Stat310: Test 1

Answer the questions in the spaces provided on the question sheets. If you run out of room for an answer, continue on the back of the page.
You have **two hours** and you may use one double-sided page of notes. You may not use your text book or a computer (calculators are fine), or communicate anything about the contents of this test to anyone else. Please pledge on the last page.

Name: Hadley Wickham

1. On average, Ron Artest of the Houston rockets makes an attempt at a shot about every 90 seconds.
   (a) What distribution might you use to model the waiting time for his next shot? (Make sure to include the parameter values) (2)

   $W \sim \text{Exponential}(\theta = 90)$ or $\text{Exponential}(\lambda = 1/90)$

   1 pt 1 pt

   (b) What's the probability you wait for more than 5 minutes for his next attempt? (2)

   $P(W > 5 \times 60) = 1 - P(W \leq 300)$ 1 pt

   $= 1 - (1 - e^{-300/90}) = 0.036$ 1 pt (correct computation)

   (c) Instead of modelling the waiting time between attempts, I could model the **number** of attempts in a fixed amount of time. Let S be the number of shots he attempts in a five minute period. Assume that this is a Poisson distribution. What is the value of λ? (1)

   $\lambda = \frac{300}{90} = 3.33$ 1 pt

   (d) What's the probability that he attempts zero shots in the next 5 minutes? (2)

   $P(S = 0) = \frac{(3.33)^0 e^{-3.33}}{1} = 0.036$

   1 pt correct conversion to math problem

   1 pt correct calculation

(e) Compare your answers from b and d. Why are they different or the same? (1)

Same. Because waiting > 5 mins & no attempts in 5 mins are the same event! Or similar

(f) List two questionable assumptions that I made when I assumed that this variable had a Poisson distribution. (2)

1 pt Rate is constant throughout game
- depends on being on-court

1 pt Independence of time periods
- ball returned to centre after a success

2. In last month, I received 997 spam emails and 2409 real emails. 11% of the spam messages contained the word bank, but only 1.5% of non-spam email contained it.

(a) Convert the word problem to a mathematical problem by setting up the events that you will use to answer the next questions. (1)

$S^+ = \{\text{email is spam}\}$ $S^- = \{\text{email not spam}\}$

$B^+ = \{\text{email contains bank}\}$ $B^- = \{\text{doesn't contain bank}\}$

doesn't matter what events are called as long as they get the relationships correct

$$P(S^+) = \frac{997}{997 + 2409} = 0.29$$

$$P(B^+ \mid S^+) = 0.11 \qquad P(B^+ \mid S^-) = 0.015$$

(b) What proportion of my email contains the word "bank"? (3)

$$P(B^+) = P(B^+ \mid S^+)P(S^+) + P(B^+ \mid S^-)P(S^-)$$

$$= 0.11 \times 0.29 + 0.015 \times (1 - 0.29)$$

$$= 0.043$$

1 pt for correct calculation

1 pt for correct answer

1 pt for correct conversion to stat problem

(c) Given that an email contains "bank" what is the probability it's spam? (3)

$$P(S^+|B^+) = \frac{P(B^+|S^+)\,P(S^+)}{P(B^+)} = \frac{0.11 \times 0.29}{0.043} = 0.74$$

1 pt 1 pt 1 pt

(d) If it doesn't contain "bank", what's the probability it's not spam? (3)

$$P(S^+|B^-) = \frac{P(B^-|S^+)\,P(S^+)}{P(B^-)} = \frac{(1-0.11)\times 0.29}{(1-0.043)} = 0.27$$

1 pt 1 pt 1 pt

3. (a) Show that $P(A|B) + P(\bar{A}|B) = 1$. Explain your working. (5)

$$P(A|B) + P(\bar{A}|B) = \frac{P(A \cap B)}{P(B)} + \frac{P(\bar{A} \cap B)}{P(B)}$$ 1 pt definition of cond. prob. 1 pt

$$= \frac{P(A \cap B) + P(\bar{A} \cap B)}{P(B)}$$ 1 pt

$$= \frac{P(B)}{P(B)}$$ law of total prob 1 pt $= 1 \;\square$ 1 pt

(b) Does $P(A|B) + P(A|\bar{B}) = 1$? Why/why not? (3)

$$= \frac{P(A \cap B)}{P(B)} + \frac{P(A \cap \bar{B})}{P(\bar{B})} = \frac{P(A \cap B)}{P(B)} + \frac{P(A \cap \bar{B})}{1-P(B)}$$ 1 pt

$$= \frac{P(A \cap B)(1-P(B)) + P(B)\,P(A \cap \bar{B})}{P(B)(1-P(B))}$$ 1 pt

$$= P(A \cap B) + P(B) * (P(A \cap \bar{B}) - P(A \cap B)) \stackrel{?}{=} P(B)\,P(\bar{B})$$

No – no relation why it should 1 pt

(c) If $P(A) = .5$, $P(B) = 0.5$ and $P(A \cap B) = 0.25$ what does that imply about A and B? (1)

A & B are independent

(d) Using the values above, what is $P(\overline{A \cup B})$? (1)

$P(\overline{A \cup B}) = \cancel{P(\bar{A} \cap \bar{B})}$ ~~now~~ $P(\bar{A} \cap \bar{B}) = 0.25$ ← know $\bar{A}$ & $\bar{B}$ indep.

$= \cancel{1 - 0.5}$ either strategy is ok

$= 1 - P(A \cup B) = 1 - (0.5 + 0.5 - 0.25) = 0.25$

4. (a) For the range $x \in [0, 2]$, which of the following functions is a valid pdf. If it's not a valid pdf, indicate why. (5)

(a) $f_1(x) = c_1 x^2$

+ve? ✓ ½

Valid 1 pt

$\int_0^2 c_1 x^2 dx = \left[c_1 \frac{x^3}{3}\right]_0^2 = c_1\left[\frac{8}{3} - 0\right]$ so can integrate to 1 ✓ ½

(b) $f_2(x) = c_2/x$

+ve? ✓

Not valid 1 pt

$\int_0^2 c_2/x \, dx = \left[c_1 \ln x\right]_0^2 =$ undefined! ½

(c) $f_3(x) = c_3(1 - x)$

Not valid! 1 pt

+ve? $x = 2 \Rightarrow f_3 = -c_3$ ½

1 pt for correct answer

½ pt for each correct component

(d) Fill in the blanks below to find the MGF for the exponential distribution. (5)

$$M_X(t) = \underline{E(e^{Xt})} \quad \tfrac{1}{2}$$

by definition of mgf

$$= \underline{\int_0^\infty \frac{1}{\theta} e^{-x/\theta} e^{xt}\, dx} \quad \tfrac{1}{2}$$

by definition of E and the pdf of the exponential

$$= \frac{1}{\theta}\int_0^\infty e^{x(t-1/\theta)}\, dx$$

algebra $\tfrac{1}{2}$

or similar justification

$$= \underline{\frac{1}{\theta}\left[\frac{e^{x(t-1/\theta)}}{t-1/\theta}\right]_0^\infty} \quad 1$$

: $\int e^{ax}\,dx = \frac{1}{a}e^{ax}$

$$= \left[\frac{e^{x(t-1/\theta)}}{\theta t-1}\right]_0^\infty$$

algebra $\tfrac{1}{2}$

definite integral

$$= \lim_{x\to\infty} \frac{e^{x(t-1/\theta)}}{\cancel{\theta t-1}} \;\tfrac{1}{2} - \frac{e^0}{\cancel{\theta t-1}} \quad \tfrac{1}{2}$$

$\lim e^{xa} = 0$ iff $a < 0 \Rightarrow 1/\theta > t$

$$= \underline{\frac{-1}{\theta t-1}} \;\tfrac{1}{2} = \underline{\frac{1}{\theta t-1}} \;\tfrac{1}{2} \quad \square \quad \text{if } \frac{1}{\theta} > t$$

Pledge:

Question	Points	Score
1	10	
2	10	
3	10	
4	10	
Total:	40	