

Conditional random variables

Discrete random variables

Suppose that X and Y have a joint discrete distribution. If $\mathbb{P}(Y = y) \neq 0$ then we can define the conditional probability

$$\mathbb{P}(X = x \mid Y = y) = \frac{\mathbb{P}(X = x \text{ and } Y = y)}{\mathbb{P}(Y = y)}.$$

Definition The conditional random variable $X \mid Y = y$, called “ X given $Y = y$ ”, has probability mass function

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

for values x of X .

In particular, we obtain the *conditional expectation*

$$\mathbb{E}(X \mid Y = y) = \sum_x x \mathbb{P}(X = x \mid Y = y).$$

Example Let $Y_1, Y_2, \dots$, be an infinite sequence of independent Bernoulli(p) random variables, where $0 < p < 1$. Put

$$X = i \quad \text{if } Y_i = 1 \text{ but } Y_j = 0 \text{ whenever } j < i,$$

so that $X \sim \text{Geom}(p)$. Then

$$\begin{aligned} X \mid Y_1 = 1 & \quad \text{is the constant 1;} \\ X \mid Y_1 = 0 & \quad \text{is } 1 + Z \text{ where } Z \sim \text{Geom}(p). \end{aligned}$$

We get a conditional random variable $X | Y = y$, and its conditional expectation $\mathbb{E}(X | Y = y)$ for each value y of Y . Put $g(y) = \mathbb{E}(X | Y = y)$ and $h(y) = \text{Var}(X | Y = y)$. Then $g(Y)$ is the random variable $\mathbb{E}(X | Y)$ and $h(Y)$ is the random variable $\text{Var}(X | Y)$: the randomness in these last two is in Y , not in X .

Theorem 20 $\mathbb{E}[\mathbb{E}(X | Y)] = \mathbb{E}(X)$.

Proof

$$\begin{aligned}
 \mathbb{E}[\mathbb{E}(X | Y)] &= \sum_y \mathbb{E}(X | Y = y) \mathbb{P}(Y = y) \\
 &= \sum_y \sum_x x \mathbb{P}(X = x | Y = y) \mathbb{P}(Y = y) \\
 &= \sum_x x \sum_y \mathbb{P}(X = x | Y = y) \mathbb{P}(Y = y) \\
 &= \sum_x x \mathbb{P}(X = x), \quad \text{by the Theorem of Total Probability,} \\
 &= \mathbb{E}(X). \quad \blacksquare
 \end{aligned}$$

Example In our example, $\mathbb{P}(Y_1 = 1) = p$ and $\mathbb{P}(Y_1 = 0) = q = 1 - p$. Also,

$$\mathbb{E}(X | Y_1 = 1) = \mathbb{E}(1) = 1$$

and

$$\begin{aligned}
 \mathbb{E}(X | Y_1 = 0) &= \mathbb{E}(1 + Z), \quad \text{where } Z \sim \text{Geom}(p), \\
 &= 1 + \mathbb{E}(Z) \\
 &= 1 + \mathbb{E}(X).
 \end{aligned}$$

Therefore

$$\begin{aligned}
 \mathbb{E}(X) &= \mathbb{E}(X | Y_1 = 1) \mathbb{P}(Y_1 = 1) + \mathbb{E}(X | Y_1 = 0) \mathbb{P}(Y_1 = 0) \\
 &= 1 \cdot p + (1 + \mathbb{E}(X)) \cdot q \\
 &= p + q + q \mathbb{E}(X)
 \end{aligned}$$

so

$$(1 - q) \mathbb{E}(X) = p + q = 1$$

and so

$$\mathbb{E}(X) = \frac{1}{1 - q} = \frac{1}{p}.$$

Thus conditional random variables give us an alternative way of finding the expectation of a geometric random variable.

Theorem 21 $\text{Var}(X) = \mathbb{E}[\text{Var}(X | Y)] + \text{Var}[\mathbb{E}(X | Y)]$.

Proof Wait until next year! $\blacksquare$

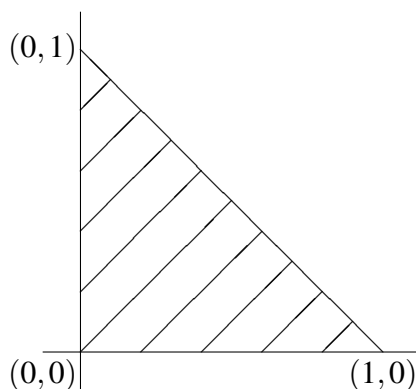
Continuous random variables

If X and Y have a joint continuous distribution and $f_Y(y) \neq 0$ then the conditional random variable $X | Y = y$ is defined to have probability density function

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

Theorems 20 and 21 still hold.

Example Consider the triangle in the plane $\mathbb{R}^2$ whose vertices are at the points $(0,0)$, $(0,1)$ and $(1,0)$.



Let (X,Y) be the random variable whose value (x,y) is equally likely to be anywhere in this triangle. Then $f_{X,Y}(x,y)$ is equal to some constant c if (x,y) is in the triangle, and to zero otherwise. Therefore

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x,y) dx dy = c \times (\text{area of this triangle}) = \frac{c}{2}.$$

But we know that

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x,y) dx dy = 1,$$

so $c = 2$ and

$$f_{X,Y}(x,y) = \begin{cases} 2 & \text{if } 0 \leq x \leq 1, 0 \leq y \leq 1 \text{ and } x+y \leq 1, \\ 0 & \text{otherwise.} \end{cases}$$

Hence

$$f_X(x) = \int_0^{1-x} 2 dy = \begin{cases} 2(1-x) & \text{if } 0 \leq x \leq 1 \\ 0 & \text{otherwise.} \end{cases}$$

Similarly,

$$f_Y(y) = \begin{cases} 2(1-y) & \text{if } 0 \leq y \leq 1 \\ 0 & \text{otherwise.} \end{cases}$$

If $0 \leq y \leq 1$ then,

$$f_{X|Y=y}(x) = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \begin{cases} \frac{2}{2(1-y)} = \frac{1}{1-y} & \text{if } 0 \leq x \leq 1-y \\ 0 & \text{otherwise} \end{cases}$$

and so $X | Y = y \sim U(0, 1-y)$. Therefore

$$\mathbb{E}(X | Y = y) = \frac{1-y}{2}.$$

Theorem 20 gives

$$\mathbb{E}(X) = \mathbb{E}\left(\frac{1-Y}{2}\right). \quad (1)$$

Hence

$$\begin{aligned} \mathbb{E}(X) &= \int_0^1 \left(\frac{1-y}{2}\right) f_Y(y) \, dy \\ &= \int_0^1 \left(\frac{1-y}{2}\right) 2(1-y) \, dy \\ &= \int_0^1 (1-y)^2 \, dy \\ &= \left[\frac{(1-y)^3}{-3} \right]_0^1 = \frac{1}{3}. \end{aligned}$$

Working directly from the probability density function for X gives

$$\begin{aligned} \mathbb{E}(X) &= \int_0^1 x f_X(x) \, dx \\ &= \int_0^1 x 2(1-x) \, dx \\ &= \int_0^1 (2x - 2x^2) \, dx \\ &= \left[x^2 - \frac{2x^3}{3} \right]_0^1 = \frac{1}{3}. \end{aligned}$$

Alternatively, X and Y have the same distribution, so $\mathbb{E}(X) = \mathbb{E}(Y)$. Substituting this into Equation (1) gives

$$2 \mathbb{E}(X) = 1 - \mathbb{E}(X)$$

and so $\mathbb{E}(X) = 1/3$.