6.8 Eve's law and conditional variance

Unit 6: Joint Distributions and Conditional Expectation

Adapted from Blitzstein-Hwang Chapters 7 and 9.

Once we've defined conditional expectation given an r.v., we have a natural way to define conditional variance given an r.v.: replace all instances of $E(\cdot)$ in the definition of unconditional variance with $E(\cdot|X)$.

Definition 6.8.1 (Conditional variance).

The *conditional variance of* Y *given* X is

$$\mathrm{Var}(Y|X) = E((Y - E(Y|X))^2|X).$$

This is equivalent to

$$\mathrm{Var}(Y|X) = E(Y^2|X) - (E(Y|X))^2.$$

Warning 6.8.2.

Like $E(Y|X)$, $\mathrm{Var}(Y|X)$ is a random variable, and it is a function of X.

Since conditional variance is defined in terms of conditional expectations, we can use results about conditional expectation to help us calculate conditional variance.

Example 6.8.3.

Let $Z \sim \mathcal{N}(0,1)$ and $Y = Z^2$. Find $\mathrm{Var}(Y|Z)$ and $\mathrm{Var}(Z|Y)$.

Solution

Without any calculations we can see that $\text{Var}(Y|Z) = 0$: conditional on Z, Y is a known constant, and the variance of a constant is 0. By the same reasoning, $\text{Var}(h(Z)|Z) = 0$ for any function h. To get $\text{Var}(Z|Y)$, apply the definition:

$$\text{Var}(Z|Z^2) = E(Z^2|Z^2) - (E(Z|Z^2))^2.$$

The first term equals Z^2. The second term equals 0 by symmetry, as we found in Example 6.7.4. Thus $\text{Var}(Z|Z^2) = Z^2$, which we can write as $\text{Var}(Z|Y) = Y$.

We learned in the previous section that Adam's law relates conditional expectation to unconditional expectation. A companion result for Adam's law is *Eve's law*, which relates conditional variance to unconditional variance.

Theorem 6.8.4 (Eve's law).

For any r.v.s X and Y,

$$\text{Var}(Y) = E(\text{Var}(Y|X)) + \text{Var}(E(Y|X)).$$

The ordering of E's and Var's on the right-hand side spells EVVE, whence the name Eve's law. Eve's law is also known as the law of total variance.

Proof

Let $g(X) = E(Y|X)$. By Adam's law, $E(g(X)) = E(Y)$. Then

$$E(\text{Var}(Y|X)) = E(E(Y^2|X) - g(X)^2) = E(Y^2) - E(g(X)^2),$$

$$\text{Var}(E(Y|X)) = E(g(X)^2) - (Eg(X))^2 = E(g(X)^2) - (EY)^2.$$

Adding these equations, we have Eve's law.

To visualize Eve's law, imagine a population where each person has a value of X and a value of Y. We can divide this population into subpopulations, one for each possible value of X. For example, if X represents age and Y represents height, we can group people based on age. Then there are two sources contributing to the variation in people's heights in the overall population. First, within each age group, people have different heights. The average amount of variation in height within each age group is the *within-group variation*, $E(\text{Var}(Y|X))$.

Second, across age groups, the average heights are different. The variance of average heights across age groups is the *between-group variation*, $\text{Var}(E(Y|X))$. Eve's law says that to get the total variance of Y, we simply add these two sources of variation.

Figure 6.8.5 illustrates Eve's law in the simple case where we have three age groups. The average amount of scatter within each of the groups $X = 1$, $X = 2$, and $X = 3$ is the within-group variation, $E(\text{Var}(Y|X))$. The variance of the group means $E(Y|X = 1)$, $E(Y|X = 2)$, and $E(Y|X = 3)$ is the between-group variation, $\text{Var}(E(Y|X))$.

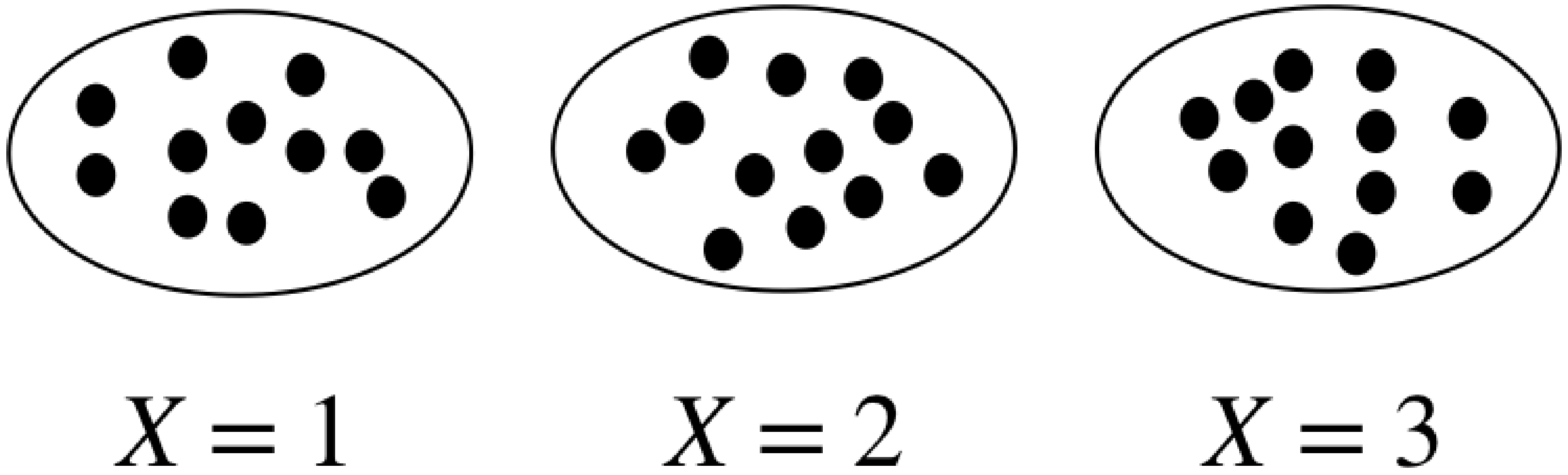

Figure 6.8.5: Eve's law says that total variance is the sum of within-group and between-group variation.

View Larger Image
Image Description

Warning 6.8.6.

Let Y be an r.v. and A be an event. It is wrong to say "$\text{Var}(Y) = \text{Var}(Y|A)P(A) + \text{Var}(Y|A^c)P(A^c)$", even though this looks analogous to the law of total expectation. Rather, we should use Eve's law: letting I be the indicator of A,

$$\text{Var}(Y) = E(\text{Var}(Y|I)) + \text{Var}(E(Y|I)).$$

Example 6.8.7 (Random sum).

A store receives N customers in a day, where N is an r.v. with finite mean and variance. Let X_j be the amount spent by the jth customer at the store. Assume that each X_j has mean μ and variance σ^2, and that N and all the X_j are independent of one another. Find the mean and variance of the random sum $X = \sum_{j=1}^{N} X_j$, which is the store's total revenue in a day, in terms of μ, σ^2, $E(N)$, and $\text{Var}(N)$.

Solution

Since X is a sum, our first impulse might be to claim "$E(X) = N\mu$ by linearity". Alas, this would be a category error, since $E(X)$ is a number and $N\mu$ is a random variable. The key is that X is not merely a sum, but a *random sum*; the number of terms we are adding up is itself random, whereas linearity applies to sums with a *fixed* number of terms. Yet this category error actually suggests the correct strategy: if only we were allowed to treat N as a constant, then linearity would apply. So let's condition on N! By linearity of *conditional* expectation,

$$E(X|N) = E(\sum_{j=1}^{N} X_j|N) = \sum_{j=1}^{N} E(X_j|N) = \sum_{j=1}^{N} E(X_j) = N\mu.$$

We used the independence of the X_j and N to assert $E(X_j|N) = E(X_j)$ for all j. Note that the statement "$E(X|N) = N\mu$" is not a category error because both sides of the equality are r.v.s that are functions of N. Finally, by Adam's law,

$$E(X) = E(E(X|N)) = E(N\mu) = \mu E(N).$$

This is a pleasing result: the average total revenue is the average amount spent per customer, multiplied by the average number of customers.

For $\mathrm{Var}(X)$, we again condition on N to get $\mathrm{Var}(X|N)$:

$$\mathrm{Var}(X|N) = \mathrm{Var}(\sum_{j=1}^{N} X_j|N) = \sum_{j=1}^{N} \mathrm{Var}(X_j|N) = \sum_{j=1}^{N} \mathrm{Var}(X_j) = N\sigma^2.$$

Eve's law then tells us how to obtain the unconditional variance of X:

$$\begin{aligned}\mathrm{Var}(X) &= E(\mathrm{Var}(X|N)) + \mathrm{Var}(E(X|N)) \\ &= E(N\sigma^2) + \mathrm{Var}(N\mu) \\ &= \sigma^2 E(N) + \mu^2 \mathrm{Var}(N).\end{aligned}$$

Learn About Verified Certificates