LPS ECON-103-601, Fall 2014

Midterm 10.1.14

Exam Instructions:

- You will have two hours to complete this exam (17.00-19.00).
- There are 120 marks. Marks are indicated on each question.
- No writing is permitted once exam time ends.
- No calculators or notes are permitted.
- Write your name on each bluebook. Show all work clearly and labelled in your bluebooks. Use one bluebook as scratch paper. Do not write on this exam. Submit all bluebooks and this exam.
- Clearly indicate your final answer, but marks will be awarded based on a correct solution. Answers without solutions will not receive full credit.
- Your solutions or proofs should justify the steps taken and clearly explain the logic used in each step.
- State all theorems, propositions or definitions you apply. Prove any propositions you use not presented in class.

Notice: There are some easy, medium, and hard questions. Marks are suggestive of how to allocate your time (1 min per mark). Illegible writing or failure to follow instructions will result in a loss of at least 20 marks and violations of academic integrity will result in a 0. No exams written in pencil or erasable ink will be considered for regrade.

Q1 [20 marks]

(a) [2 marks]

In the UK, a postal code consists of four letters and two numbers (e.g. SE8 4QD). How many unique postal codes are possible? The first number in a postal code cannot be zero.

By the Fundamental Counting Principle,

$$\underline{24}^4 \cdot \underline{9} \cdot \underline{10} = 90 \cdot 24^4$$

(b) [4 marks]

Selima and Josh's Ice Cream is bankrupt! They used to have 33 flavors of ice cream but now can only afford 7. Fortunately, they still offer a special cup that allows the buyer to mix and match any of the flavours. The buyer can repeat flavours as long as the total number of scoops is exactly four. How many different cups of ice cream can one customer possibly order with the special cup? Express your answer as an integer.

This is a combination with replacement, so ${}_{7+4-1}C_4 = {}_{10}C_4 = \frac{10!}{6! \cdot 4!} = 10 \cdot 3 \cdot 7 = 210$

(c) [6 marks]

Felix drove a car load of statistics review journals to the university library. He needs at least four people to carry the boxes into the library in just one trip. He finds Luke, John, Tom, Sean, and three other friends studying in the library. In how many ways can he select *at least* four of the students to help carry the boxes of statistics journals? Express your answer as an integer.

This is just the total number of combinations. Felix can choose 4, 5, 6, or 7 students to help, hence

$${}_7C_4 + {}_7C_5 + {}_7C_6 + {}_7C_7 = 35 + 21 + 7 + 1 = 64$$

(d) [8 marks]

Aziz places 7 beads on a bracelet chain, then tied the ends together. The beads are identical except for color. If 3 of the beads are red, 2 are green, and 2 are blue, how many distinct bracelets are possible? Two bracelets are identical if one can be rotated and/or flipped to produce the other. Also, beads may pass over the knot created by tying the two ends together. Express your answer as an integer.

There are $\frac{7!}{2!2!3!}$ ways to arrange the beads distinctly in a row. There are $\frac{1}{7} \cdot \frac{7!}{2!2!3!}$ ways to arrange them in a circle, because we need to discount from rotating the bracelet. This leaves 30 possible bracelets. If we have a red bead in the middle, then we can have a symmetric pattern BRGRGRB which can be formed $3! = 6$ ways. Flipping the bracelet for these six ways does not produce a new bracelet pattern, because the bracelet pattern is symmetric. There are 24 possible bracelets left. We can flip each these bracelets to get another bracelet, so there are really only $24/2 = 12$ distinct bracelets in this case. So the total number of distinct bracelets is $6 + 12 = 18$.

Q2 [30 marks]

(a) [5 marks]

Consider the following data $\{100, 50, 25, 12.5, 6.25, 3.125\}$

- [2 marks] Find the mean of the data. You may use the fact that $1 - (0.5^6) \approx 1$.
- [2 marks] Find the median of the data.
- [1 mark] In direction what is the data skewed? Why?

First note that the data is a geometric sequence with $a_1 = 100$, $n = 6$ and $r = 0.5$. So the sum of this seqeuence is

$$S_n = \frac{100(1-(0.5)^6)}{1-0.5} \approx 200 \Rightarrow mean \approx 200/6 = 33.\overline{3}$$

Now we know the median is $\frac{25+12.5}{2} = \frac{37.5}{2} = 18.75$. Since $median < mean$, the data is right or positively skewed.

(b) [5 marks]

What are two problems with making inference from observational data? Give an example for each of the problems you list.

Selection Bias, Non-response, Behavioural changes. See Chapter 2 for an example.

(c) [20 marks]

Statistically, the average social circle of a white American is only 1% black. Is this evidence of racism? Why or why not? Define any terms you use in your argument. Marks will be awarded based on the economic and statistical content of your answer and its coherence. Plan your response, because an incoherent response will receive zero marks even if it contains relevant content. Responses longer than 3 bluebook pages or using small font will be penalized so be concise.

The problems listed above may make the evidence unreliable. You should discuss how these problems come into play and possible confounders as we discussed in class. A complete answer will define a cofounder. An obvious confounder is residential segregation. We usually do not interact with people we do not live by. However, this residential segregation itself may be considered evidence of racism. Your answer should exhibit this type of common sense, because it shows you actually thought about the question rather than just went the motions of citing the above problems.

Q3 [25 marks]

(a) [5 marks]

If Scarlett selects six cards at random (without replacement) from a standard deck of 52 cards, what is the probability there will be no pairs?

We want the probability of drawing 6 cards from different ranks, which is

$$\frac{52}{52} \cdot \frac{48}{51} \cdot \frac{44}{50} \cdot \frac{40}{49} \cdot \frac{36}{48} \cdot \frac{32}{47}$$

(b) [5 marks]

Dustin and Paul-Julien take turns tossing a fair six-sided die. Dustin goes first. The winner is the first player to throw a 4. What is the probability that Paul-Julien wins?

$P(\textbf{PJ wins}) = 1 - P(\textbf{Dustin wins})$. **The probability of Dustin winning is the probability that he rolls a four on his turn. Suppose Dustin rolls a 4 on his** 3^{rd} **try. It must be the case that Paul-Julien missed his first two tries along with Dustin, which are independent events that happen with probability** $5/6$**. So**

$$P(\textbf{Dustin wins on } 1^{st} \textbf{ try}) = \frac{1}{6}$$

$$P(\textbf{Dustin wins on } 2^{nd} \textbf{ try}) = \left(\frac{5}{6}\right)\left(\frac{5}{6}\right)\frac{1}{6}$$

$$P(\textbf{Dustin wins on } 3^{rd} \textbf{ try}) = \left(\frac{5}{6}\right)\left(\frac{5}{6}\right)\left(\frac{5}{6}\right)\left(\frac{5}{6}\right)\frac{1}{6}$$

$$\vdots$$

$$P(\textbf{Dustin wins on } n^{th} \textbf{ try}) = \underbrace{\left(\frac{5}{6}\right)\left(\frac{5}{6}\right)\left(\frac{5}{6}\right) \cdot \ldots \cdot \left(\frac{5}{6}\right)}_{=\left(\frac{5}{6}\right)^{2(n-1)}}\frac{1}{6}$$

Since winning on any given try are all independent events,

$$\begin{aligned}
P(\textbf{Dustin wins}) &= P(\textbf{Dustin wins } 1^{st} \textbf{ try}) + P(\textbf{Dustin wins } 2^{nd} \textbf{ try}) + ... \\
&= \frac{1}{6} + \left(\frac{5}{6}\right)\left(\frac{5}{6}\right)\frac{1}{6} + \left(\frac{5}{6}\right)\left(\frac{5}{6}\right)\left(\frac{5}{6}\right)\left(\frac{5}{6}\right)\frac{1}{6} + ... \\
&= \frac{1}{6}\left(1 + \left(\frac{5}{6}\right)^2 + \left(\frac{5}{6}\right)^4 + \left(\frac{5}{6}\right)^6 + ...\right) \\
&= \frac{1}{6}\left(\frac{1}{1-\frac{5^2}{6^2}}\right) \Rightarrow P(\textbf{PJ wins}) = 1 - \frac{1}{6}\left(\frac{1}{1-\frac{5^2}{6^2}}\right)
\end{aligned}$$

(c) [3 marks]

Define a conditional probability space $(Y, \mathcal{B}_Y, P_Y)$.

See Chapter 3 of lecture notes

(d) [2 marks]

State and prove Bayes' Rule.

See Chapter 3 of lecture notes

(e) [3 marks]

Consider a probability space $(X, \mathcal{B}, P)$. State the restrictions we impose on $P : \mathcal{B} \rightarrow [0, 1]$.

See Chapter 3 of lecture notes

(f) [2 marks]

Prove $P(A^c) = 1 - P(A)$ from the properties stated in (e).

<u>Proof</u>: We know by definition,

$$\begin{aligned}
P(X) &= 1 \\
P(X) = P(A \cup A^c) &= 1
\end{aligned}$$

By additive countability, we know

$$P(A \cup A^c) = P(A) + P(A^c) = 1 \Rightarrow P(A^c) = 1 - P(A) \ \blacksquare$$

(g) [5 marks]

Suppose $P(A) = 0.3$, $P(B) = 0.5$, $P(A^c|B^c) = 0.8$, and $P(B^c|A^c) = \frac{5}{7}$.
What is $P((A \cap B^c) \cup (A^c \cap B))$?

First note that $P((A \cap B^c) \cup (A^c \cap B)) = P(A \cap B^c) + P(A^c \cap B)$, because the sets are disjoint. From $P(A^c|B^c)$ and $P(B^c|A^c)$, we know $P(A|B^c) = 0.2$ and $P(B|A^c) = \frac{2}{7}$. From the defintion of conditional probability, we know that $P(A|B^c) = \frac{P(A \cap B^c)}{P(B^c)}$ $\left(0.2 = \frac{P(A \cap B^c)}{0.5}\right)$ and $P(B|A^c) = \frac{P(A^c \cap B)}{P(A^c)}$ $\left(\frac{2}{7} = \frac{P(A \cap B^c)}{0.7}\right)$. Now we just solve for $P(A \cap B^c)$ $(= 0.1)$ and $P(A^c \cap B)$ $(= 0.2)$ to see that $P((A \cap B^c) \cup (A^c \cap B)) = 0.3$.

Q4 [20 marks]

(a) [4 marks]

Kathy has two puppies. Joyce tells us that one puppy is a girl, Sassy. What is the probability that Sassy's sibling is a girl?

$$P(\{GG\}|G) = \frac{P(\{GG\} \cap G)}{P(\{GG\}) + P(\{BG\}) + P(\{GB\})} = \frac{1/4}{3/4} = \frac{1}{3}$$

(b) [6 marks]

Kathy has two puppies. She picks one of them at random and brings it to class. She brings a girl puppy named Sassy. What is the probability that Sassy's sibling is a girl?

Let G' be event that the girl is chosen randomly. Note that this event is different from the previous event (G), because $G' \subseteq G$ but not $G \subseteq G'$. Just because Kathy has a girl puppy does not mean that Joyce told us she has a girl randomly.

$$\begin{aligned}
P(\{GG\}|G') &= \frac{P(\{GG\} \cap G')}{P(G')} \\
&= \frac{P(G'|\{GG\})}{P(G'|\{GG\})P(\{GG\}) + P(G'|\{BG\})P(\{BG\}) + P(G'|\{GB\})P(\{GB\})} \\
&= \frac{1/4}{1/2} = \frac{1}{2}
\end{aligned}$$

(c) [10 marks]

There are 10 cups on Uzair's desk: N_B of them are Blue Cups and the remaining $10 - N_B$ are Red Cups. Each cup contains five balls: Blue Cups contain 4 blue balls and 1 red ball while Red Cups contain 4 red balls and 1 blue ball. Lindsey choses a cup at random so that each cup is equally likely to be selected. Then Bill draws three balls at random with replacement from the chosen cup. Let C_B be the event that he chose a Blue Cup and let RRB be the event that represents the three draws: a red ball, followed by another red ball, followed by a blue ball.

- [6 marks] Calculate $P(C_B|RRB)$.
- [2 marks] How large would N_B have to be for it to be more likely that Bill drew from a blue cup given that the event RRB has occurred? Prove your answer.
- [2 marks] Suppose that he made draws without replacement. What is $P(C_B|RRB)$ in this case? Explain your answer.

<u>Solution:</u> By Bayes' Rule and the Law of Total Probability,

$$\begin{aligned} P(C_B|RRB) &= \frac{P(C_B)P(RRB|C_B)}{P(RRB)} \\ &= \frac{P(C_B)P(RRB|C_B)}{P(RRB|C_B)P(C_B) + P(RRB|C_R)P(C_R)} \end{aligned}$$

Now,

$$\begin{aligned} P(C_B) &= \frac{N_B}{10} \\ P(C_R) &= 1 - \frac{N_B}{10} = \frac{10 - N_B}{10} \\ P(RRB|C_B) &= \frac{1}{5} \cdot \frac{1}{5} \cdot \frac{4}{5} = \frac{4}{5^3} \\ P(RRB|C_R) &= \frac{4}{5} \cdot \frac{4}{5} \cdot \frac{1}{5} = \frac{4^2}{5^3} \end{aligned}$$

Hence,

$$\begin{aligned} P(C_B|RRB) &= \frac{\frac{N_B}{10} \cdot \frac{4}{5^3}}{\frac{4}{5^3}\left(\frac{N_B}{10}\right) + \frac{4^2}{5^3}\left(\frac{10-N_B}{10}\right)} \\ &= \frac{N_B}{N_B + 4(10 - N_B)} \\ &= \frac{1}{\frac{40}{N_B} - 3} \end{aligned}$$

How large would N_B have to be for it to be more likely that Bill drew from a blue cup given that the event RRB has occurred? We want N_B to be such that $N_B \leq 10$ and

$$\frac{1}{\frac{40}{N_B} - 3} > \frac{1}{2} \Leftrightarrow N_B > 8 \Rightarrow N_B = 9 \text{ at least.}$$

Suppose that he made draws without replacement. What is $P(C_B|RRB)$ in this case? There is only 1 red ball in the blue cup, so it is not possible to draw RRB without replacement, hence $P(C_B|RRB) = 0$ in this case.

Q5 [25 marks]

(a) [3 marks]

Prove that for a discrete random variable X with a probability mass function $p(x)$, $Var(X) = \sum_{i=1}^{n} x_i^2 \cdot p(x_i) - \mu^2$ where $\mu = \mathbb{E}[X]$.

Proof:

$$
\begin{aligned}
Var(X) &= \mathbb{E}\left[(X-\mu)^2\right] \\
&= \sum_{i=1}^{n}(x_i-\mu)^2 \cdot p(x_i) \\
&= \sum_{i=1}^{n}(x_i^2 - 2x_i\mu + \mu^2) \cdot p(x_i) \\
&= \sum_{i=1}^{n}\left[x_i^2 \cdot p(x_i) - 2x_i\mu \cdot p(x_i) + \mu^2 p(x_i)\right] \\
&= \sum_{i=1}^{n} x_i^2 \cdot p(x_i) - \sum_{i=1}^{n} 2x_i\mu \cdot p(x_i) + \sum_{i=1}^{n} \mu^2 p(x_i) \\
&= \sum_{i=1}^{n} x_i^2 \cdot p(x_i) - \sum_{i=1}^{n} 2x_i\mu \cdot p(x_i) + \sum_{i=1}^{n} \mu^2 \cdot p(x_i) \\
&= \sum_{i=1}^{n} x_i^2 \cdot p(x_i) - 2\mu \underbrace{\sum_{i=1}^{n} x_i \cdot p(x_i)}_{=\mu} + \mu^2 \underbrace{\sum_{i=1}^{n} \cdot p(x_i)}_{=1} \\
&= \sum_{i=1}^{n} x_i^2 \cdot p(x_i) - 2\mu^2 + \mu \\
&= \sum_{i=1}^{n} x_i^2 \cdot p(x_i) - \mu^2 \ \blacksquare
\end{aligned}
$$

(b) [2 marks]

Prove that for a continuous random variable X, $\mathbb{E}[aX+b] = a \cdot \mathbb{E}[X] + b$ where $a, b \in \mathbb{R}$.
<u>**Proof:**</u>

$$
\begin{aligned}
\mathbb{E}[aX+b] &= \int_{-\infty}^{\infty} (ax+b) f(x) \mathrm{d}x \\
&= \int_{-\infty}^{\infty} \left[a \cdot x \cdot f(x) + b \cdot f(x)\right] \mathrm{d}x \\
&= \int_{-\infty}^{\infty} a \cdot x \cdot f(x) \mathrm{d}x + \int_{-\infty}^{\infty} b \cdot f(x) \mathrm{d}x \\
&= a \underbrace{\int_{-\infty}^{\infty} x \cdot f(x) \mathrm{d}x}_{=\mathbb{E}[X]} + b \underbrace{\int_{-\infty}^{\infty} f(x) \mathrm{d}x}_{=1} \\
&= a \cdot \mathbb{E}[X] + b \; \blacksquare
\end{aligned}
$$

(c) [10 marks]

Consider the probability density function $f(x) = c \cdot e^{-x/10}$ for the non-negative random variable X.

- [2 marks] What is the support of X?
- [3 marks] Find c. Express your answer as an integer.
- [3 marks] Find the CDF of X.
- [2 marks] Find $P(X > x_0)$ where $x_0 \in \mathbb{R}$.

<u>**Solution:**</u>
i) $supp(X) = [0, \infty)$

ii)

$$
\begin{aligned}
\int_{-\infty}^{\infty} f(x)\mathrm{d}x &= 1 \\
\Leftrightarrow \int_{0}^{\infty} c \cdot e^{-x/10}\mathrm{d}x &= 1 \\
\Leftrightarrow c\int_{0}^{\infty} e^{-x/10}\mathrm{d}x &= 1 \\
\Leftrightarrow \int_{0}^{\infty} e^{-x/10}\mathrm{d}x &= \frac{1}{c} \\
\int_{0}^{\infty} e^{-x/10}\mathrm{d}x &= -10 \cdot e^{-x/10}\Big|_{0}^{\infty} \\
&= -\lim_{x\to\infty}\left(10 \cdot e^{-x/10}\right) + \frac{1}{10} \cdot e^{0} \\
&= -1 \cdot 0 + 10 = 10 \Rightarrow c = \frac{1}{10}
\end{aligned}
$$

iii)

$$
\begin{aligned}
F(x_0) &= \int_{0}^{x_0} \frac{e^{-x/10}}{10}\mathrm{d}x \\
&= -e^{-x/10}\Big|_{0}^{x_0} \\
&= 1 - e^{-x_0/10}
\end{aligned}
$$

iv)

$$P(X > x_0) = 1 - P(X \leq x_0) = 1 - F(x_0) = e^{-x_0/10}$$

(d) [10 marks]

Suppose $\mathbb{E}[X] = \frac{3}{5}$ and the density function of X is

$$f(x) = \begin{cases} a + \frac{b}{x^2} & 1 \leq x \leq e \\ 0 & otherwise \end{cases}$$

Find a and b.

$$\int_{-\infty}^{\infty} f(x)\mathrm{d}x = 1$$

$$\int_{1}^{e}\left(a+\frac{b}{x^2}\right)\mathrm{d}x = a(e-1) - b\cdot\left.\frac{1}{x}\right|_1^e$$

$$= a(e-1) + b\left(\frac{e-1}{e}\right)$$

$$\Rightarrow a(e-1) + b\left(\frac{e-1}{e}\right) = 1 \qquad (*)$$

$$\int_{-\infty}^{\infty} x\cdot f(x)\mathrm{d}x = \frac{3}{5}$$

$$\int_{1}^{e} x\cdot\left(a+\frac{b}{x^2}\right)\mathrm{d}x = \int_{1}^{e} x\cdot\left(a+\frac{b}{x^2}\right)\mathrm{d}x$$

$$= \int_{1}^{e}\left(ax+\frac{b}{x}\right)\mathrm{d}x$$

$$= a\cdot\frac{(e-1)^2}{2} + b$$

$$\Rightarrow a\cdot\frac{(e-1)^2}{2} + b = \frac{3}{5} \qquad (**)$$

Now $(*)$ and $(**)$ are two equations with two unknowns. Hence, a and b are

$$\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} e-1 & \frac{e-1}{e} \\ \frac{(e-1)^2}{2} & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 \\ \frac{3}{5} \end{pmatrix}$$

$$= \frac{1}{e-1-\frac{(e-1)^3}{2e}} \begin{pmatrix} 1 & -\frac{e-1}{e} \\ -\frac{(e-1)^2}{2} & e-1 \end{pmatrix} \begin{pmatrix} 1 \\ \frac{3}{5} \end{pmatrix}$$