Lecture Notes for Math 210 – 22 October 2007

Shannon Starr

22 October 2007

Chapter 5: Tools in Probability Theory

Conditional Probability and Conditional Expectation

Remember that for one sample space Ω, there can be different event spaces (σ-fields) $\mathcal{F}$.

A random variable $\mathsf{X} : \Omega \to \mathbb{R}$ is said to be $\mathcal{F}$-measurable if all the sets $\{\omega \in \Omega : \mathsf{X}(\omega) = x\}$ are in $\mathcal{F}$ for all $x \in \mathbb{R}$.

This is an abstract idea, but consider the following simple example.

Example 1. Let X be the outcome of rolling 1 fair die, and let Y be the outcome of rolling another fair die.

So we could say

$$\Omega = \{(i,j) : i,j = 1,2,\dots,6\} .$$

Let $\mathcal{F}_1$ be the full event space consisting of all possible subsets of Ω.

Let $\mathcal{F}_2$ be the event space generated by X, by itself.

So, some events in $\mathcal{F}_2$ are

$$\begin{aligned} E_1 &= \{(1,j) \,:\, j = 1,\ldots,6\}\,, \\ E_2 &= \{(2,j) \,:\, j = 1,\ldots,6\}\,, \\ &\vdots \\ E_6 &= \{(6,j) \,:\, j = 1,\ldots,6\}\,. \end{aligned}$$

All other events in $\mathcal{F}_2$ are unions of some of these sets, or the empty set $\emptyset$.

(The empty set $\emptyset$ is always in every event space).

Note that using events in $\mathcal{F}_2$ we can ask the question "Is $\mathsf{X} = 4$?" or even "What is X?" but we cannot specify anything about Y.

So X is $\mathcal{F}_2$-measurable, but Y is not $\mathcal{F}_2$-measurable.

But both X and Y are $\mathcal{F}_1$-measurable.

<u>*Fact:*</u> Suppose Ω is some sample space

and that $\mathcal{F}_1$ and $\mathcal{F}_2$ are two event spaces

such that $\mathcal{F}_2 \subseteq \mathcal{F}_1$.

If X is a random variable which is $\mathcal{F}_1$-measurable,

then there is always another random variable, called $\mathbf{E}[\mathsf{X} \,|\, \mathcal{F}_2]$,

which is $\mathcal{F}_2$-measurable, and such that for any other $\mathcal{F}_2$ measurable function, Y, we have

$$\mathbf{E}[\mathsf{Y} \cdot \mathbf{E}[\mathsf{X} \,|\, \mathcal{F}_2]] \;=\; \mathbf{E}[\mathsf{Y} \cdot \mathsf{X}]\,.$$

This is also a somewhat abstract definition, but in particular cases calculating $\mathbf{E}[\mathsf{X} \,|\, \mathcal{F}_2]$ is usually a very concrete problem.

In case $\mathcal{F}_2$ is the event space generated by some other random variable Y, we usually write $\mathbf{E}[\mathsf{X}\,|\,\mathsf{Y}]$ instead of $\mathcal{F}_2$.

Usually, when we calculate $\mathbf{E}[\mathsf{X}\,|\,\mathsf{Y}]$, we find that it is an explicit function of Y.

Example 1 (cont.) Let us reconsider example 1 where X and Y are the results of rolling independent fair dice.

So,

$$\mathbf{P}\{\mathsf{X} = i\,,\ \mathsf{Y} = j\} = \frac{1}{36} \quad \text{for all} \quad i, j = 1, \ldots, 6\,.$$

Let $\mathsf{Z} = \mathsf{X} + \mathsf{Y}$.

Note that the possible outcomes for Z are $k = 2, 3, \ldots, 12$.

Also,

$$\mathbf{P}\{\mathsf{Z} = k\} = \frac{n(k)}{36}\,,$$

where $n(k) = 6 - |7 - k|$.

k	2	3	4	5	6	7	8	9	10	11	12
$n(k)$	1	2	3	4	5	6	5	4	3	2	1

Let us calculate

(a) $\mathbf{E}[\mathsf{X}\,|\,\mathsf{Z}]$,

(b) $\mathbf{E}[\mathsf{X}^2\,|\,\mathsf{Z}]$.

A: The answer to (a) is actually easy using a trick:

By symmetry we should have $\mathbf{E}[\mathsf{X}\,|\,\mathsf{Z}] = \mathbf{E}[\mathsf{Y}\,|\,\mathsf{Z}]$.

But we also know that

$$\mathbf{E}[\mathsf{X}\,|\,\mathsf{Z}] + \mathbf{E}[\mathsf{Y}\,|\,\mathsf{Z}] = \mathbf{E}[\mathsf{X} + \mathsf{Y}\,|\,\mathsf{Z}] = \mathbf{E}[\mathsf{Z}\,|\,\mathsf{Z}]\,.$$

But Z is already measurable with respect to the event space generated by Z, by definition.

So $\mathbf{E}[\mathsf{Z}\,|\,\mathsf{Z}] = \mathsf{Z}$.

Therefore,

$$\mathbf{E}[\mathsf{X}\,|\,\mathsf{Z}] \;=\; \frac{1}{2}\,\mathbf{E}[\mathsf{X}+\mathsf{Y}\,|\,\mathsf{Z}] \;=\; \frac{\mathsf{Z}}{2}\,.$$

So that is the answer.

To solve (b) is not as trivial.

Actually, what we have to do is first calculate the conditional probability.

We have

$$\mathbf{P}(\{\mathsf{X}=i\}\,|\,\mathsf{Z}) \;=\; \mathbf{E}[I_{\mathsf{X}=i}\,|\,\mathsf{Z}]\,,$$

where $I_{\{\mathsf{X}=i\}}$ is the indicator random variable:

$$I_{\{\mathsf{X}=i\}} \;=\; \begin{cases} 1 & \text{if } \mathsf{X}=i\ , \\ 0 & \text{otherwise}\ . \end{cases}$$

We remember the famous formula for conditional probability

$$\mathbf{P}(\{\mathsf{X}=i\}\,|\,\{\mathsf{Z}=k\}) \;=\; \frac{\mathbf{P}\{\mathsf{X}=i\,,\ \mathsf{Z}=k\}}{\mathbf{P}\{\mathsf{Z}=k\}}\,.$$

In our new, abstract notation, we actually *define* the conditional probability as

$$\mathbf{P}(\{\mathsf{X}=i\}\,|\,\{\mathsf{Z}=k\}) \;=\; g(k)\,,$$

where g is the function such that $\mathbf{P}(\{\mathsf{X}=i\}\,|\,\mathsf{Z}) = g(\mathsf{Z})$.

But we do not have to get caught up in notation.

What we get is

$$\mathbf{P}(\{\mathsf{X} = i\} \mid \{\mathsf{Z} = k\}) \;=\; \begin{cases} 1/n(k) & \text{if } i = \frac{k-n(k)+1}{2}, \ldots, \frac{k+n(k)-1}{2} \;, \\ 0 & \text{otherwise .} \end{cases}$$

E.g., if $k = 5$ then we have $n(k) = 6 - 2 = 4$. So

$$\mathbf{P}(\{\mathsf{X} = i\} \mid \{\mathsf{Z} = 5\}) \;=\; \frac{1}{4} \quad \text{for} \quad i \;=\; \frac{5-4+1}{2}, \ldots, \frac{5+4-1}{2} \;=\; 1, \ldots, 4\,.$$

Using this and the formula

$$\sum_{i=1}^{n} i^2 \;=\; \frac{n(n+1)(2n+1)}{6}\,,$$

it is straightforward, but tedious, to calculate

$$\mathbf{E}[\mathsf{X}^2 \mid \{\mathsf{Z} = k\}] \;=\; \frac{k(k^2 + 3(n(k))^2 - 3)}{12n(k)}\,.$$

So technically, we should have

$$\mathbf{E}[\mathsf{X}^2 \mid \mathsf{Z}] \;=\; \frac{\mathsf{Z}(\mathsf{Z}^2 + 3(6 - |7 - \mathsf{Z}|)^2 - 3)}{12(6 - |7 - \mathsf{Z}|)}\,.$$

(There may be a simpler formula. Also, there is some possibility that this formula has a small error. The point is not to derive the formula for this complicated problem, but to illustrate the notation using this problem.)