STAB52, An Introduction to Probability - Tutorial 9

Chong Chen

University of Toronto, Scarborough

First recall the that $\mathbb{P}(A|B) = \mathbb{P}(A \cap B)/\mathbb{P}(B)$ whenever $\mathbb{P}(B) > 0$. That is, the probability of A given B is "the probability of A restricted to the set B relative to the probability of the event B". A similar idea exists of expectation: given a random variable X and an event B with $\mathbb{P}(B) > 0$, the conditional expectation of X given B is $E(X|B) = \mathbb{E}(X\mathbb{I}_B)/\mathbb{P}(B)$ i.e. "the expected value of X restricted to B relative to the probability of the event B". Furthermore, just as there was a law of total probability, there is a law of total expectation: given a random variable X and a partition $\{B_n\}$ of the sample space, we have

$$\mathbb{E}(X) = \sum_n \mathbb{E}(X|B_n) \cdot \mathbb{P}(B_n).$$

The law of total expectation works in the same way that the law of total probability does. Instead of outright computing the expectation of a random variable, we compute its expected value conditional on partitions of the sample space (which is always easier to do). This is an extremely useful tool! Let us see an example of its utility:

Problem 1. (Evans & Rosenthal, Exercise 3.3.15)
Suppose you roll one fair six-sided die and then flip as many coins as the number showing on the die. Let X be the number showing on the die and Y be the number of heads obtained. Compute $\mathrm{Cov}(X, Y)$.

Solution: We first write $Y = Z_1 + Z_2 + ... + Z_X$ where Z_i are i.i.d. Bernoulli(1/2) random variables (and hence $Y \sim \mathrm{Binomial}(X, 1/2)$). Now, we must compute $\mathrm{Cov}(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$. Let us compute each part individually: Clearly, $\mathbb{E}(X) = 7/2$. Next, using the law of total expectation we have

$$\mathbb{E}(Y) = \sum_{x=1}^{6} \mathbb{E}(Y|X = x) \cdot \mathbb{P}(X = x) = \frac{1}{6}\sum_{x=1}^{6} \mathbb{E}(Y|X = x) = \frac{1}{6}\sum_{x=1}^{6} x \cdot \frac{1}{2} = \frac{21}{12}$$

where the third equality uses that fact that the expectation of a Binomial$(X, 1/2)$ distribution is $X \cdot 1/2$. Similarly, we compute

$$\begin{aligned}
\mathbb{E}(XY) &= \sum_{k=1}^{6} \mathbb{E}(XY|X = k) \cdot \mathbb{P}(X = k) \\
&= \frac{1}{6}\sum_{k=1}^{6} \mathbb{E}(XY|X = k) \\
&= \frac{1}{6}\sum_{k=1}^{6} k\mathbb{E}(Y|X = k) \\
&= \frac{1}{6}\sum_{k=1}^{6} k^2 \cdot \frac{1}{2} = \frac{91}{12}.
\end{aligned}$$

Collecting the above results, we conclude, $\mathrm{Cov}(X, Y) = 91/12 - 7/2 \cdot 21/12 = 35/24$.

$\square$

Remark: If we did not have the law of total expectation at our disposal, we would have to compute the joint p.m.f. of X and Y (not a fun task to do).

Problem 2. (Evans & Rosenthal, Exercise 3.3.29)
Let Y be a non-negative random variable. Prove that $\mathbb{E}(Y) = 0$ if and only if $\mathbb{P}(Y = 0) = 1$. (You may assume for simplicity that Y is discrete, but the result is true for any Y).

Solution: If we assume that Y is discrete, then

$$\mathbb{E}(Y) = \sum_{y \geq 0} y \cdot \mathbb{P}(Y = y) = 0 + \sum_{y > 0} y \cdot \mathbb{P}(Y = y).$$

It is then obvious to see that if $\mathbb{E}(Y) = 0$, then we must have $\mathbb{P}(Y = y) = 0$ for all $y > 0$ and hence

$$\mathbb{P}(Y = 0) = 1 - \sum_{y > 0} \mathbb{P}(Y = y) = 1 - \sum_{y > 0} 0 = 1$$

by disjoint additivity. Conversely, if $\mathbb{P}(Y = 0) = 1$, we must have $\mathbb{P}(Y = y) = 0$ for all $y > 0$ by monotonicity and hence $\mathbb{E}(Y) = 0$ by the first line in the solution. To show the result for a general random variable Y, we must use Definition 3.7.1. (Evans & Rosenthal). The proof is then very similar to the discrete case.

$\square$

Problem 3. (Evans & Rosenthal, Exercise 3.3.30)

Solution: Prove that $\mathrm{Var}(X) = 0$ if and only if there is a real number c such that $\mathbb{P}(X = c) = 1$.

Solution: We first observe that $(X - \mathbb{E}X)^2$ is a non-negative random variable and thus the identity

$$0 = \mathbb{E}((X - \mathbb{E}X)^2) = \mathrm{Var}(X)$$

holds if and only if $\mathbb{P}((X - \mathbb{E}X)^2 = 0) = 1$ by the previous exercise i.e. if and only if $\mathbb{P}(X - \mathbb{E}X = 0) = \mathbb{P}(X = EX) = 1$. Therefore, if $\mathrm{Var}(X) = 0$, then the desired constant is $c = \mathbb{E}X$. To prove the converse direction of implication, assume no such constant c exists. Then we immediately have a contradiction by the above argument.

$\square$

Problem 4. (Evans & Rosenthal, Exercise 3.3.31)
Give an example of a random variable X such that $\mathbb{E}X = 5$ and $\mathrm{Var}X = \infty$.

Solution: Notice that if we can find a random variable Y such that $\mathbb{E}Y = 0$ and $\mathrm{Var}Y = \infty$, then we are done since $\mathbb{E}(Y + 5) = 5$ and $\mathrm{Var}(Y + 5) = \mathrm{Var}(Y) = \infty$. So we may w.l.o.g. assume that $\mathbb{E}X = 0$. To construct such a random variable, consider X such that

$$\begin{aligned}
\mathbb{P}(X = 1) = \mathbb{P}(X = -1) = 1/4 \\
\mathbb{P}(X = 2) = \mathbb{P}(X = -2) = 1/16 \\
\mathbb{P}(X = 4) = \mathbb{P}(X = -4) = 1/64 \\
\mathbb{P}(X = 8) = \mathbb{P}(X = -8) = 1/256 \\
\vdots \\
\mathbb{P}(X = 2^n) = \mathbb{P}(X = -2^n) = 1/(2^{(n+1)^2}.)
\end{aligned}$$

Then clearly $\mathbb{E}X = 0$ but

$$\begin{aligned}
\mathrm{Var}(X) &= \mathbb{E}X^2 - (\mathbb{E}X)^2 \\
&= 1 \cdot \frac{1}{2} + 4 \cdot \frac{1}{8} + 16 \cdot \frac{1}{32} + 64 \cdot \frac{1}{128} + ... \\
&= \frac{1}{2} + \frac{1}{2} + \frac{1}{2} + \frac{1}{2} + \frac{1}{2} + ... = \infty.
\end{aligned}$$

Notice that chose the support of X such that

$$\mathbb{E}|X| = 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{8} + 4 \cdot \frac{1}{32} + 8 \cdot \frac{1}{128} + ... = \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} = 1 < \infty$$

i.e. X is integrable (this is also written as $X \in L^1$). If X was not integrable, its expectation would not be well-defined. An calculus analog to this idea would be evaluating $\int_{\mathbb{R}} \sin(x)$; the integral to this is not 0, the function itself is not integrable since $\int_{\mathbb{R}} |\sin(x)| = \infty$

$\square$