Caltech Department of Mathematics

Ma 3/103
Introduction to Probability and Statistics

KC Border
Winter 2020

Midterm Exam

Due Tuesday, February 11 by 4:00 p.m.
in the dropbox in Linde Hall.
Some of the questions may relate to material not covered until Friday, February 7 (Lecture 14).

Instructions

This exam has eight questions. Many of them are short. You have **four and a half hours in one sitting** to answer them.

You may use scientific/statistical software (or a calculator) to compute some numerical answers.

- No collaboration is allowed.
- You may use the textbooks by Pitman and by Larsen and Marx), homework solutions, lecture notes, and other handouts from the current Ma 3 web site and auxiliary web site, your own notes and homework, and TA notes. You may also use someone else's notes that you have copied by hand.
- **You are allowed to use statistical software, and often the software's documentation resides on-line, so you may use the on-line software documentation. You may use Wolfram Alpha for computations only, and you may refer to the course web site or the auxiliary course web site. You may not use any other internet resources.**
- Please indicate clearly any work done in **overtime**. Points will be recorded separately and may be considered informally in course grades.
- Write legibly in complete sentences and explain yourself.
- When asked for a numerical result, explain how you calculated it.
- If you have any questions about these instructions, consult a TA for the course (they are on the course web page), the lead TA Josh Lieber <jlieber@caltech.edu>, or the professor <kcborder@caltech.edu>.

Exercise 1 (20 pts) Alex and Blair play the following card game with a well-shuffled deck. Alex turns over cards one-by-one until a red Ace appears. Blair takes the remainder of the deck and continues until the other red ace appears. The winner is the player who turns over fewer cards. (It's a tie if they turn over the same number.) Which player has the greater probability of winning? Explain your answer.

Hint: Think before you compute. □

Exercise 2 (10 pts) Let X be the result of rolling an ordinary six-sided die. Find $\boldsymbol{E}\,X$, $\boldsymbol{E}(X^2)$, and $\boldsymbol{E}(1/X)$. □

Exercise 3 The random variable X is generated by the following two-stage procedure: First a fair coin is flipped. If the coin comes up Heads, then X is drawn from a uniform distribution on the interval $[0, 1]$. If the coin comes up Tails, then X is drawn from a uniform distribution on the interval $[0, 2]$.

1. (20 pts) Find the cumulative distribution function of X.

2. (10 pts) Find the density of X. □

Exercise 4 Two urns A and B contain only black and white balls. Let f_i be the fraction of black balls in urn i. Assume $0 < f_A < f_B < 1$. An urn is chosen at random, and a sample of size n is drawn *with replacement*. Remarkably, all the balls are black.

1. (15 pts) Given the outcome of the experiment, what is the probability that urn A was selected? Explain you reasoning.

2. (5 pts) What happens to this probability as $n \to \infty$? □

Exercise 5 (20 pts) A building has ten floors above the basement. Eight people get into an elevator in the basement, and each chooses a floor at random to get out, independently of the others.

What is the expectation of the number of floors where the elevator stops to let out at least one of the people? (You may assume no one else gets on the elevator on the way up.) □

Exercise 6 This question has been discarded. □

Exercise 7 (30 pts) According to the National Collegiate Athletic Association's National Center for Catastrophic Sport Injury Research,[1] for the years 1992–2012 (inclusive) there were 23 instances of athletes, coaches, officials, or spectators being struck by either a discus, shot (as in shot put), or javelin at high school and college track meets or practices in the U.S. Of these injuries, 4 were fatal.

Based on these facts, choose and justify a probability model for the number of such strikes annually, and use it to answer the following questions. (Assume that no safety changes were made, so that the probabilities are representative of 2020.)

[1] http://nccsir.unc.edu/files/2014/06/NCCSIR-30th-Annual-All-Sport-Report-1982_2012.pdf, Table 1, pp. 25–26.

1. In 2002, three athletes were struck by a shot and one by a discus; a coach was struck by a shot. In how many of these years do you predict there would be five strikes that year?

2. In how many of these years do you predict there would be two fatalities that year? □

Exercise 8 There are m urns. For each $i = 1, \dots, m$, urn i contains n_i balls, and each ball in urn i has a label that says i. Let $N = \sum_{i=1}^{m} n_i$ denote the total number of balls. To make sense of the following question, assume $N > 0$.

1. (5 pts) An urn is selected at random. Let X be the number of balls in the selected urn. (So if urn i is selected, then $X = n_i$.)

   (a) Write down a parsimonious sample space for this random experiment.

   (b) What is $\boldsymbol{E}\, X$?

2. (5 pts) The urns are emptied in to a giant "meta-urn," containing all N balls. A ball is selected at random and its label is observed. Write down a parsimonious sample space for this random experiment. Let Y denote the number of balls in the urn to which the selected ball belongs. (So if the ball is labeled i, then $Y = n_i$.)

   (a) Write down a parsimonious sample space for this random experiment.

   (b) What is $\boldsymbol{E}\, Y$?

3. (20 pts) Which is larger, $\boldsymbol{E}\, X$ or $\boldsymbol{E}\, Y$? Justify your answer. (Hint: What is the relation between $\boldsymbol{E}\, Y$ and $\boldsymbol{E}(X^2)$?) □