Week 11 notes

Recall some items related to conditional probability.

Conditioning definition:

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

Multiplication rule:

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

Division into Cases / Total Probability:

$$P[B] \quad = \quad P[B \mid A_1]\, P[A_1] + \cdots + P[B \mid A_n]\, P[A_n]$$

Conditional distribution

01 Theory

Conditional distribution, by a fixed event

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

Discrete case:

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

Continuous case:

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

We can also condition the CDF directly and derive the PDF from the CDF:

$$F_{X|A}(x) \;=\; P[X \leq x \mid A], \qquad f_{X|A}(x) \;=\; \frac{dF_{X|A}(x)}{dx}$$

We can also translate Division into Cases / Total Probability into distributional terms:

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

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

Conditional distribution, by a 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 the hypothesis that $Y = y$ is known.

Discrete case:

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

Continuous case:

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

Notice:

- $P_{X,Y}(k,\ell)$ is the probability of "$X = k$ and $Y = \ell$."
- $P_{X|Y}(k|\ell)$ is the probability of $X = k$, given the hypothesis that $Y = \ell$ is known.

Sometimes it is useful to rewrite the formulas this way, for example to describe a "continuous probability tree:"

$$
\begin{aligned}
P_{X,Y}(k,\ell) &= P_{X|Y}(k|\ell)\, P_Y(\ell) \\
f_{X,Y}(x,y) &= f_{X|Y}(x|y)\, f_Y(y)
\end{aligned}
$$

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 $X \in [x, x+dx]$, on the hypothesis that $Y \in [y, y+dy]$ is known. Calculate this probability:

$$
P\Big[x \leq X \leq x + dx \;\Big|\; y \leq Y \leq y + dy\Big]
$$

$$
\gg\gg \quad \frac{P\Big[x \leq X \leq x + dx,\; y \leq Y \leq y + dy\Big]}{P\Big[y \leq Y \leq y + dy\Big]}
$$

$$
\gg\gg \quad \frac{f_{X,Y}(x,y)\,dx\,dy}{f_Y(y)\,dy}
$$

$$
\gg\gg \quad \frac{f_{X,Y}(x,y)}{f_Y(y)}\,dx
$$

Conditional expectation

02 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 of value of X given the hypothesis that $X \in A$ is known.

Discrete case:

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

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

Continuous case:

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

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

Conditional variance:

$$\mathrm{Var}[\,X \mid A\,] \quad = \quad E\Big[(X - \mu_{X|A})^2 \mid A\Big] \quad = \quad E\big[X^2 \mid A\big] - \mu_{X|A}^2$$

Division into Cases / Total Probability applied to expectation:

$$E[X] \quad = \quad E[\,X \mid A_1\,]\, P[A_1] + \cdots + E[\,X \mid A_n\,]\, P[A_n]$$

Linearity of conditional expectation:

$$E[\,aX_1 + bX_2 + c \mid Y = y\,] = a\, E[\,X_1 \mid Y = y\,] + b\, E[\,X_2 \mid Y = y\,] + c$$

Extra - Proof: Division of Expectation into Cases

We prove the discrete case only.

1. Expectation formula:

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

2. Division into Cases for the PMF:

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

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

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

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

$$\gg\gg \quad \sum_{i=1}^{n} P[A_i]\, E[\,X \mid A_i\,]$$

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:

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

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

Continuous case:

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

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

03 Illustration

Example - Conditioning on a fixed event

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

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

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

(a) Find the PMF of X conditioned on L.

(b) Find the conditional expected value and variance of X given L.

Solution

(a)

1. By the definition:

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

2. Adding exclusive probabilities:

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

3. Note that $0.1/0.4 = 0.25$. Therefore:

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

(b)

1. Find $E[\,X \mid L\,]$:

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

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

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

2. Find $E[\,X^2 \mid L\,]$:

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

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

$$\gg\gg \quad 43.5\,\text{min}^2$$

3. Find $\text{Var}[\,X \mid L\,]$:

$$\text{Var}[\,X \mid L\,] \quad = \quad E[\,X^2 \mid L\,] - E[\,X \mid L\,]^2 \quad \gg\gg \quad 1.25\,\text{min}^2$$

Example - Conditioning on variable events, discrete PMF function

Suppose X and Y have joint PMF given by:

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

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

Solution

First compute the marginal PMFs:

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

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

Therefore, assuming $\ell = 1$ or 2, then for any $k=1,2,3$ we have:

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

And, assuming $k = 1$, 2, or 3, then for any $\ell = 1,2$ we have:

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

04 Theory

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.

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

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

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

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

Considered as a random variable, $E[\,X \mid Y\,]$ 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 $E[\,X \mid Y\,]$ depends only on $Y(s)$, and not on any additional information about s, it is common to *represent* a conditional expectation $E[\,X \mid Y\,]$ using only the function g.

Iterated Expectation

$$E[\,E[X \mid Y]\,] = E[X]$$

Proof of Iterated Expectation, discrete case

$P[AB] = P[A] \cdot P[B|A] = P[B] \cdot P[A|B]$

$P_{X|Y} = \frac{P_{X,Y}}{P_Y} \rightsquigarrow P_{X,Y} = P_Y \cdot P_{X|Y} = P_X \cdot P_{Y|X}$

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

$\sum_\ell g(\ell)\, P_Y(\ell)$

$= \sum_\ell \sum_k k\, P_{X,Y}(k,\ell)$

$= E[X]$

$X = g(X, Y)$

05 Illustration

Exercise - Proof of Iterated Expectation, continuous case

Prove Iterated Expectation for the continuous case.

Example - Conditional expectations from joint density

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

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

Find $E[\,X \mid Y = y\,]$. Use this to compute $E[X]$.

view $X = g(X,Y)$

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

$>>>$ $E[X] = \iint x\, f_{X,Y}\, dx\, dy$

$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x \frac{1}{y} e^{-x/y} e^{-y}\, dx\, dy$

$>>>$?? Can you do this?

Solution

First derive the marginal density $f_Y(y)$:

$$f_Y(y) \quad \gg\gg \quad \int_0^{+\infty} \tfrac{1}{y} e^{-x/y} e^{-y}\, dx$$

$$\gg\gg \quad -e^{-x/y} e^{-y} \Big|_{x=0}^{\infty} \quad \gg\gg \quad e^{-y}$$

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

$$f_{X|Y}(x|y) \quad \gg\gg \quad \frac{f_{X,Y}(x,y)}{f_Y(y)}$$

$$\gg\gg \quad \tfrac{1}{y} e^{-x/y} e^{-y} \cdot (e^{-y})^{-1} \quad \gg\gg \quad \tfrac{1}{y} e^{-x/y}$$

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

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

$$\gg\gg \quad \int_0^{\infty} \tfrac{x}{y} e^{-x/y}\, dx \quad \gg\gg \quad y \qquad = g(y)$$

e.g.: "assuming $Y = 2$, expect $X = 2$"

Thus $E[X \mid Y] = Y$.

So, set $g(y) = y$. By Iterated Expectation, we know that $E[X] = E[g(Y)]$.

Therefore:

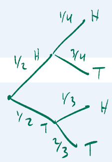
$E[E[X|Y]] >>> E[Y] =$

$$E[X] = E[g(Y)] \quad = \quad \int_{-\infty}^{+\infty} g(y)\, f_Y(y)\, dy$$

$$\gg\gg \quad \int_0^{+\infty} y\, e^{-y}\, dy \quad \gg\gg \quad 1$$

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

☰ Example - Flip coin, choose RV



$E[Z] = \sum z P_Z(z)$

$1 \cdot \left(\frac{1}{2}\cdot\frac{1}{4} + \frac{1}{2}\cdot\frac{1}{3}\right)$

$+ 0 \cdot \left(\frac{1}{2}\cdot\frac{3}{4} + \frac{1}{2}\cdot\frac{2}{3}\right) = \frac{7}{24}$

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 $G \sim \mathrm{Ber}(1/2)$ describe the fair coin.

Then:

$$E[Z] \quad = \quad E[\,E[Z \mid G\,]\,]$$

$$\gg\gg \quad E[\,Z \mid G = 0\,]\,P_G(0) + E[\,Z \mid G = 1\,]\,P_G(1)$$

$$\gg\gg \quad E[Y]\,P_G(0) + E[X]\,P_G(1)$$

$$\gg\gg \quad \frac{1}{4}\cdot\frac{1}{2} + \frac{1}{3}\cdot\frac{1}{2} \quad \gg\gg \quad \frac{7}{24}$$

☰ 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?

$X = X_1 + X_2 + \dots + X_N$

Solution

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

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

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

$g(X,N) \rightsquigarrow g(X,n)$

$$E[X \mid N = n] \quad \gg\gg \quad E\left[\left(\sum_{i=1}^{N} X_i\right) \mid N = n\right]$$

verified as HW question

$$\gg\gg \quad E\left[\left(\sum_{i=1}^{n} X_i\right) \mid N = n\right]$$

$$\gg\gg \quad \sum_{i=1}^{n} E[X_i \mid N = n]$$

$E[X \mid N=n]$

$$\gg\gg \quad \sum_{i=1}^{n} E[X_i] \quad \gg\gg \quad 8n$$

$\Rightarrow E[X \mid N] = 8N$

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

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