# Math 340 / 640 Fall 2024
# Midterm Examination One

Professor Adam Kapelner

September 26, 2024

Full Name ______________________________________________

## Code of Academic Integrity

Since the college is an academic community, its fundamental purpose is the pursuit of knowledge. Essential to the success of this educational mission is a commitment to the principles of academic integrity. Every member of the college community is responsible for upholding the highest standards of honesty at all times. Students, as members of the community, are also responsible for adhering to the principles and spirit of the following Code of Academic Integrity.

Activities that have the effect or intention of interfering with education, pursuit of knowledge, or fair evaluation of a student's performance are prohibited. Examples of such activities include but are not limited to the following definitions:

**Cheating** Using or attempting to use unauthorized assistance, material, or study aids in examinations or other academic work or preventing, or attempting to prevent, another from using authorized assistance, material, or study aids. Example: using an unauthorized cheat sheet in a quiz or exam, altering a graded exam and resubmitting it for a better grade, etc.

I acknowledge and agree to uphold this Code of Academic Integrity.

______________________________________________ ____________________

signature date

## Instructions

This exam is 110 minutes (variable time per question) and closed-book. You are allowed **one** page (front and back) of a "cheat sheet", blank scrap paper (provided by the proctor) and a graphing calculator (which is not your smartphone). Please read the questions carefully. Within each problem, I recommend considering the questions that are easy first and then circling back to evaluate the harder ones. No food is allowed, only drinks.

Problem 1 This problem is about a random phenomenon called Benford's Law, it represents the distribution of leading digit in data across a wide variety of data sets that measure natural phenomenon such as street addresses, stock prices, population numbers, etc. Letting $x$ be the digit in base 10, we have the following rv which has mean and variance:

$$X \sim \text{Benford} := \log_{10}\left(\frac{x+1}{x}\right) \mathbb{1}_{x \in \{1,2,\ldots,9\}}, \quad \mathbb{E}[X] = 3.44, \quad \mathbb{V}\text{ar}[X] = 6.06$$

(a) [3 pt / 3 pts] What is the support of $X$? $\mathbb{S}_X = \{1, 2, \ldots, 9\}$

(b) [3 pt / 6 pts] Circle one: this rv is... discrete / continuous

(c) [3 pt / 9 pts] Circle one: the PMF (or PDF) of $X$ is in... old-style / new style

(d) [3 pt / 12 pts] Circle one: $X$ ... has parameter(s) / does not have parameter(s)

(e) [6 pt / 18 pts] Find $\mathbb{P}(X \leq 3)$ exactly and then approximate to the nearest 3 decimals.

$$\mathbb{P}(X = 1) + \mathbb{P}(X = 2) + \mathbb{P}(X = 3) = \log_{10}\left(\frac{2}{1}\right) + \log_{10}\left(\frac{3}{2}\right) + \log_{10}\left(\frac{4}{3}\right) = 0.602$$

(f) [7 pt / 25 pts] Verify the Humpty-Dumpty Identity for the PMF (or PDF). Hint: remember the precalculus rule that $\log_{10}(a/b) = \log_{10}(a) - \log_{10}(b)$.

$$\sum_{x \in \{1,2,\ldots,9\}} \log_{10}\left(\frac{x+1}{x}\right) = \sum_{x \in \{1,2,\ldots,9\}} \log_{10}(x+1) - \sum_{x \in \{1,2,\ldots,9\}} \log_{10}(x)$$

$$= \sum_{x \in \{2,3,\ldots,10\}} \log_{10}(x) - \sum_{x \in \{1,2,\ldots,9\}} \log_{10}(x) = \cancelto{1}{\log_{10}(10)} - \cancelto{0}{\log_{10}(1)} = 1 \checkmark$$

(g) [7 pt / 32 pts] Find an expression for $F_X(x)$, the CDF of $X$, that is valid for all $x \in \mathbb{R}$. Hint: one possible answer has $|\mathbb{S}_X|$ terms and each includes an indicator function.

$$\begin{aligned} F_X(x) &= \log_{10}\left(\frac{2}{1}\right) \mathbb{1}_{x \geq 1} + \log_{10}\left(\frac{3}{2}\right) \mathbb{1}_{x \geq 2} + \ldots + \log_{10}\left(\frac{10}{9}\right) \mathbb{1}_{x \geq 9} \\ &= \sum_{i \in \{1,2,\ldots,9\}} \log_{10}\left(\frac{i+1}{i}\right) \mathbb{1}_{x \geq i} \end{aligned}$$

(h) [10 pt / 42 pts] Find an upper bound for $\mathbb{E}\left[X^4\right]$ to the nearest two decimals. Hint: use the Cauchy-Schwartz inequality.

This problem's original solution was incorrect. It is removed from the exam.

(i) [10 pt / 52 pts] Let $Y_n := \frac{1}{n}X$. Show $Y_n \xrightarrow{d} 0$. Hint: $\phi_X(t) = \sum_{x \in \{1,2,\ldots,9\}} e^{itx} \log_{10}\left(\frac{x+1}{x}\right)$.

$$\begin{aligned}
\lim_{n\to\infty} \phi_{Y_n}(t) &= \lim_{n\to\infty} \phi_{X_n}(t/n) = \lim_{n\to\infty} \sum_{x \in \{1,2,\ldots,9\}} e^{i\frac{t}{n}x} \log_{10}\left(\frac{x+1}{x}\right) \\
&= \sum_{x \in \{1,2,\ldots,9\}} \log_{10}\left(\frac{x+1}{x}\right) e^{ix \lim_{n\to\infty} \frac{t}{n}} = \sum_{x \in \{1,2,\ldots,9\}} \log_{10}\left(\frac{x+1}{x}\right) e^{ix(0)} \\
&= \sum_{x \in \{1,2,\ldots,9\}} \log_{10}\left(\frac{x+1}{x}\right) = 1 = e^{it(0)} = \phi_Y(t) \quad \Rightarrow \quad Y_n \xrightarrow{d} Y \sim \text{Deg}(0) = 0 \checkmark
\end{aligned}$$

(j) [10 pt / 62 pts] Benford's Law is used by the Internal Revenue Service (IRS) to catch people committing tax fraud. The "1040 form" the IRS uses has about 100 numeric entries. When people commit fraud, they may fabricate numbers by drawing iid from $U(\{1, 2, \ldots, 9\})$ which has mean 5 and thus an average probability of first digit greater than 5 of 4/9.

If the 100 first digits on the IRS form were distributed according to Benford's Law, what is the approximate probability the average value of the first digit on the IRS 1040 form is greater than 5?

$$\begin{aligned}
\bar{X}_{100} &\stackrel{\cdot}{\sim} \mathcal{N}\left(\mathbb{E}\left[X\right], \frac{\mathbb{V}\text{ar}\left[X\right]}{100}\right) = \mathcal{N}\left(3.44, \frac{6.06}{100}\right) \quad \text{by the CLT} \\
\mathbb{P}\left(\bar{X}_{100} > 5\right) &\approx \mathbb{P}\left(Z > \frac{5 - 3.44}{\sqrt{\frac{6.06}{100}}}\right) = \mathbb{P}\left(Z > 6.34\right) \approx 0
\end{aligned}$$

Consider $X_1, X_2 \overset{iid}{\sim}$ Benford and $T_2 := X_1 + X_2$

(k) [3 pt / 65 pts] Find the covariance, $\mathbb{C}\text{ov}[X_1, X_2] =$ 0 (due to independence)

(l) [3 pt / 68 pts] What is the support of $T_2$? $\mathbb{S}_{T_2} = \{2, 3, \ldots, 18\}$

## Problem 2

Consider the following rv: $X, Y \overset{iid}{\sim} \dfrac{\lambda - 1}{(x+1)^\lambda} \mathbb{1}_{x \in (0,\infty)}, \quad T = X + Y$

(a) [3 pt / 71 pts] What is the support of $X$? $\mathbb{S}_X = (0, \infty)$

(b) [3 pt / 74 pts] Circle one: this rv is... discrete / continuous

(c) [3 pt / 77 pts] Circle one: the PMF (or PDF) of $X$ is in... old-style / new style

(d) [3 pt / 80 pts] Circle one: $X$ ... has parameter(s) / does not have parameter(s)

(e) [10 pt / 90 pts] Find the PMF (or PDF) of $T$ valid for all $t \in \mathbb{R}$. Leave in sum (or definite integral) format but factor out all constants and simplify *as much as possible.*

$$
\begin{aligned}
f_T(t) &= \int_{\mathbb{S}_X} f(x) f(t-x) \mathbb{1}_{t-x \in \mathbb{S}_X} \, dx = \int_{x \in (0,\infty)} \frac{\lambda - 1}{(x+1)^\lambda} \frac{\lambda - 1}{(t - x + 1)^\lambda} \mathbb{1}_{t-x \in (0,\infty)} \, dx \\
&= (\lambda - 1)^2 \int_{x \in (0,\infty)} \frac{1}{((x+1)(t-x+1))^\lambda} \mathbb{1}_{x \in (-\infty, t)} \, dx \\
&= (\lambda - 1)^2 \mathbb{1}_{t \in (0,\infty)} \int_{x \in (0,t)} \frac{1}{((x+1)(t-x+1))^\lambda} \, dx
\end{aligned}
$$

(f) [10 pt / 100 pts] Find $\mathbb{P}(X > Y)$. Leave in sum (or definite integral) format but factor out all constants and simplify *as much as possible.*

$$
\begin{aligned}
\mathbb{P}(X > Y) &= \int_{y \in \mathbb{R}} \int_{x \in \mathbb{R}} f_{X,Y}(x,y) \mathbb{1}_{x > y} \, dx dy \\
&= \int_{y \in \mathbb{R}} \int_{x \in \mathbb{R}} \frac{\lambda - 1}{(x+1)^\lambda} \mathbb{1}_{x \in (0,\infty)} \frac{\lambda - 1}{(y+1)^\lambda} \mathbb{1}_{y \in (0,\infty)} \mathbb{1}_{x \in (y,\infty)} \, dx dy \\
&= (\lambda - 1)^2 \int_{y \in (0,\infty)} \frac{1}{(y+1)^\lambda} \int_{x \in (0,\infty)} \frac{1}{(x+1)^\lambda} \mathbb{1}_{x \in (y,\infty)} \, dx dy \\
&= (\lambda - 1)^2 \int_{y \in (0,\infty)} \frac{1}{(y+1)^\lambda} \int_{x \in (y,\infty)} \frac{1}{(x+1)^\lambda} \, dx dy
\end{aligned}
$$