Expectation of functions of multiple random variables and some related theorems

March 30, 2020

1 The Expectation of a function of Random Variables

Recall if you're given one cts. random variable X, with pdf $f_X(x)$, and a function $g:\mathbb{R}\to\mathbb{R}$ can compute $\mathbb{E}[g(X)]$ by

$$\mathbb{E}[X] = \int g(x) f_X(x) dx.$$

(in the discrete case it's $\mathbb{E}[X] = \sum_x g(x) p_X(x)$).

Now that we have a notion of a joint distribution of random variables, we can instead consider a function $g:\mathbb{R}^2\to\mathbb{R}$ (this is just notation for a function that takes 2 real numbers as inputs and outputs 1 real number) and compute its expectation when applied to a pair of random variables. This is the content of following definition:

Definition 1.1 (Expectation of 2 random variables- Continuous Case). *Let X_1, X_2 be continuous random variables with joint pdf $f_{X_1,X_2}(x_1,x_2)$, and let $g:\mathbb{R}^2\to\mathbb{R}$, then the **expected value** of $g(X_1,X_2)$ is*

$$\mathbb{E}[g(X_1,X_2)] = \iint g(x_1,x_2) f_{X_1,X_2}(x_1,x_2)\ dx_1 dx_2.$$

The discrete case looks similar, integrals and pdfs are switched with sumsand pmfs.

Definition 1.2 (Expectation of 2 random variables- Discrete Case). *Let Y_1, Y_2 be discrete random variables with joint pmf $p_{Y_1,Y_2}(y_1,y_2)$, and let $g:\mathbb{R}^2\to\mathbb{R}$, then the **expected value** of $g(Y_1,Y_2)$ is*

$$\mathbb{E}[g(Y_1,Y_2)] = \sum_{y_1,y_2} g(y_1,y_2) p_{Y_1,Y_2}(y_1,y_2).$$

Let's do some examples.

Example 1.3. *Let Y_1 and Y_2 be discrete random variables with joint pmf:*

$$p_{Y_1,Y_2}(-1,0) = 1/10\ p_{Y_1,Y_2}(1,0) = 1/4\ p_{Y_1,Y_2}(0,1) = 1/5\ p_{Y_1,Y_2}(0,-1) = 3/10\ p_{Y_1,Y_2}(0,0) = 3/20$$

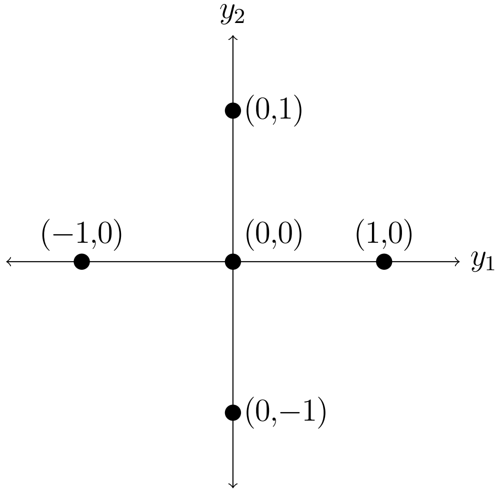
and $p_{Y_1,Y_2}(y_1,y_2) = 0$ otherwise.

(a) What is $\mathbb{E}[Y_1]$?

(b) What is $\mathbb{E}[Y_1Y_2]$?

(c) What is $\mathbb{E}[Y_1^2 + 3Y_2Y_1 - 5Y_2^5]$? $\mathbb{E}[Y_1^2 - Y_1Y_2]$.

Solution: When working with joint distributions of random variables, stating by drawing the support of the random variables is usually a good idea:



You should then imagine a spike coming out of each of these points with height given by the pmf. If you want to know the probability that (Y_1,Y_2) take values in a set, you add up the heights of the all the spikes in that set.

(a) Now we compute compute $\mathbb{E}[Y_1]$.

$$\mathbb{E}[Y_1] = \sum_{y_1=-1,0,1} \sum_{y_2=-1,0,1} y_1 p_{Y_1,Y_2}(y_1,y_2)$$

The sums are over $-1,0,1$ because those are all the possible values Y_1 and Y_2 can take. Now we expand and substitute in the probabilities.

$$\begin{aligned}\mathbb{E}[Y_1] &= -1(p_{Y_1,Y_2}(-1,-1)+p_{Y_1,Y_2}(-1,0)+p_{Y_1,Y_2}(-1,1)) \\ &\quad +0(p_{Y_1,Y_2}(0,-1)+p_{Y_1,Y_2}(0,0)+p_{Y_1,Y_2}(0,1)) \\ &\quad +1(p_{Y_1,Y_2}(1,-1)+p_{Y_1,Y_2}(1,0)+p_{Y_1,Y_2}(1,1)) \\ &= -1(0+1/10+0) \\ &\quad +0(3/10+3/20+1/5) \\ &\quad +1(0+1/4+0) \\ &\quad = -1/10+1/4\end{aligned}$$

You could have also just written this as the sum

$$\mathbb{E}[Y_1] = -1(1/10)+0(3/10+3/20+1/5)+1(1/4) = -1/10+1/4$$

or you could have first computed the marginal pmf of Y_1 and then computed $\mathbb{E}[Y_1]$ from the marginal pmf just as you did in Chapter 2.

(b) $\mathbb{E}[Y_1Y_2]=0$ because at least one of Y_1 and Y_2 is always zero, so every term in the sum is zero.

(c) In the next section, we'll see theorems on how to simplify the calculation of a term like this, but for now we'll just directly apply the formula.

$$\mathbb{E}[Y_1^2+3Y_2Y_1-5Y_2^5]=\left((-1)^2+0-5(0)^5\right)1/10+\left((1)^2+0-5(0)^5\right)1/4+\left((0)^2+0-5(1)^5\right)1/5$$
$$+\left((0)^2+0-5(-1)^5\right)3/10+\left((0)^2+0-5(0)^5\right)3/20$$
$$=1/10+1/4-5/5+15/10+0$$

where the term inside the first parenthesis is the evaluation of $y_1^2+3y_2y_1-5y_2^5$ at each of the points the pair of random variables can take values, and the term after the parenthesis is the probability Y_1,Y_2 takes that particular value.

EXAMPLE 1.4. *You have recently become infected with a contagious disease, but are asymptomatic, so you don't know it. Each day you meet Y people where Y is a Poisson random variable with mean λ. The number of people you infect is X, where conditional on the event $Y=y$, X is uniformly distributed on the integers* $0,1,...y$.

(a) *What is $\mathbb{E}[X]$, the expected number of people you infect?*

(b) *What is $\mathbb{E}[XY]$?.*

Solution: We first need to determine the joint pmf of X,Y. We know the marginal distribution of Y (it's Poisson) and the conditional distribution of X given Y (it's uniform on the $y+1$ integers between 0 and y) so we have:

$$p_{X,Y}(x,y)=p_{X|Y}(x|y)p_Y(y)=\frac{1}{y+1}e^{-\lambda}\frac{\lambda^y}{y!} \text{ for } y=0,1,2,..., x=0,1,...,y.$$

(a) Now we compute $\mathbb{E}[X]$

$$\mathbb{E}[X]=\sum_{y=0}^{\infty}\sum_{x=0}^{y}x\frac{1}{y+1}e^{-\lambda}\frac{\lambda^y}{y!}$$

we begin with the sum $\sum_{x=0}^{y}x$ (the other terms don't depend on x and can be ignored for now)

$$\sum_{x=0}^{y}x=\frac{(y+1)(y)}{2}$$

so

$$\mathbb{E}[X]=\sum_{y=0}^{\infty}\frac{1}{y+1}e^{-\lambda}\frac{\lambda^y}{y!}\frac{(y+1)(y)}{2}=\frac{1}{2}\sum_{y=1}^{\infty}e^{-\lambda}\frac{\lambda^y}{(y-1)!}=\frac{\lambda}{2}\sum_{y=1}^{\infty}e^{-\lambda}\frac{\lambda^{y-1}}{(y-1)!}$$

but $\sum_{y=1}^{\infty}e^{-\lambda}\frac{\lambda^{y-1}}{(y-1)!}=e^{-\lambda}\sum_{z=0}^{\infty}\frac{\lambda^z}{z!}=1$, where the last equality uses the Taylor expansion of e^x. So we have

$$\mathbb{E}[X]=\lambda/2$$

(b) To compute $\mathbb{E}[XY]$ we follow a similar procedure

$$\mathbb{E}[XY]=\sum_{y=0}^{\infty}\sum_{x=0}^{y}xy\frac{1}{y+1}e^{-\lambda}\frac{\lambda^y}{y!}$$

Once again we begin with $\sum_{x=0}^{y} x = \frac{(y+1)(y)}{2}$, to get:

$$\mathbb{E}[XY] = \sum_{y=0}^{\infty} y^2 \frac{1}{2} e^{-\lambda} \frac{\lambda^y}{y!} = \frac{\lambda^2+\lambda}{2}$$

where the last equality uses identities and/or techniques from the section on Poisson rv (or just a computer or whatever, the important of this problem is understanding the first line).

EXAMPLE 1.5. *Let Y_1 and Y_2 be continuous random variables with jpdf:*

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

(a) Compute $\mathbb{E}[Y_1]$.

(b) Compute $\mathbb{E}[Y_1^2 - Y_1 Y_2]$.

Solution: Normally we would begin by sketching the support of the joint pdf, but we've seen this example, so remind yourself of the triangle the jpdf is positive on.

(a) Since we have previously computed the marginal pdf of Y_1, we could use that to compute $\mathbb{E}[Y_1]$, but let's try with the new formula

$$\mathbb{E}[Y_1] = \int_0^2 \int_0^{y_2} y_1 \left(\frac{3}{4} y_1^2\right) dy_1 dy_2$$

The term in the parenthesis is the jpdf, and the bounds of integration are where the jpdf is positive. Now we can proceed just as in Multivariable Calc.

$$\begin{aligned} \mathbb{E}[Y_1] &= \int_0^2 \int_0^{y_2} y_1 \left(\frac{3}{4} y_1^2\right) \, dy_1 dy_2 \\ &= \frac{3}{4} \int_0^2 \left(\frac{y_1^4}{4}\right)_0^{y_2} \, dy_2 \\ &= \frac{3}{4} \int_0^2 \frac{y_2^4}{4} \, dy_2 \\ &= \frac{3}{4} \left(\frac{y_2^5}{5*4}\right)_0^2 \, dy_2 \\ &= \frac{3}{4} \frac{2^5}{5*4} \\ &= \frac{6}{5} \end{aligned}$$

(b) The computation of $\mathbb{E}[Y_1^2 - Y_1 Y_2]$ is similar.

$$
\begin{aligned}
\mathbb{E}[Y_1^2 - Y_1 Y_2] &= \int_0^2 \int_0^{y_2} (y_1^2 - y_1 y_2)\left(\frac{3}{4} y_1^2\right)\, dy_1 dy_2 \\
&= \frac{3}{4} \int_0^2 \int_0^{y_2} \left(y_1^4 - y_1^3 y_2\right)\, dy_1 dy_2 \\
&= \frac{3}{4} \int_0^2 \left(\frac{y_1^5}{5} - \frac{y_1^4}{4} y_2\right)_0^{y_2}\, dy_2 \\
&= \frac{3}{4} \int_0^2 \frac{y_2^5}{5} - \frac{y_2^5}{4} dy_2 \\
&= -\frac{3}{4} \frac{1}{20} \left(\frac{y_2^6}{6}\right)_0^2 \\
&= -\frac{3}{4} \frac{1}{20} \frac{2^6}{6} \\
&= -\frac{2}{5}
\end{aligned}
$$

There is nothing particularly special about having 2 random variables instead of n random variables, for any integer bigger than 1, other than the bounds of integration begin a bit easier to visualize. The formula for the expectation of a function of n random variables is:

Definition 1.6 (Expectation of n random variables- Continuous Case). *Let $X_1, X_2, ..., X_n$ be continuous random variables with joint pdf $f_{X_1, X_2, ..., X_n}(x_1, x_2, ..., x_n)$, and let $g : \mathbb{R}^n \to \mathbb{R}$, then the* ***expected value*** *of $g(X_1, X_2, ..., X_n)$ is*

$$\mathbb{E}[g(X_1, X_2, ..., X_n)] = \int ... \iint g(x_1, x_2, ..., x_n) f_{X_1, X_2, ..., X_n}(x_1, x_2, ..., x_n)\ dx_1 dx_2 ... dx_n.$$

Definition 1.7 (Expectation of n random variables- Discrete Case). *Let $Y_1, Y_2, ..., Y_n$ be discrete random variables with joint pmf $p_{Y_1, Y_2, ..., Y_n}(y_1, y_2, ..., y_n)$, and let $g : \mathbb{R}^n \to \mathbb{R}$, then the* ***expected value*** *of $g(Y_1, Y_2, ..., Y_n)$ is*

$$\mathbb{E}[g(Y_1, Y_2, ..., Y_n)] = \sum_{y_1, y_2, ..., y_n} g(y_1, y_2, ..., y_n) p_{Y_1, Y_2, ..., Y_n}(y_1, y_2, ..., y_n).$$

2 Special Theorems

Now we'll look at some theorems concerning the Expectation of functions of several random variables. These can give us shortcuts to computing the above expectations. I'll state all theorems in the case of 2 random variables, but they generalize in a straight-forward way to any number of random variables.

The first theorem concerns expectations of scalar multiples of random variables. This should look familiar from the one variable case.

Theorem 2.1. *Let Y_1, Y_2 be random variables, let $g : \mathbb{R}^2 \to \mathbb{R}$ be a function and c be a real number, then*

$$\mathbb{E}[c\ g(Y_1, Y_2)] = c\ \mathbb{E}[g(Y_1, Y_2)]$$

.

In other words, we can pull constants outside of expectations. This theorem, is fairly short to verify and I recommend thinking through the proof the first few times you use it

Proof. I'll consider the discrete case. You can do the continuous case.

$$\mathbb{E}[c\ g(Y_1,Y_2)] = \sum_{y_1,y_2} c\ g(y_1,y_2) = c \sum_{y_1,y_2} g(y_1,y_2) = c\ \mathbb{E}[g(Y_1,Y_2)]$$

as desired. □

The next theorem concerns sums of random variables.

THEOREM 2.2. *Let Y_1,Y_2 be random variables, let $g_1,g_2,\ldots,g_n : \mathbb{R}^2 \to \mathbb{R}$ be functions , then*

$$\mathbb{E}[g_1(Y_1,Y_2)+g_2(Y_1,Y_2)+\ldots+g_n(Y_1,Y_2)] = \mathbb{E}[g_1(Y_1,Y_2)]+\mathbb{E}[g_2(Y_1,Y_2)]+\ldots+\mathbb{E}[g_n(Y_1,Y_2)].$$

In other words, we can pull sums outside of expectations. This theorem, is also fairly short to verify and I recommend thinking through the proof the first few times you use it

Proof. I'll consider the continuous case. You can do the discrete case.

$$\begin{aligned}
\mathbb{E}[g_1(Y_1,Y_2)+g_2(Y_1,Y_2)+\ldots+g_n(Y_1,Y_2)] &= \iint (g_1(y_1,y_2)+g_2(y_1,y_2)+\ldots+g_n(y_1,y_2)) f_{Y_1,Y_2}(y_1,y_2)\ dy_1 dy_2 \\
&= \iint g_1(y_1,y_2) f_{Y_1,Y_2}(y_1,y_2)\ dy_1 dy_2 \\
&+ \iint g_2(y_1,y_2) f_{Y_1,Y_2}(y_1,y_2)\ dy_1 dy_2 + \ldots \\
&+ \iint g_n(y_1,y_2)) f_{Y_1,Y_2}(y_1,y_2)\ dy_1 dy_2 \\
&= \mathbb{E}[g_1(Y_1,Y_2)]+\mathbb{E}[g_2(Y_1,Y_2)]+\ldots+\mathbb{E}[g_n(Y_1,Y_2)]
\end{aligned}$$

as desired. □

These theorems tell us, for example, that earlier when we computed $\mathbb{E}[Y_1^2 + 3Y_2Y_1 - 5Y_2^5]$ or $\mathbb{E}[Y_1^2 - Y_1Y_2]$, we could have instead computed $\mathbb{E}[Y_1^2] + 3\mathbb{E}[Y_2Y_1] - 5\mathbb{E}[Y_2^5]$ or $\mathbb{E}[Y_1^2] - \mathbb{E}[Y_1Y_2]$. This amounts breaking the problem into smaller parts, that can be a bit easier to approach.

Note that we made no assumptions on the functions $g_1,g_2,\ldots,g_n$ above, sums can allows be broken up. One often useful case will be $g_1(X_1,X_2) = X_1$ and $g_2(X_1,X_2) = X_2$, and its higher dimensional variants. In case we have

$$E[X_1+X_2] = \mathbb{E}[X_1]+\mathbb{E}[X_2].$$

We will look more at this special case in future sections. It may be tempting to apply a similar for the variance replacing the expectation, so I will warn you now, in general, $\text{Var}(X+Y) \neq \text{Var}(X)+\text{Var}(Y)$.

The final special theorem is for the special case that X_1 and X_2 are independent random variable, and we are applying a function that factorizes, meaning $g(x_1,x_2) = g_1(x_1)g_2(x_2)$, for some functions $g_1,g_2 : \mathbb{R} \to \mathbb{R}$.

THEOREM 2.3. *Let Y_1,Y_2 be **independent** random variables. Then*

$$\mathbb{E}[g_1(Y_1)g_2(Y_2)] = \mathbb{E}[g_1(Y_1)]\mathbb{E}[g_2(Y_2)].$$

Proof. I will write the discrete case, you can think about the continuous case. The main point is if Y_1,Y_2 are independent, then $p_{Y_1,Y_2}(y_1,y_2) = p_{Y_1}(y_1)p_{Y_2}(y_2)$

$$\begin{aligned}
\mathbb{E}[g_1(Y_1)g_2(Y_2)] &= \sum_{y_1,y_2} g_1(y_1)g_2(y_2)p_{Y_1,Y_2}(y_1,y_2) \\
&= \sum_{y_1,y_2} g_1(y_1)g_2(y_2)p_{Y_1}(y_1)p_{Y_2}(y_2) \\
&= \sum_{y_1} g_1(y_1)p_{Y_1}(y_1)\sum_{y_2} g_2(y_2)p_{Y_2}(y_2) \\
&= \mathbb{E}[g_1(Y_1)]\mathbb{E}[g_2(Y_2)]
\end{aligned}$$

□

This theorem is only true for independent random variables, if they are not independent then you simply use the formula at the beginning of these notes:

$$\mathbb{E}[g_1(Y_1)g_2(Y_2)] = \iint g_1(y_1)g_2(y_2)f_{Y_1,Y_2}(y_1,y_2)\ dy_1dy_2 \ \textit{or} \sum_{y_1,y_2} g_1(y_1)g_2(y_2)p_{Y_1,Y_2}(y_1,y_2)$$

So you can break up expectations when doing sums, but not in general with products, unless the random variables are independent.