Introduction to Statistics (MAT 283)

Dipti Dubey

Department of Mathematics
Shiv Nadar University

Feburary 5, 2021

Table of Contents

EXAMPLE: A fair coin is tossed two times; let X denote the number of heads on the first toss and Y the total number of heads. Sample space of this random experiment is

$$\Omega = \{HH, HT, TH, TT\}.$$

The joint PMF f of X and Y is as given in the following table:

	Y=0	Y=1	Y=2	P(X=x)
X= 0	$\frac{1}{4}$	$\frac{1}{4}$	0	$\frac{2}{4}$
X= 1	0	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{2}{4}$
P(Y=y)	$\frac{1}{4}$	$\frac{2}{4}$	$\frac{1}{4}$	

Then find the conditional mean of X when $Y = 0$.

$$E(X|Y=0) = \sum_{x \in R_X} x \, g(x|y=0)$$

$$g(x \mid y=0) = \frac{f(x,0)}{f_y(0)}$$

$$g(x=0 \mid y=0) = \frac{f(0,0)}{f_y(0)} = \frac{1/4}{1/4} = 1$$

$$g(x=1 \mid y=0) = \frac{f(1,0)}{f_y(0)} = 0$$

$$E(X \mid y=0) = 0 \times 1 + 1 \times 0 = 0$$

EXAMPLE: Let the random variables X and Y have the joint pdf

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

What is the conditional mean $E(Y|X = \frac{1}{3})$?

$$h(y \mid x = \tfrac{1}{3}) = \frac{f(1/3, y)}{f_X(1/3)}$$

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

$$= \int_0^1 (x + y)\, dy$$

$$= \left[xy + \frac{y^2}{2} \right]_0^1 = x + \frac{1}{2}$$

$$h(y \mid x) = \frac{f(x, y)}{f_X(x)} = \frac{x + y}{x + 1/2}$$

$$E\left(Y \mid X = \frac{1}{3}\right) = \int_0^1 y\, h(h \mid x = 1/3)\, dy$$

$$= \int_0^1 y \, \frac{1/3 + y}{1/3 + 1/2}\, dy$$

$$= \int_0^1 y \, \frac{1/3 + y}{5/6}\, dy$$

$$= \frac{6}{5} \int_0^1 \left(\frac{1}{3} y + y^2\right) dy$$

$$= \frac{6}{5} \times \frac{3}{6} =$$

Let X and Y be random variables defined on a probability space $(\Omega, \mathcal{F}, P)$, let ψ be a Borel-measurable function. Then the conditional expectation of $\psi(X)$, given Y, written as $E(\psi(X)|Y)$, is a random variable that takes the value $E(\psi(X)|y)$, defined as

$$E(\psi(X)|y) = \begin{cases} \sum_{x \in R_X} \psi(x)\, g(x|y) & \text{if X is discrete} \\ \\ \int_{-\infty}^{\infty} \psi(x)\, g(x|y) dx & \text{if X is continuous,} \end{cases}$$

a similar definition may be given for the conditional expectation $E(\psi(Y)|X)$.

The conditional mean of Y given $X = x$, that is $E(Y|x)$, is a function $\phi(x)$ of the variable x.

The conditional mean of Y given $X = x$, that is $E(Y|x)$, is a function $\phi(x)$ of the variable x.

Define

$$\phi(X) = E(Y|X),$$

this function $\phi(X)$ is a random variable.

Theorem. The expected value of the random variable $E(Y|X)$ is equal to the expected value of Y , that is

$$E_X(E(Y|X)) = E_Y(Y),$$

where $E_x(X)$ stands for the expectation of X with respect to the distribution of X and $E_Y(Y)$ stands for the expectation of Y with respect to the distribution of Y.

Proof. We prove this theorem for continuous case.

$$\begin{aligned} E_X(E(Y|X)) &= \int_{-\infty}^{\infty} E(Y|x) f_X(x) dx \\ &= \int_{-\infty}^{\infty} \Big(\int_{-\infty}^{\infty} y \ h(y|x) dy \Big) f_X(x) dx \end{aligned}$$

$$
\begin{aligned}
&= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} y \ h(y|x) f_X(x) dy \right) dx \\
&= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} \underline{h(y|x) f_X(x)} dx \right) y dy \\
&= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} f(x, y) dx \right) y \ dy \\
&= \int_{-\infty}^{\infty} y f_Y(y) dy \\
&= E_Y(Y).
\end{aligned}
$$

$\left[\because h(y|x) = \dfrac{f(x,y)}{f_X(x)} \right.$

EXAMPLE: An insect lays Y number of eggs, where Y has a Poisson distribution with parameter λ. If the probability of each egg surviving is p, then on the average how many eggs will survive?

Solution: Let X denote the number of surviving eggs.

EXAMPLE: An insect lays Y number of eggs, where Y has a Poisson distribution with parameter λ. If the probability of each egg surviving is p, then on the average how many eggs will survive?

Solution: Let X denote the number of surviving eggs. Then, given that Y = y (that is given that the insect has laid y eggs) the random variable X has a binomial distribution with parameters y and p. Thus

$$X|Y \sim BIN(Y, p)$$

and

$$Y \sim POI(\lambda)$$

Therefore, the expected number of survivors is given by

$$\begin{aligned} E_X(X) &= E_Y(E(X|Y)) \\ &= E_Y(pY) \quad (\text{since } X|Y \sim BIN(Y, p)) \\ &= pE_Y(Y) \\ &= p\lambda \quad (\text{since } Y \sim POI(\lambda)). \end{aligned}$$

Theorem: Let X and Y be two random variables with mean μ_X and μ_Y , and standard deviation σ_X and σ_Y , respectively. If the conditional expectation of Y given $X = x$ is linear in x, then

$$E(Y|X = x) = \mu_Y + \rho \frac{\sigma_X}{\sigma_Y}(x - \mu_X)$$

where ρ denotes the correlation coefficient of X and Y.

Proof: Probability and Mathematical Statistics by Sahoo, Page No. 243-244.

Table of Contents

Theorem: Let X and Y be two random variables with mean μ_X and μ_Y , and standard deviation σ_X and σ_Y , respectively. If the conditional expectation of Y given $X = x$ is linear in x, then

$$E(Y|X = x) = \mu_Y + \rho \frac{\sigma_X}{\sigma_Y}(x - \mu_X)$$

where ρ denotes the correlation coefficient of X and Y.

Proof: Probability and Mathematical Statistics by Sahoo, Page No. 243-244.

CONDITIONAL VARIANCE :

Let X and Y be two random variables with joint density f and $h(y|x)$ be the conditional density of Y given $X = x$. The conditional variance of Y given $X = x$, denoted by $Var(Y|x)$, is defined as

$$Var(Y|x) = E(Y^2|x) - (E(Y|x))^2$$

where $E(Y|x)$ denotes the conditional mean of Y given $X = x$.

EXAMPLE: A fair coin is tossed two times; let X denote the number of heads on the first toss and Y the total number of heads. Sample space of this random experiment is

$$\Omega = \{HH, HT, TH, TT\}.$$

The joint PMF f of X and Y is as given in the following table:

	Y=0	Y=1	Y=2	P(X=x)
X= 0	$\frac{1}{4}$	$\frac{1}{4}$	0	$\frac{2}{4}$
X= 1	0	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{2}{4}$
P(Y=y)	$\frac{1}{4}$	$\frac{2}{4}$	$\frac{1}{4}$	

Find the conditional variance of Y given $x = 1$.

EXAMPLE: If the joint probability density of X and Y is given by

$$f(x,y)=\begin{cases}\frac{2}{3}(x+2y) & \text{for} \quad 0<x<1;\ 0<y<1\\ \\ 0 & \text{otherwise,}\end{cases}$$

find the conditional mean and the conditional variance of X given $Y=\frac{1}{2}$.