

W10 Notes

Review: True/False

TRUE or FALSE:

- (a) Suppose $\text{Cov}[X, Y] = 0.05$. It is possible that the correlation coefficient $\rho_{X,Y} = 0.05$.
- (b) Suppose $\text{Cov}[X, Y] = 0.05$. It is possible that X and Y have a strong linear relationship.
- (c) Suppose $\text{Cov}[X, Y] = 0$. It is possible that X and Y are *not* independent.
- (d) Suppose X and Y are *not* independent. It is possible that $\text{Cov}[X, Y]$ is equal to 0.
- (e) Suppose X and Y *are* independent. $\text{Cov}[X, Y]$ must be equal to 0.

Review: Conditional probability

Recall some items related to conditional probability.

Conditioning definition:

$$P[- \mid A] = \frac{P[- \cap A]}{P[A]}$$

Multiplication rule:

$$P[AB] = P[B \mid A] P[A]$$

Division into Cases / Total Probability:

$$P[B] = P[B \mid A_1] P[A_1] + \dots + P[B \mid A_n] P[A_n]$$

Conditional distribution

01 Theory

Conditional distribution - fixed event

Suppose X is a random variable, and suppose $A \subset \mathbb{R}$. The distribution of X **conditioned on A** describes the probabilities of values of X given knowledge that $X \in A$.

Discrete case:

$$P_{X|A}(k) = \begin{cases} \frac{1}{P[A]} P_X(k) & k \in A \\ 0 & k \notin A \end{cases}$$

Continuous case:

$$f_{X|A}(x) = \begin{cases} \frac{1}{P[A]} f_X(x) & x \in A \\ 0 & x \notin A \end{cases}$$

There is also a conditional CDF, of which this conditional PDF is the derivative:

$$F_{X|A}(x) = P[X \leq x | A], \quad f_{X|A}(x) = \frac{d}{dx} F_{X|A}(x)$$

The Law of Total Probability has versions for distributions:

$$P_X(k) = P_{X|A_1}(k) P[A_1] + \cdots + P_{X|A_n}(k) P[A_n]$$

$$f_X(x) = f_{X|A_1}(x) P[A_1] + \cdots + f_{X|A_n}(x) P[A_n]$$

Conditional distribution - variable event

Suppose X and Y are any two random variables. The **distribution of X conditioned on Y** describes the probabilities of values of X in terms of y , given knowledge that $Y = y$.

Discrete case:

$$\begin{aligned} P_{X|Y}(k|\ell) &= P[X = k | Y = \ell] \\ &= \frac{P_{X,Y}(k, \ell)}{P_Y(\ell)} \quad (\text{assuming } P_Y(\ell) \neq 0) \end{aligned}$$

Continuous case:

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x, y)}{f_Y(y)} \quad (\text{assuming } f_Y(y) \neq 0)$$

Remember: $P_{X,Y}(k, \ell)$ is the probability that “ $X = k$ and $Y = \ell$.”

Sometimes it is useful to have the formulas rewritten like this:

$$P_{X,Y}(k, \ell) = P_{X|Y}(k|\ell) P_Y(\ell)$$

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

Extra - Deriving $f_{X|Y}(x|y)$

The density $f_{X|Y}$ ought to be such that $f_{X|Y}(x|y) dx$ gives the probability of $X \in [x, x + dx]$, given knowledge that $Y \in [y, y + dy]$. Calculate this probability:

$$\begin{aligned} &P[x \leq X \leq x + dx \mid y \leq Y \leq y + dy] \\ \gg \gg &\frac{P[x \leq X \leq x + dx, y \leq Y \leq y + dy]}{P[y \leq Y \leq y + dy]} \\ \gg \gg &\frac{f_{X,Y}(x, y) dx dy}{f_Y(y) dy} \\ \gg \gg &\frac{f_{X,Y}(x, y)}{f_Y(y)} dx \end{aligned}$$

02 Illustration

≡ Example - Conditional PMF, variable event, via joint density

Suppose X and Y have joint PMF given by:

$$P_{X,Y}(k, \ell) = \begin{cases} \frac{k+\ell}{21} & k = 1, 2, 3; \ell = 1, 2 \\ 0 & \text{otherwise} \end{cases}$$

Find $P_{X|Y}(k|\ell)$ and $P_{Y|X}(\ell, k)$.

Solution

Marginal PMFs:

$$P_X(k) = \frac{2k+3}{21}, \quad k = 1, 2, 3$$

$$P_Y(\ell) = \frac{\ell+2}{7}, \quad \ell = 1, 2$$

Assuming $\ell = 1$ or 2 , for each $k = 1, 2, 3$ we have:

$$P_{X|Y}(k|\ell) = \frac{P_{X,Y}(k, \ell)}{P_Y(\ell)} \gg \gg \frac{k+\ell}{3\ell+6}$$

Assuming $k = 1, 2$, or 3 , for each $\ell = 1, 2$ we have:

$$P_{Y|X}(\ell|k) = \frac{P_{Y,X}(\ell, k)}{P_X(k)} \gg \gg \frac{k+\ell}{2k+3}$$

Conditional expectation

03 Theory

▣ Expectation conditioned by a fixed event

Suppose X is a random variable and $A \subset \mathbb{R}$. The **expectation of X conditioned on A** describes the typical value of X given the hypothesis that $X \in A$ is known.

Discrete case:

$$E[X | A] = \sum_k k P_{X|A}(k)$$

$$E[g(X) | A] = \sum_k g(k) P_{X|A}(k)$$

Continuous case:

$$E[X | A] = \int_{-\infty}^{+\infty} x f_{X|A}(x) dx$$

$$E[g(X) | A] = \int_{-\infty}^{+\infty} g(x) f_{X|A}(x) dx$$

Conditional variance:

$$\text{Var}[X | A] = E[(X - \mu_{X|A})^2 | A] = E[X^2 | A] - \mu_{X|A}^2$$

Division into Cases / Total Probability applied to expectation:

$$E[X] = E[X | A_1] P[A_1] + \dots + E[X | A_n] P[A_n]$$

Linearity of conditional expectation:

$$E[aX_1 + bX_2 + c | Y = y] = a E[X_1 | Y = y] + b E[X_2 | Y = y] + c$$

☞ Extra - Proof: Division of Expectation into Cases

We prove the discrete case only.

1. Expectation formula:

$$E[X] = \sum_k k P_X(k)$$

2. Division into Cases for the PMF:

$$P_X(k) = \sum_{i=1}^n P_{X|A_i}(k) P[A_i]$$

3. Substitute in the formula for $E[X]$:

$$\begin{aligned} \sum_k k P_X(k) &\ggg \sum_k k \sum_{i=1}^n P_{X|A_i}(k) P[A_i] \\ &\ggg \sum_{i=1}^n P[A_i] \sum_k k P_{X|A_i}(k) \\ &\ggg \sum_{i=1}^n P[A_i] E[X | A_i] \end{aligned}$$

☞ Expectation conditioned by a variable event

Suppose X and Y are any two random variables. The **expectation of X conditioned on $Y = y$** describes the typical of value of X in terms of y , given the hypothesis that $Y = y$ is known.

Discrete case:

$$E[X | Y = y] = \sum_k k P_{X|Y}(k|y) \quad (k \text{ over all poss. vals.})$$

$$E[g(X, Y) | Y = y] = \sum_k g(k, y) P_{X|Y}(k|y)$$

Continuous case:

$$E[X | Y = y] = \int_{-\infty}^{+\infty} x f_{X|Y}(x|y) dx$$

$$E[g(X, Y) | Y = y] = \int_{-\infty}^{+\infty} g(x, y) f_{X|Y}(x|y) dx$$

05 Illustration

≡ Example - Conditional PMF, fixed event, expectation

Suppose X measures the lengths of some items and has the following PMF:

$$P_X(k) = \begin{cases} 0.15 & k = 1, 2, 3, 4 \\ 0.1 & k = 5, 6, 7, 8 \\ 0 & \text{otherwise} \end{cases}$$

Let L be the event that $X \geq 5$.

- (a) Find the conditional PMF of X given that L is known.
- (b) Find the conditional expected value and variance of X given L .

Solution

(a)

Conditional PMF formula with $x \in L$ plugged in:

$$P_{X|L}(x) = \begin{cases} \frac{P_X(x)}{P[L]} & x = 5, 6, 7, 8 \\ 0 & \text{otherwise} \end{cases}$$

Compute $P[L]$ by adding cases:

$$P[L] = \sum_{k=5}^8 P_X(k) \gg \gg 0.4$$

Divide nonzero PMF entries by 0.1:

$$P_{X|L}(k) = \begin{cases} 0.25 & k = 5, 6, 7, 8 \\ 0 & \text{otherwise} \end{cases}$$

(b)

Find $E[X | L]$:

$$E[X | L] = \sum_{k=5}^8 k P_{X|L}(k)$$

$$\gg \gg 5 \cdot (0.25) + 6 \cdot (0.25) + 7 \cdot (0.25) + 8 \cdot (0.25)$$

$$\gg \gg 6.5 \text{ min}$$

Find $E[X^2 | L]$:

$$E[X^2 | L] = \sum_{k=5}^8 k^2 P_{X|L}(k)$$

$$\ggg 5^2 \cdot (0.25) + 6^2 \cdot (0.25) + 7^2 \cdot (0.25) + 8^2 \cdot (0.25)$$

$$\ggg 43.5 \text{ min}^2$$

Find $\text{Var}[X | L]$ using “short form” with conditioning:

$$\text{Var}[X | L] = E[X^2 | L] - E[X | L]^2 \ggg 1.25 \text{ min}^2$$

04 Theory - extra

□ Expectation conditioned by a random variable

Suppose X and Y are any two random variables. The **expectation of X conditioned on Y** is a random variable giving the typical value of X on the assumption that Y has value determined by an outcome of the experiment.

$$E[X | Y] = g(Y) \quad \text{where } g(y) = E[X | Y = y]$$

In other words, start by defining a function $g(y)$:

$$\begin{aligned} g : \mathbb{R} &\rightarrow \mathbb{R} \\ y &\mapsto E[X | Y = y] \end{aligned}$$

Now $E[X | Y]$ is defined as the composite random variable $g(Y)$.

Considered as a random variable, $E[X | Y]$ takes an outcome $s \in S$, computes $Y(s)$, sets $y = Y(s)$, then returns the expectation of X conditioned on $Y = y$.

Notice that X is *not* evaluated at s , only Y is.

Because the value of $E[X | Y]$ depends only on $Y(s)$, and not on any additional information about s , it is common to *represent* a conditional expectation $E[X | Y]$ using only the function g .

📄 Iterated Expectation

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

📄 Proof of Iterated Expectation, discrete case

$$\begin{aligned}
E[E[X | Y]] &= \sum_{\ell} E[X | Y = \ell] P_Y(\ell) \\
&= \sum_{\ell} \sum_k k P_{X|Y}(k|\ell) P_Y(\ell) \\
&= \sum_k k \sum_{\ell} P_{X,Y}(k, \ell) \\
&= \sum_k k P_X(k) = E[X]
\end{aligned}$$

05 Illustration - extra

≡ Example - Conditional expectations from joint density

Suppose X and Y are random variables with joint density given by:

$$f_{X,Y}(x, y) = \begin{cases} \frac{1}{y} e^{-x/y} e^{-y} & x, y \in (0, \infty) \\ 0 & \text{otherwise} \end{cases}$$

Find $E[X | Y = y]$. Use this to compute $E[X]$.

Solution

(1) Derive the marginal density $f_Y(y)$:

$$\begin{aligned}
f_Y(y) &\ggg \int_0^{+\infty} \frac{1}{y} e^{-x/y} e^{-y} dx \\
&\ggg -e^{-x/y} e^{-y} \Big|_{x=0}^{\infty} \ggg e^{-y}
\end{aligned}$$

(2) Use $f_Y(y)$ to compute $f_{X|Y}(x|y)$:

$$\begin{aligned}
f_{X|Y}(x|y) &\ggg \frac{f_{X,Y}(x, y)}{f_Y(y)} \\
&\ggg \frac{1}{y} e^{-x/y} e^{-y} \cdot (e^{-y})^{-1} \ggg \frac{1}{y} e^{-x/y}
\end{aligned}$$

(3) Use $f_{X|Y}(x|y)$ to calculate expectation conditioned on the variable event:

$$\begin{aligned}
E[X | Y = y] &\ggg \int_{-\infty}^{+\infty} x f_{X|Y}(x|y) dx \\
&\ggg \int_0^{\infty} \frac{x}{y} e^{-x/y} dx \ggg y
\end{aligned}$$

(4) Apply Iterated Expectation:

Set $g(y) = y$. By Iterated Expectation, we know that $E[X] = E[g(Y)]$. Therefore:

$$\begin{aligned} E[X] = E[g(Y)] &= \int_{-\infty}^{+\infty} g(y) f_Y(y) dy \\ &\ggg \int_0^{+\infty} y e^{-y} dy \ggg 1 \end{aligned}$$

Notice that $g(Y) = Y$, so $E[X | Y] = Y$, and Iterated Expectation says that $E[X] = E[Y]$.

≡ Example - Flip coin, choose RV

Suppose $X \sim \text{Ber}(1/3)$ and $Y \sim \text{Ber}(1/4)$ represent two biased coins, giving 1 for heads and 0 for tails.

Here is the experiment:

1. Flip a fair coin.
2. If heads, flip the X coin; if tails, flip the Y coin.
3. Record the outcome as Z .

What is $E[Z]$?

Solution

Let $G \sim \text{Ber}(1/2)$ describe the fair coin. Then:

$$\begin{aligned} E[Z] &= E[E[Z | G]] \\ &\ggg E[Z | G = 0] P_G(0) + E[Z | G = 1] P_G(1) \\ &\ggg E[Y] P_G(0) + E[X] P_G(1) \\ &\ggg \frac{1}{4} \cdot \frac{1}{2} + \frac{1}{3} \cdot \frac{1}{2} \ggg \frac{7}{24} \end{aligned}$$

≡ Example - Sum of random number of RVs

Let N denote the number of customers that enter a store on a given day.

Let X_i denote the amount spent by the i^{th} customer.

Assume that $E[N] = 50$ and $E[X_i] = \$8$ for each i .

What is the expected total spend of all customers in a day?

Solution

A formula for the total spend is $X = \sum_{i=1}^N X_i$.

By Iterated Expectation, we know $E[X] = E[E[X | N]]$.

Now compute $E[X | N]$ as a function of N :

$$\begin{aligned}
E[X \mid N = n] &\ggg E\left[\left(\sum_{i=1}^N X_i\right) \mid N = n\right] \\
&\ggg E\left[\left(\sum_{i=1}^n X_i\right) \mid N = n\right] \\
&\ggg \sum_{i=1}^n E[X_i \mid N = n] \\
&\ggg \sum_{i=1}^n E[X_i] \ggg 8n
\end{aligned}$$

Therefore $g(n) = 8n$ and $g(N) = 8N$ and $E[X \mid N] = 8N$.

Then by Iterated Expectation, $E[X] = E[8N] = 8E[N] = \$400$.