Prelims Probability & Statistics

Hilary Term Collection

January 14th, 2022

Instructions: You have 90 minutes to complete this closed-book, closed-notes collection. Graphing/programmable calculators are not permitted, but scientific or ordinary four-function calculators are allowed. Answer all questions in the space provided, continuing on the back of the page if you run out of space. Show your work for full credit. A table of the standard normal distribution is attached to this document.

Question:	1	2	3	4	5	Total
Points:	25	15	20	25	15	100
Score:						

☺☺☺ **Good Luck!** ☺☺☺

All parts of this question are taken from your assigned practice problems from last term.

3 1. (a) State the complement rule.

Solution: $P(A^C) = 1 - P(A)$

3 (b) State the definition of a conditional probability.

Solution: $P(A|B) \equiv P(A \cap B)/P(B)$

3 (c) True or false: If X is a continuous variable then $\mathbb{P}(X \leq 0.3) = \mathbb{P}(X < 0.3)$. Explain briefly.

Solution: TRUE. By the third axiom of probability $P(X \leq 0.3) = P(X = 0.3) + P(X < 0.3)$. For a continuous random variable $P(X = x) = 0$ for any x, so $P(X = 0.3) = 0$.

3 (d) Suppose that $X \sim N(\mu = -2, \sigma^2 = 3)$. What is $\text{Cov}(X, X)$?

Solution:

$$Cov(X, X) \equiv \mathbb{E}\left[\{X - \mathbb{E}(X)\}\{X - \mathbb{E}(X)\}\right] = \mathbb{E}\left[\{X - \mathbb{E}(X)\}^2\right] \equiv \text{Var}(X) = 3$$

3 (e) Let a and b be constants. If $U \sim N(0, \sigma^2)$, what is the distribution of $a + bU$?

Solution: $N(a, b^2\sigma^2)$

4 (f) Suppose that X and Y are two random variables with correlation $\rho = 0.3$, and standard deviations $\sigma_X = 4$ and $\sigma_Y = 5$. Find $\text{Var}(Z)$ for $Z \equiv (X + Y)/2$.

Solution: First, $\text{Cov}(X, Y) = \rho\sigma_X\sigma_Y = 0.3 \times 4 \times 5 = 6$. Thus, we obtain

$$\begin{aligned}\text{Var}\left(\frac{X}{2} + \frac{Y}{2}\right) &= \left(\frac{1}{2}\right)^2 \text{Var}(X) + \left(\frac{1}{2}\right)\text{Var}(Y) + 2 \times \frac{1}{2} \times \frac{1}{2} \times \text{Cov}(X, Y) \\ &= \frac{1}{4}\left[4^2 + 5^2 + 12\right] = \frac{53}{4} = 13.25\end{aligned}$$

6 (g) Find the following for $X \sim N(0, 1)$. You may use the standard normal tables.

(i) $\mathbb{P}(X \leq -1)$

(ii) the value of t such that $\mathbb{P}(X > t) = 0.8$

(iii) the value of $t > 0$ such that $\mathbb{P}(-t \leq X \leq t) = 0.9$

Solution:

(i) $\mathbb{P}(X \leq -1) \approx 0.1587$

(ii) $\mathbb{P}(X > t) \approx 0.8$ for $t \in [-0.84, -0.85]$

(iii) $\mathbb{P}(-t \leq X \leq t) \approx 0.9$ for $t \in [1.64, 1.65]$.

15 2. Ruth wants to know the fraction of undergraduates who smoke marijuana. Because she worries that her subjects may be unwilling to admit to breaking the law, Ruth conducts her interviews as follows:

> *Hi, I'm carrying out a survey on marijuana use. In a moment I'm going to ask if you smoke marijuana, but before I do I'd like you to flip this fair coin in secret. If you get heads, answer my question truthfully. If you get tails, answer YES regardless of whether you smoke marijuana. Because I don't know the outcome of your coin flip, if you answer YES I'll never know whether you really smoke marijuana or just flipped a tails.*

Ruth interviews a large number of subjects who all follow her instructions exactly: 85% answer YES. Calculate the fraction of undergraduates who smoke marijuana.

Solution: Let Y be the event that someone answers YES, M be the event she smokes marijuana, and T be the event that she flips a tails. Since $Y = M \cup T$, by the addition rule we have

$$P(Y) = P(M) + P(T) - P(M \cap T)$$

Since the coin is fair, $P(T) = 0.5$. Since the outcome of a coin toss doesn't depend on whether someone does or doesn't smoke marijuana, M and T are independent which implies $P(M \cap T) = P(M)P(T)$. Substituting this information along with $P(Y) = 0.85$,

$$0.85 = P(M) + 0.5 - 0.5 \times P(M)$$

and solving we see that $P(M) = 0.7$.

Alternatively, you could solve this using the Law of Total Probability:

$$\begin{aligned} P(Y) &= P(Y|T)P(T) + P(Y|T^c)P(T^c) \\ 0.85 &= 1 \times 1/2 + P(Y|T^c) \times 1/2 \\ 0.7 &= P(Y|T^c) \end{aligned}$$

using the fact that $P(Y|T^c) = P(M)$ since anyone who flips heads answers truthfully, and the outcome of the coin flip has no bearing on whether someone smokes marijuana.

3. A plane has crashed in one of three possible locations: the mountains (M), the desert (D), or the sea (S). Based on its flight path, experts have calculated the following prior probabilities that the plane is in each location: $P(M) = 0.5$, $P(D) = 0.3$ and $P(S) = 0.2$. If we search the mountains then, given that the plane is actually there, we have a 30% chance of *failing* to find it. If we search the desert then, given that the plane is actually there, we have a 20% chance of *failing* to find it. Finally, if we search the sea then, given that the plane is actually there, we have a 90% chance of *failing* to find it. Naturally if the plane is *not* in a particular location but we search for it there, we will not find it. You may assume that searches in each location are independent given the location of the plane. Let F_M be the event that we *fail* to find the plane in the mountains. Define F_D and F_S analogously.

10 (a) We started by searching the mountains. We failed to find the plane. What is the conditional probability that the plane is nevertheless in the mountains? Explain.

Solution: By Bayes' Rule: $P(M|F_M) = P(F_M|M)P(M)/P(F_M)$. We first calculate the denominator using the Law of Total Probability:

$$\begin{aligned} P(F_M) &= P(F_M|M)P(M) + P(F_M|M^C)P(M^C) \\ &= 0.3 \times 0.5 + 1 \times 0.5 = 0.15 + 0.5 = 0.65 \end{aligned}$$

Hence, the desired probability is $15/65 = 3/13 \approx 0.23$.

10 (b) After failing to find the plane in the mountains, we searched the desert, and the sea. We did not find the plane in either location. After this more exhaustive search what is the conditional probability that the plane is in the mountains? Explain.

Solution: We are asked to calculate $P(M|F_M \cap F_D \cap F_S)$. By Bayes' rule,

$$P(M|F_M \cap F_D \cap F_S) = \frac{P(F_M \cap F_D \cap F_S|M)P(M)}{P(F_M \cap F_D \cap F_S)}$$

Define the shorthand $A = F_M \cap F_D \cap F_S$. By the Law of Total Probability

$$\begin{aligned} P(A) &= P(A|M)P(M) + P(A|D)P(D) + P(A|S)P(S) \\ &= (0.3 \times 1 \times 1) \times 0.5 + (1 \times 0.2 \times 1) \times 0.3 + (1 \times 1 \times 0.9) \times 0.2 \\ &= 0.15 + 0.06 + 0.18 = 0.39 \end{aligned}$$

using independence. Hence, the desired probability is $15/39 \approx 0.38$.

4. To answer this question you will need to refer to the following joint pmf:

		Y		
		1	2	3
X	0	0	0.2	0.4
	1	0.1	0.3	0

5 (a) Are X and Y independent? Briefly explain why or why not.

Solution: No: if $Y = 1$ then X must be 1.

5 (b) Write out the marginal pmf of Y and use it to calculate $E[Y^3]$.

Solution: The support set of Y is $\{1, 2, 3\}$ and its marginal pmf is $p_Y(1) = 0.1, p_Y(2) = 0.5, p_Y(3) = 0.4$. We know that $E[g(X)] = \sum_{\text{all } x} g(x)p(x)$. Hence,

$$E[Y^3] = 1^3 \times 0.1 + 2^3 \times 0.5 + 3^3 \times 0.4 = 0.1 + 4 + 10.8 = 14.9$$

5 (c) Write out the marginal CDF of Y.

Solution:

$$F_Y(y) = \begin{cases} 0 & y < 1 \\ 0.1 & 1 \leq y < 2 \\ 0.6 & 2 \leq y < 3 \\ 1 & y \geq 3 \end{cases}$$

5 (d) Derive the conditional pmf of X given $Y = 2$.

Solution: $p_{X|Y}(0|2) = 0.2/0.5 = 0.4$ and $p_{X|Y}(1|2) = 0.3/0.5 = 0.6$

5 (e) Calculate the expected value of the *ratio* of X and Y, namely $E\left[\frac{X}{Y}\right]$.

Solution: We know that $E[g(X,Y)] = \sum_{\text{all } x,y} g(x,y)p_{XY}(x,y)$, hence

$$E[X/Y] = 0.1 \times (1/1) + 0.3 \times (1/2) = 0.25$$

5. Let X be a variable with support $[-1, 1]$ and PDF $f(x) = \frac{3}{4}(1 - x^2)$.

5 (a) Show that $E[X] = 0$.

Solution:

$$E[X] = \int_{-\infty}^{\infty} x f(x)\,dx = \frac{3}{4}\int_{-1}^{1} (x - x^3)\,dx = \frac{3}{4}\left(\frac{x^2}{2} - \frac{x^4}{4}\right)\Big|_{-1}^{1}$$
$$= \frac{3}{4}\left[\left(\frac{1}{2} - \frac{1}{4}\right) - \left(\frac{1}{2} - \frac{1}{4}\right)\right] = 0$$

5 (b) Calculate $\mathrm{Var}(X)$.

Solution: Since $E[X] = 0$, by the shortcut formula $\mathrm{Var}(X) = E[X^2]$:

$$E[X^2] = \int_{-\infty}^{\infty} x^2 f(x)\,dx = \frac{3}{4}\int_{-1}^{1} (x^2 - x^4)\,dx = \frac{3}{4}\left(\frac{x^3}{3} - \frac{x^5}{5}\right)\Big|_{-1}^{1}$$
$$= \frac{3}{4}\left[\left(\frac{1}{3} - \frac{1}{5}\right) - \left(-\frac{1}{3} + \frac{1}{5}\right)\right] = \frac{3}{4}\left[\frac{2}{3} - \frac{2}{5}\right] = \frac{3}{4} \times \frac{4}{15} = 1/5$$

5 (c) Calculate the CDF of X.

Solution:

$$F(x) = \int_{-\infty}^{x} f(t)\,dt = \frac{3}{4}\int_{-1}^{x} (1 - t^2)\,dt = \frac{3}{4}\left(t - \frac{t^3}{3}\right)\Big|_{-1}^{x}$$
$$= \frac{3}{4}\left[\left(x - \frac{x^3}{3}\right) - \left(-1 + \frac{1}{3}\right)\right] = \frac{3}{4}\left(\frac{2}{3} + x - \frac{x^3}{3}\right)$$
$$= \frac{1}{4}(2 + 3x - x^3)$$

The preceding expression is valid for $x \in [-1, 1]$. For $x > 1$, $F(x) = 1$. For $x < -1$, $F(x) = 0$.