MATH 4710 **PRELIM 2** **FALL 2020**

Instructions

(1) There are 4 independent problems. The point total is 30. Start early so that you have time to come back to those questions that require you to work harder and read some documents. Take care to solve correctly those questions that are easiest for you. Even when it is not required, ask yourself: what probability space am I using? what are the elementary outcomes used to describe the problem?

(2) You can write your answers on any reasonable media you that is convenient to you as long as you can produce a clean pdf file to upload on gradescope. Write your name and Cornell NetID on the top of the first page before you begin.

(3) Make sure you clearly indicate the questions you are addressing and separate them neatly from each other. Write clearly using a black or blue pen or pencil if you write on paper. Use extra pages when needed.

When you upload your exam on gradescope, please assign problems to your pages.

(4) Provide reasons for your answers and explain your computations. For numerical answers, give either a simplified fraction or a decimal answer, which ever comes more easily. You can use a basic calculator (e.g., Desmos) if needed.

(5) You can use your notes, our canvas website including all documents provided there and the book. Do not use other websites or the internet (except for Desmos or a simple electronic calculator). Do not discuss prelim problems with other students. Do not discuss prelim problems with anyone except Pr. Saloff-Coste (ask Professor Saloff-Coste privately on Piazza, by email, or in office hours if you have questions. It is Ok to do so).

(6) Academic integrity is expected of all Cornell University students at all times, whether in the presence or absence of members of the faculty.

Problem 1: (4 pts)

(a- 2pts) If a random variable X has expectation μ and variance σ^2, what is the expected value of $Y = (3 + X)^2$?

We compute $Y = 9 + 6X + X^2$ so $E(Y) = 9 + 6E(X) + E(X^2) = 9 + 6\mu + \sigma^2 + \mu^2$ because $\mathrm{Var}(X) = E(X^2) - E(X)^2$.

(b- 2pts) X is a continuous random variable with a continuous density function f. Let F be the cumulative distribution function of X and $Y = F(X)$. What is the expectation of Y?

Write $E(Y) = \int_{-\infty}^{+\infty} F(x)f(x)dx$. Because $F' = f$, $(\frac{1}{2}F^2)' = Ff$. This means that for any $A > 0$,

$$\int_{-A}^{+A} F(x)f(x)d(x) = \left[\frac{1}{2}F^2\right]_{-A}^{+A} = \frac{1}{2}(F(A) - F(-A)).$$

When A tends to infinity, $F(A)$ tends to 1 and $F(-A)$ tends to 0. It follows that $E(Y) = 1/2$.

Problem 2: (6 pts). Let X be uniform on $[-1,2]$ and set $Y = X^2, Z = |X|$

(a-2 pts) Compute $E(X)$, $E(Y)$ and $E(Z)$.

$E(X) = \frac{1}{3}\int_{-1}^{2} x dx = \frac{1}{3}\left[\frac{x^2}{2}\right]_{-1}^{2} = \frac{1}{2}$.

$E(Y) = \frac{1}{3}\int_{-1}^{2} x^2 dx = \frac{1}{3}\left[\frac{x^3}{3}\right]_{-1}^{2} = 1$.

$E(Z) = \frac{1}{3}\int_{-1}^{2} |x| dx = -\frac{1}{3}\left(\int_{-1}^{0} x dx + \int_{0}^{2} x dx\right) = \frac{1}{3}(\frac{1}{2} + 2) = \frac{5}{6}$.

(b-4 pts) Find the density function of the random variable Y.

We first compute the cumulative distribution function of Y.
For $s < 0$, $F_Y(s) = 0$.

$$\begin{aligned} F_Y(s) &= P(Y \le s) = P(X^2 \le s) = P(-\sqrt{s} \le X \le \sqrt{s}) \\ &= \frac{1}{3}\int_{-\min\{1,\sqrt{s}\}}^{\min\{2,\sqrt{s}\}} dx = \frac{1}{3}(\min\{2,\sqrt{s}\} + \min\{1,\sqrt{s}\}). \end{aligned}$$

All together,

$$F_Y(s) = \frac{1}{3}\begin{cases} 0 & \text{if } s > 0, \\ 2\sqrt{s} & \text{if } 0 \le s \le 1 \\ 1 + \sqrt{s} & \text{if } 1 \le s \le 4 \\ 3 & \text{if } s \ge 4. \end{cases}$$

Then we take the derivative to find the density function of Y, $f_Y(s) = F_Y'(s)$, except at the 3 points where F_Y is not differentiable (Fact 3.13). We find that (the values at the break-points can be set arbitrarily)

$$f_Y(s) = \begin{cases} 0 & \text{if } s < 0, \\ 1/3\sqrt{s} & \text{if } 0 \le s < 1 \\ 1/6\sqrt{s} & \text{if } 1 \le s < 4 \\ 0 & \text{if } s \ge 4. \end{cases}$$

Problem 3: (8 pts) In this problem, give numerical answers to 2 decimals. Every evening, Mark rolls a dice until the dice shows a number greater than 1.

(a-4 pts) Give your best estimate for the probability that, in the next 120 days, there will be exactly 2 days when Mark needs more than 3 rolls.

Let X be the number of days Mark needs more than 3 rolls. We observe that this problem concerns a sequence of $n = 120$ identical, independent, experiment with success probability p equal to the probability that Mark needs more than 6 rolls to see a number greater than 1. This probability p is equal to $\sum_{k=4}^{\infty}(1/6)^{k-1}(5/6) = (5/6)(1/6)^3\sum_{0}^{\infty}(1/6)^i = (5/6)(1/6)^3(6/5) = (1/6)^3 = 4.6 \times 10^{-3}$. Because np^2 is very small, we use the Poisson approximation with $\lambda = np = 120 \times 4.6 \times 10^{-3} = 0.56$

$$P(X = 2) \approx e^{-0.56}\frac{(0.56)^2}{2!} \approx .09.$$

(b-4pts) Give your best estimate for the probability that, in the next 360 days, there will be at least 70 days when Mark needs more than one roll.

The probability that Mark needs more than one roll is $p = 1 - 5/6 = 1/6$.. There are $n = 360$ experiments and $np(1-p) = 50$. So we use the Normal approximation with continuity correction. Let X be the number of days Mark needs more than one roll. We want compute $P(X \geq 70) = P(X \geq 69.5)$. Note that $np = 60$ and $\sqrt{np(1-p)} = \sqrt{50} \approx 7.07$ so that

$$P(X \geq 69.5) = P\left(\frac{X - 60}{\sqrt{50}} \geq \frac{9.5}{\sqrt{50}}\right)$$

and $\frac{9.5}{\sqrt{50}} \approx 1.34$. It follows that $P(X \geq 70) \approx 1 - \Phi(1.34) \approx .09$.

Problem 4: (12 pts) An urn contains m green balls and n blue balls. Jessica picks balls in the urn without replacement one by one, each times uniformly at random, until she gets a blue ball. Let X the number of balls she has to pick.

(a-3 pts) Imagine that the green balls are marked $1, 2, \dots, m$ so that they are distinguishable and let $X_i = 1$ if Jessica picks the green ball marked i before any blue ball and $X_i = 0$ otherwise. Show that the probability that $X_i = 1$ is $1/(1+n)$ (if you cannot show this, you can still use it for the rest of the problem). What is $E(X_i)$?

Consider the set of the n blue balls together with the green ball marked i. Each one of these $(1+n)$ balls as the same probability to come first among this set of balls when Jessica picks balls in the urn repeatedly without replacement. It follows that the probability that the green ball marked i comes before the n blue balls is $1/(1+n)$. Because $E(X_i) = 1 \times P(X_i = 1) + 0 \times P(X_i = 0)$, $E(X_i) = 1/(1+n)$.

(b- 3pts) Compute $E(X)$.

Observe that $X = 1 + X_1 + \cdots + X_m$. It follows that

$$E(X) = E\left(1 + \sum_1^m X_i\right) = 1 + \sum_1^m E(X_i) = 1 + m/(1+n) = \frac{1+n+m}{1+n}.$$

(c- 3pts) For $k \in \{1, \dots, m\}$, what is the probability that $X = k$?

Solution 1: Imagine picking all the balls one by one without replacement. The result of doing so can be described as a subset of size n of $\{1, 2, \dots, m+n\}$ which tells us the position of the blue balls and there is a total of $\binom{n+m}{n}$ outcomes, each equally likely. For X to be equal to k, the first blue ball must come at position k. There are $\binom{n+m-k}{n-1}$ such different outcomes. It follows that

$$P(X = k) = \frac{\binom{n+m-k}{n-1}}{\binom{n+m}{n}}.$$

Solution 2: Think of all the balls being distinguishable. The set of the first $k-1$ balls picked is equally likely to be any fix set of $k-1$ balls and there are $\binom{n+m}{k-1}$ such sets. For X to be equal to k, these first $k-1$ ball must be green balls (there are $\binom{m}{k-1}$ such possibilities). Given the first $k-1$ balls picked are green, the probability that the k-th ball is blue is $n/(n+m-k+1)$). This gives

$$P(X = k) = \frac{\binom{m}{k-1}}{\binom{n+m}{k-1}} \frac{n}{m+n-k+1}.$$

The answer is the same in both solutions!

(d- 3pts) A deck of 52 cards is perfectly shuffled and I turn the cards up one by one until I find a spade. On average, how many cards do I turn to find a spade?

This is the same problem has the urn with m green balls and n blue balls. The blue balls are the 13 spades and the green balls are the other cards. The question ask for the expectation of X (the average of the number of cards turned to see a spade). The answer is $E(X) = \frac{1+n+m}{1+n}$ with $n + m = 52$ and $n = 13$, that is, $53/14$.