APPM 4570/5570, Statistical Methods, Summer 2017, Exam #1 Review

NAME:

Instructions: On the exam, you will be permitted to bring one sheet of notes (8.5 x 11) and a calculator. The following questions are of the style that may be asked on the exam. Some of these questions might help you answer questions on the exam. I don't plan to provide a full list of solutions for this review, but I will answer any question you have about it (either in office hours or on D2L).

1. Medical records show that, among patients suffering from disease D, 75% will die of it. For each situation described below: (i) define an appropriate random variable to model the situation; (ii) give the values that the random variable can take on; (iii) find the probability that the random variable is greater than 3; (iv) state any assumptions you need to make. **Be sure to label parts (i)—(iv) clearly for full credit!**

   (a) Out of 10 people suffering from D, x people will survive.

   (b) People die from D at a rate of 4 people per day.

   (c) The average amount of time, t (in hours), until the first person dies on a given day is 3.

2. Please answer 'True in all cases' or 'False for at least one case'. Briefly justify your answers for full credit.

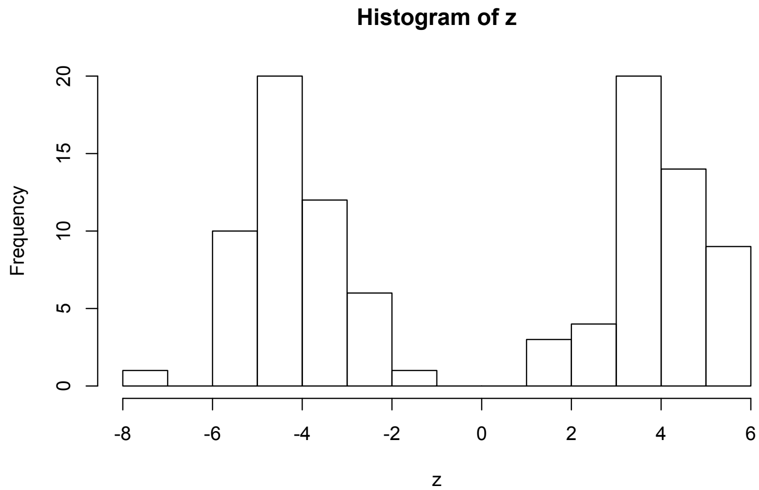
   (a) Let $f(x)$ be the probability distribution function for a random variable X. Then $f(x) > 0$ for all x.

   (b) If $P(A) = 0.5$ and $P(B) = 0.5$, then $P(A \cap B) = 0.25$.

   (c) If X is a continuous random variable then for any real number x, $P(X = x) = 0$.

   (d) The median for the following dataset is 6: 3, 4, 12, 13, 4, 10, 6, 15, 1, 2, 2, 12.

   (e) Suppose a sample is taken, and the variable of interest z is measured for each unit in the sample. If the distribution of z is bimodal as shown below, the sample median would be the best measure of center.



(f) For *any* data set $x_1, x_2, ..., x_n$, the variance can be written as $\sum_{i=1}^{n} x_i^2 - n\bar{x}^2$ (Prove or disprove).

(g) For *any* data set $x_1, x_2, ..., x_n$, $\sum_{i=1}^{n}(x_i - \bar{x}) = 0$.

(h) $E(aX^2 + b) = aE(X^2) + b$.

(i) For any set B, $P(B|\neg B) = 0$.

(j) For sets A and B, $P(A|B) = P(B|A)$.

3. Show that for a continuous random variable X:

   (a) $E(X + a) = E(X) + a$.

   (b) $\text{Var}(X + a) = \text{Var}(X)$.

   (c) If $X \geq 0$ then $E(X) \geq 0$.

   (d) X is a constant if and only if $\text{Var}(X)$ =0.

4. The axioms of probability state that, for any event E, $P(E) \geq 0$, but they do not explicitly state that $P(E) \leq 1$. Prove the claim that, for any event E, $P(E) \leq 1$. Why do you think that this claim is omitted from the axioms?

5. A box contains 5 defective bulbs, 10 partially defective bulbs (the bulb starts to fail after 10 hours of use) and 25 perfect bulbs. If a bulb is selected at random from the box, tested, and it does not fail immediately, what is the probability that it is perfect?

6. Show that the definition of conditional probability ($P(A|B) = \dfrac{P(A \cap B)}{P(B)}$, where $P(B) > 0$) satisfies the axioms of probability theory.

7. Suppose that in the field of exercise science, 1 out of 100 *claims* are true (let C = 'a given claim is true'. An example of a claim in exercise science: "jogging 3 times a week for 20 minutes increases cardiovascular health."). In practice, the only way we *know* whether a claim is true is if we investigate/research the claim. Claims that are investigated and thought to be true are published as a "research finding". Of course, research findings can be wrong. Given that a research *claim* in exercise science is true, a research *finding* correctly reports it 50% of the time (let R = 'research finding says the claim is true'). Given that a research *claim* in exercise science is false, a research *finding* incorrectly reports it 40% of the time.

    (a) What is the probability that a research finding says that a research claim in exercise science is true?

    (b) What is the probability that a research claim is true given that a research finding says that it is?

    (c) Interpret your answer for (b)? Is it alarming? Why? (Note that this scenario is not implausible!)