Econometrics

Homework 2

Problem 1. The LS objective function discussed in class is

$$Q(a,b) = \sum_{i=1}^{n} (Y_i - a - bX_i)^2 .$$

Now consider a modification of it:

$$\widetilde{Q}(a,b) = \left[\sum_{i=1}^{n} (Y_i - a - bX_i)\right]^2 .$$

Let $\left(\widetilde{\alpha}, \widetilde{\beta}\right)$ be the minimizer of $\widetilde{Q}$. Show that $\widetilde{Q}\left(\widetilde{\alpha}, \widetilde{\beta}\right) = 0$. Hint: You do not need to derive the first order conditions.

Solution. Define $\overline{Y} = n^{-1}\sum_{i=1}^{n} Y_i$. Note that since $\left(\widetilde{\alpha}, \widetilde{\beta}\right)$ is the minimizer, we have

$$0 \leq \widetilde{Q}\left(\widetilde{\alpha}, \widetilde{\beta}\right) \leq \widetilde{Q}\left(\overline{Y}, 0\right) = 0,$$

where the equality is due to the fact $\sum_{i=1}^{n}\left(Y_i - \overline{Y}\right) = 0$.

Problem 2. Suppose that you had a new battery for your camera, and the life of the battery is a random variable X, with PDF

$$f_X(x) = k \times \exp\left(-\frac{x}{\beta}\right),$$

where $x > 0$ and β is a parameter. Assume now that t and s are non-negative real numbers.
(a). Use the properties of a PDF to determine the value of k.
(b). Find an expression for $\Pr[X \geq t]$.
(c). Find an expression for the conditional probability: $\Pr[X \geq t + s \mid X \geq s]$. Hint: Use $\Pr(A \mid B) = \frac{\Pr(A \cap B)}{\Pr(B)}$.
(d). Suppose that your battery has already lasted for s weeks without dying. Based on your above answers, are you more concerned that the battery is about to die than you were when you first put it in the camera?

Solution.

(a) We know that the PDF must integrate to 1:

$$\begin{aligned}\int_0^\infty f_X(x)\,\mathrm{d}x =& \int_0^\infty k \times \exp\left(-\frac{x}{\beta}\right)\mathrm{d}x\\ 1 =& k\int_0^\infty \exp\left(-\frac{x}{\beta}\right)\mathrm{d}x\\ \frac{1}{k} =& -\beta\cdot\exp\left(-\frac{x}{\beta}\right)\Big|_0^\infty\\ k =& \frac{1}{\beta}.\end{aligned}$$

(b) This expression is straightforward now that we have the integration constant, k:

$$P(X \geq t) = \int_t^\infty \frac{1}{\beta}\exp\left(-x/\beta\right)\mathrm{d}x$$

which can be simplified to

$$P(X \geq t) = \exp\left(-t/\beta\right).$$

(c)

$$P(X \geq t+s \mid X \geq s) = \frac{\exp\left(-(t+s)/\beta\right)}{\exp\left(-s/\beta\right)} = \exp\left(-t/\beta\right).$$

(d) If my batteries have lasted s weeks without dying, based on my answer to part 3, I should be just as worried as I was before, since survival of the battery tells me nothing new about its likelihood of dying. The exponential distribution (which this is) has this very special property, that no matter how long something has lasted, its rate/probability of failure is constant at any given time.

Problem 3. Suppose that X is a continuous random variable with a strictly increasing and differentiable CDF F_X and PDF $f_X = F_X'$. (a) Show that $\mathrm{E}\left[(X-a)^2\right]$ is minimized at $a = \mathrm{E}[X]$. (b) Show that $\mathrm{E}[|X-a|]$ is minimized at $a = F_X^{-1}(1/2)$ (the median). Hint: $\mathrm{E}[|X-a|] = \int_{-\infty}^a (a-x) f_X(x)\,dx + \int_a^\infty (x-a) f_X(x)\,dx$.

Solution.
(a) Note

$$\mathrm{E}\left[(X-a)^2\right] = \mathrm{E}\left[X^2\right] - 2a\mathrm{E}[X] + a^2.$$

The first-order condition is:

$$\frac{\partial}{\partial a}\mathrm{E}\left[(X-a)^2\right] = -2\mathrm{E}[X] + 2a = 0.$$

(b) The first-order condition:

$$\frac{\partial}{\partial a}\mathrm{E}[|X-a|] = \int_{-\infty}^a f_X(x)\,dx - \int_a^\infty f_X(x)\,dx = 0.$$

It implies that $F_X(a) - (1 - F_X(a)) = 0$ and then, $a = F_X^{-1}(1/2)$.

Problem 4. Consider a simple regression model with no intercept:

$$Y_i = \beta X_i + U_i,$$

and assume that for all $i = 1, \ldots, n$:

$$\begin{aligned}
\mathrm{E}\left[U_i \mid X_1, \ldots, X_n\right] &= 0, \\
\mathrm{E}\left[U_i^2 \mid X_1, \ldots, X_n\right] &= \sigma^2, \\
\mathrm{E}\left[U_i U_j \mid X_1, \ldots, X_n\right] &= 0 \text{ for } i \neq j.
\end{aligned}$$

1. Show that the OLS estimator of β is

$$\hat{\beta} = \frac{\sum_{i=1}^n X_i Y_i}{\sum_{i=1}^n X_i^2}.$$

2. Show that for the fitted residuals $\hat{U}_i = Y_i - \hat{\beta} X_i$,

$$\sum_{i=1}^n \hat{U}_i X_i = 0.$$

3. Is it necessarily true that $\sum_{i=1}^n \hat{U}_i = 0$? Explain.
4. Show that $\hat{\beta}$ in part (a) is unbiased.
5. Show that conditionally on X's:

$$\operatorname{Var}\left[\hat{\beta}\right] = \frac{\sigma^2}{\sum_{i=1}^n X_i^2}.$$

Solution.
(1) The sum of squared residuals is given by

$$Q(b) = \sum_{i=1}^n \left(Y_i - b X_i\right)^2.$$

The derivative of $Q(b)$ is given by

$$\frac{dQ(b)}{db} = -2 \sum_{i=1}^n \left(Y_i - b X_i\right) X_i.$$

The first-order condition that the OLS estimator $\hat{\beta}$ has to satisfy is

$$\sum_{i=1}^n \left(Y_i - \hat{\beta} X_i\right) X_i = 0.$$

Re-arranging terms gives

$$\begin{aligned} 0 &= \sum_{i=1}^{n} \left(Y_i - \hat{\beta} X_i\right) X_i \\ &= \sum_{i=1}^{n} \left(Y_i X_i - \hat{\beta} X_i^2\right) \\ &= \sum_{i=1}^{n} Y_i X_i - \hat{\beta} \sum_{i=1}^{n} X_i^2. \end{aligned}$$

Solving for $\hat{\beta}$:

$$\hat{\beta} = \frac{\sum_{i=1}^{n} Y_i X_i}{\sum_{i=1}^{n} X_i^2}.$$

(2) Since the fitted residuals are defined as $\hat{U}_i = Y_i - \hat{\beta} X_i$, we have

$$\sum_{i=1}^{n} \hat{U}_i X_i = \sum_{i=1}^{n} \left(Y_i - \hat{\beta} X_i\right) X_i.$$

This is the same expression as the first-order condition. Therefore $\sum_{i=1}^{n} \hat{U}_i X_i = 0$.

(3) No. When the estimated regression has an intercept, the least squares problem has another first-order condition corresponding to the derivative of the sum of squared residuals function with respect to the intercept: $\sum_{i=1}^{n} \left(Y_i - \hat{\alpha} - \hat{\beta} X_i\right) = 0$. This additional condition ensures that the sum of the fitted residuals $\hat{U}_i$ is zero. When the estimated regression has no intercept, the equation is absent and nothing guarantees that $\sum_{i=1}^{n} \hat{U}_i = 0$.

(4) We have

$$\begin{aligned} \hat{\beta} &= \frac{\sum_{i=1}^{n} Y_i X_i}{\sum_{i=1}^{n} X_i^2} \\ &= \frac{\sum_{i=1}^{n} X_i \left(\beta X_i + U_i\right)}{\sum_{i=1}^{n} X_i^2} \\ &= \frac{\sum_{i=1}^{n} \left(\beta X_i^2 + U_i X_i\right)}{\sum_{i=1}^{n} X_i^2} \\ &= \frac{\beta \sum_{i=1}^{n} X_i^2 + \sum_{i=1}^{n} U_i X_i}{\sum_{i=1}^{n} X_i^2} \\ &= \beta + \frac{\sum_{i=1}^{n} U_i X_i}{\sum_{i=1}^{n} X_i^2}. \end{aligned}$$

In the following derivations, expectations are understood as conditional expectations given $X_1, ..., X_n$. We have

$$\begin{aligned} E\left(\hat{\beta}\right) &= E\left(\beta + \frac{\sum_{i=1}^{n} U_i X_i}{\sum_{i=1}^{n} X_i^2}\right) \\ &= \beta + E\left(\frac{\sum_{i=1}^{n} U_i X_i}{\sum_{i=1}^{n} X_i^2}\right) \\ &= \beta + \frac{\sum_{i=1}^{n} X_i E\left(U_i\right)}{\sum_{i=1}^{n} X_i^2}, \end{aligned}$$

where the last equation holds because once we condition on X's, they are treated as non-random. By the assumption, we have that for all i, $E(U_i \mid X_1, ..., X_n) = 0$ and therefore

$$E\left(\hat{\beta}\right) = \beta + \frac{\sum_{i=1}^{n} X_i \cdot 0}{\sum_{i=1}^{n} X_i^2} = \beta.$$

By the law of iterated expectation, the unconditional expectation is

$$E\left(\hat{\beta}\right) = E\left(E\left(\hat{\beta} \mid X_1, ..., X_n\right)\right) = E(\beta) = \beta.$$

(5) Conditional on X's,

$$\begin{aligned} Var\left(\hat{\beta}\right) &= E\left(\hat{\beta} - E\hat{\beta}\right)^2 \\ &= E\left(\hat{\beta} - \beta\right)^2 \\ &= E\left(\frac{\sum_{i=1}^{n} U_i X_i}{\sum_{i=1}^{n} X_i^2}\right)^2 \\ &= \left(\frac{1}{\sum_{i=1}^{n} X_i^2}\right)^2 E\left(\sum_{i=1}^{n} U_i X_i\right)^2, \end{aligned}$$

where the last equality holds because $\sum_{i=1}^{n} X_i^2$ is not random when we condition on X's. Next,

$$\begin{aligned} E\left(\sum_{i=1}^{n} U_i X_i\right)^2 &= E\left(\sum_{i=1}^{n}\sum_{j=1}^{n} X_i U_i X_j U_j\right) \\ &= E\left(\sum_{i=1}^{n} X_i^2 U_i^2 + \sum_{i=1}^{n}\sum_{j\neq i} X_i X_j U_i U_j\right) \\ &= \sum_{i=1}^{n} X_i^2 E\left(U_i^2\right) + \sum_{i=1}^{n}\sum_{j\neq i} X_i X_j E\left(U_i U_j\right) \\ &= \sum_{i=1}^{n} X_i^2 \sigma^2 + \sum_{i=1}^{n}\sum_{j\neq i} X_i X_j \cdot 0 \\ &= \sigma^2 \sum_{i=1}^{n} X_i^2, \end{aligned}$$

where the last equality holds because $E\left(U_i^2\right) = \sigma^2$ and $E(U_i U_j) = 0$ by the assumptions. Now we have

$$Var\left(\hat{\beta}\right) = \left(\frac{1}{\sum_{i=1}^{n} X_i^2}\right)^2 \sigma^2 \sum_{i=1}^{n} X_i^2 = \sigma^2 \left(\frac{1}{\sum_{i=1}^{n} X_i^2}\right).$$

Problem 5. The following table gives the joint probability distribution between employment status and college graduation among those either employed or unemployed.

	$Y = 1$ (employed)	$Y = 0$ (unemployed)
$X = 0$ (no college)	0.05	0.6
$X = 1$ (college graduate)	0.33	0.02

1. What is the mean of Y?
2. What is the mean of X?
3. What is the conditional mean of Y given $X = 0$?
4. What is the covariance of X and Y?
5. Are X and Y independent?
6. What is the probability of being employed?
7. What is the variance of X?
8. Suppose that you had a sample (X_i, Y_i), $i = 1, ..., n$ drawn from the joint distribution in the table. And now suppose that you estimate the following model by using this sample:
$$Y_i = \alpha + \beta X_i + U_i,\ \mathrm{E}[U_i \mid X_i] = 0.$$
What is the true value of the parameter β?

Solution.

1.
$$E(Y) = 1 \times P(Y = 1) + 0 \times P(Y = 0) = 0.38.$$

2.
$$E(X) = 1 \times P(X = 1) + 0 \times P(X = 0) = 0.35.$$

3.
$$E(Y \mid X = 0) = 1 \times P(Y = 1 \mid X = 0) + 0 \times P(Y = 0 \mid X = 0) = \frac{P(Y = 1, X = 0)}{P(X = 0)} = \frac{0.05}{0.65}.$$

4.
$$Cov(X, Y) = E(XY) - E(X) E(Y) = 1 \times P(X = 1, Y = 1) - 0.38 \times 0.35 = 0.197.$$

5. No, since $Cov(X, Y) \neq 0$.
6. $P(Y = 1) = 0.38$.
7.
$$Var(X) = E(X^2) - E(X)^2 = 0.35 - (0.35)^2 = 0.2275.$$

8.
$$\beta = \frac{Cov(X, Y)}{Var(X)} = \frac{0.197}{0.2275} = 0.86593.$$

Problem 6. Suppose you have a sample (X_i, Y_i), $i = 1, ..., n$ and estimate the linear model
$$Y_i = \beta_0 + \beta_1 X_i + U_i,\ \mathrm{E}[U_i \mid X_i] = 0$$
by OLS. The OLS estimator for the slope is $\widehat{\beta}_1$. Now estimate another linear model
$$X_i = \gamma_0 + \gamma_1 Y_i + V_i,\ \mathrm{E}[V_i \mid Y_i] = 0.$$
The OLS estimator for the slope is $\widehat{\gamma}_1$ Is it true that$\widehat{\gamma}_1 = \widehat{\beta}_1^{-1}$?

Solution. Note:

$$\widehat{\beta}_1 = \frac{\sum_{i=1}^n \left(X_i - \overline{X}\right)\left(Y_i - \overline{Y}\right)}{\sum_{i=1}^n \left(X_i - \overline{X}\right)^2}$$

and

$$\widehat{\gamma}_1 = \frac{\sum_{i=1}^n \left(Y_i - \overline{Y}\right)\left(X_i - \overline{X}\right)}{\sum_{i=1}^n \left(Y_i - \overline{Y}\right)^2} \neq \frac{1}{\widehat{\beta}_1}.$$

Problem 7. (Wooldridge Problem 2.8) Consider the standard simple regression model $Y_i = \beta_0 + \beta_1 X_i + U_i$ under the following assumptions: (1) (X_i, Y_i), $i = 1, ..., n$ are i.i.d. observations (2) $\mathrm{E}[U_i \mid X_i] = 0$ and $\mathrm{Var}[U_i \mid X_i] = \sigma^2$ (conditional homoskedasticity). The usual OLS estimators $\hat{\beta}_0$ and $\hat{\beta}_1$ are unbiased for β_0 and β_1. Let $\tilde{\beta}_1$ be the estimator of β_1 obtained by assuming the intercept is zero: $\tilde{\beta}_1$ is the minimizer of

$$\min_b \sum_{i=1}^n (Y_i - bX_i)^2 .$$

Hint: The estimator $\tilde{\beta}_1$ is constructed under the assumption that $\beta_0 = 0$. When answering the question, keep in mind that this assumption can be false and the true value of β_0 can be different from zero.

1. Find $\mathrm{E}\left[\tilde{\beta}_1\right]$ (the conditional expectation) in terms of the X's, β_0, and β_1. Verify that $\tilde{\beta}_1$ is unbiased for β_1 when the population intercept (β_0) is zero. Are there other cases where $\tilde{\beta}_1$ is unbiased?
2. Find the variance of $\tilde{\beta}_1$. (Hint: The variance does not depend on β_0.)
3. Show that $\mathrm{Var}\left[\tilde{\beta}_1\right] \leq \mathrm{Var}\left[\hat{\beta}_1\right]$ by showing the following fact: for any sample of data, $\sum_{i=1}^n X_i^2 \geq \sum_{i=1}^n \left(X_i - \bar{X}\right)^2$, with strict inequality unless $\bar{X} = 0$.

Solution.
(1) Easy to check:

$$\widetilde{\beta}_1 = \left(\sum_{i=1}^n X_i Y_i\right) / \left(\sum_{i=1}^n X_i^2\right).$$

Plugging in $Y_i = \beta_0 + \beta_1 X_i + U_i$ gives

$$\widetilde{\beta}_1 = \left(\sum_{i=1}^n X_i \left(\beta_0 + \beta_1 X_i + U_i\right)\right) / \left(\sum_{i=1}^n X_i^2\right).$$

After standard algebra, the numerator can be written as

$$\beta_0 \sum_{i=1}^n X_i + \beta_1 \sum_{i=1}^n X_i^2 + \sum_{i=1}^n X_i U_i.$$

Putting this over the denominator shows we can write $\widetilde{\beta}_1$ as

$$\widetilde{\beta}_1 = \beta_0 \left(\sum_{i=1}^n X_i\right) / \left(\sum_{i=1}^n X_i^2\right) + \beta_1 + \left(\sum_{i=1}^n X_i U_i\right) / \left(\sum_{i=1}^n X_i^2\right).$$

Conditional on the X's, we have

$$E(\widetilde{\beta}_1) = \beta_0 \left(\sum_{i=1}^{n} X_i \right) / \left(\sum_{i=1}^{n} X_i^2 \right) + \beta_1.$$

because $E(U_i) = 0$ for all i. Therefore, the bias in $\widetilde{\beta}_1$ is given by the first term in this equation. This bias is obviously zero when $\beta_0 = 0$. It is also zero when $\sum_{i=1}^{n} X_i = 0$, which is the same as $\overline{X} = 0$. In the latter case, regression through the origin is identical to regression with an intercept.
(2) From the last expression for $\widetilde{\beta}_1$ in part (i) we have, conditional on the X's,

$$\begin{aligned} Var(\widetilde{\beta}_1) &= \left(\sum_{i=1}^{n} X_i^2 \right)^{-2} Var\left(\sum_{i=1}^{n} X_i U_i \right) \\ &= \left(\sum_{i=1}^{n} X_i^2 \right)^{-2} \left(\sum_{i=1}^{n} X_i^2 Var(U_i) \right) \\ &= \left(\sum_{i=1}^{n} X_i^2 \right)^{-2} \left(\sigma^2 \sum_{i=1}^{n} X_i^2 \right) \\ &= \sigma^2 / \left(\sum_{i=1}^{n} X_i^2 \right) \end{aligned}$$

(3) The variance of the OLS estimator is $Var(\hat{\beta}_1) = \sigma^2 / \left(\sum_{i=1}^{n} \left(X_i - \overline{X} \right)^2 \right)$. From the hint, $\sum_{i=1}^{n} X_i^2 \geq \sum_{i=1}^{n} \left(X_i - \overline{X} \right)^2$, and so $Var(\widetilde{\beta}_1) \leq Var(\hat{\beta}_1)$. A more direct way to see this is to write $\sum_{i=1}^{n} \left(X_i - \overline{X} \right)^2 = \sum_{i=1}^{n} X_i^2 - n(\overline{X})^2$, which is less than $\sum_{i=1}^{n} X_i^2$ unless $\overline{X} = 0$.