Unit 4: Homework

John Andrus

September 22, 2020

1 Safety First through Statistics

Suppose the strength of a particular metal beam is given by,

$$S = 10 + 0.5TP^2$$

Where T is a random variable representing the forging temperature, P is a random variable representing purity. Assume T has a normal distribution with mean 12 and standard deviation 2, P is uniform on [0,1], and T and P are independent.

Compute the expectation of S.

Hint: First compute the expectation conditional on one of the variables. Then apply the law of iterated expectations.

To compute S using the Law of Iterated Expectations, first compute $E[S|T]$.

$$E[S|T] = E[10 + 0.5TP^2|T]$$

Apply linearity,

$$= E[10|T + 0.5TE[P^2|T]$$

Given that T and P are independent,

$$= E[10] + 0.5TE[P^2]$$

Computing each term,

$$E[10] = 10$$

$$0.5T = 0.5(12) = 6$$

$$E[P^2] = \int_{-\infty}^{\infty} P^2 f_P(p)dp = \frac{1}{3}$$

Therefore,

$$E[S|T] = 2$$

Apply Law of Iterated Expectations,

$$E[S] = E_T[E_S[S|T]] = E_T[2] = 2$$

2 Reasoning about a BLP

Suppose that discrete random variables X and Y have joint probability mass function given by:

$$f(x,y) = \begin{cases} \frac{1}{2} & (x,y) \in \{(0,0),(2,1)\} \\ 0 & otherwise \end{cases}$$

Let $g(x) = \beta_0 + \beta_1 x$ be a predictor for y and define the error, $\epsilon = Y - g(x)$

1. If you impose the moment condition $E[\epsilon] = 0$, what one point in the plane must the predictor pass through? (In some places, this point is referred to as the grand mean.)

   Given that,

   $$\epsilon = Y - g(X)$$

   $$E[\epsilon] = 0$$

   The predictor must pass through $(E[X], E[Y])$.

   $$E[X] = 0.5(0) + 0.5(2) = 1$$

   $$E[Y] = 0.5(0) + 0.5(1) = 0.5$$

   The grand mean is therefore $(1, 0.5)$

2. Assuming $E[\epsilon] = 0$, write an expression for $cov[X, \epsilon] = cov[X, Y - g(x)]$ in terms of β_1

Substitute g(x) in the covariance expression,

$$cov[X, Y - g(X)] = cov[X, Y - \beta_0 - \beta_1 x]$$

Apply linearity,

$$= cov[X, Y] - cov[X, \beta_0] - cov[X, \beta_1 x]$$

Applying properties of covariance,

$$= cov[X, Y] - \beta_1 Var[X]$$

Apply the definition of variance and covariance,

$$= (E[XY] - E[X]E[Y]) - \beta_1(E[X^2] - E^2[X])$$

Compute $E[X]$, $E[Y]$, $E[XY]$, $E[X^2]$.

$$E[X] = 1$$

$$E[Y] = 0.5$$

$$E[XY] = 0.5(0)(0) + 0.5(2)(1) = 1$$

$$E[X^2] = 0.5(0^2) + 0.5(2^2) = 2$$

Substitute in to the previous expression for $cov[X, Y - g(x)]$ and solve.

$$cov[X, Y - g(x)] = 1.5 - \beta_1$$

3. In you own words, explain how the sign of $cov[X, \epsilon]$ is related to the angle of the line.

   In this case $cov[X, \epsilon] = 0$. The sign of covariance informs whether the angle of the line is positive or negative, as a positive covariance indicates the two variables change in the same direction and negative covariance indicates that the two variables change in opposite directions.

4. What predictor fulfills both $E[\epsilon] = 0$ and $cov[X, \epsilon] = 0$

   For $E[\epsilon] = 0$ the slope of the predictor must be

$$\beta_1 = \frac{E[Y]}{E[X]} = 0.5$$

   The best linear predictor will pass through the grand mean. Solving for β_0, g(x) can be expressed as.

$$g(x) = 0.5x$$