CSL003P1M : Probability and Statistics Lecture 22 (Conditional Expectation and Covariance)

Sumit Kumar Pandey

October 27, 2021

Conditional Expectation

Conditional Expectation

Recall that if X and Y are jointly discrete random variables, then the conditional probability mass function of X, given that $Y = y$, is defined, for all y such that $P\{Y = y\} > 0$, by

$$p_{X|Y}(x|y) = P\{X = x|Y = y\} = \frac{p(x, y)}{p_Y(y)}$$

It is therefore natural to define, in this case, the conditional expectation of X given that $Y = y$, for all values of y such that $p_Y(y) > 0$, by

$$\begin{aligned} E[X|Y = y] &= \sum_x xP\{X = x|Y = y\} \\ &= \sum_x xp_{X|Y}(x|y) \end{aligned}$$

Exercise-1

If X and Y are independent binomial random variables with identical parameter n and p, calculate the conditional expected value of X given that $X + Y = m$.

Solution:

- We first calculate the conditional probability mass function of X given that $X + Y = m$. For $k \leq \min(n, m)$,

$$
\begin{aligned}
P\{X = k | X + Y = m\} &= \frac{P\{X = k, X + Y = m\}}{P\{X + Y = m\}} \\
&= \frac{P\{X = k, Y = m - k\}}{P\{X + Y = m\}} \\
&= \frac{P\{X = k\}P\{Y = m - k\}}{P\{X + Y = m\}}
\end{aligned}
$$

Exercise-1

$$\frac{P\{X=k\}P\{Y=m-k\}}{P\{X+Y=m\}} =$$

$$\frac{\binom{n}{k}p^k(1-p)^{n-k}\binom{n}{m-k}p^{m-k}(1-p)^{n-m+k}}{\binom{2n}{m}p^m(1-p)^{2n-m}}$$

$$= \frac{\binom{n}{k}\binom{n}{m-k}}{\binom{2n}{m}}$$

which is a hypergeometric distribution with parameters $(m, 2n, n)$. So,

$$E[X|X+Y=m] = \frac{mn}{2n} = \frac{m}{2}.$$

Computing Expectations by Conditioning

- Let us denote by $E[X|Y]$ that function of the random variable Y whose value at $Y = y$ is $E[X|Y = y]$.
- Note that $E[X|Y]$ is itself a random variable.

Proposition

$$E[X] = E[E[X|Y]]$$

If Y is a discrete random variable, then above equation states

$$E[X] = \sum_{y} E[X|Y = y]P\{Y = y\}$$

Proof:

- We must show that

$$E[X] = \sum_y E[X|Y = y]P\{Y = y\}.$$

Now,

$$\begin{aligned}
\sum_y E[X|Y = y]P\{Y = y\} &= \sum_y \sum_x xP\{X = x|Y = y\}P\{Y = y\} \\
&= \sum_y \sum_x x\frac{P\{X = x, Y = y\}}{P\{Y = y\}}P\{Y = y\} \\
&= \sum_y \sum_x xP\{X = x, Y = y\} \\
&= \sum_x x\sum_y P\{X = x, Y = y\} \\
&= \sum_x xP\{X = x\} = E[X]
\end{aligned}$$

Exercise-2

A miner is trapped in a **mine** containing 3 doors. The first door leads to a tunnel that will take him to **safety** after under 3 hours of travel. The second door leads to a tunnel that will return him to the **mine** after 5 hours of travel. The third door leads to a tunnel that will return him to the **mine** after 7 hours. If we assume that the miner is at all times equally likely to choose any one of the doors, what is the expected length of time until he reaches **safety**?

Solution:

- Let Xdenote the amount of time (in hours) until the miner reaches safety.
- Let Y denote the door initially he chooses.
- Now,

$$E[X] = \sum_{r=1}^{3} E[X|Y = r]P\{Y = r\}$$

Exercise-2

$$\sum_{r=1}^{3} E[X|Y=r]P\{Y=r\} = \frac{1}{3}\left(\sum_{r=1}^{3} E[X|Y=r]\right)$$

However,

$$\begin{aligned} E[X|Y=1] &= 3 \\ E[X|Y=2] &= 5+E[X] \\ E[X|Y=3] &= 7+E[X] \end{aligned}$$

Hence,

$$E[X] = \frac{1}{3}\left(\sum_{r=1}^{3} E[X|Y=r]\right) = \frac{1}{3}(3+5+E[X]+7+E[X])$$

So,

$$E[X] = 15.$$

Covariance

Covariance

Covariance

The covariance between X and Y, denoted by $Cov(X, Y)$, is defined by

$$Cov(X, Y) = E[(X - E[X])(Y - E[Y])]$$

Covariance

If X and Y are independent random variables, $Cov(X, Y) = 0$.

Solution:

$$
\begin{aligned}
Cov(X, Y) &= E[(X - E[X])(Y - E[Y])] \\
&= E[XY - XE[Y] - YE[X] + E[Y]E[X]] \\
&= E[XY] - E[X]E[Y] - E[Y]E[X] + E[X]E[Y] \\
&= E[XY] - E[X]E[Y] \\
&= E[X]E[Y] - E[X]E[Y] \quad (X \text{ and } Y \text{ are independent}) \\
&= 0.
\end{aligned}
$$

The converse is not true.

$\mathit{Example}$-1

Let X be a random variable such that

$$P\{X = 0\} = P\{X = 1\} = P\{X = -1\} = \frac{1}{3}$$

and defining

$$Y = \begin{cases} 0 & \text{if } X \neq 0 \\ 1 & \text{if } X = 0 \end{cases}$$

Now, $XY = 0$, so $E[XY] = 0$. Also $E[X] = 0$. Thus,

$$Cov(X, Y) = E[XY] - E[X]E[Y] = 0.$$

However, X and Y are clearly not independent.

Covariance

Proposition

1. $Cov(X, Y) = Cov(Y, X)$
2. $Cov(X, X) = Var(X)$
3. $Cov(aX, Y) = aCov(X, Y)$
4. $Cov\left(\sum_{i=1}^{n} X_i, \sum_{j=1}^{n} Y_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} Cov(X_i, Y_j)$

(Exercise!!!)

$$Var\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} Var(X_i) + 2\sum\sum_{i<j} Cov(X_i, X_j)$$

Proof:

- It follows from 2 and 4, upon taking $Y_j = X_j$, $j = 1, \ldots, n$, that

$$\begin{aligned} Var\left(\sum_{i=1}^{n} X_i\right) &= Cov\left(\sum_{i=1}^{n} X_i, \sum_{j=1}^{n} X_j\right) \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} Cov(X_i, X_j) \\ &= \sum_{i=1}^{n} Var(X_i) + \sum\sum_{i\neq j} Cov(X_i, X_j) \end{aligned}$$

Covariance

Since each pair of indices i, j, $i \neq j$, appears twice in the double summation, the preceding formula is equivalent to

$$Var\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} Var(X_i) + 2\sum\sum_{i<j} Cov(X_i, X_j).$$

If $X_1, \ldots, X_n$ are pairwise independent, in that X_i and X_j are independent for $i \neq j$, then the above equation reduces to

$$Var\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} Var(X_i).$$

Thank You