which we study in the next section.

Example: Let X and Y be independent random variables each of which has the standard normal distribution. Find the density of $Z = X + Y$.

We need to compute the convolution

$$\begin{aligned} f_Z(z) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp(-\frac{1}{2}x^2 - \frac{1}{2}(z-x)^2)\, dx \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp(-x^2 - \frac{1}{2}z^2 + xz)\, dx \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp(-(x-z/2)^2 - \frac{1}{4}z^2)\, dx \\ &= e^{-z^2/4} \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp(-(x-z/2)^2)\, dx \end{aligned}$$

Now the substitution $u = x - z/2$ shows

$$\int_{-\infty}^{\infty} \exp(-(x-z/2)^2)\, dx = \int_{-\infty}^{\infty} \exp(-u^2)\, du$$

This is a constant - it does not depend on z. So $f_Z(z) = ce^{-z^2}$. Another simple substitution allows one to evaluate the constant, but there is no need. We can already see that Z has a normal distribution with mean zero and variance 2. The constant is whatever is needed to normalize the distribution.

6.5 Moment generating functions

This will be very similar to what we did in the discrete case.

Definition 3. *For a continuous random variable X, the moment generating function (mgf) of X is*

$$M_X(t) = \mathbf{E}[e^{tX}] = \int_{-\infty}^{\infty} e^{tx} f_X(x)\, dx$$

Example: Compute it for exponential. Should find $M(t) = \frac{\lambda}{\lambda - t}$.
Example: In the homework you will compute it for the gamma distribution and find (hopefully)

$$M(t) = \left(\frac{\lambda}{\lambda - t}\right)^w$$

Proposition 3. *(1) Let X be a continuous random variable with mgf $M_X(t)$. Then*

$$\mathbf{E}[X^k] = \frac{d^k}{dt^k} M_X(t)|_{t=0}$$

(2) If X and Y are independent continuous random variables then

$$M_{X+Y}(t) = M_X(t) M_Y(t)$$

(3) If the mgf of X is $M_X(t)$ and we let $Y = aX + b$, then

$$M_Y(t) = e^{tb} M_X(at)$$

Proof. For (1)

$$\begin{aligned}
\frac{d^k}{dt^k} M_X(t)|_{t=0} &= \frac{d^k}{dt^k} \int_{-\infty}^{\infty} f_X(x)\, e^{tx}|_{t=0}\, dx \\
&= \int_{-\infty}^{\infty} f_X(x)\, \frac{d^k}{dt^k} e^{tx}|_{t=0}\, dx \\
&= \int_{-\infty}^{\infty} f_X(x)\, x^k\, e^{tx}|_{t=0}\, dx \\
&= \int_{-\infty}^{\infty} f_X(x)\, x^k\, dx = \mathbf{E}[X^k]
\end{aligned}$$

If X and Y are independent, then

$$\begin{aligned}
M_{X+Y}(t) &= \mathbf{E}[\exp(t(X+Y))] = \mathbf{E}[\exp(tX)\exp(tY)] \\
&= \mathbf{E}[\exp(tX)]\mathbf{E}[\exp(tY)] = M_X(t) M_Y(t)
\end{aligned}$$

This calculation assumes that since X and Y are independent, then $\exp(tX)$ and $\exp(tY)$ are independent random variables. We have not shown this.

Part (3) is just

$$M_Y(t) = \mathbf{E}[e^{tY}] = \mathbf{E}[e^{t(aX+b)}] = e^{tb}\mathbf{E}[e^{taX}] = e^{tb} M_X(at)$$

□

End of October 26 lecture

As an application of part (3) we have

Example: Find the mgf of the standard normal and use part (3) to find the mgf of the general normal.

Let Z have a standard normal distribution. We complete the square and get

$$M(t) = \exp(\frac{1}{2}t^2)$$

Now let $X = \mu + \sigma Z$. Then X has a normal distribution with mean μ and variance σ^2. By (3)

$$M_X(t) = \exp(\mu t) M_Z(\sigma t) = \exp(\mu t + \frac{1}{2}\sigma^2 t^2)$$

Proposition 4. *(a) If $X_1, X_2, \cdots, X_n$ are independent and each is normal with mean μ_i and variance σ_i^2, then $Y = X_1 + X_2 + \cdots + X_n$ has a normal distribution with mean μ and variance σ^2 given by*

$$\begin{aligned}
\mu &= \sum_{i=1}^{n} \mu_i, \\
\sigma^2 &= \sum_{i=1}^{n} \sigma_i^2
\end{aligned}$$

(b) If $X_1, X_2, \cdots, X_n$ are independent and each is exponential with parameter λ, then $Y = X_1 + X_2 + \cdots + X_n$ has a gamma distribution with parameters $\lambda = \lambda$ and $w = n$.
(c) If $X_1, X_2, \cdots, X_n$ are independent and each is gamma with parameters λ, w_i, then $Y = X_1 + X_2 + \cdots + X_n$ has a gamma distibution with parameters λ and $w = w_1 + \cdots + w_n$.

We will prove the theorem by proving statements about generating functions. For example, for part (a) what we will really prove is that the moment generating function of Y is that of a normal with the stated parameters. To complete the proof we need to know that if two random variables have the same moment generating functions then they have the same densities.

This is a theorem but it is a hard theorem and it requires some technical assumptions on the random variables. We will ignore these subtleties and just assume that if two RV's have the same mgf, then they have the same density.

Proof. We prove all three parts by simply computing the mgf's involved. $\square$

6.6 Cumulative distribution functions and more independence

Recall that for a discrete random variable X we have a probability mass function $f_X(x)$ which is just $f_X(x) = \mathbf{P}(X = x)$. And for a continuous random variable X we have a probability density function $f_X(x)$. It is a density in the sense that if $\epsilon > 0$ is small, then $\mathbf{P}(x \leq X \leq x + \epsilon) \approx f(x)\epsilon$.

For both types of random variables we have a cumulative distribution function and its definition is the same for all types of RV's.

Definition 4. *Let X, Y be random variables (discrete or continuous). Their joint (cumulative) distribution function is*

$$F_{X,Y}(x, y) = \mathbf{P}(X \leq x, Y \leq y)$$

If X and Y are jointly continuous then we can compute the joint cdf from their joint pdf:

$$F_{X,Y}(x, y) = \int_{-\infty}^{x} \left[\int_{-\infty}^{y} f(u, v)\, dv \right] du$$

If we know the joint cdf, then we can compute the joint pdf by taking partial derivatives of the above :

$$\frac{\partial^2}{\partial x \partial y} F_{X,Y}(x, y) = f(x, y)$$

Calc review : partial derivatives

The joint cdf has properties similar to the cdf for a single RV.

Proposition 5. *Let $F(x, y)$ be the joint cdf of two continuous random variables. Then $F(x, y)$ is a continuous function on $\mathbb{R}^2$ and*

$$\lim_{x,y \to -\infty} F(x, y) = 0, \qquad \lim_{x,y \to \infty} F(x, y) = 1,$$

$$F(x_1, y) \leq F(x_2, y) \; if \; x_1 \leq x_2, \quad F(x, y_1) \leq F(x, y_2) \; if \; y_1 \leq y_2$$

$$\lim_{x \to \infty} F(x, y) = F_Y(y) \quad \lim_{y \to \infty} F(x, y) = F_X(x)$$

We will use the joint cdf to prove more results about independent of RV's.

Theorem 3. *If X and Y are jointly continuous random variables then they are independent if and only if $F_{X,Y}(x, y) = F_X(x)F_Y(y)$.*

The theorem is true for discrete random variables as well.

Proof. □

Example: Suppose that the joint cdf of X and Y is

$$F(x, y) = \begin{cases} \frac{1}{2}(1 - e^{-2x})(y + 1) & \text{if } -1 \leq y \leq 1, x \geq 0 \\ (1 - e^{-2x}) & \text{if } y \geq 0, x > 1 \\ 0 & \text{if } y < 0 \\ 0 & \text{if } y \geq 0, x < -1 \end{cases}$$

Show that X and Y are independent and find their joint density.

Theorem 4. *If X and Y are independent jointly continuous random variables and g and h are functions from $\mathbb{R}$ to $\mathbb{R}$ then $g(X)$ and $h(Y)$ are independent random variables.*

Proof. We will only prove a special case. We assume that g and h are increasing. We also assume they are differentiable. Let $W = g(X)$, $Z = h(Y)$. By the previous theorem we can show that W and Z are independent by showing that $F_{W,Z}(w, z) = F_W(w)F_Z(z)$. We have

$$F_{W,Z}(w, z) = \mathbf{P}(g(X) \leq w, h(Y) \leq z)$$

Because g and h are increasing, the event $\{g(X) \leq w, h(Y) \leq z\}$ is the same as the event $\{X \leq g^{-1}(w), Y \leq h^{-1}(z)\}$. So

$$\begin{aligned} F_{W,Z}(w, z) &= \mathbf{P}(X \leq g^{-1}(w), Y \leq h^{-1}(z)) \\ &= F_{X,Y}(g^{-1}(w), h^{-1}(z)) = F_X(g^{-1}(w))F_Y(h^{-1}(z)) \end{aligned}$$

where the last equality comes from the previous theorem and the independence of X and Y. The individual cdfs of W and Z are

$$\begin{aligned} F_W(w) &= \mathbf{P}(X \le g^{-1}(w)) = F_X(g^{-1}(w)) \\ F_Z(z) &= \mathbf{P}(Y \le h^{-1}(z)) = F_Y(h^{-1}(z)) \end{aligned}$$

So we have shown $F_{W,Z}(w,z) = F_W(w)F_Z(z)$. □

Suppose we have two random variables X and Y and we know their joint density. We have two function $g : \mathbb{R}^2 \to \mathbb{R}$ and $g : \mathbb{R}^2 \to \mathbb{R}$, and we define two new random variables by $W = g(X,Y)$, $Y = h(X,Y)$. Can we find the joint density of W and Z. In principle we can do this by computing their joint cdf and then taking partial derivatives. In practice this can be a mess. There is a another way involving Jacobians which we will study in the next section. We illustrate the cdf approach with an example.

Example Let X and Y be independent standard normal RV's. Let $W = X + Y$ and $Z = X - Y$. Find the joint density of W and Z.

Theorem 5. *If X and Y are independent and jointly continuous, then*

$$var(X+Y) = var(X) + var(Y)$$

X,Y independent iff expectation factors for all functions of X,Y
statistic paradigm - iid

6.7 Change of variables

calc review - Jacobians