FINAL EXAM WINTER 2022 - DO NOT CLICK ON THIS PAGE UNTIL YOU ARE READY TO START THE EXAM. ONCE YOU CLICK YOU MUST START THE EXAM AND BE READY TO SPEND 5 CONSECUTIVE HOURS DOING IT. READ THE FINAL EXAM INSTRUCTIONS TWO PAGES ABOVE BEFORE STARTING.

Started: Jun 5 at 12:20pm

Quiz Instructions

Note: if you can see this page you must start the exam right away and finish in 5 consecutive hours.

Note: Reproduction of this exam or sharing with anything or anybody during or after the exam time window is prohibited and will result in an F in the exam and a trip to the Dean of Students office.

The following are the instructions you received days ago regarding the final exam. I copy paste them here.

Time window within which you can do the exam: June 3rd 11:55 AM and June 6th, 8AM Select the 5 hours that best work for you within the time window allowed. Make sure you have all the technology needed, space and energy to do it.

All questions regarding the course must have been asked before the time window for the final exam starts, that is, before 11:55 AM on 6/3. There will be no consultations on course matters during the exam time window.

- **Final exam time limit: 5 consecutive hours maximum. The exam is a three hour exam, but to help you, I am allowing 2 more hours. You must do the exam in one sitting of 5 hours and you must SUBMIT before the 5 hours pass.**
- **You may only consult the material in our course web site, Cognella and the textbook (that includes the apps and calculators in our webpage, but that is it, unless I put a URL in the questions that you must use to answer, and when asked for work you must show the complete work, not just refer to one of the calculators). Indication that you have**

consulted with anyone or anything else will result in 0 points and a visit to the Dean of Students Office. All calculations must be done by you, not by external sources.

- **Late exams will not be accepted. Answers to questions emailed or sent outside the quiz space will not be accepted. You must start early to guarantee that you will have time to complete the exam in five consecutive hours before the time window closes. CAE students can take the time allowed, but it must be within the time window for everyone. Because students have the flexibility of choosing any 5 hours within the time window, you will not have a clock ticking for you in the final exam. But Canvas keeps track of the time you take and that will be checked. And you will not be able to submit your exam after 8 AM Monday June 6th. Read the instructions given below. They all are required. Students lose an average of 6 points per exam for not reading and not following the directions.**
- **Once you access the final exam, your time starts being recorded and you must complete it in the next 5 hours immediately after you access it. So do not access it until you are ready.**

ONLINE PROBLEM SOLVERS other than the apps in our bruinlearn course website are prohibited during the exam. Indication that you have used anything outside our bruinlearn course web site materials will result in an F in the exam and a visit to the Dean of Students office.

Tips to prepare for the exam and penalties for not following the instructions.

- The final is 25% of your course grade, more than the midterm because it is comprehensive, including material up to and including week 10, has material for the whole quarter, and the last material is more difficult. It is a summative assessment that will assess whether you have been engaged with the course, what you have learned in it, and your ability to use it in new contexts.
- The final covers all the course content and work done in assessments in modules 1 to 10, inclusive.
- **Tip to prepare:** If you have not gone over all the materials that we have covered so far, you will not have time to do that during the exam, because you will be asked questions about the lecture videos and all the material that has been posted and assignments done. If you have to start figuring out where things are during the exam, you will not have time at all. The exam requires familiarity with all the course content, and having practiced all the assessments given, because all concepts from the beginning reappear over and over again with probability distributions. If you have followed the advice I have been giving you every week, you will have summaries written in your own words, your cheat sheet that you can quickly consult. Prepare a summary before you start the exam if you have not already done so.
- **Everyone that has been studying the videos, doing the reading, the supplementary materials, assignments and homeworks, TA session material, and has been engaged with the class can do this exam. The time window of the exam is not the right time to start looking at the course.**

You will not have time to do the exam if you have not been engaged with the course. Prepare as if this was going to be an in-person exam in a classroom. Be very familiar with everything.

- You will have to complete a form (a question within the final) indicating that all the quiz was done on your own and you did not consult with anyone, or anything except the course materials in our course web site. Indication to the contrary and indication that you did not adhere to UCLA Student Code of Conduct will give you a 0 in the exam and a visit to the Dean of Students Office. Familiarize yourself with the **UCLA Student Code of Conduct. (https://deanofstudents.ucla.edu/individual-student-code)**
- **The signed form can not be submitted by email or late. If you do not submit it completely filled out with your name and hand-written signature and before the deadline your exam will receive 0 points.**
- There will be penalties for not following the instructions. Thus, it is good that you read them ahead of time, and ask if you have any question before the period of the final starts.
- **All questions regarding course content must have been asked before the final time window starts. No questions on the course can be asked during the final exam time window. Make sure you read all the instructions below and also read the ones in the head of the final (in case some additional one is added).**

Instructions

- Your final's format will be a Bruinlearn quiz format. It is meant to be a three hour exam, but you are given 5 hours.
- There is a time limit to do the final. You have five hours to do the final once you open it (once you view it for the first time).
- You may choose any five hours within the final exam time window. The final exam time window is Friday, 11:55 AM 6/3/2022 to Monday 8 AM 6/6/2022 but you must complete the final exam in 5 consecutive hours of your choice, meaning that you can not start, spend one hour and then come back the next day and do another hour. That is not allowed. For example, if you start on Friday at 1:00 PM you must have finished by Friday 6 PM. Students taking more than five hours will suffer a grade penalty as follows:
  - For every 5 minutes that your midterm time taking is larger than 5 hours, your grade will be reduced by 2% of the total maximum score.
- Because you have the flexibility of choosing any five hours within the time window, the quiz itself will show a deadline of 8:00 AM on 6/6/2022. But the quiz keeps track of when you start and when you submit. So we will know whether you took more than five hours. Do not get nervous if it takes 5 minutes more or less, but keep in mind the penalty given above.
- Within the five hours that you choose to take the final, you can save, come back, continue. Everything you enter is kept. Whatever you have entered and saved is submitted automatically, were you to forget to submit.

- Your not having submitted in five hours is no excuse for you to come back to the final after your 5 hours to submit it. If you come back after the five hours because of that, you will also be subject to the penalty. It is better to leave the final alone and let it submit itself when the deadline closes it. Please, be sure that no excuse of the sort "I came back because I forgot to submit" will be accepted. See the penalties above.
- Do not submit until you are sure that you are completely done (within your five hours). Once you submit we have no way of getting you back in, and we will not take any excuses.
- Best strategy: Always enter material as you do the final and save. We will not accept excuses of the sort "I did it on paper and did not have time to enter my answers." Your answers must be entered since you start. Enter something, you can change it later if you change your mind. But do not wait to enter all until the end, because then you could end up with 0 points if you run out of time. We will not accept excuses of the sort " I took longer because I had to enter my answers." Penalties given above will apply.
- Access to the exam too late is not excuse to turn it in late or to use technical problems as an excuse to submit it. You are given a pretty wide time window to allow you to prevent to do it at the last minute. You will get 0 points if you have not submitted before the deadline. But again, if you enter things from the moment you start, even if you are late, something will be there. So follow the advice given above and enter things from the start.
- You must work on your own. During the exam time window, no group work allowed, no consulting with anyone or anything allowed. See the note at the top. No sharing of information allowed. You may use all the material in our Bruinlearn course web site. Communication with other students, external sources of information, or anything out there that helps students during the time window of the exam, even if you have submitted your exam, will have serious consequences.
- You may not contact the TA during the time window of the final exam. There will be a Q&A File that Dr. Sanchez will update if needed, in case some unexpected typo pops up. Consult that before you email Dr. Sanchez but only if you have started the exam. You may email Dr. Sanchez if you catch a typo or need some clarification, but first you must look at the Q&A. If you do not receive an answer, it could be because the question should not have been asked.
- The exam can not be shared with anyone or anything during or after the final exam time period. It is copyrighted for this course. Any indication that the exam has been shared with anything or anyone will result in serious consequences for all the persons or things involved.

Question 1

3 pts

Suppose X has range 1, 2,3, ..., n and P(X=x) = 1/n, for all x. Find the moment generating function and use it to calculate the expected value and standard deviation of X. Show work in the space provided. You may type your answer using the equation editor ("insert-> equation" in the menu) or you may insert a screenshot or a pdf file.

Edit View Insert Format Tools Table

12pt Paragraph

$$E(X^2) = M''(0)$$

$$M''(t) = \frac{d}{dt}\left[\left(\frac{n+1}{2}\right)e^{\frac{n(n+1)}{2}t}\right] = n\left(\frac{n+1}{2}\right)^2 e^{\frac{n(n+1)}{2}t}$$

$$E(X^2) = M''(0) = n\left(\frac{n+1}{2}\right)^2 e^{\frac{n(n+1)}{2}(0)} = n\left(\frac{n+1}{2}\right)^2$$

$$Var(X) = E(X^2) - [E(X)]^2 = n\left(\frac{n+1}{2}\right)^2 - \left(\frac{n+1}{2}\right)^2 = (n-1)\left(\frac{n+1}{2}\right)^2$$

p 13 words </>

Question 2

1 pts

The following problem was given in a separate question. What is the distribution that can be used to answer the question? (choose all that applies)

====

The brown recluse spider often lives in human houses throughout central North America. This spider is a moderate health threat, as its bite causes nasty, slow-healing wounds. Bites are very rarely fatal, but the resulting wounds are disgusting. Information on the spider's diet is useful for developing effective pest management strategies. A diet-preference study gave each of 41 brown recluse spiders a choice between two crickets, one live and one dead. Thirty one of the 41 spiders chose the dead cricket over the live one. Researchers wonder whether this represents evidence for a diet preference. So they phrase their research question as: what would be the probability that of 41 spiders 31 choose a dead cricket or more if there is no diet preference, that is, if the chance of choosing one or the other (a binary choice) is

0.5.? Then they calculate that probability of that happening under the assumption of no diet preference. If that probability is smaller than 0.05 then they conclude that the spider has a diet preference.

So what will be the conclusion of this study using that methodology?

==========

- [x] The binomial distribution Bin(n=41, p=0.5)
- [x] The normal distribution N(μ=20.5, σ=3.201)
- [ ] The hypergeometric distribution
- [ ] The negative binomial distribution

Question 3 — 1 pts

Which of the following statements applies to the Law of Large Numbers? (choose all that applies)

- [x] An insurance company, by managing car insurance for many housands of people can get an estimate of car accident very close to the true probability of a car accident with probability 1.
- [x] The average weight of a large number of people randomly chosen converges in probability to the true average for the whole population from where these people were sampled.
- [x] The limit of the sample mean when n goes to infinity is the population mean
- [ ]

Question 4 — 1 pts

This question is about the properties of moment generating functions, in particular the property that the moment generating function of a sum of independent and identically distributed random variables is the product of the moment generating functions.

Suppose that the time, X, between inquiries in a interactive system has an exponential distribution with average value of one second. Let t be an arbitrary point in time and T the elapsed time until the fifth inquiry arrives (after time t). The expected value and standard deviation of T are, respectively, and the distribution of T is

(note: assume that the inquiries come sequentially, one after the other, not all at once).

- (selected) 5 and 5 seconds and Gamma
- 5 and 5 seconds and exponential
- 1/5 and 1 seconds and gamma
- 1 and 1/5 seconds and geometric

Question 5

3 pts

Below, in the box, is the statement on academic honesty that you must complete for your exam to be graded.

Please, copy paste it in your tablet, or write it identically as it is in your notebook, if you want by hand, scan and insert it as an image or as pdf file, signed by hand at the bottom and with your printed name at the beginning.

Not turnint in this statement will result in an F in your final exam.

====== **The statement that you submit must contain everything in the box below, exactly, but substitute the bolded with your name (at the top, as it appears in MyUCLA) and your hand-written signature (In English) at the bottom. Notice that the url must also be included.The name on the first line and the hand-written signature (in English)**

I, ______________**In this space, replace the bolded sentence with your printed name.** ___sign below to confirm that the final exam that I submitted for Spring 2022 Stat 100A with Dr. Sanchez reflects my work and only my work, that I have not consulted with anyone or anything except the class material posted in Bruinlearn, the textbook, Cognella active learning and my own summaries of the material, and that I have taken the time specified in the instructions or very close to that time to complete the exam from the moment that I first clicked on the exam file until it was submitted. I also confirm that I have adhered to the UCLA Student Code of Conduct at **https://deanofstudents.ucla.edu/individual-student-code (https://deanofstudents.ucla.edu/individual-student-code)** and that I have not shared or will share the exam with anyone or anything. I also confirm that I have not spoken and will not speak to anyone about this exam, even if I submitted it, until after the submission deadline.

YOUR SIGNATURE (In English, and hand-written). Typed signatures not allowed. You must sign by hand, a real signature.

Edit View Insert Format Tools Table

12pt Paragraph

CamScanner 06-05-2022 12.35-1.pdf

p 4 words </>

Question 6

3 pts

Borkiewicz in 1898 fitted he Poisson distribution to the number of deaths from horse kicks in the Prussian Cavalry per corps-year for each of 200 corps-years. His data is given in the following table.

Write your work in the space provided. You may type, or insert a screenshot image or a file.

Number of deaths	Observed number of Corps-years
0	109
1	65
2	22
3	3
4	1
≥5	0
Total	200

(a) Estimate the Poisson parameter using the method used in class examples, calculate the Chi-square statistic used to determine whether the Poisson distribution fits this dataset

(b) Calculate the probability that X is at least 4 using the model you estimated.

(c) Calculate the Chi-square statistic and determine the probability that the chi-square random variable is larger than that value. You may use the Chi-square app posted in the course web site to calculate that probability.

Write your work in the space provided. You may type, or insert a screenshot image or a file.

Edit View Insert Format Tools Table

12pt Paragraph

2	22	$\frac{22}{200} = 0.110$	$\frac{0.61^2 e^{-0.61}}{2!} = 0.101$	$0.101 \times 200 = 20.22$	$(22 - 20.22)^2 = 3.168$	$\frac{3.168}{20.22} = 0.157$
3	3	$\frac{3}{200} = 0.015$	$\frac{0.61^3 e^{-0.61}}{3!} = 0.021$	$0.021 \times 200 = 4.11$	$(3 - 4.11)^2 = 1.232$	$\frac{1.232}{4.11} = 0.300$
4	1	$\frac{1}{200} = 0.005$	$\frac{0.61^4 e^{-0.61}}{4!} = 0.003$	$0.003 \times 200 = 0.63$	$(1 - 0.63)^2 = 0.137$	$\frac{0.137}{0.63} = 0.217$
≥5	0	$\frac{0}{200} = 0$	$1 - P(X \leq 4) = 0.001$	$0.001 \times 200 = 0.16$	$(0 - 0.16)^2 = 0.026$	$\frac{0.026}{0.2} = 0.128$
Total	200	1.000	1.000			

b) We can calculate $P(X \geq 4)$ from the table in a) by calculating

$P(4) + P(X \geq 5)$ as follows:

5

p 315 words </>

Question 7

2 pts

We pick a 6-sided die at random and roll it three times, getting three consecutive sixes. It is known that the probability of picking a loaded die is 0.01, and we know that the probability of 3 sixes given that the die is loaded is $0.5^3 = 0.125.$

We are suspicious that the die we picked is loaded and we want to find the probability that the die is loaded given the evidence that we got three consecutive sixes in the three rolls. Find out the answer showing work.

Enter your work in the space provided. Do not insert screenshots or files, please. Just type the work.

$P(S|L^C) = \left(\frac{1}{6}\right)^3 = \frac{1}{216}$

Apply the Complement Rule:

$P(L^C) = 1 - P(L) = 1 - 0.01 = 0.99$

Plug values in:

$P(S) = 0.00125 + (\frac{1}{216} \times 0.99) = 0.00583$

$P(L|S) = \frac{0.00125}{0.00583} = \boxed{0.214}$

p 60 words </>

Question 8

3 pts

The brown recluse spider often lives in human houses throughout central North America. This spider is a moderate health threat, as its bite causes nasty, slow-healing wounds. Bites are very rarely fatal, but the resulting wounds are disgusting. Information on the spider's diet is useful for developing effective pest management strategies. A diet-preference study gave each of 41 brown recluse spiders a choice between two crickets, one live and one dead. Thirty one of the 41 spiders chose the dead cricket over the live one. Researchers wonder whether this represents evidence for a diet preference. So they phrase their research question as: what would be the probability that of 41 spiders 31 choose a dead cricket or more if there is no diet preference, that is, if the chance of choosing one or the other (a binary choice) is 0.5.? Then they calculate that probability of that happening under the assumption of no diet preference. If that probability is smaller than 0.05 then they conclude that the spider has a diet preference.

So what will be the conclusion of this study using that methodology? Do the spiders show or do not show a diet preference? Explain showing work.

Type answer or insert screenshot image or pdf file.

Edit View Insert Format Tools Table

12pt Paragraph

Bernoulli trials occurring in a sample size of 41, using a binomial model to model X seems appropriate.

Given information:

$P(D) = p = 0.5$

$n = 41$

$x = 31$

Apply the Binomial Distribution App:

p 82 words </>

Question 9

4 pts

Suppose the joint density function of the continuous consumption of corn (X) in a certain region of the world and the daily destruction by war in the region (Y) is given by

$$f(x,y) = 2 - x - y, \quad 0 < x < 1, \quad 0 < y < 1$$

(both X and Y are in billions of tons) .

Showing detail work for partial credit,

(a) Find the marginal density function of the consumption of corn and the marginal density function of the destruction of corn. Are these two random variables independent?

(b) Find the conditional density functions of X and Y.

(c) Calculate the conditional expectations E(X | Y=y) and E(Y | X=x). Notice that we are asking for the general formulas, without assigning a specific value to the variable we are conditioning on.

(d) Show that the E[E(X|y=y)] = E(X).

Answer in the space provided below. You may type your answer using the "insert" then "equation" editor, or insert a screenshot image or insert a pdf file.

Edit View Insert Format Tools Table

12pt Paragraph

$$E[E(X|Y)] = \int_y E(X|Y) f(y)$$

$$E[E(X|Y)] = \int_0^1 \left(\frac{3y-4}{6y-9}\right)\left(-\frac{2y-3}{2}\right) dy = -\frac{3y^2-8y}{12}\Big|_0^1 = \frac{5}{12}$$

Calculate $E(X)$:

$$E(X) = \int_x x f(x) dx$$

$$E(X) = \int_0^1 x\left(-\frac{2x-3}{2}\right)dx = -\frac{4x^3-9x^2}{12}\Big|_0^1 = \frac{5}{12}$$

p

4

66 words </>

Question 10

3 pts

Suppose a discrete random variable X can assume only the values -2, 0, 2 with p(-2)=p(2)=1/8 and p(0) = 3/4.

(a) Showing work, compare the exact value of

$$P[\mid X - E(X) \mid \geq 2]$$

with the value that would be obtained using Chebyshev's inequality.

(b) Give an example seen in class where Chebyshev's is a poor approximation to the true probability. Provide details.

Insert your answers in the space provided below. You may write or insert a screenshot image, or a pdf file.

Edit View Insert Format Tools Table

12pt Paragraph

Plug in:

$$k = \frac{2}{1} = 2$$

$$\boxed{P[|X - E(X)| \geq k\sigma] \leq \frac{1}{4}}$$

b) I couldn't find an example where Chebyshev's was a poor approximation, but there was an example in Lecture 29 where we saw that Markov's was a poor approximation. In this example, we wanted to estimate the probability that there p ▸ strong 130 words </>

Question 11

1 pts

Engineers of the XYZ Engineering corporation use an online terminal to make routine engineering calculations. If the time each engineer spends in a session at a terminal has an exponential distribution with an average value of 36 minutes, find the value of R such that ninety percent of the sessions end in less than R minutes.

- (selected) 82.8 minutes
- 65 minutes
- 91 minutes
- 45 minutes

Question 12

1 pts

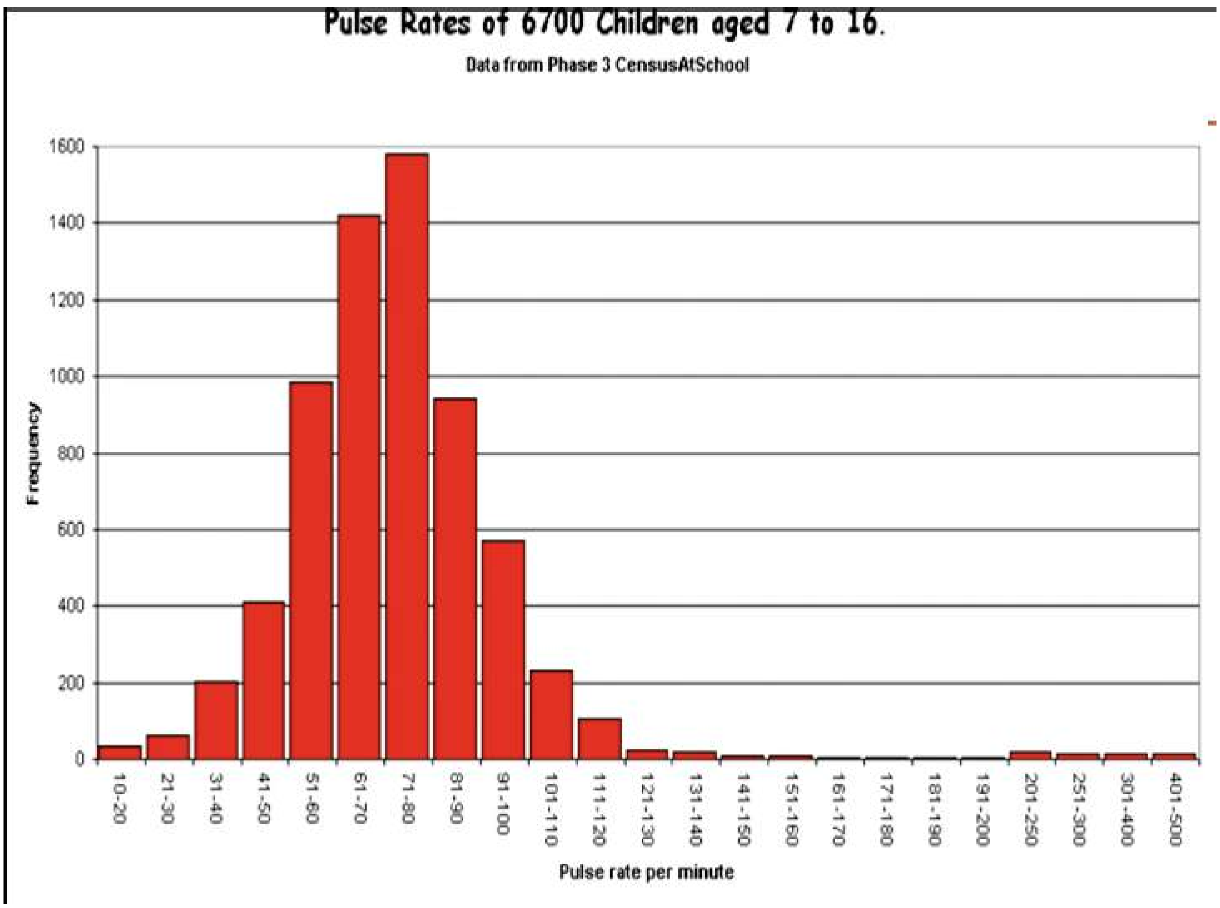
Suppose it has been determined that the number of customers that arrive per second at the central bank 's web site can be described by a Poisson random variable with an average rate of 10 messages per second. What is the probability that 15 or fewer inquiries arrive in a one second period?

- (selected) 0.95126
- 0.13543
- 0.27
- 0.7813

Question 13

2 pts

If we wanted to write down what the sample space is for the experiment "measuring pulse rate of children aged 7 to 16", would all the values observed in 6700 children, as depicted in the image below, be a potential outcome in the sample space? Answer yes or no, in the space provided, and justify your answer.



Edit View Insert Format Tools Table

12pt Paragraph

No, these values would not be a potential outcome in the sample space. The existence of the outliers with a 200+ pulse rate tells us that this data is somehow flawed. By the Central Limit Theorem, a sample of this size should fall into a standard Gaussian distribution, which would make the existence of those outliers so improbable that this dataset could not realistically occur in the sample space.

p

70 words </>

Question 14

1 pts

A professor offers two non-mutually exclusive options for the final assessment of students in a class: a project or a final exam. Students that select both would get the average of the grade in both. A total of 24% of the students select a project, 61% select a final exam and 11% select both. What proportion of the students select only one of the options?

- (selected) 0.63
- 0.74
- 0.85
- 0.26

Question 15

1 pts

If P(A)=1/3, then the odds for A (in favor of A) are

- (selected) 1 to 2
- 2 to 1
- 7 to 3
- 3 to 7

Question 16

1 pts

Consider the experiment of drawing three adults at random from Brazil's population to determine whether they know capoeira or not. The event "at least two of them know capoeira" has how many outcomes in it?

- (•) 4
- ( ) 3
- ( ) 2
- ( ) 5

Question 17

1 pts

A random experiment consists of polling the terminals in a communication line in sequence until either one of the seven terminals on the line is found to be ready to transmit or all the terminals have been polled. Let A be the event that at least five polls are required and B the event that not more than four polls are required. In this example, which of the following is true?

- [x] The probability of A can not be calculated. We need more information
- [ ] The probability of A is 0.5
- [ ] The probability of A is 0.5714
- [x] A and B are mutually exclusive (disjoint)

Question 18

1 pts

Engineers of the XYZ Engineering corporation use an online terminal to make routine engineering calculations. If the time each engineer spends in a session at a terminal has an exponential distribution with an average value of 36 minutes what would be the answer to the following question?

Q. If an engineer has already been at the terminal for 30 minutes, what is the probability that this engineer will spend more than another hour at the terminal?

- (selected) 0.1889
- 0.5654
- 0.9
- 0.1

Question 19

1 pts

Suppose we roll a 3-4 flat dice 5 times and are interested in the sum of the rolls. What is the probability that the sum is 10?

- (selected) 0.010
- 0.016
- 0.023
- 0.015

Question 20

1 pts

$$P(A \cup B) = P(A \cap B) + P(A \cap B^c) + P(A^c \cap B)$$

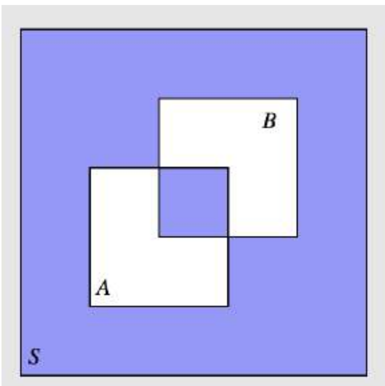
Which of the following would help me, by substitution, go from that formula above to the following result:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

- [x] $P(A \cap B^c)$
- [x] $P(A^c \cap B)$
- [ ] $P(A^c \cap B^c)$
- [ ] $P(A^c)$

Question 21

1 pts



The event shaded in darker color (not white) is

- [x] $(A \cap B) \cup (A^c \cap B^c)$
- [x] $(A \cap B) \cup (A \cup B)^c$
- [ ] $(A \cup B)^c \cap (A^c \cap B)$
- [ ] $(A \cap B) \cup (A \cap B)^c$

Question 22

1 pts

Which of the following collection of events form a partition of the event $A \cup B$?

- (selected) $(A \cap B), (A \cap B^c), (A^c \cap B)$

- ○ $A, (A \cap B), (A \cup B)^c$
- ○ $(A \cap B)^c, (A \cap B)$
- ○ $(A \cup B^c), A, B^c$

Question 23

1 pts

A Stat 100A class has 50 students from the physical sciences or the social sciences. It is known that

25 students are from the physical sciences

25 students are seniors

For 20 students, this is the first statistics class taken

10 students are physical sciences seniors

4 students are physical sciences taking the first stats course

15 students are seniors taking the first stats course

3 students are physical sciences seniors taking the first stats course

What is the probability that a randomly chosen student was a senior not in Physical sciences taking stats for the first time?

- ◉ 0.24
- ○ 0.3
- ○ 0.2
- ○ 0.36

Question 24

2 pts

This question was asked in another question in this exam. Show how you found the answer showing your work.

Suppose that the time, X, between inquiries in a interactive system has an exponential distribution with average value of one second. Let t be an arbitrary point in time and T the elapsed time until the fifth inquiry arrives (after time t). The expected value and standard deviation of T are, respectively, and the distribution of T is

(note: assume that the inquiries come sequentially, one after the other, not all at once).

Edit View Insert Format Tools Table

12pt Paragraph

$$M_T(t) = \left(\frac{1}{1-t}\right)^5$$

We identify this as a *gamma distribution*.

Calculate the expected value of T:

$$M_T'(t) = \frac{d}{dt}\left[\left(\frac{1}{1-t}\right)^5\right] = \frac{5}{(t-1)^6}$$

$$E(T) = M_T'(0) = \frac{5}{(0-1)^6} = \boxed{5}$$

p ▸ strong 47 words </>

Question 25

3 pts

How much mathematical work are you able to show, and show it, to answer the following question mathematically ? Explain why you can or can not calculate the things that you can or can not calculate.

Suppose we roll a 3-4 flat dice 5 times and are interested in the sum of the rolls. What is the probability that the sum is 10?

Be specific about distribution, expectation and variance and assumptions made.

Edit View Insert Format Tools Table

12pt Paragraph

Define events:

X = the result of a single dice roll

Given:

$P(X=1)=P(X=2)=P(X=5)=P(X=6)=\frac{1}{8}$

$P(X=3)=P(X=4)=\frac{1}{4}$

$S_n=\sum_{i=1}^{n} X_i$

p 63 words </>

Quiz saved at 4:05pm Submit Quiz

Find the moment generating function:

$$M(t) = E(e^{tx}) = \sum_{i=1}^{n} e^{tx} P(X)$$

$$M(t) = \sum_{i=1}^{n} e^{ti} \left(\frac{1}{n}\right) = \frac{1}{n} \sum_{i=1}^{n} e^{ti} = \frac{1}{n} e^{t \sum_{i=1}^{n} i} = \frac{1}{n} e^{\frac{tn(n+1)}{2}}$$

Calculate the expected value:

$$E(X) = M'(0)$$

$$M'(t) = \frac{d}{dt}\left(\frac{1}{n} e^{\frac{tn(n+1)}{2}}\right) = \frac{1}{n}\left[\frac{d}{dt}\left(e^{\frac{n(n+1)}{2}t}\right)\right] = \frac{1}{n}\left(\frac{n(n+1)}{2}\right) e^{\frac{n(n+1)}{2}t} = \left(\frac{n+1}{2}\right) e^{\frac{n(n+1)}{2}t}$$

$$E(X) = M'(0) = \left(\frac{n+1}{2}\right) e^{\frac{n(n+1)}{2}(0)} = \boxed{\frac{n+1}{2}}$$

Calculate the standard deviation:

$$E(X^2) = M''(0)$$

$$M''(t) = \frac{d}{dt}\left[\left(\frac{n+1}{2}\right) e^{\frac{n(n+1)}{2}t}\right] = n\left(\frac{n+1}{2}\right)^2 e^{\frac{n(n+1)}{2}t}$$

$$E(X^2) = M''(0) = n\left(\frac{n+1}{2}\right)^2 e^{\frac{n(n+1)}{2}(0)} = n\left(\frac{n+1}{2}\right)^2$$

$$Var(X) = E(X^2) - [E(X)]^2 = n\left(\frac{n+1}{2}\right)^2 - \left(\frac{n+1}{2}\right)^2 = (n-1)\left(\frac{n+1}{2}\right)^2$$

$$\sigma = \sqrt{Var(X)} = \sqrt{(n-1)\left(\frac{n+1}{2}\right)^2} = \sqrt{n-1}\left(\frac{n+1}{2}\right) = \boxed{\frac{\sqrt{n-1}(n+1)}{2}}$$

a) We first calculate λ by taking the total number of deaths observed and dividing it by the number of corps-years:

$$\lambda = \frac{0(109)+1(65)+2(22)+3(3)+4(1)}{200} = 0.61$$

We then calculate the empirical probability of each death count by dividing the number of corps years in which that death count was observed by the total number of corps-years observed. For instance, for 0 deaths, we calculate:

$$\text{Empirical Probability} = \frac{109}{200} = 0.545$$

We then calculate the theoretical probability of each death count by plugging in the death count as n in the following Poisson model:

$$\text{Theoretical Probability} = P(X = n) = \frac{\lambda^n e^{-\lambda}}{n!} = \frac{0.61^n e^{-0.61}}{n!}$$

For instance, for 0 deaths, we calculate:

$$P(X = 0) = \frac{0.61^0 e^{-0.61}}{0!} = e^{-0.61} = 0.543$$

We keep in mind that, by to the axioms of probability, $P(S) = 1$, so when we calculate $P(X \geq 5)$, we don't plug into the Poisson distribution and instead calculate $1 - P(X \leq 4)$, using the complement rule to find the theoretical probability.

Next, we calculate the expected tally, by multiplying the theoretical probability with the total number of corps years we observed. For 0 deaths, this gives us:

$$E = 0.543 \times 200 = 108.67$$

Then, we find the difference between this expected tally and the observed tally, O, and we square the result. For 0 deaths, we get:

$$(O - E)^2 = (109 - 108.67)^2 = 0.109$$

Finally, we take that value and divide it by the expected tally. For 0 deaths, we get:

$$\frac{(O-E)^2}{E} = \frac{0.109}{108.67} = 0.001$$

We apply this process to each death count to derive the following table:

# of deaths	Observed # of Corps-years (O)	Empirical Probability	Theoretical Probability	Expected # of Corps-years (E)	$(O-E)^2$	$\frac{(O-E)^2}{E}$
0	109	$\frac{109}{200} = 0.545$	$\frac{0.61^0 e^{-0.61}}{0!} = 0.543$	$0.543 \times 200 = 108.67$	$(109 - 108.67)^2 = 0.109$	$\frac{0.109}{108.67} = 0.001$

1	65	$\frac{65}{200} = 0.325$	$\frac{0.61^1 e^{-0.61}}{1!} = 0.331$	$0.331 \times 200 = 66.29$	$(65 - 66.29)^2 = 1.664$	$\frac{1.664}{66.29} = 0.025$
2	22	$\frac{22}{200} = 0.110$	$\frac{0.61^2 e^{-0.61}}{2!} = 0.101$	$0.101 \times 200 = 20.22$	$(22 - 20.22)^2 = 3.168$	$\frac{3.168}{20.22} = 0.157$
3	3	$\frac{3}{200} = 0.015$	$\frac{0.61^3 e^{-0.61}}{3!} = 0.021$	$0.021 \times 200 = 4.11$	$(3 - 4.11)^2 = 1.232$	$\frac{1.232}{4.11} = 0.300$
4	1	$\frac{1}{200} = 0.005$	$\frac{0.61^4 e^{-0.61}}{4!} = 0.003$	$0.003 \times 200 = 0.63$	$(1 - 0.63)^2 = 0.137$	$\frac{0.137}{0.63} = 0.217$
≥5	0	$\frac{0}{200} = 0$	$1 - P(X \leq 4) = 0.001$	$0.001 \times 200 = 0.16$	$(0 - 0.16)^2 = 0.026$	$\frac{0.026}{0.2} = 0.128$
Total	200	1.000	1.000			

b) We can calculate $P(X \geq 4)$ from the table in a) by calculating $P(4) + P(X \geq 5)$ as follows:

$$P(X \geq 4) = P(4) + P(X \geq 5) = 0.003 + 0.001 = \boxed{0.004}$$

c) From the table in a), we calculate the Chi-square statistic by taking the sum of $\frac{(O-E)^2}{E}$, or in other words, the sum of the values in the last column:

$$\text{Chi-square statistic} = 0.001 + 0.025 + 0.157 + 0.300 + 0.217 + 0.128 = \boxed{0.828}$$

Using the Chi-square app, we see that the probability that the Chi-square random variable is larger than this value is:

$$P(\text{"Chi-square with 5 degrees of freedom"} > 0.828) = \boxed{0.975}$$

Since the P-square statistic is larger than 0.05. a statistician would conclude that the Poisson model with $\lambda = 0.61$ is a good fit to the horse kick data.

Define events:

Let S = the event in which 3 rolls result in 3 consecutive 6s

Let L = the event in which the loaded die is picked

Given information:

$$P(L) = 0.01$$

$$P(S|L) = 0.125$$

Assume each of the dice rolls are independent

Find $P(L|S)$

Apply Bayes' Theorem:

$$P(L|S) = \frac{P(S|L)P(L)}{P(S)} = \frac{0.125 \times 0.01}{P(S)} = \frac{0.00125}{P(S)}$$

Apply the Law of Total Probability:

$$P(S) = P(S|L)P(L) + P(S|L^C)P(L^C) = (0.125 \times 0.01) + P(S|L^C)P(L^C)$$

Apply the Product Rule for Independent Events:

$$P(S|L^C) = \left(\frac{1}{6}\right)^3 = \frac{1}{216}$$

Apply the Complement Rule:

$$P(L^C) = 1 - P(L) = 1 - 0.01 = 0.99$$

Plug values in:

$$P(S) = 0.00125 + (\frac{1}{216} \times 0.99) = 0.00583$$

$$P(L|S) = \frac{0.00125}{0.00583} = \boxed{0.214}$$

Define events:

Let D = the event in which a brown recluse chooses a dead cricket

Let X = the number of brown recluses that choose a dead cricket in the experiment

Identify model:

Since we are looking to find the probability of a given number of successful Bernoulli trials occurring in a sample size of 41, using a binomial model to model X seems appropriate.

Given information:

$P(D) = p = 0.5$

$n = 41$

$x = 31$

Apply the Binomial Distribution App:

$P(X \geq 31) = 0.00073$

Since $0.00073 < 0.05$, the experiment would conclude that the spider *has a diet preference*.

a) **Find the marginal density function of X:**

$f(x) = \int_y f(x,y)dy$

$f(x) = \int_0^1 (2-x-y)dy = 2y - xy - \frac{y^2}{2}|_0^1 = \boxed{-\frac{2x-3}{2}, \quad 0 < x < 1}$

Find the marginal density function of Y:

$f(y) = \int_x f(x,y)dx$

$f(y) = \int_0^1 (2-x-y)dx = 2x - \frac{x^2}{2} - xy|_0^1 = \boxed{-\frac{2y-3}{2}, \quad 0 < y < 1}$

Determine independence using the definition of independence:

For independent variables X and Y: $f(x,y) = f(x)f(y)$

$2 - x - y = \left(-\frac{2x-3}{2}\right)\left(-\frac{2y-3}{2}\right) = \frac{4xy-6x-6y+9}{4} = xy - \frac{3}{2}x - \frac{3}{2}y + \frac{9}{4}$

Since the two expressions are not equal, we know variables X and Y are *not independent*.

b) **Find the conditional density function of X given Y:**

$f(X|Y) = \frac{f(x,y)}{f(y)}$

$f(X|Y) = \frac{2-x-y}{-\frac{2y-3}{2}} = \boxed{-\frac{2(2-x-y)}{2y-3}, \quad 0 < x < 1, 0 < y < 1}$

Find the conditional density function of Y given X:

$f(Y|X) = \frac{f(x,y)}{f(x)}$

$f(Y|X) = \frac{2-x-y}{-\frac{2x-3}{2}} = \boxed{-\frac{2(2-x-y)}{2x-3}, \quad 0 < x < 1, 0 < y < 1}$

c) **Calculate the conditional expectation $E(X|Y=y)$:**

$E(X|Y) = \int_x xf(X|Y)dx$

$$E(X|Y) = \int_0^1 -x\frac{2(2-x-y)}{2y-3}dx = \frac{2x^3+3x^2y-6x^2}{6y-9}\Big|_0^1 = \boxed{\frac{3y-4}{6y-9}, \quad 0<y<1}$$

Calculate the conditional expectation $E(Y|X=x)$:

$$E(Y|X) = \int_y yf(Y|X)dy$$

$$E(Y|X) = \int_0^1 -y\frac{2(2-x-y)}{2x-3}dy = \frac{2y^3+3xy^2-6y^2}{6x-9}\Big|_0^1 = \boxed{\frac{3x-4}{6x-9}, \quad 0<x<1}$$

d) **Calculate $E[E(X|Y=y)]$:**

$$E[E(X|Y)] = \int_y E(X|Y)f(y)$$

$$E[E(X|Y)] = \int_0^1 \left(\frac{3y-4}{6y-9}\right)\left(-\frac{2y-3}{2}\right)dy = -\frac{3y^2-8y}{12}\Big|_0^1 = \frac{5}{12}$$

Calculate $E(X)$:

$$E(X) = \int_x xf(x)dx$$

$$E(X) = \int_0^1 x\left(-\frac{2x-3}{2}\right)dx = -\frac{4x^3-9x^2}{12}\Big|_0^1 = \frac{5}{12}$$

Compare:

$$\boxed{E[E(X|Y)] = E(X) = \frac{5}{12}}$$

a) **Calculate the exact value of $P[|X - E(X)| \geq 2]$:**

$|X - E(X)| \geq 2 \Rightarrow (X - E(X)) \geq 2$ or $(X - E(X)) \leq -2$

Calculate $E(X)$:

$E(X) = -2(P(-2)) + 0(P(0)) + 2(P(2)) = 0$

Plug in:

$$P[|X - E(X)| \geq 2] = P[|X| \geq 2] = P(-2) + P(2) = \tfrac{1}{8} + \tfrac{1}{8} = \boxed{\frac{1}{4}}$$

Use Chebyshev's inequality:

$P\{|X - E(X)| < k\sigma\} \geq 1 - \frac{1}{k^2}$

Apply Complement Rule:

$P[|X - E(X)| \geq k\sigma] = 1 - P[|X - E(X)| < k\sigma] = 1 - (1 - \frac{1}{k^2})$

$P[|X - E(X)| \geq k\sigma] \leq \frac{1}{k^2}$

Solve for k:

$k\sigma = 2 \Rightarrow k = \frac{2}{\sigma}$

Solve for σ:

$\sigma^2 = E(X^2) - [E(X)]^2$

$E(X^2) = (-2)^2 P(-2) + (0)^2 P(0) + (2)^2 P(2) = 1$

$\sigma^2 = 1 - 0^2 = 1$

$\sigma = 1$

Plug in:

$k = \frac{2}{1} = 2$

$$\boxed{P[|X - E(X)| \geq k\sigma] \leq \frac{1}{4}}$$

b) I couldn't find an example where Chebyshev's was a poor approximation, but there was an example in Lecture 29 where we saw that Markov's was a poor approximation. In this example, we wanted to estimate the probability that there will be 3+ error crashes in a week in any given week, knowing that the average number of system crashes in a week is 2. We compared this estimate to the actual probability, obtained assuming the variable is Poisson. Here, we saw that Markov's Theorem estimated that $P(X \geq 3) \leq \frac{2}{3}$, but the exact probability found using Poisson was $P(X \geq 3) = 0.323$. Clearly, there was a very large difference between the estimate and the actual.

Given:

$\mu_X = 1$

Calculate λ:

$\lambda = \frac{1}{\mu} = \frac{1}{1} = 1$

Use the formula for MGF of an Exponential Distribution:

$M_X(t) = \frac{\lambda}{\lambda - t}$

$M_X(t) = \frac{1}{1-t}$

Use the property that the MGF of a sum of IID random variables is the product of the MGFs:

$T = \sum_{i=1}^{5} X_i \Rightarrow M_T(t) = \prod_{i=1}^{5} M_X(t)$

$M_T(t) = \left(\frac{1}{1-t}\right)^5$

We identify this as a *gamma distribution*.

Calculate the expected value of T:

$M_T'(t) = \frac{d}{dt}\left[\left(\frac{1}{1-t}\right)^5\right] = \frac{5}{(t-1)^6}$

$E(T) = M_T'(0) = \frac{5}{(0-1)^6} = \boxed{5}$

Calculate the standard deviation of T:

$M_T''(t) = \frac{d}{dt}\left(\frac{5}{(1-t)^6}\right) = -\frac{30}{(t-1)^7}$

$E(X^2) = M_T''(0) = -\frac{30}{(0-1)^7} = 30$

$Var(X) = E(X^2) - [E(X)]^2 = 30 - 5^2 = 5$

$\sigma_T = \sqrt{Var(X)} = \sqrt{5} = \boxed{2.24}$

Define events:

X = the result of a single dice roll

Given:

$P(X=1) = P(X=2) = P(X=5) = P(X=6) = \frac{1}{8}$

$P(X=3) = P(X=4) = \frac{1}{4}$

$S_n = \sum_{i=1}^{n} X_i$

$n = 5$

Find $P(S_n = 10)$

Define model for S_n:

By Central Limit Theorem, we can estimate that S_n can be modeled by a standard Gaussian distribution. This may be inaccurate, as the sample size $n = 5$ is relatively small, but we assume it is accurate enough.

Find expectation and variance for X:

$E(X) = 1(P(1)) + 2(P(2)) + 3(P(3)) + 4(P(4)) + 5(P(5)) + 6(P(6))$

$E(X) = 1(\frac{1}{8}) + 2(\frac{1}{8}) + 3(\frac{1}{4}) + 4(\frac{1}{4}) + 5(\frac{1}{8}) + 6(\frac{1}{8}) = 3.5$

$Var(X) = E(X^2) - E(X)^2$

$E(X^2) = 1(\frac{1}{8}) + 4(\frac{1}{8}) + 9(\frac{1}{4}) + 16(\frac{1}{4}) + 25(\frac{1}{8}) + 36(\frac{1}{8}) = 14.5$

$Var(X) = 14.5 - 3.5^2 = 2.25$

Apply Central Limit Theorem:

$E(S_n) = nE(X) = 5 \times 3.5 = 17.5$

$Var(S_n) = nVar(X) = 5 \times 2.25 = 11.25$

$\sigma_{S_n} = \sqrt{Var(S_n)} = \sqrt{11.25} = 3.354$

$P(S_n) \sim N(\mu = 17.5, \sigma = 3.354)$

Use the Normal Distribution App:

$\boxed{P(S_n = 10) \approx 0.01}$