20210 Summer 2025: Problem Set 1
Econometric
Instructor(s): Murilo Ramos

Agustín Esteva
aesteva@uchicago.edu
Due Date: July 29, 2025

Problem 1

Suppose $X \sim \text{Bernoulli}(p)$.

(a) Show $\mathbb{E}[X^3] = p$

> Solution: Using the following part, we let $k = 3$ and conclude ♡

(b) Show $\mathbb{E}[X^k] = p$

> Solution: By definition,
>
> $$\begin{aligned}\mathbb{E}[X^k] &= 1^k\mathbb{P}\{X = 1\} + 0^k\mathbb{P}\{X = 0\} \\ &= \boxed{p}\end{aligned}$$
>
> ♡

(c) Suppose that p = 0.3. Compute the mean, variance, skewness, and kurtosis of X.

> Solution: From previous part, $\mathbb{E}[X] = p = \boxed{0.3}$. $\text{Var}[X] = p(1-p) = \boxed{0.21}$. To calculate the skewness, we compute
>
> $$\begin{aligned}\mathbb{E}\left[\left(\frac{X - \mathbb{E}[X]}{\sqrt{\mathbb{V}[X]}}\right)^3\right] &= \frac{1}{\mathbb{V}[X]^{\frac{3}{2}}}\mathbb{E}[(X-\mu)^3] \\ &= \frac{1}{\mathbb{V}[X]^{\frac{3}{2}}}(\mathbb{E}[X^3] - 3\mathbb{E}[X^2]\mathbb{E}[X] + 2\mathbb{E}[X]^3 \\ &= \frac{p - 3p^2 + 2p^3}{(p(1-p))^{\frac{3}{2}}} \\ &= \frac{p(1-p)(1-2p)}{(p(1-p))^{\frac{3}{2}}} \\ &= \frac{1-2p}{\sqrt{1-p}} \\ &\approx \boxed{0.87}\end{aligned}$$

Using similar algebra, we skip a few steps:

$$\mathbb{E}\left[\left(\frac{X-\mathbb{E}[X]}{\sqrt{\mathbb{V}[X]}}\right)^4\right] = \frac{1-6p(1-p)}{p(1-p)} + 3 = \boxed{1.762}$$

♡

In a population, $\mu_Y = 100$ and $\sigma_Y^2 = 43$. Use the Central Limit Theorem to answer the following questions:

(a) In a random sample of size $n = 100$, find

$$\Pr\left(\overline{Y} \le 101\right).$$

Solution: We know by the CLT

$$\frac{\sqrt{100}(\overline{Y}-100)}{\sqrt{43}} \sim N(0,1)$$

Hence

$$\mathbb{P}\{\overline{Y} \le 101\} = \mathbb{P}\{\frac{10(\overline{Y}-100)}{\sqrt{43}} \le \frac{10(101-100)}{\sqrt{43}}\} = \mathbb{P}\{Z \le \frac{10}{\sqrt{43}}\} \approx \boxed{0.94}$$

♡

(b) In a random sample of size $n = 165$, find

$$\Pr\left(\overline{Y} > 98\right).$$

Solution: Using similar logic to the above problem, we find that

$$Z = \frac{\sqrt{165}(\overline{Y}-100)}{\sqrt{43}} \sim N(0,1)$$

so then

$$\mathbb{P}\{Z \ge \frac{\sqrt{165}(98-100)}{\sqrt{43}}\} \approx \boxed{1}$$

♡

(c) In a random sample of size $n = 64$, find

$$\Pr\left(101 \le \overline{Y} \le 103\right).$$

Solution:

$$Z = \frac{8(\overline{Y} - 100)}{\sqrt{43}} \sim N(0,1)$$

and thus

$$\mathbb{P}\{\frac{8(101-100)}{\sqrt{43}} \leq Z \leq \frac{8(103-100)}{\sqrt{43}}\} = \boxed{0.1111}$$

♡

Problem 2

Show that $\mathbb{E}[Y \mid X]$ is minimizes

$$\min_{g(X)} \mathbb{E}[(Y - g(X))^2]$$

Solution: We claim that $Y - \mathbb{E}[Y \mid X] \perp \mathbb{E}[Y \mid X] - g(X)$. To see this, we note that

$$\begin{aligned}
\langle Y - \mathbb{E}[Y \mid X], \mathbb{E}[Y \mid X] - g(X)\rangle &= \langle Y, \mathbb{E}[Y \mid X]\rangle - \langle Y, g(X)\rangle - \langle \mathbb{E}[Y \mid X], \mathbb{E}[Y \mid X]\rangle + \langle \mathbb{E}[Y \mid X], g(X)\rangle \\
&= \mathbb{E}[Y\mathbb{E}[Y \mid X]] - \mathbb{E}[Yg(X)] - \mathbb{E}[\mathbb{E}[Y \mid X]^2] + \mathbb{E}[\mathbb{E}[Y \mid X]g(X)] \\
&= \mathbb{E}[\mathbb{E}[Y\mathbb{E}[Y \mid X] \mid X]] - \mathbb{E}[\mathbb{E}[Yg(X) \mid X]] - \mathbb{E}[\mathbb{E}[Y \mid X]^2] + \mathbb{E}[\mathbb{E}[Y \mid X]g(X)] \\
&= \mathbb{E}[\mathbb{E}[Y \mid X]^2] - \mathbb{E}[g(X)\mathbb{E}[Y \mid X]] - \mathbb{E}[\mathbb{E}[Y \mid X]^2] + \mathbb{E}[g(X)\mathbb{E}[Y \mid X]] \\
&= 0
\end{aligned}$$

Here, we use the fact that $g(X)$ is $X-$measurable, and thus we pull it out of the conditional expectation. We also made heavy use of LIE. By orthogonality, we can use the Pythagorean theorem

Computing,

$$\begin{aligned}
\mathbb{E}[(Y - g(X))^2] &= \mathbb{E}[(Y - \mathbb{E}[Y \mid X] + \mathbb{E}[Y \mid X] - g(X))^2] \\
&= \|(Y - \mathbb{E}[Y \mid X]) + (\mathbb{E}[Y \mid X] - g(X))\|^2 \\
&= \|Y - \mathbb{E}[Y \mid X]\|^2 + \|\mathbb{E}[Y \mid X] - g(X)\|^2
\end{aligned}$$

Clearly, the left hand side will be minimized when

$$\|\mathbb{E}[Y \mid X] - g(X)\|^2 = 0 \iff \mathbb{E}[Y \mid X] - g(X) = 0 \iff g(X) = \mathbb{E}[Y \mid X]$$

♡

Problem 3

Prove that

$$\mathrm{Cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$$

SOLUTION: By definition,

$$\begin{aligned}
\mathrm{Cov}(X, Y) &\equiv \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])] \\
&= \mathbb{E}[XY - Y\mathbb{E}[X] - X\mathbb{E}[Y] + \mathbb{E}[X]\mathbb{E}[Y] \\
&= \mathbb{E}[XY] - \mathbb{E}[Y\mathbb{E}[X]] - \mathbb{E}[X\mathbb{E}[Y]] + \mathbb{E}[\mathbb{E}[X]\mathbb{E}[Y]] \\
&= \mathbb{E}[XY] - \mathbb{E}[Y]\mathbb{E}[X] + (-\mathbb{E}[X]\mathbb{E}[Y] + \mathbb{E}[X]\mathbb{E}[Y]) \\
&= \mathbb{E}[XY] - \mathbb{E}[Y]\mathbb{E}[X]
\end{aligned}$$

Here we make heavy use of the fact that $\mathbb{E}[Z]$ is a constant and can thus be pulled out of the expectation. ♡

Problem 4

Suppose that in the State of Illinois the written exam for a drivers license consists of 10 multiple-choice questions. Each question has 4 possible choices, only one of which is correct. Passing requires answering at least 5 questions correctly. What is the probability a student-driver passes his exam by "guessing randomly" on each question?

SOLUTION: Let $X \sim \text{Binomial}(10, \frac{1}{4})$. It should be clear that the solution to this problem is the same as finding

$$\mathbb{P}[X \geq 5] = 1 - (\sum_{k=0}^{4} \binom{10}{k} (\frac{1}{4})^k (\frac{3}{4})^{10-k}) = \boxed{0.0781}$$

where the numerical answer was found using software.

♡

Problem 5

Let $X \sim \text{Bernoulli}(p)$. Define $Z = 3^X - 1$.

- Is Z a random variable. Why?

SOLUTION: Yes, functions of random variabls are clearly random variables. To be precise, we assum that $(\Omega, \mathcal{F}, \mathbb{P})$ is probability space, then $X : \Omega \to \mathbb{R}$ is $\mathcal{F}$ measurable. Since $Y = g(X)$ where g is continuous, then clearly g is $\mathcal{F}$ measurable and thus a random variable. ♡

- Show that $\mathbb{E}[Z] = 2p$.

SOLUTION:

$$\begin{aligned}
\mathbb{E}[Z] &= \mathbb{E}[3^X - 1] \\
&= \mathbb{E}[3^X] - 1 \\
&= 3^1 p + 3^0(1-p) - 1 \\
&= 3p + 1 - p - 1 \\
&= 1 + 2p - 1 \\
&= \boxed{2p}
\end{aligned}$$

♡

- Show that $\mathbb{E}[Z^2] = 4p$

SOLUTION: We use work from the previous part to find that

$$\begin{aligned}
\mathbb{E}[Z^2] &= \mathbb{E}[(3^X - 1)^2] \\
&= \mathbb{E}[3^{2X} - 2 \cdot 3^X + 1] \\
&= \mathbb{E}[3^{2X}] - 2\mathbb{E}[3^X] + 1 \\
&= 1 + 2p - 2(1 + 2p) + 1 \\
&= \boxed{4p}
\end{aligned}$$

♡

- Find $\text{Var}(Z)$

SOLUTION: We use the identity and the previous problems to find that

$$\text{Var}(Z) = \mathbb{E}[Z^2] - \mathbb{E}[Z]^2 = \boxed{4p - 4p^2}$$

♡

Problem 6

Let $\texttt{GPA}$ denote a random variable for a college student's grade point average, and $\texttt{SAT}$ denote a random variable for the college student's SAT score. Suppose that there is the following relationship between $\texttt{GPA}$ and $\texttt{SAT}$:

$$\mathbb{E}[\texttt{GPA} \mid \texttt{SAT}] = 0.70 + 0.002 \cdot \texttt{SAT}.$$

(a) What is the expected GPA when $\texttt{SAT} = 750$? What is the expected GPA when $\texttt{SAT} = 1500$?

SOLUTION: Plugging in,

$$\mathbb{E}[\texttt{GPA} \mid 750] = 0.70 + 0.002 \cdot 750 = \boxed{2.20}$$

$$\mathbb{E}[\texttt{GPA} \mid 1500] = 0.70 + 0.002 \cdot 1500 = \boxed{3.70}$$

♡

(b) If $\mathbb{E}[\texttt{SAT}] = 1000$, what is $\mathbb{E}[\texttt{GPA}]$?

SOLUTION: Using the Law of Total Expectation:

$$\mathbb{E}[\texttt{GPA}] = \mathbb{E}[\mathbb{E}[\texttt{GPA} \mid \texttt{SAT}] = \mathbb{E}[0.70 + 0.002 \cdot \texttt{SAT}] = 0.70 + 0.002\mathbb{E}[\texttt{SAT}] = \boxed{2.70}$$

♡

Problem 7

Suppose $X \sim \text{Unif}(-1,1)$, and let $Y = X^2$. Show that $\text{Cov}[X,Y] = 0$, but that X and Y are not independent.

SOLUTION: By Problem 3,

$$\text{Cov}(X,Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = \mathbb{E}[X^3] - \mathbb{E}[X]\mathbb{E}[X^2]$$

We compute, noting that the pdf of X is

$$f_X(x) = \frac{1}{2}$$

$$\mathbb{E}[X] = \frac{-1+1}{2} = 0$$

$$\mathbb{E}[X^2] = \int_{\mathbb{R}} X^2 \, d\mathbb{P} = \int_{-1}^{1} \frac{1}{2}x^2 \, dx = \frac{1}{3}$$

$$\mathbb{E}[X^3] = \int_{\mathbb{R}} X^3 \, d\mathbb{P} = \int_{-1}^{1} \frac{1}{2}x^3 \, dx = \frac{1}{4}x^4\Big|_{-1}^{1} = \frac{1}{4} - \frac{1}{4} = 0$$

Hence, the covariance is zero. To see that X and Y are not independent, simply compare the joint densities. We know that X and Y are independent iff

$$f_{X,Y}(x,y) = f_X(x) f_Y(y) \tag{1}$$

We know that $f_X(x) = \frac{1}{2}$. Moreover, supposing that $y \in (-1,1)$,

$$F_Y(y) = \mathbb{P}\{Y \le y\} = \mathbb{P}\{X^2 \le y\} = \mathbb{P}\{-\sqrt{y} \le X \le \sqrt{y}\} = \mathbb{P}\{X \le \sqrt{y}\} - \mathbb{P}\{X \le -\sqrt{y}\} = \sqrt{y}.$$

Hence, $f_Y(y) = \frac{1}{2\sqrt{y}}$. Thus the right hand side of (1) is $\frac{1}{4\sqrt{y}}$. To compute the left hand side, first we compute the conditional density

$$f_{Y|X}(y)$$

♡

Problem 8

Let X be a random variable such that

$$\mathbb{P}(X=-1)=\mathbb{P}(X=0)=\mathbb{P}(X=1)=\frac{1}{3}.$$

Let Y be another random variable such that

$$\mathbb{E}[Y \mid X]=3+3X \quad \text{and} \quad \operatorname{Var}(Y \mid X)=3.$$

For each question below, provide a numerical answer and show your work.

(a) What is $\mathbb{E}[X]$?

SOLUTION:

$$\mathbb{E}[X]=-1\cdot\frac{1}{3}+0\cdot\frac{1}{3}+1\cdot\frac{1}{3}=\boxed{0}$$

♡

(b) What is $\mathbb{E}[X^2]$?

SOLUTION:

$$\mathbb{E}[X^2]=2(1\cdot\frac{1}{3})+0\cdot\frac{1}{3}=\boxed{\frac{2}{3}}$$

♡

(c) What is $\mathbb{E}[Y]$?

SOLUTION:

$$\mathbb{E}[Y]=\mathbb{E}[\mathbb{E}[Y \mid X]]=\mathbb{E}[3+3X]]=3+3\mathbb{E}[X]=\boxed{3}$$

♡

(d) What is $\operatorname{Var}[2+\mathbb{E}[X]]$?

SOLUTION: Bro what the variance of a number is just zero. ♡

(e) What is $\operatorname{Var}[X]$?

SOLUTION:

$$\operatorname{Var}(X)=\mathbb{E}[X^2]-\mathbb{E}[X]^2=\boxed{\frac{2}{3}}$$

♡

(f) What is $\operatorname{Var}[Y]$?

SOLUTION:

$$\begin{aligned}
\mathrm{Var}(Y) &= \mathbb{E}[\mathrm{Var}(Y \mid X)] + \mathrm{Var}(\mathbb{E}[Y \mid X]) \\
&= \mathbb{E}[3] + \mathrm{Var}(3 + 3X) \\
&= 3 + 9 \cdot \mathrm{Var}(X) \\
&= \boxed{9}
\end{aligned}$$

♡

(g) What is $\mathrm{Cov}[X, Y]$?

SOLUTION:

$$\begin{aligned}
\mathrm{Cov}(X, Y) &= \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] \\
&= \mathbb{E}[3X + 3X^2] - \mathbb{E}[X]\mathbb{E}[Y] \\
&= 3\mathbb{E}[X^2] \\
&= \boxed{2}
\end{aligned}$$

♡

(h) What is $\mathrm{Corr}[X, Y]$?

SOLUTION:

$$\mathrm{Corr}(X, Y) = \frac{\mathrm{Cov}(X, Y)}{\sqrt{\mathrm{Var}(X)}\sqrt{\mathrm{Var}(Y)}} = \frac{2}{\sqrt{6}}$$

♡

(i) Is Y mean independent of X? Explain briefly.

SOLUTION: Nah. Assume, for the sake of contradiction, that it is mean independent. Then $\mathrm{Corr}(X, Y) = 0$. This is a contradiction to part h. ♡

(j) Is Y independent of X? Explain briefly.

SOLUTION: Nah, same reason as part i. ♡