Solutions to Selected Problems from Homework # 2

Jason Matthew Klusowski

1. Consider the case of a simple linear regression (no intercept) with a random design; specifically assume that $\{x_i\}_{i=1}^n$ and $\{\epsilon_i\}_{i=1}^n$ are each collections of independent normally distributed random variables with mean zero and variance 1 and σ^2, respectively. Furthermore, assume that $\mathbb{E}(\epsilon_i|X) = 0$ and $y_i = x_i\beta + \epsilon_i$, for $i = 1, 2, \ldots, n$. Define $\tilde{\beta} = \frac{1}{n}\sum_{i=1}^n x_i y_i$.

(a) Give both an argument and counterargument for using $\tilde{\beta}$ in lie of $\hat{\beta}$.

(b) Calculate $\mathbb{E}(\tilde{\beta}|X)$. Is this estimator unbiased when conditions on X? Is it unbiased when calculating the unconditional expectation?

(c) Compute the unconditional variance of $\tilde{\beta}$ and compare to the unconditional variance of $\hat{\beta}$. Which estimator which you rather use?

Solution. (a) The estimator $\hat{\beta}$ is equal to

$$\frac{\frac{1}{n}\sum_{i=1}^n x_i y_i}{\frac{1}{n}\sum_{i=1}^n x_i^2}.$$

Note that the quantity $\frac{1}{n}\sum_{i=1}^n x_i^2$ is an estimate of the variance of a normally distributed random variable having mean zero and variance 1. However, since we know this variance is 1, it seems better to use the actual value rather than an estimated one. So we can replace $\frac{1}{n}\sum_{i=1}^n x_i^2$ with 1, in which case we obtain $\hat{\beta}$.

On the other hand, even though both estimators are unbiased (see part (b)) and linear in y, $\hat{\beta}$ has the attractive property of minimizing the mean squared error, conditional on X.

(b) Observe that

$$
\begin{aligned}
\mathbb{E}(\tilde{\beta}|X) &= \mathbb{E}(\frac{1}{n}\sum_{i=1}^{n} x_i y_i|X) \\
&= \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}(x_i y_i|X) \\
&= \frac{1}{n}\sum_{i=1}^{n} x_i \mathbb{E}(y_i|X) \\
&= \frac{1}{n}\sum_{i=1}^{n} x_i \mathbb{E}(x_i\beta + \epsilon_i|X) \\
&= \frac{1}{n}\sum_{i=1}^{n} x_i [\mathbb{E}(x_i\beta|X) + \mathbb{E}(\epsilon_i|X)] \\
&= \frac{1}{n}\sum_{i=1}^{n} x_i \mathbb{E}(x_i\beta|X) \\
&= \frac{\beta}{n}\sum_{i=1}^{n} x_i^2.
\end{aligned}
$$

Clearly, $\mathbb{E}(\tilde{\beta}|X)$ isn't unbiased. But since $\frac{1}{n}\sum_{i=1}^{n} x_i^2$ is an unbiased estimator of 1 and $\mathbb{E}(\tilde{\beta}) = \mathbb{E}\mathbb{E}(\tilde{\beta}|X)$, it follows that $\mathbb{E}(\tilde{\beta}) = \beta$.

(c) We will use the fact that $\mathbb{V}(\tilde{\beta}) = \mathbb{E}(\tilde{\beta} - \beta)^2 = \mathbb{E}\mathbb{E}((\tilde{\beta} - \beta)^2|X)$.

Now,

$$
\begin{aligned}
\mathbb{E}((\tilde{\beta} - \beta)^2|X) &= \mathbb{E}((\tilde{\beta} - \mathbb{E}(\tilde{\beta}|X) + \mathbb{E}(\tilde{\beta}|X) - \beta)^2|X) \\
&= \mathbb{E}((\tilde{\beta} - \mathbb{E}(\tilde{\beta}|X))^2|X) + \mathbb{E}((\tilde{\beta}|X) - \beta)^2|X) + \\
&\quad + 2\mathbb{E}((\tilde{\beta} + \mathbb{E}(\tilde{\beta}|X))(\mathbb{E}(\tilde{\beta}|X) - \beta)|X) \\
&= \mathbb{E}((\tilde{\beta} - \mathbb{E}(\tilde{\beta}|X))^2|X) + \mathbb{E}((\mathbb{E}(\tilde{\beta}|X) - \beta)^2|X) + \\
&\quad 2(\mathbb{E}(\tilde{\beta}|X) - \beta)\mathbb{E}(\tilde{\beta} - \mathbb{E}(\tilde{\beta}|X)|X) \\
&= \mathbb{E}((\tilde{\beta} - \mathbb{E}(\tilde{\beta}|X))^2|X) + (\mathbb{E}(\tilde{\beta}|X) - \beta)^2
\end{aligned}
$$

Note that $R = \frac{n}{\beta}\mathbb{E}(\tilde{\beta}|X) = \sum_{i=1}^{n} x_i^2$ follows a chi-squared distribution with n degrees of freedom. This distribution has mean n and variance $2n$. Therefore,

$$
\begin{aligned}
\mathbb{E}(\mathbb{E}(\tilde{\beta}|X) - \beta)^2 &= \frac{\beta^2}{n^2}\mathbb{E}(R - n)^2 \\
&= \frac{\beta^2}{n^2}\mathbb{V}(R) \\
&= \frac{2\beta^2}{n}.
\end{aligned}
$$

Also, observe that

$$\mathbb{E}((\tilde{\beta} - \mathbb{E}(\tilde{\beta}|X))^2|X) = \mathbb{E}\left(\left(\frac{1}{n}\sum_{i=1}^{n}(x_i y_i - \beta x_i^2)\right)^2 |X\right).$$

Conditional on X, the collection $\{x_i y_i - \beta x_i^2\}_{i=1}^n$ is uncorrelated and has mean zero. Therefore,

$$\begin{aligned}\mathbb{E}\left(\left(\frac{1}{n}\sum_{i=1}^{n}(x_i y_i - \beta x_i^2)\right)^2 |X\right) &= \frac{1}{n^2}\sum_{i=1}^{n}\mathbb{E}(x_i y_i - \beta x_i^2)^2|X) \\ &= \frac{1}{n^2}\sum_{i=1}^{n}\mathbb{E}(\epsilon_i^2|X)\end{aligned}$$

Thus, $\mathbb{E}\mathbb{E}((\tilde{\beta} - \mathbb{E}(\tilde{\beta}|X))^2|X) = \dfrac{\sigma^2}{n}$. This shows that $\mathbb{V}(\tilde{\beta}) = \dfrac{\sigma^2 + 2\beta^2}{n}$.

To find the variance of $\hat{\beta}$, we use the fact that $\mathbb{E}((\hat{\beta} - \beta)^2|X) = \dfrac{\sigma^2}{\sum_{i=1}^n x_i^2}$. This means that

$$\begin{aligned}\mathbb{V}(\hat{\beta}) &= \mathbb{E}(\hat{\beta} - \beta)^2 \\ &= \mathbb{E}\mathbb{E}((\hat{\beta} - \beta)^2|X) \\ &= \sigma^2 \mathbb{E}\left(\frac{1}{\sum_{i=1}^n x_i^2}\right).\end{aligned}$$

Note that $\dfrac{1}{\sum_{i=1}^n x_i^2}$ has an inverse-chi-squared distribution with mean $\dfrac{1}{n-2}$. Therefore $\mathbb{V}(\hat{\beta}) = \dfrac{\sigma^2}{n-2}$.

For n sufficiently large, $\mathbb{V}(\hat{\beta}) < \mathbb{V}(\tilde{\beta})$. If we choose to measure the efficacy of an estimator in terms of its variability, it seems better to use $\hat{\beta}$ to estimate β.