LECTURE - 32

Agenda:

1. Conditional expectation of random variables
2. Properties involving conditional expectations

CONDITIONAL EXPECTATION OF RANDOM VARIABLES

Let us recollect that for two random variables X and Y, the conditional distribution of X given $Y = y$, describes the probability behaviour of the random variable X given the information that $Y = y$. In the same spirit, we defined the notion of the conditional expectation of X given $Y = y$.

Definition: If X and Y are two continuous random variables with joint probability density $f_{X,Y}$, then the conditional expectation of X given $Y = y$ is defined as

$$E(X \mid Y = y) = \int_{-\infty}^{\infty} x \underbrace{f_{X \mid Y = y}(x)}_{\text{conditional probability density of } X \text{ given } Y = y.} \, dx .$$

Definition: If X and Y are two discrete random variables with joint probability mass function $p_{X,Y}$, then the conditional expectation of X given $\cancel{X} = \cancel{y}$ is defined as

$$E(X|Y=y) = \sum_{x\in\mathcal{X}} x\, \underbrace{p_{X|Y=y}(x)}_{\text{Conditional probability mass function of } X \text{ given } Y=y}$$

Example: A soft-drink machine has a random supply Y at the beginning of a given day and dispenses a random amount X during the day (with measurements in gallons). It has been observed that X and Y have joint density

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

If it is known that on a particular day 1 gallon was supplied at the beginning of the day, find the expected value of the amount of soft-drink consumed in that day.

Note that we need to find $E(X|Y=1)$.

Let us first find the marginal density of Y.

$$f_Y(y) = \begin{cases} \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dx & \text{if } 0 \le y \le 2, \\ 0 & \text{otherwise.} \end{cases}$$

$$= \begin{cases} \int_{0}^{y} \frac{1}{2}\,dx & \text{if } 0 \le y \le 2, \\ 0 & \text{otherwise.} \end{cases}$$

$$= \begin{cases} \frac{y}{2} & \text{if } 0 \le y \le 2, \\ 0 & \text{otherwise.} \end{cases}$$

Hence, the conditional density of X given $Y = y$ is given by

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

$$= \begin{cases} \frac{1/2}{y/2} & \text{if } 0 \le x \le y, \\ 0 & \text{otherwise.} \end{cases}$$

$$= \begin{cases} \frac{1}{y} & \text{if } 0 \le x \le y. \\ 0 & \text{otherwise.} \end{cases}$$

Hence,

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

$$= \int_{0}^{1} x \cdot 1 \; dx$$

$$= \frac{1}{2}.$$

Hence, the expected amount of the amount of soft drink consumed in that day is $\frac{1}{2}$ gallon.

PROPERTIES INVOLVING CONDITIONAL EXPECTATIONS

Result: Let X and Y ~~[illegible]~~ be two random variables. Then,

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

where on the right-hand side, the inside expectation is stated with respect to the conditional distribution of X given Y, and the outside expectation is stated with respect to the distribution of Y.

CONFUSED ??

Sometimes, the way the distribution of the random variable X is specified, it is not possible to directly evaluate its expectation, and the previous identity comes in very handy.

HERE IS AN ALTERNATE WAY OF EVALUATING $E(X)$ BASED ON THE PREVIOUS IDENTITY.

① For every y, find $E(X|Y=y)$.

② Define the function h as $h(y) \triangleq E(X|Y=y)$.

③ $E(X) = E(h(Y))$.

Hence, the way to interpret the previous relationship is as follows.

$$E(X) = E(h(Y)), \quad \text{where } h(y) = E(X|Y=y).$$

Example: A quality control plan for an assembly line involves sampling n finished items per day and counting X, the number of defective items. If p denotes the probability that an item is defective, then given p, X has a binomial distribution with parameters n and p. However, it is observed that p is a random quantity, and has a uniform distribution on the interval $[0, \frac{1}{4}]$.

(a) Find $E(X)$.

Note that given $p = p_0$, X has a binomial distribution with parameters n and p_0. Hence,

$$E(X \mid p = p_0) = n p_0 .$$

Hence, $E(X) = E(E(X \mid p))$
$= E(h(p))$, where $h(p_0) = E(X \mid p = p_0)$.

Hence,
$$\begin{aligned} E(X) &= E(np) \\ &= n\,E(p) \\ &= n\,\frac{\left(0 + \frac{1}{4}\right)}{2} \end{aligned}$$

($\because p$ is uniform on the interval $[0, \frac{1}{4}]$)

$$= \frac{n}{8} .$$

(b) Find $SD(X)$.

Result: Let X and Y be two random variables. Then,
$$V(X) = E(V(X \mid Y)) + V(E(X \mid Y)).$$

HERE IS AN ALGORITHM BASED ON THIS IDENTITY.

TO EVALUATE $V(X)$.

(1) For every y, find $E(X \mid Y=y)$ and $V(X \mid Y=y)$

(2) Define the function h as $h(y) \triangleq E(X \mid Y=y)$.

(3) Define the function $\tilde{h}$ as $\tilde{h}(y) \triangleq V(X \mid Y=y)$.

(4) $V(X) = V(h(Y)) + E(\tilde{h}(Y))$.

(Note that for any function g, $E(g(X) \mid Y=y)$ is $\int_{-\infty}^{\infty} g(x) f_{X \mid Y=y}(x)\, dx$ if X is continuous, and

$\sum_{x \in \mathcal{X}} g(x)\, p_{X \mid Y=y}(x)$ if X is discrete.)

Hence, the way to interpret the previous relationship is as follows.

$V(X) = V(h(Y)) + E(\tilde{h}(Y))$, where $h(y) = E(X \mid Y=y)$ and $\tilde{h}(y) = V(X \mid Y=y)$.

Returning to the example,

$$
\begin{aligned}
V(X) &= V(E(X \mid p)) + E(V(X \mid p)) \\
&= V(np) + E[np(1-p)] \\
&= n^2 V(p) + nE(p) - nE(p^2)
\end{aligned}
$$

$$= (n^2 - n) V(p) + n E(p) - n(E(p))^2$$

$$(\because \; E(p^2) = V(p) + (E(p))^2)$$

$$= (n^2 - n) . \frac{\left(\frac{1}{4}\right)^2}{12} + \frac{n}{8} - \frac{n}{64}$$

$$= \frac{n^2}{192} + \frac{5n}{48}$$

Hence, $SD(x) = \sqrt{V(x)} = \sqrt{\frac{n^2}{192} + \frac{5n}{48}}$.