Problem Set 2[1]
Math Camp 2024, UCSB
Instructor: Seonmin (Will) Heo (sheo@ucsb.edu)
Due: 11:59 PM on Wednesday, Sep 4, 2024

This problem set will help you review the key concepts from the course so far. You are free to use any software but make sure to type the answers and include the code and the results.

1. Prove that

   (a) $S \in \mathcal{B}$ and $\emptyset \in \mathcal{B}$ in a sample space S with a σ-algebra $\mathcal{B}$ on S.
   (Hint: start with that $\mathcal{B}$ should be nonempty.)

   - $S \in \mathcal{B}$

$$\begin{aligned}
\mathcal{B} \neq \emptyset \quad &\Rightarrow \quad E \in \mathcal{B} && (\mathcal{B} \text{ is nonempty})\\
&\Rightarrow \quad E^c \in \mathcal{B} && (\text{closed under complements})\\
&\Rightarrow \quad E \cup E^c \in \mathcal{B} && (\text{closed under countable unions})\\
&\Rightarrow \quad S \in \mathcal{B}
\end{aligned}$$

   - $\emptyset \in \mathcal{B}$

$$S \in \mathcal{B} \quad \Rightarrow \quad \emptyset \in \mathcal{B} \qquad (\text{closed under complements})$$

   (b) σ-algebra $\mathcal{B}$ on S is closed under countable intersections.

   - $\mathcal{B}$ is closed under countable intersections.

$$\begin{aligned}
E_1, E_2, \cdots \in \mathcal{B} \quad &\Rightarrow \quad E_1^c, E_2^c, \cdots \in \mathcal{B} && (\text{closed under complements})\\
&\Rightarrow \quad \bigcup_{i=1}^{\infty} E_i^c \in \mathcal{B} && (\text{closed under countable unions})\\
&\Rightarrow \quad \left(\bigcap_{i=1}^{\infty} E_i\right)^c \in \mathcal{B} && (\text{DeMorgan's Laws})\\
&\Rightarrow \quad \bigcap_{i=1}^{\infty} E_i \in \mathcal{B} && (\text{closed under complements})
\end{aligned}$$

2. Prove Theorem 2.3.

   - $\mathbb{P}(A^c) = 1 - P(A)$

$$\begin{aligned}
\mathbb{P}(S) &= \mathbb{P}(A) + \mathbb{P}(A^c) && (S = A \cup A^c \quad \wedge \quad A \cap A^c = \emptyset)\\
1 &= \mathbb{P}(A) + \mathbb{P}(A^c) && (\mathbb{P}(S) = 1)
\end{aligned}$$

[1] This problem set was created by Woongchan Jeon and modified by Seonmin (Will) Heo.

- $\mathbb{P}(A) \leq 1$

$$\begin{aligned} 1 &= \mathbb{P}(A) + \mathbb{P}(A^c) \\ \mathbb{P}(A) &= 1 - \mathbb{P}(A^c) \\ \mathbb{P}(A) &\leq 1 && (\mathbb{P} : \mathcal{B} \mapsto [0, \infty)) \end{aligned}$$

- $\mathbb{P}(B \cap A^c) = \mathbb{P}(B) - \mathbb{P}(A \cap B)$

$$\mathbb{P}(B) = \mathbb{P}(B \cap A^c) + \mathbb{P}(B \cap A)$$
$$(B = (B \cap A^c) \cup (B \cap A) \quad \wedge \quad (B \cap A^c) \cap (B \cap A) = \emptyset)$$

- $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B \cap A^c) \qquad (A \cup B = A \cup (B \cap A^c) \quad \wedge \quad A \cap (B \cap A^c) = \emptyset)$$

- If $A \subseteq B$, then $\mathbb{P}(A) \leq \mathbb{P}(B)$

$$\begin{aligned} \mathbb{P}(B \cap A^c) &= \mathbb{P}(B) - \mathbb{P}(A \cap B) \\ \mathbb{P}(B \cap A^c) &= \mathbb{P}(B) - \mathbb{P}(A) && (A \subseteq B) \\ 0 &\leq \mathbb{P}(B) - \mathbb{P}(A) && (\mathbb{P} : \mathcal{B} \mapsto [0, \infty)) \end{aligned}$$

3. Consider testing for the presence of a disease. The test is very accurate in the sense that if a patient has the disease, the test always comes back positive, i.e., $\mathbb{P}(\text{positive}|\text{disease}) = 1$. Sometimes the test is inaccurate, however, in the sense that the test gives a false positive (a positive value when a person doesn't have the disease) with probability 0.005, i.e., $\mathbb{P}(\text{positive}|\text{no disease}) = 0.005$. If the probability of having the disease is 0.001, i.e., $\mathbb{P}(\text{disease}) = 0.001$, what is the probability a patient has the disease, given they have a positive test?

   Recall the Bayes' rule:

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(B|A)\mathbb{P}(A)}{\mathbb{P}(B)} = \frac{\mathbb{P}(B|A)\mathbb{P}(A)}{\sum_{j \in I} \mathbb{P}(B|A_j)\mathbb{P}(A_j)}.$$

   Assign the event of having the disease to A and the event of a positive test result to B. We first use the Law of Total Probability to obtain the probability of having a positive result, $\mathbb{P}(B)$:

$$\begin{aligned} \mathbb{P}(\text{positive}) &= \mathbb{P}(\text{disease})\mathbb{P}(\text{positive}|\text{disease}) + \mathbb{P}(\text{no disease})\mathbb{P}(\text{positive}|\text{no disease}) \\ &= 0.001 \times 1 + 0.999 \times 0.005 = 0.005995. \end{aligned}$$

   Then by Bayes' rule, we have that:

$$\begin{aligned} \mathbb{P}(\text{disease}|\text{positive}) &= \frac{\mathbb{P}(\text{positive}|\text{disease})\mathbb{P}(\text{disease})}{\mathbb{P}(\text{positive})} \\ &= \frac{1 \times 0.001}{0.005995} \approx 0.1668 \end{aligned}$$

4. A variable X is lognormally distributed if $Y = \ln(X)$ is normally distributed with μ and σ^2, i.e. $f_Y(y) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2}$. Let $x = g(y) = e^y$ and $y = g^{-1}(x) = \ln(x)$.

   (a) Derive $f_X(x)$.

$$f_X(x) = f_Y\left(g^{-1}(x)\right)\left|\frac{dg^{-1}(x)}{dx}\right| = \frac{1}{\sqrt{2\pi\sigma^2}x} e^{-\frac{1}{2}\left(\frac{\ln(x)-\mu}{\sigma}\right)^2}$$

   (b) Derive $\mathbb{E}[X^t]$ using $M_Y(t)$. What are $\mathbb{E}[X]$ and $V(X)$?

   We know that $M_Y(t) = e^{\mu t + \frac{1}{2}\sigma^2 t^2}$. Then

$$\begin{aligned}
\mathbb{E}[X^t] &= \mathbb{E}[e^{tY}] = M_Y(t) = e^{\mu t + \frac{1}{2}\sigma^2 t^2} \\
\mathbb{E}[X] &= e^{\mu + \frac{1}{2}\sigma^2} \\
\mathbb{E}[X^2] &= e^{2\mu + 2\sigma^2} \\
V(X) &= (e^{\sigma^2} - 1) \cdot e^{2\mu + \sigma^2}
\end{aligned}$$

5. In class, we have seen the proof of the Law of Iterated Expectations in the discrete case. Suppose that $\mathbb{E}[Y] < \infty$. Prove the following in the continuous case.

   (a) $\mathbb{E}[Y] = \mathbb{E}\Big[\mathbb{E}[Y|X]\Big]$

$$\begin{aligned}
\mathbb{E}[Y] &= \int y f(y) dy \\
&= \int y \left(\int f(y,x) dx\right) dy \\
&= \int y \left(\int f(y|x) f(x) dx\right) dy \\
&= \int \left(\int y f(y|x) dy\right) f(x) dx \\
&= \int \mathbb{E}[Y|X = x] f(x) dx \\
&= \mathbb{E}[\mathbb{E}[Y|X]],
\end{aligned}$$

   where the second line uses the fact that the joint probability is the sum of all the marginal probabilities, the third line uses the fact that the joint probability is equal to the product of the conditional probability and the marginal probability, the fourth line uses the fact that the ordering in double integration can be switched, and the fifth line follows from the definition for the conditional expectation.

(b) $\mathbb{E}[Y|X] = \mathbb{E}\Bigg[\mathbb{E}[Y|X,Z] \mid X\Bigg]$

Note that

$$\mathbb{E}[Y|X=x, Z=z] = \int_{\mathbb{R}} y f(y|x,z) dy.$$

In addition, note that

$$f(y|x,z)f(z|x) = \frac{f(y,x,z)}{f(x,z)} \frac{f(x,z)}{f(x)} = \frac{f(y,x,z)}{f(x)} = f(y,z|x).$$

Then we find that

$$\begin{aligned}
\mathbb{E}\Bigg[\mathbb{E}[Y|X,Z] \mid X\Bigg] &= \int \mathbb{E}[Y|X=x, Z=z]\, f(z|x) dz \\
&= \int \left(\int y f(y|x,z) dy\right) f(z|x) dz \\
&= \int \int y f(y|x,z) f(z|x) dy dz \\
&= \int \int y f(y,z|x) dy dz \\
&= \mathbb{E}[Y|X].
\end{aligned}$$

6. Assume that the n units are volunteers to receive the treatment. Given any $i \in \{1, \cdots, n\}$, $D_i = 1$ if treated, and $D_i = 0$ otherwise. Let $(D_1, \cdots, D_n)$ be a vector stacking the treatment indicators of all units. Treatments have capacity constraints and only $n_1 (< n)$ units can be treated: $\sum_{i=1}^n D_i = n_1$.

   (a) What is the number of possible values $(D_1, \cdots, D_n)$ can take?
   We are choosing n_1 out of n units, unordered, so there are $\binom{n}{n_1}$ possible ways.

   We say that treatment is randomly assigned if $(D_1, \cdots, D_n)$ are random variables, and if for any vector of n numbers $(d_1, \cdots, d_n) \in \{0,1\} \times \cdots \times \{0,1\}$ such that $\sum_{i=1}^n d_i = n_1$,

$$P(D_1 = d_1, \cdots, D_n = d_n) = \frac{1}{\binom{n}{n_1}}$$

   That is, random assignment generates uniform treatment probabilities across units.

   (b) If full randomization is satisfied, then for every $i \in \{1, \cdots, n\}$, what is $P(D_i = 1)$?

$$P(D_i = 1) = \frac{\binom{n-1}{n_1-1}}{\binom{n}{n_1}} = \frac{n_1}{n}$$

   (c) If full randomization is satisfied, then for every $i \neq j$, what is $P(D_i = 1 \ \wedge \ D_j = 1)$?

Is it true that unit i getting treated is independent from unit j getting treated?

$$P(D_i = 1 \ \wedge \ D_j = 1) = \frac{\binom{n-2}{n_1-2}}{\binom{n}{n_1}} = \frac{n_1(n_1-1)}{n(n-1)}.$$

Note that $P(D_i = 1 \ \wedge \ D_j = 1) \neq P(D_i = 1) P(D_j = 1)$ and $P(D_i = 1 \mid D_j = 1) \neq P(D_i = 1)$. The intuition is that if $D_j = 1$, D_i is less likely to be equal to 1 than if $D_j = 0$. If $D_j = 1$, then there are only $n_1 - 1$ treatment seats left for $n-1$ units, while if $D_j = 0$, then there are still n_1 treatment seats left for $n-1$ units.

7. Download the `lbw.dta` dataset from Stata.

   (a) Run a linear regression of low birthweight (`bwt`) on smoking (`smoke`), controlling for age (`age`) weight at the last menstrual period (`lwt`), and history of hypertension (`ht`). Report the coefficients, standard errors, and confidence intervals.

   The coefficients, standard errors, and confidence intervals are provided below in Table 1.

Table 1: The OLS regression result

	(1) birthweight
Smoked during pregnancy	-261.851
	$[-467.160, -56.542]$
	(0.013)
Age of mother	5.451
	(0.578)
Weight at last menstrual period	5.169
	(0.003)
Has history of hypertension	-579.031
	(0.008)
Constant	2285.929
	(0.000)
Observations	189
R^2	0.106

   (b) Estimate the bootstrapped standard error for this OLS. Use the following procedure:

      i. Get a bootstrapped sample, which is a random sample *with replacement*.

      ii. Run the OLS based using this bootstrapped sample b_1. Let's call the coefficient on smoking $\widehat{\beta}_{b_1}$.

      iii. Repeat the process (i)-(ii) $B = 1000$ times.

      iv. Calculate the bootstrapped variance using the following formula:

$$\widehat{\text{Var}}(\hat{\theta}) = \frac{1}{B-1} \sum_{b=1}^{B} \left(\widehat{\theta}_b - \bar{\widehat{\theta}}_b \right)^2.$$

v. Construct confidence intervals by obtaining the 97.5th and the 2.5th percentiles of the B bootstrapped coefficients $(\widehat{\beta}_{b_1}, \ldots, \widehat{\beta}_{b_B})$. Are the bootstrapped confidence intervals similar to those obtained in (a)?

Table 2: Bootstrap results of the OLS regression

	(1) birthweight
Smoked during pregnancy	-261.851
	$[-467.510, -56.192]$
	(0.013)
Observations	189

Yes, the bootstrapped confidence intervals are similar to those obtained in (a). The bootstrapped standard error is slightly smaller than the normal standard error.