Econ 709 Problem Set 1

Sarah Bass *

November 9, 2020

Question 2.1

By the law of iterated expectations,

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

Question 2.2

Using the conditioning theorem,

$$\begin{aligned}\mathbb{E}[XY] &= \mathbb{E}[X\mathbb{E}[Y|X]] \\ &= \mathbb{E}[X(a + bX)] \\ &= \mathbb{E}[Xa + bX^2] \\ &= \mathbb{E}[Xa] + \mathbb{E}[bX^2] \\ &= a\mathbb{E}[X] + b\mathbb{E}[X^2]\end{aligned}$$

Question 2.3

If $\mathbb{E}|Y| < \infty$, then for some function $h(x)$ such that $\mathbb{E}|h(X)e| < \infty$,

$$\mathbb{E}[h(X)e] = \mathbb{E}[h(X)\mathbb{E}[e|X]] = \mathbb{E}[h(X) * 0] = 0$$

*I have discussed this problem set with Emily Case, Michael Nattinger, Alex Von Hafften, and Danny Edgel.

Question 2.4

$$\begin{aligned}
\mathbb{E}[Y|X=0] &= 0.8 \\
\mathbb{E}[Y|X=1] &= 0.6 \\
\mathbb{E}[Y^2|X=0] &= 0.8 \\
\mathbb{E}[Y^2|X=1] &= 0.6 \\
var[Y|X=0] &= 0.8 - 0.64 = 0.16 \\
var[Y|X=1] &= 0.6 - 0.36 = 0.24
\end{aligned}$$

Question 2.5 - Part C

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

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

However, note that

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

So,

$$E[(e^2 - S(X))^2] = E[(e^2 - \sigma^2)^2] + E[(\sigma^2(X) - S(X))^2].$$

$E[(e^2-\sigma^2)^2]$ is not dependent on $S(X)$ and $E[(\sigma^2(X)-S(X))^2]$ is minimized when $S(X) = \sigma^2(X)$.

Question 2.8

$$\begin{aligned}
\mathbb{E}[Y|X] &= X'\beta \\
[Y|X] &= X'\beta
\end{aligned}$$

Note that $\mathbb{E}[e|X] = \mathbb{E}[Y - X'\beta|X] = \mathbb{E}[Y|X] - X'\beta = X'\beta - X'\beta = 0$. So this justifies a linear regression model of the form $Y = X'\beta + e$.

Question 2.10

True

$$\mathbb{E}[X^2 e] = \mathbb{E}[X^2 \mathbb{E}[e|X]] = \mathbb{E}[X^2 * 0] = 0$$

Question 2.11

False. Consider $Y = X^2$ with $X \sim N(0,1)$. Then $\beta = 0, e = Y - X\beta = X^2$ $\mathbb{E}[Xe] = 0$ by symmetry, but $\mathbb{E}[X^2 e] = \mathbb{E}[X^4] > 0$.

Question 2.12

False. Consider the following probabilities:

$$\begin{aligned}
P(X = 0, e = 0) &= \frac{1}{4} \\
P(X = 0, e = -1) &= \frac{1}{8} \\
P(X = 0, e = 1) &= \frac{1}{8} \\
P(X = 1, e = 0) &= \frac{1}{2} \\
P(X = 1, e = -1) &= 0 \\
P(X = 1, e = 1) &= 0
\end{aligned}$$

Then $\mathbb{E}[e|X] = 0$. However $P(e = 1|X = 1) = 0 \neq \left(\frac{1}{8}\right)\left(\frac{1}{2}\right) = P(e = 1)P(X = 1)$ so e is not independent of X.

Question 2.13

False. Using the example from 2.11, $\mathbb{E}[Xe] = 0$ by symmetry, but $\mathbb{E}[e|X = 1] = \mathbb{E}[X^2|X = 1] = 1$.

Question 2.14

False. Let $X_i \sim N(0,1)$, and consider Z_i such that $E[Z_i|X_i] = 1$ and $Var(Z_i|X_i) = \sigma^2/(X_i^2)$. Let $Y_i = X_i Z_i$ and $e_i = Y_i - E[Y_i|X_i]$. Then since $E[e_i|X_i] = 0$

$$\begin{aligned}
E[e_i^2|X_i] &= E[X_i^2(Z_i - 1)^2|X_i] \\
&= X_i^2 E[(Z_i - 1)^2|X_i] \\
&= X_i^2 Var(Z_i|X_i) \\
&= \sigma^2
\end{aligned}$$

However e_i and X_i are not independent.

Question 2.16

We'll first compute the marginal density of X, then use that to calculate the conditional density of Y given X, then calculate the expectation of Y given X.

$$
\begin{aligned}
f_X(x) &= \int_0^1 \frac{3}{2}(x^2+y^2)dy \\
&= \frac{3}{2}x^2y + \frac{1}{2}y^3\Big|_0^1 \\
&= \frac{3}{2}x^2 + \frac{1}{2} \\
f_{Y|X=x}(y) &= \frac{\frac{3}{2}(x^2+y^2)}{\frac{3}{2}x^2+\frac{1}{2}} \\
\mathbb{E}[Y|X=x] &= \int_0^1 \frac{y(\frac{3}{2}(x^2+y^2))}{\frac{3}{2}x^2+\frac{1}{2}}dy \\
&= \frac{1}{x^2+\frac{1}{3}}\left(\int_0^1 x^2 y dy + \int_0^1 y^3 dy\right) \\
&= \frac{1}{x^2+\frac{1}{3}}\left(\frac{1}{2}x^2+\frac{1}{4}\right) \\
&= \frac{x^2+\frac{1}{2}}{2x^2+\frac{2}{3}}
\end{aligned}
$$

To calculate the best linear predictor, we'll first write $\tilde{X} = \begin{pmatrix} 1 & X \end{pmatrix}$. So we can calculate the best linear predictor as follows.

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

Note that

$$
\begin{aligned}
\mathbb{E}[X] &= \int_0^1 \frac{3}{2}x^3 + \frac{x}{2}dx = \frac{5}{8} \\
\mathbb{E}[Y] &= \int_0^1 \frac{3}{2}y^3 + \frac{y}{2}dy = \frac{5}{8} \\
\mathbb{E}[X^2] &= \int_0^1 \frac{3}{2}x^4 + \frac{x^2}{2}dx = \frac{7}{15} \\
\mathbb{E}[XY] &= \int_0^1 \int_0^1 \frac{3}{2}(x^3y + y^3x)dydx = \frac{3}{8}
\end{aligned}
$$

Using this, we can solve for $\hat{\beta}$ as

$$
\begin{aligned}
\hat{\beta} &= \frac{1}{\frac{7}{15} - \frac{25}{64}} \begin{pmatrix} \frac{7}{15}(\frac{5}{8}) - \frac{5}{8}(\frac{3}{8}) \\ -\frac{5}{8}(\frac{5}{8}) + \frac{3}{8} \end{pmatrix} \\
&= \begin{pmatrix} 55/73 \\ -15/73 \end{pmatrix}
\end{aligned}
$$

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

Question 4.1

(a)

Let $\hat{\mu}_k = \frac{1}{n}\sum_{i=1}^n X_i^k$.

(b)

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

(c)

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

This is finite if $|\mu_{2k}| < \infty$.

(d)

Let $v\hat{a}r(\bar{\mu}_k) = \frac{1}{n}\left(\frac{1}{n}\sum_{i=1}^{n} X_i^{2k} - \left(\frac{1}{n}\sum_{i=1}^{n} X_i^k\right)^2\right)$

Question 4.2

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

The skew is 0 when the third central moment is 0.

Question 4.3

$\bar{Y}$ is the sample mean of Y, while μ is the true population mean. $\bar{Y}$ is an unbiased estimator of μ. Similarly, $n^{-1}\sum_{i=1}^{n} X_i X_i'$ is the unbiased sample estimator of $\mathbb{E}[X_i X_i']$

Question 4.4

False.

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

In general this expression does not equal 0.

Question 4.5

4.15

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

4.16

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

Question 4.6

Let A be any $n \times k$ unbiased function of X such that $A'X = I_k$. The estimator has variance $Var[A'Y|X] = A'var[e|X]A = A'\Omega A$.

Let $C = A - \Omega^{-1}X(X'\Omega^{-1}X)^{-1}$ and note that $X'C = 0$. Then we have the following:

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

Thus, $A'\Omega A - (X'\Omega^{-1}X)^{-1}$ is positive semidefinite.