E 370 – Fall 2004-05
Exam Two—Practical—Friday, White
Statement of Academic Integrity:

"I swear that I have neither given nor received assistance on this exam and that I will not discuss this exam until all sections have completed it, that is until 14:30 hours on Friday, November 5, 2004."

I have read and agree with the above statement.

Signature:____

Name: (please print)____

Team Number___________________

Instructions

1) Write your name on every page of your exam.

2) Answer the questions in the spaces provided on the exam.

3) **You will have exactly 50 minutes to complete this exam**.

4) **The value of each question is given by each question. Budget your time accordingly.**

5) **Absolutely ALL cell phones must be turned off and out of reach.**

6) You must write any EXCEL functions you choose to use, with arguments, as well as the numerical output. For example, =PERCENTILE(A1:A100,20) = 47. Include 4 decimal places in your answer where applicable.
7) For any question that does not use an Excel command, you must show your work and explanation to receive full credit. This means we must see what your thought process was as you solved the problem. Don't write volumes; just show your work.

8) You may only use EXCEL for this exam. No other calculators or electronic devices of any kind may be on your desk.

9) Only the exam, pencils, erasers, and the gold tool cards may be on your desk. Put all the rest of your belongings along the wall or at the front of the room.

10) Remember, a student is to avoid even the appearance of cheating. Keep your eyes on your exam or on your computer screen. Any questionable behavior on your part is sufficient reason for your coach to confiscate your exam and ask you to leave the room.

11) You may only leave your seat to leave the room. Once you leave your seat, you must turn in your exam.

12) When you are finished, you may turn in all pages of the exam and leave the room as soon as you are able to without disturbing your classmates.

13) Stay calm and do your best.

1. A report issued by the Centers for Disease Control and Prevention (CDC) and the World Health Organization (WHO) stated that one in every seven students in the world, between the ages of 13 and 15, smokes cigarettes. The Monroe County Community Schools Corporation (MCCSC) decided to survey 21 high school students in Bloomington between the ages of 13 and 15 in order to assess these findings.

(a) How many smokers should the MCCSC expect to find? What is its standard deviation? Sketch a graph for this distribution and comment on its shape.

E(X) = 21*1/7 = 3
S(X) = SQRT(21*1/7*6/7) = 1.603567

RIGHT SKEWED (WITH 22 COLUMNS)

(b) What is the probability that at least one of the students will be a smoker?

=1-BINOM.DIST(0,21,1/7,1) = 0.96075

(c) What is the probability that at least 6 but no more than 15 students DON'T smoke cigarettes?

=BINOM.DIST(15,21,6/7,1)-BINOM.DIST(5,21,6/7,1) = 0.06817

(d) What is the probability that no student or every student smoke cigarettes?

=BINOM.DIST(0,21,1/7,0)+BINOM.DIST(21,21,1/7,0) = 0.039275111

(e) What is the minimum number of students you need to sample in order to use the Normal approximation? Explain your answer and use this number to approximate the probability that you will find more than 10 students smoke.

N*pi ≥5 → N*1/7 ≥ 5 → N ≥ 5*7 → N ≥ 35
N*(1-pi) ≥5 → N*6/7 ≥ 5 → N ≥ 5*7/6 → N ≥ 5.833 ≈ 6
Therefore the minimum N is 35
P(X>10) = 1-NORM.DIST(10.5,35*1/7,SQRT(35*1/7*6/7),1) = 0.003944868

[EXTRA CREDIT] The MCCSC will receive an award if at least 12 out of the 15 high schools have less than 5 out of 21 students that smoke (assuming that a sample of 21 students was taken from each school). What is the probability that the MCCSC will receive an award?

P(X<5) =BINOM.DIST(4,21,1/7,1) = 0.829050311
P(Y≥12) =1-BINOM.DIST(11,15,0.82905,1) = 0.753763225

2. The distribution of last year's E370 final exam scores is known to be normal with mean score of 72 and standard deviation 12.

(a) What is the probability that a randomly selected student has a score greater than 80?

$$P[X > \mathbf{80}] = \mathbf{1} - P[X < \mathbf{80}] = \mathbf{1} - norm.dist(\mathbf{80,72,12,1}) = \mathbf{0.2525}$$

(b) What is the probability that a student who is picked randomly has a score between 70 and 80?

$P[\mathbf{70} < X < \mathbf{80}] = P[X < \mathbf{80}] - P[X < \mathbf{70}]$
= NORM.DIST(80,72,12,1) - NORM.DIST(70,72,12,1)
= 0.3137

(c) What is the probability that a randomly selected student has a score of 79?

Since we are dealing with a continuous distribution (the scores are assumed to follow normal distribution which is a continuous distribution), the probability of any observation taking a specific value is zero, i.e., $P[X = 79] = 0$ ***.***

(d) According to the official E370 preferences, it was preferred that the 60% of the students get at least 72. However, this was not so in the initial grading so it was proposed that graders should use more lenient grading. In doing so the mean score will increase to satisfy the stated preference above. If the variability of the scores remains the same, what should be the mean score in order for the E370 official preferences to be met? (Hint: Use Z score)

First find the Z-value above which we expect have 60% of probability, or find a Z-value below which there is 40% probability i.e., Z=norm.s.inv(.4)=-0.25. Then solve the following equation for mean:
(72-mean)/12=-0.25335 so that we get mean = 75.04
Or, norm.inv(.6,72,12)=75.04

(e) Now suppose the standard deviation of the E370 final scores is not known but the mean is still 72. A random sample of 20 students from that class reveals that the sample standard deviation is 15. What is the probability that a randomly selected student from the class has a score more than 74?

Since the variance of the population is not given, we use t-distribution. Letting X to be the random variable representing scores in E370 final exam, we need to find the probability of $P[X > 74]$ ***.***

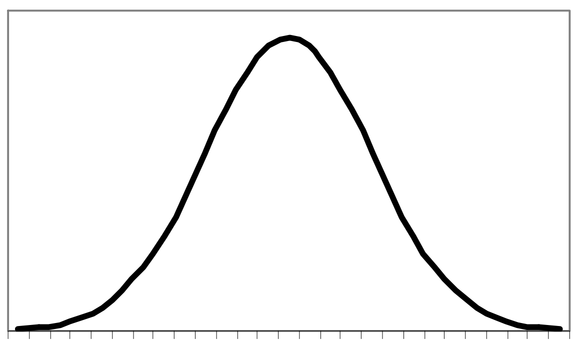
Define $t_{\text{more}} = \frac{74 - 72}{15} = 0.13$ ***. Then,*** $P[\text{X} > 74] = P[\text{t} > 0.13] = \mathbf{1} - T.DIST(\mathbf{0.13,19,1}) = \mathbf{0.448966}$

3. Some parameters for two variables are listed in the tables below. "Incomes" refers to the annual income of a population of blue-collar workers working for Delco, a manufacturer of automobile parts and is measured in dollars. "Bulbs" refers to the burn time of GE's brand of "long-life" light bulbs and is measured in hours. Use this information to answer the questions below.

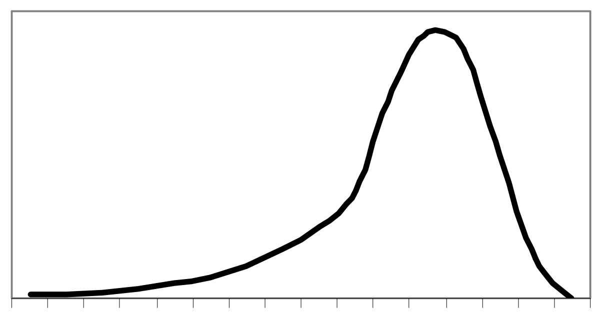
Incomes	
Mean	29120
Median	29120
Mode	29110
Standard Deviation	8700
Variance	75690000

Bulbs	
Mean	5025
Median	5095
Mode	5098
Standard Deviation	362
Variance	131044

Delco Blue Collar Worker Salaries



Long-Life Light Bulb Lifetimes



(a) Suppose that you worked as a statistician for GE. GE knows that consumer groups are going to try to check if its claim that the bulbs burn at least 5,000 hours is valid. Such groups are notorious for using small samples, in this case a sample of 16. How would means of samples of this size be distributed?

$\bar{X}$ ***~?(5025,362/√16), standard error is 362/4 = 90.5***

(b) To counter possible "negative" outcomes from consumer groups tests, you prepare a set of parameters for samples of size 49. How would means of samples of this size be distributed?

$\bar{X}$ ***~N(5025,362/√49), standard error is 362/7 = 51.7143***

(c) Calculate the probability that a Delco blue-collar worker's annual income will exceed $30,000.

=1-NORM.DIST(30000,29120,8700,1) = 0.4597

(d) Calculate the probability that the mean annual income of groups of 25 Delco blue-collar workers will exceed $30,000.

=1-NORM.DIST(30000,29120,8700/√25,1) = 0.3065

(e) Will the probability that the mean annual income of groups of 100 Delco workers will exceed $30,000 be larger than/equal to/smaller than the probability you calculated in d)? Why? (In 5 words or less.)

smaller than; Standard error is smaller; sample size is larger; distribution is more focused.