Econometrics HW1

Michael B. Nattinger*

November 9, 2020

1 Question 2.1

By the law of iterated expections,

$$\begin{aligned} E[E[E[Y|X_1, X_2, X_3]|X_1, X_2]|X_1] &= E[E[Y|X_1, X_2]|X_1] \\ &= E[Y|X_1] \end{aligned}$$

2 Question 2.2

If $E[Y|X] = a + bX$, then, by the conditioning theorem:

$$\begin{aligned} E[XY] &= E[XE[Y|X]] = E[X(a + bX)] \\ &= E[aX] + E[bX^2] = aE[X] + bE[X^2] \end{aligned}$$

3 Question 2.3

Let $h(x)$ be such that $E[h(X)e] < \infty$. Then, by the conditioning theorem, $E[h(X)e] = E[h(X)E[e|X]] = E[h(X) * 0] = E[0] = 0$.

4 Question 2.4

$$\begin{aligned} E[Y|X = 0] &= (1/5)(0) + (4/5)(1) = 4/5 \\ E[Y|X = 1] &= (2/5)(0) + (3/5)(1) = 3/5 \\ E[Y^2|X = 0] &= (1/5)(0^2) + (4/5)(1^2) = 4/5 \\ E[Y^2|X = 1] &= (2/5)(0^2) + (3/5)(1^2) = 3/5 \\ Var[Y|X = 0] &= E[Y^2|X = 0] - (E[Y|X = 0])^2 = (4/5) - (16/25) = 4/25 \\ Var[Y|X = 1] &= E[Y^2|X = 1] - (E[Y|X = 1])^2 = (3/5) - (9/25) = 6/25 \end{aligned}$$

*I worked on this assignment with my study group: Alex von Hafften, Andrew Smith, and Ryan Mather. I have also discussed problem(s) with Emily Case, Sarah Bass, and Danny Edgel.

5 Question 2.5 (c)

Let $S(x)$ be some predictor of e^2 given X.

$$\begin{aligned}
E[(e^2 - S(X))^2] &= E[(e^2 - \sigma^2(X) + \sigma^2(X) - S(X))^2] \\
&= E[(e^2 - \sigma^2)^2] + 2E[(e^2 - \sigma^2(X))(\sigma^2(X) - S(X))] + E[(\sigma^2(X) - S(X))^2]. \\
E[(e^2 - \sigma^2(X))(\sigma^2(X) - S(X))] &= E[E[(e^2 - \sigma^2(X))(\sigma^2(X) - S(X))|X] \\
&= E[(\sigma^2(X) - S(X))E[(e^2 - \sigma^2(X))|X]] \\
&= E[(\sigma^2(X) - S(X))(E[e^2|X] - \sigma^2(X))] \\
&= E[(\sigma^2(X) - S(X))(\sigma^2(X) - \sigma^2(X))] \\
&= 0. \\
\Rightarrow E[(e^2 - S(X))^2] &= E[(e^2 - \sigma^2)^2] + E[(\sigma^2(X) - S(X))^2].
\end{aligned}$$

The first expectation is not dependent on $S(X)$ and the second is minimized when $S(X) = \sigma^2(X)$.

6 Question 2.8

Let Y be poisson conditional on X. Then, from our hint, clearly $E[Y|X] = X'\beta = Var[Y|X]$ and $E[e|X] = E[Y - X'\beta|X] = E[Y|X] - E[X'\beta|X] = E[Y|X] - X'\beta = E[Y|X] - E[Y|X] = 0$. Therefore, it does justify a linear model with conditional error expected to be 0.

7 Question 2.10

True. By the conditioning theorem,

$$E[X^2 e] = E[X^2 E[e|X]] = E[X^2 * 0] = E[0] = 0.$$

8 Question 2.11

False. Let $Y = X^2$. Then, for $X \sim N(0,1)$, $\beta = 0, e = Y - X\beta = X^2$, $E[Xe] = 0$ by symmetry but $E[X^2 e] = E[X^4] = 3 \neq 0$.

9 Question 2.12

False. Let $p(X = 0, e = 0) = 1/4, P(X = 0, e = 1) = 1/8, P(X = 0, e = -1) = 1/8, P(X = 1, e = 0) = 1/2$. Then, $P(e = 1|X = 1) = 0 \neq (1/8)(1/2) = P(e = 1)P(X = 1)$.

10 Question 2.13

False. Use our example from before in question 2.11. Then, $E[Xe] = 0, E[e|X = 1]$ $= E[X^2|X = 1] = 1^2 = 1 \neq 0$.

11 Question 2.14

False. Let $X_i \sim N(0,1)$, and let Z_i be such that $E[Z_i|X_i] = 1, Var(Z_i|X_i) = \sigma^2/(X_i^2)$. Define $Y_i = X_i Z_i, e_i = Y_i - E[Y_i|X_i]$. Then, $E[e_i|X_i] = 0, E[e_i^2|X_i] = E[X_i^2(Z_i - 1)^2|X_i] = X_i^2 E[(Z_i - 1)^2|X_i] = X_i^2 Var(Z_i|X_i) = \sigma^2$. Note however that e_i, X_i are not independent.

12 Question 2.16

To compute the expectation of Y conditional on X we first should compute the marginal density of X and use that, along with the joint density, to compute the conditional density of Y given X. Then we can find the expectation of Y given X:

$$\begin{aligned}
f_X(x) &= \int_0^1 (3/2)(x^2+y^2)dy = (3/2)x^2 + 1/2, \\
f_{Y|X=x}(y) &= \frac{(3/2)(x^2+y^2)}{(3/2)x^2+1/2}, \\
E[Y|X=x] &= \int_0^1 y f_{Y|X=x}(y)dy = \frac{1}{x^2+1/3}(x^2\int_0^1 ydy + \int_0^1 y^3dy) \\
&= \frac{1}{x^2+1/3}(x^2(1/2)+(1/4)) \\
&= \frac{x^2+1/2}{2x^2+2/3}.
\end{aligned}$$

This is different from the best linear predictor, which we will derive below:

Write $\tilde{X} = \begin{pmatrix} 1 \\ X \end{pmatrix}$. Then,

$$\begin{aligned}
\tilde{\beta} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} &= (E[\tilde{X}\tilde{X}'])^{-1}E[\tilde{X}Y] \\
&= \left(E\begin{pmatrix} 1 & X \\ X & X^2 \end{pmatrix}\right)^{-1} E\begin{pmatrix} Y \\ XY \end{pmatrix} \\
&= \frac{1}{E(X^2)-E(X)^2}\begin{pmatrix} E(X^2) & -EX \\ -EX & 1 \end{pmatrix} E\begin{pmatrix} Y \\ XY \end{pmatrix} \\
&= \frac{1}{E(X^2)-E(X)^2}\begin{pmatrix} E(X^2)EY - EXE(XY) \\ E(XY) - EXEY \end{pmatrix}
\end{aligned}$$

$$\begin{aligned}
EX = EY &= \int_0^1 xf_X(x)dx = \int_0^1 (3/2)x^3 + x/2dx = (3/8)+(1/4) = 5/8, \\
EX^2 &= \int_0^1 x^2 f_X(x)dx = \int_0^1 (3/2)x^4 + x^2/2dx = (3/10)+(1/6) = (9/30)+(5/30) = 7/15, \\
E[XY] &= \int_0^1\int_0^1 f_{X,Y}(x,y)dydx = \int_0^1\int_0^1 (3/2)(x^3y+y^3x)dydx = (3/2)\int_0^1 x^3/2 + x/4dx \\
&= (3/4)((1/4)+(1/4)) = 3/8, \\
\Rightarrow \tilde{\beta} &= \frac{1}{(7/15)-(25/64)}\begin{pmatrix} (7/15)(5/8)-(5/8)(3/8) \\ (3/8)-(5/8)(5/8) \end{pmatrix} \\
&= \begin{pmatrix} 55/73 \\ -15/73 \end{pmatrix}.
\end{aligned}$$

Thus, the best linear predictor $L(x) = (55/73) - (15/73)x$ is different from the best predictor of Y, $m(x) = E[Y|X=x] = \frac{x^2+1/2}{2x^2+2/3}$.

13 Question 4.1

Define $\hat{\mu}_k := \frac{1}{n}\sum_{i=1}^n X_i^k$. We will show that this is unbiased.

$$\begin{aligned} E[\hat{\mu}_k] &= E\left[\frac{1}{n}\sum_{i=1}^n X_i^k\right] = \frac{1}{n}\sum_{i=1}^n E[X_i^k] \\ &= \frac{1}{n}\sum_{i=1}^n \mu_k \\ &= \mu_k. \end{aligned}$$

Thus, $\hat{\mu}_k$ is an unbiased estimator for μ_k.

$$\begin{aligned} Var(\hat{\mu}_k) &= \frac{1}{n^2}\sum_{i=1}^n Var(X_i^k) = \frac{1}{n}(E[X_i^{2k}] - E[X_i^k]^2) \\ &= \frac{1}{n}(\mu_{2k} - \mu_k^2). \end{aligned}$$

This is finite if $|\mu_{2k}| < \infty$.
An estimator of the variance can be found by the plug-in estimator:

$$\hat{Var}(\hat{\mu}_k) = \frac{1}{n}\left(\left(\frac{1}{n}\sum_{i=1}^n X_i^{2k}\right) - \left(\frac{1}{n}\sum_{i=1}^n X_i^k\right)^2\right).$$

14 Question 4.2

$$\begin{aligned} E[(\bar{Y}-\mu)^3] &= \frac{1}{n^3}E\left[\left(\sum_{i=1}^n (y_i-\mu)\right)^3\right] \\ &= \frac{1}{n^3}\left(\sum_{i=1}^n E(y_i-\mu)^3 + 3\sum_{i\neq j} E((y_i-\mu)^2 E(y_j-\mu) + 6\sum_{1\le i<j<k\le n} E(y_i-\mu)E(y_j-\mu)E(y_l-\mu)\right) \\ &= \frac{1}{n^3}\sum_{i=1}^n E(y_i-\mu)^3 = \frac{1}{n^2}E[(y_i-\mu)^3]. \end{aligned}$$

This is zero if the third central moment of Y is 0.

15 Question 4.3

$\bar{Y}$ is the sample mean of Y and is a consistent and unbiased estimator of the mean of Y. μ is the true mean of Y. Similarly, $n^{-1}\sum_{i=1}^n x_i x_i'$ is a consistent estimator of $E[x_i x_i']$ which is the true population value.

16 Question 4.4

$$\sum_{i=1}^{n} X_i^2 \hat{e}_i = \sum_{i=1}^{n} X_i^2 (Y_i - X_i \hat{\beta}) = \sum_{i=1}^{n} X_i^2 Y_i - \sum_{i=1}^{n} X_i^3 \hat{\beta}$$
$$= \sum_{i=1}^{n} X_i^2 Y_i - \sum_{i=1}^{n} \left(X_i^3 \left(\sum_{j=1}^{n} X_j^2 \right)^{-1} \sum_{j=1}^{n} X_j Y_j \right)$$

In general this expression is not equal to 0 so the general answer is false. It is trivial to show this via simulated data in Matlab, example code for which is written in a footnote.[1]

17 Question 4.5

$$\begin{aligned} E[\hat{\beta}|X] &= E[(X'X)^{-1}X'Y|X] = (X'X)^{-1}X'E[Y|X] = (X'X)^{-1}X'(X\beta + E[e|X]) \\ &= (X'X)^{-1}X'X\beta = \beta. \\ Var[\hat{\beta}|X] &= E[(\hat{\beta}-\beta)(\hat{\beta}-\beta)'|X] \\ &= E[(X'X)^{-1}X'ee'X(X'X)^{-1}|X] \\ &= (X'X)^{-1}X'E[ee'|X]X(X'X)^{-1} \\ &= (X'X)^{-1}X'\Omega X(X'X)^{-1} \end{aligned}$$

18 Question 4.6

Let A be any $n \times k$ function of X such that $A'X = I_k$, so that the estimator is unbiased. The estimator has variance $Var[A'Y|X] = A'\Omega A$. Proving the generalized gauss-markov inequality consists of proving that $A'\Omega A - (X'\Omega^{-1}X)^{-1}$ is positive semidefinite.

[1]Example matlab code follows:

```
clear; close all; clc
rng(99) % set seed - any seed is fine
n = 1000; % # obs for sim
e = randn(n,1); % true error
btrue = 1; % true beta
X = randn(n,1); % independent variables draws
Y = X*btrue + e; % dependent variable
bols = inv(X'*X)*(X'*Y); % OLS betahat
r = Y -X*bols; % OLS residuals
sum1 = sum(X.^2.*r); % quantity we are asked about
sum2 = sum(X.^2.*Y) - sum(X.^3 * inv(X'*X)*(X'*Y)); %rewritten version of quantity - is exactly identical
disp(num2str(sum1));
disp(num2str(sum2));
return
% This code yields the following (clearly nonzero) results in the command window:
% 58.9587
% 58.9587
```

Let $C = A - \Omega^{-1}X(X'\Omega^{-1}X)^{-1}$. Then we have the following:

$$\begin{aligned}
A'\Omega A - (X'\Omega^{-1}X)^{-1} &= (C + \Omega^{-1}X(X'\Omega^{-1}X)^{-1})'\Omega(C + \Omega^{-1}X(X'\Omega^{-1}X)^{-1}) \\
&= C'\Omega C + C'\Omega\Omega^{-1}X(X'\Omega^{-1}X)^{-1} + (\Omega^{-1}X(X'\Omega^{-1}X)^{-1})'\Omega C \\
&+ (\Omega^{-1}X(X'\Omega^{-1}X)^{-1})'\Omega\Omega^{-1}X(X'\Omega^{-1}X)^{-1}) - (X'\Omega X)^{-1} \\
&= C'\Omega C + (X'C)'(X'\Omega^{-1}X)^{-1} + (X'\Omega^{-1}X)^{-1}X'C \\
&+ (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}X(X'\Omega^{-1}X)^{-1}) - (X'\Omega X)^{-1} \\
&= C'\Omega C = (\Omega^{1/2}C)'(\Omega^{1/2}C)
\end{aligned}$$

Where the second-to-last equality holds as $X'C = 0$. Thus, $A'\Omega A - (X'\Omega^{-1}X)^{-1}$ is positive semidefinite.