ECO220Y: Homework 9

Note: First do Exercises 1 – 8 on page 4 of "Normal Table: Read it, Use it" (posted on Quercus).

Required Exercises: Chapter 9: 21, 23, 25†, 27†, 31*, 32*, 39, 66*, 71†, 73†, 78*, 81*, 99*, 101*, 106*

* Use the **one-page Normal table** and not a computer or the two-page table at the back of the textbook;

† Note the instruction to do these particular exercises *without* using tables or computers, but by using the Empirical Rule instead. (This is important in building your conceptual understanding of the Normal distribution.)

Required Problems:

(1) A U of T student gets an internship at Barilla Corporation, which makes pasta products. She learns that while the boxes say the weight of pasta contained within is 454 grams there is actually some variation. For the penne shape – her favourite – she discovers that the mean is actually 458 grams with a standard deviation of 5 grams and the distribution is Normal (Bell shaped). When she gets home she is curious about the boxes she happens to have in her pantry. She carefully weighs the pasta from each using a digital scale.

(a) If she has 1 box in her pantry of Barilla penne pasta, what is the chance that it weighs less than what is indicated on the box (i.e. is underweight)?

(b) If she has 4 boxes in her pantry of Barilla penne pasta, what is the chance that at least one of them weighs less than what is indicated on the box?

(c) What is the chance that all four boxes contain less than what is indicated on the box?

(d) Is she adds up the weight of her four boxes, what is the chance the total is less than 454*4?

(e) Suppose she finds three additional boxes deeper in her pantry. What is the chance that the total of all seven boxes is less than 454*7?

(f) Write out the final answers for parts **(a)**, **(d)** and **(e)** (e.g. $P(X_1 < 454) = \#$, …). Why aren't the answers to **(a)**, **(d)** and **(e)** the same? Explain.

(2) In a survey of adult GTA residents one question asks "Do you support Mayor John Tory? Yes or No." For all parts below suppose that in truth exactly 60% of the entire adult population of GTA residents support Tory.

(a) Suppose a random sample of 2,000 Toronto residents is asked the survey question. Draw a *fully labeled graph* of the distribution of the random variable recording the number of people who answer "Yes" to the survey.

(b) What is the probability that the percent of the sample supporting Ford is at least 50 percent but not more than 70 percent?

(c) Compare and contrast your answers with those to Required Problem 4 in HW 8.

(3) Create an "alternative" Empirical Rule where the percentages are round numbers. Find in the missing values.

About 25% of observations are within ± _______ s.d. of mean
About 50% of observations are within ± _______ s.d. of mean
About 75% of observations are within ± _______ s.d. of mean

(4) Consider a September 2019 *NBER Working Paper* "Legacy and Athlete Preferences at Harvard" by Arcidiacono et al. https://www.nber.org/papers/w26316. Here is the abstract.

ABSTRACT: The lawsuit *Students For Fair Admissions v. Harvard University* provided an unprecedented look at how an elite school makes admissions decisions. Using publicly released reports, we examine the preferences Harvard gives for recruited athletes, legacies, those on the dean's interest list, and children of faculty and staff (ALDCs). Among white admits, over 43% are ALDC. Among admits who are African American, Asian American, and Hispanic, the share is less than 16% each. Our model of admissions shows that roughly three quarters of white ALDC admits would have been rejected if they had been treated as white non-ALDCs. Removing preferences for athletes and legacies would significantly alter the racial distribution of admitted students, with the share of white admits falling and all other groups rising or remaining unchanged.

See Table 2 to the right. Recall from the abstract that A = "athlete," L = "legacy," D = "dean's interest list," and C = "children of faculty and staff." An applicant has legacy status if at least one of their parents is a Harvard alumni, but it does not extend to other relatives. Also, the paper describes the dean's interest list as "[it] contains a set of applicants that is of special importance to the dean of admissions. In particular, this list will include applicants whose parents have donated to Harvard, and applicants whose relatives have donated to Harvard." Finally, the academic rating is on scale from 1 to 5 where a 1 is for applicants with "summa potential ... near-perfect [test] scores and grades" and a 5 is for applicants with "marginal potential ... modest grades and 500 scores (25 and below ACT)."

Table 2: Admit Rates of Harvard Applicants by ALDC Status and Academic Rating

Academic Rating		Non-ALDC	LDC	Athlete
1	# of Applicants	612	60	1
	# of Admits	405	58	1
	Admit Rate (%)	66.18	96.67	100.00
2	# of Applicants	59,731	3,118	303
	# of Admits	5,986	1,528	291
	Admit Rate (%)	10.02	49.01	96.04
3	# of Applicants	57,874	2,444	821
	# of Admits	1,390	442	716
	Admit Rate (%)	2.40	18.09	87.21
4	# of Applicants	18,176	373	210
	# of Admits	3	13	167
	Admit Rate (%)	0.02	3.49	79.52
5	# of Applicants	6,335	46	8
	# of Admits	0	0	4
	Admit Rate (%)	0.00	0.00	50.00
Total	# of Applicants	142,728	6,041	1,343
	# of Admits	7,784	2,041	1,179
	Admit Rate (%)	5.45	33.79	87.79

Source: Data presented in Trial Exhibit P618. The final set of rows is computed by the authors using the information in the preceding sets of rows.

(a) In a random sample of three non-ALDC applications, what is the chance that more than one (i.e. greater than 33%) has an academic rating of 2 or better?

(b) In a random sample of 300 non-ALDC applications, what is the chance that more than 100 (i.e. greater than 33%) has an academic rating of 2 or better?

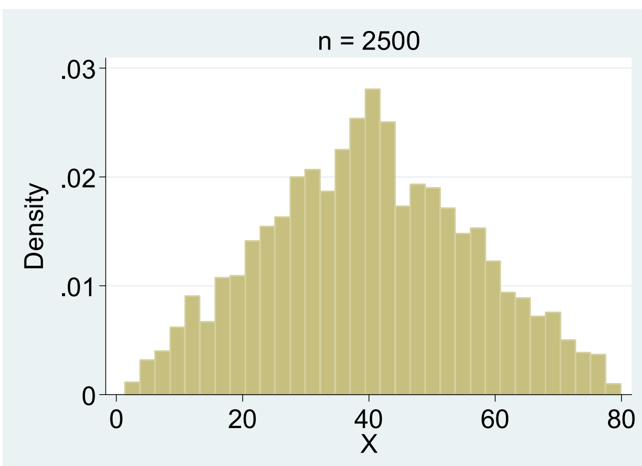
(c) In a random sample of 200 admitted students who are not athletes, what is the chance that fewer than 50 were LDC applicants?

(5) With event planning, there is uncertainty about how many people will accept an invitation. Imagine professional or other events where invitations go individuals, not families or other groups who would attend together. This allows us to treat each person's decision to attend as independent of the decisions of others. With that in mind, assess the following.

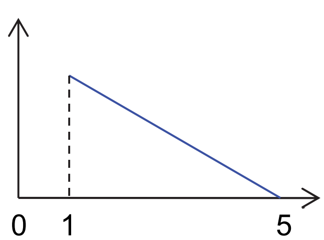
> **(a)** For a small event, the organizer plans to invite 20 people. It is expected that 15 percent of invitees will decline. The organizer can book a room to accommodate up to 18 people for $1,000. If more than 18 people accept, a larger room costing $2,000 will need to be booked. What is the expected cost of this event? If each person that accepts pays $60 for the ticket (which is non-refundable and non-transferable), what is the expected profit from this event (recalling that profit is total revenue minus total cost)?
>
> **(b)** For a large event, the organizer plans to invite 2,000 people. It is expected that 30 percent of invitees will decline. The organizer has a venue that can accommodate 1,500 people and costs $50,000. If more than 1,500 accept, then the stage and seating will need to be reconfigured costing an extra $5,000. What is the expected cost of this event?

(6) Consider a test that measures adults' memory skills and suppose that scores are bell shaped. You take the test and obtain a score of 254.21 and are told that you are in 96.1th percentile. Your friend also takes the test and obtains a score of 161.42 and is told that he is in the 22.8th percentile. What is overall mean and standard deviation of the test scores? Sketch the distribution and show where you are and where your friend is.

(7) To the nearest integer, find the approximate standard deviation of the following histogram. (Hint: Do not try to use the Empirical Rule or Chebycheff's Theorem. Instead use what you know about a distribution of this shape with the parameters that can be estimated from the graph.)



(8) Consider the following density function. Write the exact density function (i.e. formula for it).



(9) Consider rolling one die and creating a random variable that is the value shown. Find the continuous probability distribution that approximates this discrete probability distribution and give its parameters and sketch it. (Hints: Be careful with the endpoints and make sure that the probabilities you get from your proposed continuous approximation are good approximations (e.g. $P(X = 1) \approx P(0.5 < X < 1.5)$).)