Math 241 Fall 2021
Final Examination

Professor Adam Kapelner

Thursday, December 16, 2021

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.

By taking this exam, you acknowledge and agree to uphold this Code of Academic Integrity.

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 and a graphing calculator. 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 [12min] (and 12min will have elapsed) The NY state lotto consists of the following game: 6 balls are drawn from a vat of 59 balls by a mechanical ball-picking machine where each ball is labeled 1, 2, ..., 59. The 6 numbers on the 6 balls are chosen *uniformly* and *without* replacement and these numbers are called the "lotto draw". Before the balls are selected, you write down 6 unique numbers from the set $\{1, 2, \ldots, 59\}$ which is your "lottery ticket". If your numbers are the same as the 6 numbered balls, you win the lotto "jackpot". The order of your written numbers doesn't matter. If you have 5 of the 6 numbers correct (and 1 number incorrect), you win the lotto "3rd prize". If you have 3 numbers written down of the 6 numbered balls (and 3 numbers incorrect), you win the lotto "5th prize". The order of your written numbers doesn't matter.

- [16 pt / 16 pts] Record the letter(s) of all the following that are **true** in general. At least one will be true.

  (a) There are 59! ways of arranging all the balls

  (b) There are 6! / 59! possible lotto draws

  (c) There are 59! / 6! possible lotto draws

  (d) The probability of winning the jackpot is 1 in 59!

  (e) The probability of winning the jackpot is $1/\binom{59}{6}$

  (f) The probability of winning a 3rd prize is $53/\binom{59}{6}$

  (g) The probability of winning a 3rd prize is $318/\binom{59}{6}$

  (h) The probability of winning a 3rd prize is $30/\binom{59}{6}$

  (i) The probability of winning a 3rd prize is $6/\binom{59}{6}$

  (j) The probability of winning a 3rd prize is $\binom{54}{1}/\binom{59}{6}$

  (k) The probability of winning a 5th prize is 1/2

  (l) The probability of winning a 5th prize is $\binom{6}{3}\binom{59}{3}/\binom{59}{6}$

  (m) The probability of winning a 5th prize is $\binom{6}{3}\binom{53}{3}/\binom{59}{6}$

  (n) The probability of winning a 5th prize is $\binom{6}{3}/\binom{59}{6}$

  (o) The count of the numbers you write down which are in the lotto draw can be modeled as Binomial $\left(6,\ 1/\binom{59}{6}\right)$

  (p) The count of the numbers you write down which are in the lotto draw can be modeled as Hyper $(6,\ 6, 59)$

Your answer will consist of a lowercase string (e.g. `aebgd`) where the order of the letters does not matter.

Problem 2 [5min] (and 17min will have elapsed) Winning fifth prize in the lotto has a probability of $p := \binom{6}{3}\binom{53}{3}/\binom{59}{6} \approx 1\%$. Consider the scenario where you buy one ticket every day and stop buying the day you first win.

- [6 pt / 22 pts] Record the letter(s) of all the following that are **true** in general. At least one will be true.

  (a) The number of days that go by including the day you stop T in the lotto can be modeled as $T \sim \text{Binomial}(n, p)$ where $n \to \infty$

  (b) The number of days that go by including the day you stop T in the lotto can be modeled as $T \sim \text{Geometric}(p)$

  (c) The number of days that go by including the day you stop T in the lotto can be modeled as $T \sim \text{Exp}(np)$ where $n \to \infty$

  (d) The expected number of days that go by including the day you stop is approximately 100

  (e) The expected number of days that go by including the day you stop is approximately 99

  (f) The *exact* expected number of days that go by including the day you stop cannot be determined given the information available to you

Your answer will consist of a lowercase string (e.g. `aebgd`) where the order of the letters does not matter.

Problem 3 [12min] (and 29min will have elapsed) Winning fifth prize in the lotto has a probability of $p := \binom{6}{3}\binom{53}{3}/\binom{59}{6} \approx 1\%$. Consider the scenario where you buy one ticket every day. Assume $p = 1\%$ exactly going forward. The number of days that go by up to and including the day that you *first* win fifth prize T in the lotto can be modeled as $T \sim \text{Geometric}(p)$.

- [11 pt / 33 pts] Record the letter(s) of all the following that are **true** in general. At least one will be true.

(a) The probability you win 5th prize on the first day is greater than the probability you win on the 5th prize on the second day

(b) The probability you win 5th prize on the first day is equal to the probability you win on the 5th prize on the second day

(c) The probability you stop on the first day is greater than the probability you stop on the second day

(d) The probability you stop on the first day is equal to the probability you stop on the second day

(e) The probability you stop on the first day is equal to the probability you win the 5th prize on the second day

(f) The probability you don't win any 5th prizes by day 10 is $1 - .99^{10}$

(g) If you know 10 days have passed without winning, the probability you don't win any 5th prizes by day 20 is $1 - .99^{20}$

(h) If you know 10 days have passed without winning, the probability you don't win any 5th prizes by day 20 is $1 - .99^{10}$

Each lotto ticket costs \$1 and each winning ticket pays out \$1. So if you lose, you lose \$1 and if you win you get back \$2.

(i) The amount of money you have when you stop is equal to $-T$

(j) The amount of money you have when you stop is equal to $1 - T$

(k) The amount of money you have when you stop is equal to $3 - T$

Your answer will consist of a lowercase string (e.g. `aebgd`) where the order of the letters does not matter.

Problem 4 [11min] (and 40min will have elapsed) Winning fifth prize in the lotto has a probability of $p := \binom{6}{3}\binom{53}{3}/\binom{59}{6} \approx 1\%$. Consider the scenario where you buy one ticket every day. Assume $p = 1\%$ exactly going forward. Now consider we play for exactly 1,000 days. During this time period, we can win multiple times or never win at all.

- [17 pt / 50 pts] Record the letter(s) of all the following that are **true** in general. At least one will be true.

(a) The number of 5th prizes you win can be modeled as $\text{Binomial}\,(1000,\, 0.01)$

(b) The number of 5th prizes you win can be modeled as $\text{Hyper}\,(1000,\, 0.01N, N)$ where N is very large

Each lotto ticket costs \$1 and each winning ticket pays out \$2. So if you lose, you lose \$1 and if you win you get \$2. Denote the payouts by $X_1, X_2, \ldots, X_{1000}$. Let $p_{X_i}(x)$ denote the PMF of X_i, μ_i denote its expectation and σ_i^2 denote its variance. Let $\bar{X}$ denote the average of all 1,000 payouts. Let $M_{X_i}(t)$ denote the MGF of X_i.

(c) $X_1, X_2, \ldots, X_{1000}$ are identically distributed

(d) $X_1, X_2, \ldots, X_{1000}$ are independent

(e) $\mathbb{C}\text{ov}\,[X_1, X_2] = 0$

(f) $\mathbb{S}\text{upp}\,[X_1] = \{-1, 2\}$

(g) $\mathbb{S}\text{upp}\,[X_1] = [-1, 2]$

(h) $p_{X_1}(-1) = 0.99$

(i) $\mu_1 = -\$0.97$

(j) The 5th prize on the lottery is a fair game

(k) $\sigma_1^2 > 0$

(l) $\sigma_1^2 = \$0.0891$

(m) $\sigma_1 = \$0.0891$

(n) $\mathbb{E}\left[\bar{X}\right] = \mu_1$

(o) $\mathbb{V}\text{ar}\left[\bar{X}\right] = \sigma_1^2$

(p) For all realizations, $\dfrac{x_1 + x_2 + \ldots + x_{1000}}{1000} \approx \mu_1$

(q) $M_{X_1}(t) = .01e^{2t} + .99e^{-t}$

Your answer will consist of a lowercase string (e.g. `aebgd`) where the order of the letters does not matter.

Problem 5 [10min] (and 50min will have elapsed) Winning fifth prize in the lotto has a probability of $p := \binom{6}{3}\binom{53}{3}/\binom{59}{6} \approx 1\%$. Consider the scenario where you buy one ticket every day. Assume $p = 1\%$ exactly going forward. Now consider we play for exactly 1,000 days. During this time period, we can win multiple times or never win at all. Each lotto ticket costs \$1 and each winning ticket pays out \$2. So if you lose, you lose \$1 and if you win you get \$2. Denote the payouts by $X_1, X_2, \ldots, X_{1000}$. These rv's are iid distributed with support $\{-1, 2\}$, mean $\mu = -\$0.97$ and $\sigma^2 = 0.0891\$^2$. Let $W = X_1 + X_2 + \ldots + X_{1000}$ denote your *total* winnings over all 1000 games. Let $Z \sim \mathcal{N}(0, 1)$.

- [16 pt / 66 pts] Record the letter(s) of all the following that are **true** in general. At least one will be true.

  (a) $\mathbb{E}[W] = -\$0.00097$

  (b) $\mathbb{E}[W] = -\$0.97$

  (c) $\mathbb{E}[W] = -\$970$

  (d) $\mathbb{V}\mathrm{ar}[W] = 0.0000891\$^2$

  (e) $\mathbb{V}\mathrm{ar}[W] = 0.0891\$^2$

  (f) $\mathbb{V}\mathrm{ar}[W] = 89.1\$^2$

  (g) $\mathbb{P}(Z > 0) = \frac{1}{2}$

  (h) $\mathbb{P}(Z > 1) \approx 16\%$

  (i) $\mathbb{P}(Z > -970) \approx 1$

  (j) $\mathbb{P}(Z > 970) \approx 0$

  (k) You have enough information to compute the exact probability of winning, $\mathbb{P}(W > 0)$

  (l) $\mathbb{P}(W > 0) \approx \mathbb{P}(Z > 0)$

  (m) $\mathbb{P}(W > 0) \approx \mathbb{P}(Z > -970)$

  (n) $\mathbb{P}(W > 0) \approx \mathbb{P}(Z > -102.7621)$

  (o) $\mathbb{P}(W > 0) \approx \mathbb{P}(Z > 102.7621)$

  (p) $\mathbb{P}(W > 0) \approx 0$

Your answer will consist of a lowercase string (e.g. `aebgd`) where the order of the letters does not matter.

Problem 6 [9min] (and 59min will have elapsed) Winning fifth prize in the lotto has a probability of $p := \binom{6}{3}\binom{53}{3}/\binom{59}{6} \approx 1\%$. Consider the scenario where you buy one ticket every day. Assume $p = 1\%$ exactly going forward. Now consider we play for exactly 1,000 days. During this time period, we can win multiple times or never win at all. Let $\hat{P}$ denote the sampling distribution for $\hat{p}$, the proportion of 5th prize wins in the $n = 1,000$ plays of the lotto.

- [13 pt / 79 pts] Record the letter(s) of all the following that are **true** in general. At least one will be true.

  (a) $\hat{P} \sim \text{Binomial}\,(1000,\ 0.01)$

  (b) $\hat{P} \sim \mathcal{N}\,(10,\ 9.9)$

  (c) $\hat{P} \sim \mathcal{N}\,(0.01,\ 0.0000099)$

  (d) $\hat{P} \mathrel{\dot\sim} \mathcal{N}\,(10,\ 9.9)$

  (e) $\hat{P} \mathrel{\dot\sim} \mathcal{N}\,(0.01,\ 0.0000099)$

  (f) $\hat{P}$ is approximately normally distributed due to the Binomial Theorem

  (g) $\hat{P}$ is approximately normally distributed due to the Central Limit Theorem

  (h) $\hat{P}$ is approximately normally distributed due to the Law of Large Numbers

  (i) $\hat{P}$ is degenerate-distributed if $n = 1$

  (j) $\hat{P}$ is uniform-distributed if $n = 1$

  (k) $\hat{P}$ is Bernoulli-distributed if $n = 1$

  (l) $\mathbb{E}[\hat{P}] = p$

  (m) $\mathbb{E}[\hat{P}] \approx p$ but never exactly equal to p for a finite value of n

Your answer will consist of a lowercase string (e.g. `aebgd`) where the order of the letters does not matter.

Problem 7 [8min] (and 67min will have elapsed) Winning fifth prize in the lotto has a probability of $p := \binom{6}{3}\binom{53}{3}/\binom{59}{6} \approx 1\%$. Consider the scenario where you buy one ticket every day. Assume $p = 1\%$ exactly going forward. Now consider we play for exactly 1,000 days. During this time period, we can win multiple times or never win at all. Let $\hat{P}$ denote the sampling distribution for $\hat{p}$, the proportion of 5th prize wins in the $n = 1,000$ plays of the lotto. Due to the central limit theorem, $\hat{P} \dot{\sim} \mathcal{N}(0.01,\, 0.00314^2)$. We are now concerned that the lotto's 5th prize isn't "fair". This means that we are trying to prove that it is paying out less than what is expected according to probability theory, p. Thus we wish to prove that $p < 1\%$.

- [9 pt / 88 pts] Record the letter(s) of all the following that are **true** in general. At least one will be true.

  (a) This question is an example of statistical inference (which means we try to make inference for an unknown parameter using data)

  (b) The best, most honest way to prove to others that $p < 1\%$ is to assume $p < 1\%$ is true and then it is believed until it is disproved by data

  (c) The best, most honest way to prove to others that $p < 1\%$ is to assume $p \geq 1\%$ is true and then let the data disprove $p \geq 1\%$

  (d) The theory $p < 1\%$ in (c) is called the "alternative hypothesis"

  (e) The theory $p < 1\%$ in (c) is called the "null hypothesis"

  (f) The interval $[0, 1\%]$ is called the set RET, the retainment region for p for some level α

  (g) The interval $[1\%, 1]$ is called the set RET, the retainment region for p for some level α RET

  (h) The interval $[0, 1\%]$ is called the set $CI_{p,1-\alpha}$, the confidence interval for p for some level α

  (i) The interval $[1\%, 1]$ is called the set $CI_{p,1-\alpha}$, the confidence interval for p for some level α

Your answer will consist of a lowercase string (e.g. `aebgd`) where the order of the letters does not matter.

Problem 8 [8min] (and 75min will have elapsed) Winning fifth prize in the lotto has a probability of $p := \binom{6}{3}\binom{53}{3}/\binom{59}{6} \approx 1\%$. Consider the scenario where you buy one ticket every day. Assume $p = 1\%$ exactly going forward. Now consider we play for exactly 1,000 days. During this time period, we can win multiple times or never win at all. Let $\hat{P}$ denote the sampling distribution for $\hat{p}$, the proportion of 5th prize wins in the $n = 1,000$ plays of the lotto. Due to the central limit theorem, $\hat{P} \stackrel{\cdot}{\sim} \mathcal{N}(0.01, 0.00314^2)$. We are now concerned that the lotto's 5th prize isn't "fair". This means that we are trying to prove that it is paying out less than what is expected according to probability theory, p. Thus we wish to prove that $p < 1\%$. Thus, $H_a : p < 1\%$ and $H_0 : p \geq 1\%$. We set $\alpha = 2.5\%$.

- [14 pt / 102 pts] Record the letter(s) of all the following that are **true** in general. At least one will be true.

  (a) H_0 is either retained or rejected

  (b) If we reject H_0, the result is called "statistically significant".

  (c) If we reject H_0 when H_0 was true, this is called a "Type I error"

  (d) If we reject H_0 when H_0 was false, this is called a "Type I error"

  (e) If we retain H_0 when H_0 was true, this is called a "Type I error"

  (f) If we retain H_0 when H_0 was false, this is called a "Type I error"

  (g) The probability of making a "Type I error" in this test is 2.5%

  (h) The probability of making a "Type I error" cannot be computed

  (i) If α was smaller, the probability of a "Type I error" would increase

  (j) If α was smaller, the probability of a "Type II error" would increase

  (k) If the decision is to retain H_0, it is possible we made an error

  (l) If the decision is to reject H_0, it is possible we made an error

  (m) To make a decision about H_0, we first construct a set called RET, the retainment region at level α.

  (n) If $\hat{p}$ is an element of the set RET, we reject H_0

Your answer will consist of a lowercase string (e.g. `aebgd`) where the order of the letters does not matter.

Problem 9 [12min] (and 87min will have elapsed) Winning fifth prize in the lotto has a probability of $p := \binom{6}{3}\binom{53}{3}/\binom{59}{6} \approx 1\%$. Consider the scenario where you buy one ticket every day. Assume $p = 1\%$ exactly going forward. Now consider we play for exactly 1,000 days. During this time period, we can win multiple times or never win at all. Let $\hat{P}$ denote the sampling distribution for $\hat{p}$, the proportion of 5th prize wins in the $n = 1,000$ plays of the lotto. Due to the central limit theorem, $\hat{P} \mathrel{\dot\sim} \mathcal{N}(0.01,\, 0.00314^2)$. We are now concerned that the lotto's 5th prize isn't "fair". This means that we are trying to prove that it is paying out less than what is expected according to probability theory, p. Thus we wish to prove that $p < 1\%$. Thus, $H_a : p < 1\%$ and $H_0 : p \geq 1\%$. We set $\alpha = 2.5\%$. In the 1,000 days of tickets, we won only 5 times.

- [16 pt / 118 pts] Record the letter(s) of all the following that are **true** in general. At least one will be true.

(a) The sample data is the set of 1,000 wins or losses of the 1,000 days of lotto tickets

(b) The data were assumed to be sampled iid from a population of infinite size

(c) Using the CLT, RET $= [0, 0.0162]$ rounded to the nearest 3 significant digits

(d) Using the CLT, RET $= [0, 0.0145]$ rounded to the nearest 3 significant digits

(e) Using the CLT, RET $= [0.0162, 1]$ rounded to the nearest 3 significant digits

(f) Using the CLT, RET $= [0.0145, 1]$ rounded to the nearest 3 significant digits

(g) Using the CLT, RET $= [0, 0.00371]$ rounded to the nearest 3 significant digits

(h) Using the CLT, RET $= [0, 0.00554]$ rounded to the nearest 3 significant digits

(i) Using the CLT, RET $= [0.00371, 1]$ rounded to the nearest 3 significant digits

(j) Using the CLT, RET $= [0.00554, 1]$ rounded to the nearest 3 significant digits

(k) Using the CLT to compute RET, this test's decision using our data is to reject H_0

(l) Using the CLT to compute RET, this test's decision using our data is to retain H_0

(m) If (k) were to be true, a valid conclusion would be "there is sufficient evidence to suggest that NY state is cheating the public out of a fair 5th prize in the lotto"

(n) If (l) were to be true, a valid conclusion would be "there is not sufficient evidence to suggest that NY state is cheating the public out of a fair 5th prize in the lotto"

(o) Using the CLT, $CI_{p,95\%} = [0.000539, 0.00946]$

(p) Using the CLT, $CI_{p,95\%} = [0.000554, 0.0145]$

Your answer will consist of a lowercase string (e.g. `aebgd`) where the order of the letters does not matter.

Problem 10 [11min] (and 98min will have elapsed) The time you spend waiting for a bus (measured in minutes) is modeled as $T \sim \text{Exp}(0.2)$, an exponential rv with mean μ and variance σ^2.

- [13 pt / 131 pts] Record the letter(s) of all the following that are **true** in general. At least one will be true.

  (a) The PDF of the rv T, the bus waiting time is $f(t) = 0.2e^{-0.2t}$

  (b) The PDF of the rv T, the bus waiting time is $f(t) = e^{-0.2t}$

  (c) $\lim_{t\to\infty} f(t) = 0$

  (d) $\lim_{t\to-\infty} f(t) = 0$

  (e) $\lim_{t\to\infty} f(t) = 1$

  (f) $\lim_{t\to-\infty} f(t) = 1$

  (g) The expected waiting time is $\mu = 5$min

  (h) There is a 50% chance you will be waiting for less than 5min

  (i) You are e times more likely to be waiting 2min than 7min

  (j) If you were waiting for 5 minutes already, the total expected waiting time is 7.5min

  (k) The value of σ^2 can be calculated given the information provided to you

  (l) $\mathbb{S}\text{upp}[T] = [0, \mu]$

  (m) $\mathbb{S}\text{upp}[T] = \mathbb{R}$

Your answer will consist of a lowercase string (e.g. `aebgd`) where the order of the letters does not matter.

Problem 11 [12min] (and 110min will have elapsed) Let $X \sim U(0, m)$ where $m > 0$. Let μ denote the expectation of X, σ^2 denote the variance of X and $M_X(t) = (e^{mt} - 1)/(mt)$ denote the MGF of X.

- [14 pt / 145 pts] Record the letter(s) of all the following that are **true** in general. At least one will be true.

  (a) $\mathbb{S}\text{upp}\,[X] = [0, m]$

  (b) Let $h(t) := \frac{d}{dt}[M_X(t)]$. Then $h(0) = m/2$

  (c) Let $h(t) := \frac{d^2}{dt^2}[M_X(0)]$. Then $h(0) = m/2$

  (d) Let $h(t) := \frac{d}{dt}[M_X(0)]$. Then $h(0) = m^2/12$

  (e) Let $h(t) := \frac{d^2}{dt^2}[M_X(0)]$. Then $h(0) = m^2/12$

  (f) $\mathbb{E}[X^2] = m^2/12$

  (g) $\mathbb{E}[X^2] = m^2/4$

  (h) $\mathbb{E}[X^2] = m^2/3$

  Let $Y = aX + c$

  (i) $M_Y(t) = e^{ct}(e^{mt} - 1)/(mt)$

  (j) $M_Y(t) = e^{ct}(e^{mat} - 1)/(mat)$

  (k) $M_Y(t) = a(e^{mt} - 1)/(mt) + c$

  Let $X_1, \ldots, X_n \overset{iid}{\sim} U(0, m)$ and let $T = X_1 + \ldots + X_n$

  (l) $M_T(t) = (e^{mt} - 1)^n/(mt)^n$

  (m) $M_T(t) = n(e^{mt} - 1)/(mt)$

  (n) $M_T(t) = (e^{mt} - 1)/(mt)$

Your answer will consist of a lowercase string (e.g. `aebgd`) where the order of the letters does not matter.