Introductory Econometrics

Lecture 3: Review of Conditional Distribution and Expectation

Instructor: Ma, Jun

Renmin University of China

September 8, 2021

Conditional PMF

- Conditional PMF (when (Y, X) are discrete): If $\Pr[X = x_1] \neq 0$,

$$\begin{aligned} p_j^{Y|X=x_1} &= \Pr\left[Y = y_j \mid X = x_1\right] \\ &= \frac{\Pr\left[Y = y_j, X = x_1\right]}{\Pr[X = x_1]} \\ &= p_{1,j}/p_1^X. \end{aligned}$$

- If independent:

$$\begin{aligned} \Pr[Y = y \mid X = x] &= \frac{\Pr[X = x, Y = y]}{\Pr[X = x]} \\ &= \frac{\Pr[X = x]\Pr[Y = y]}{\Pr[X = x]} \\ &= \Pr[Y = y]. \end{aligned}$$

Conditional PDF

- Conditional PDF (when (Y, X) are continuous): $f_{Y|X=x}(y \mid x) = f_{X,Y}(x, y) / f_X(x)$.
- If X and Y are independent, $f_{Y|X}(y \mid x) = f_Y(y)$ for all x.

Randomness

- Suppose you know that $X = x$. You can update your expectation of Y by conditional expectation. We define conditional expectationfrom conditional PMF and PDF:

$$\begin{aligned} \mathrm{E}[Y \mid X = x] &= \sum_i y_i \Pr[Y = y_i \mid X = x] \text{ (discrete)} \\ \mathrm{E}[Y \mid X = x] &= \int y f_{Y|X}(y \mid x)\, \mathrm{d}y \text{ (continuous).} \end{aligned}$$

  $\mathrm{E}[Y \mid X = x]$ is a constant.

- Suppose that the conditional distribution of Y given $X = x$ is exponential (x), i.e. $f_{Y|X}(y \mid x) = x \cdot \exp(-xy)$, then

$$\mathrm{E}[Y \mid X = x] = \int_0^\infty y f_{Y|X}(y \mid x)\, \mathrm{d}y = \int_0^\infty y x \exp(-xy)\, \mathrm{d}y = \frac{1}{x}.$$

Conditional expectations as random variables

- A conditional expectation $\mathrm{E}[Y \mid X = x]$ is a number not a random variable. $\mathrm{E}[Y \mid X = x]$ is not random, not a function of Y. It is a function of the observed "realized" value x of the random variable X.
- We denote this function by $g(x) = \mathrm{E}[Y \mid X = x]$. Notice that g is an ordinary function of x, which is just a number.
- $g(X)$ is a random variable. If denoting $\mathrm{E}[Y \mid X] = g(X)$, $\mathrm{E}[Y \mid X]$ is a random variable and a function of X (Uncertainty about X has not been realized yet):

$$\begin{aligned} \mathrm{E}[Y \mid X] &= \sum_i y_i \Pr[Y = y_i \mid X] = g(X) \\ \mathrm{E}[Y \mid X] &= \int y f_{Y|X}(y \mid X)\,\mathrm{d}y = g(X). \end{aligned}$$

Properties of conditional expectations

- Conditional expectations satisfies all properties of unconditional expectation. E.g.

$$\mathrm{E}[Y + Z \mid X] = \mathrm{E}[Y \mid X] + \mathrm{E}[Z \mid X].$$

- Once you condition on X, you can treat any function of X as a constant:

$$\mathrm{E}[h_1(X) + h_2(X)Y \mid X] = h_1(X) + h_2(X)\mathrm{E}[Y \mid X],$$

  for any functions h_1 and h_2.
- Law of Iterated Expectation (LIE):

$$\begin{aligned}\mathrm{E}[\mathrm{E}[Y \mid X]] &= \mathrm{E}[Y],\\ \mathrm{E}[\mathrm{E}[Y \mid X, Z] \mid X] &= \mathrm{E}[Y \mid X]\\ \mathrm{E}[\mathrm{E}[Y \mid X] \mid X, Z] &= \mathrm{E}[Y \mid X].\end{aligned}$$

- Mean independence: Y and X are mean independent if

$$\mathrm{E}[Y \mid X] = \mathrm{E}[Y] = \text{constant}.$$

Relationship between different concepts of independence

$$
\begin{gathered}
X \text{ and } Y \text{ are independent} \\
\Downarrow \\
\mathrm{E}\left[Y \mid X\right] = \text{constant (mean independence)} \\
\Downarrow \\
\mathrm{Cov}\left[X, Y\right] = 0 \text{ (uncorrelatedness)}
\end{gathered}
$$

Proof of LIE

- If X and Y are continuous,

$$
\begin{aligned}
\mathrm{E}\left[\mathrm{E}\left[Y \mid X\right]\right] &= \int \mathrm{E}\left[Y \mid X = x\right] f_X\left(x\right) \mathrm{d}x \\
&= \int \left(\int y f_{Y|X}\left(y \mid x\right) \mathrm{d}y \right) f_X\left(x\right) \mathrm{d}x \\
&= \int \int y f_{X,Y}\left(x, y\right) \mathrm{d}y \mathrm{d}x \\
&= \int y f_Y\left(y\right) \mathrm{d}y \\
&= \mathrm{E}\left[Y\right].
\end{aligned}
$$

- The same is true if X and Y are discrete. We replace integrals by sums.
- The same is true if one of X and Y is discrete and the other is continuous.

Conditional variance

- Conditional variance is like variance, but defined by replacing ordinary expectation by conditional expectation:

$$\begin{aligned}\operatorname{Var}[Y \mid X] =& \operatorname{E}\left[(Y - \operatorname{E}[Y \mid X])^2 \mid X\right] \\ =& \operatorname{E}\left[Y^2 \mid X\right] - (\operatorname{E}[Y \mid X])^2 .\end{aligned}$$

- Similarly, we define conditional covariance between X and Y, conditional on Z:

$$\begin{aligned}\operatorname{Cov}[X, Y \mid Z] =& \operatorname{E}[(X - \operatorname{E}[X \mid Z])(Y - \operatorname{E}[Y \mid Z]) \mid Z] \\ =& \operatorname{E}[XY \mid Z] - \operatorname{E}[X \mid Z] \operatorname{E}[Y \mid Z] .\end{aligned}$$

Iterated variance

- We can calculate:

$$\begin{aligned}\mathrm{Var}[Y] =& \mathrm{E}\left[(Y - \mathrm{E}[Y])^2\right] \\ =& \mathrm{E}\left[(Y - \mathrm{E}[Y \mid X] + \mathrm{E}[Y \mid X] - \mathrm{E}[Y])^2\right] \\ =& \mathrm{E}\left[(Y - \mathrm{E}[Y \mid X])^2\right] + \mathrm{E}\left[(\mathrm{E}[Y \mid X] - \mathrm{E}[Y])^2\right] \\ &+ 2\mathrm{E}\left[(Y - \mathrm{E}[Y \mid X])(\mathrm{E}[Y \mid X] - \mathrm{E}[Y])\right].\end{aligned}$$

- By LIE,

$$\begin{aligned}\mathrm{E}\left[(Y - \mathrm{E}[Y \mid X])^2\right] =& \mathrm{E}\left[\mathrm{E}\left[(Y - \mathrm{E}[Y \mid X])^2 \mid X\right]\right] \\ =& \mathrm{E}[\mathrm{Var}[Y \mid X]].\end{aligned}$$

- By definition of variance,

$$\mathrm{E}\left[(\mathrm{E}[Y \mid X] - \mathrm{E}[Y])^2\right] = \mathrm{Var}[\mathrm{E}[Y \mid X]],$$

since $\mathrm{E}[\mathrm{E}[Y \mid X]] = \mathrm{E}[Y]$.

Iterated variance

- We can show

$$\mathrm{E}\left[(Y - \mathrm{E}\left[Y \mid X\right])\left(\mathrm{E}\left[Y \mid X\right] - \mathrm{E}\left[Y\right]\right)\right] = 0.$$

- In summary, we have

$$\mathrm{Var}\left[Y\right] = \mathrm{E}\left[\mathrm{Var}\left[Y \mid X\right]\right] + \mathrm{Var}\left[\mathrm{E}\left[Y \mid X\right]\right].$$

Bivariate normal distributions

- X and Y have a bivariate normal distribution if their joint PDF is given by:

$$
\begin{aligned}
f(x,y) &= \frac{1}{2\pi\sqrt{(1-\rho)^2\,\sigma_X^2\sigma_Y^2}} \exp\left[-\frac{1}{2(1-\rho)^2}\left(\frac{(x-\mu_X)^2}{\sigma_X^2}+\right.\right. \\
&\qquad \left.\left. +\frac{(y-\mu_Y)^2}{\sigma_Y^2} - 2\rho\frac{(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y}\right)\right],
\end{aligned}
$$

$\mu_X = \mathrm{E}[X], \mu_Y = \mathrm{E}[Y], \sigma_X^2 = \mathrm{Var}[X], \sigma_Y^2 = \mathrm{Var}[Y]$, and $\rho = \mathrm{Corr}[X,Y]$.

Properties of bivariate normal distributions

- If X and Y have a bivariate normal distribution:

$$\begin{aligned} a + bX + cY &\sim \mathrm{N}\left(\mathrm{E}\left[a + bX + cY\right], \mathrm{Var}\left[a + bX + cY\right]\right) \\ &= \mathrm{N}\left(a + b\mu_X + c\mu_Y, b^2\sigma_X^2 + c^2\sigma_Y^2 + 2bc\rho\sigma_X\sigma_Y\right). \end{aligned}$$

- $\mathrm{Cov}\left[X, Y\right] = 0 \implies X$ and Y are independent.
- $\mathrm{E}\left[Y \mid X\right] = \mu_Y + \frac{\mathrm{Cov}[X,Y]}{\sigma_X^2}\left(X - \mu_X\right).$
- Can be generalized to more than two random variables (multivariate normal).