STAT 217
Assignment #1

Note: answers may vary slightly due to rounding and whether or not you use the computer or the tables.

Login with your ID and pass word. If you do not have this, you need to go the 7th floor of the Math Science building and get it. You will have to wait at least 45 minutes until you can use your password after getting it.

Go to the MINITAB program

MINITAB INSTRUCTIONS
CALC ⇒Probability Distributions⇒normal.

Finding the area above and below a Z-value under the Standard Normal Curve on the computer.
For a given Z-value, we want to find a probability. In the dialog box, which corresponds to the Normal distribution, you have three choices:

Probability
Cumulative probability
Inverse cumulative probability

Click **Cumulative probability**. This will calculate the cumulative probability associated with a specific Z-value (or the area under the Standard Normal Curve to the **left of a specific Z-value**.)
The middle of the dialog box has 2 options:

Mean
Standard Deviation

The default values for the **Mean** and **Standard Deviation** are **0** and **1**, respectively. There is no need to change these values, so just leave these as is.

In the bottom portion of the dialog box, select **Input constant.** It is in this box that you enter a specific Z-value. Once you have completed this, either press return or "click" on **OK**. In the upper portion of your screen, or the command module, MINITAB will return the area to the **left** of the Z-value you have entered. Note that when this routine is employed, the probability returned is **ALWAYS THE AREA TO THE LEFT OF** the Z-value entered above or $P(Z<z)$

For practice, try question 1 using this routine. It will help if you also draw pictures of the distribution.

1. Given that Z is a standard normal random variable, compute the following probabilities:

(a) $P(Z \le -2.44)$ (.0073)
(b) $P(Z \ge 1.62)$ (.0526)
(c) $P(Z \ge -1.38)$ (.9162)
(d) $P(-0.72 \le Z \le 0)$ (.2642)
(e) $P(-0.62 \le Z \le 0.62)$ (.4648)
(f) $P(-0.45 \le Z \le 2.11)$ (.6562)
(g) $P(0.34 \le Z \le 2.33)$ (.3570)

Finding a Z-value for a given area (or probability under the Standard Normal Curve)
For a given probability, you are required to find a Z-value that corresponds to this probability. This requires the use of the **Inverse cumulative probability** routine in the dialog box employed above.

"Click" on the circle which corresponds to **Inverse cumulative probability** and just as was done previously, do not touch the box labeled **Mean** and **Standard Deviation**.

This routine needs an area, and will subsequently find the Z-value which matches up with the area entered. Just as was done above, move your mouse down to the bottom portion of the dialog box and "click" on the circle which corresponds to **Input constant**. Previously you entered a Z-value here. But now you want to find a Z-value for a given area, or probability. So the number you will enter in the **Input constant** box is a probability, or an area to the left of the Z-value in question. Once you have entered the correct probability, either press return or "click" on **OK.** MINITAB will return a Z-value in the command module on the upper portion of your screen.

A good rule-of-thumb in these types of problems is to draw your standard normal curve and piece together the areas given. The Z-value will be given when you specify the area to the left of that value. Practice this routine on question 2.

2. Given that Z is a standard normal random variable, determine $\mathbf{Z_o}$ if it is known that:

(a) $P(-Z_o \le Z \le Z_o) = 0.90$ (1.645)
(b) $P(-Z_o \le Z \le Z_o) = 0.10$ (.1257)
(c) $P(Z \ge Z_o) = 0.20$ (.842)
(d) $P(-1.66 \le Z \le Z_o) = 0.25$ (-.529)
(e) $P(Z \le Z_o) = 0.40$ (-.253)
(f) $P(Z_o \le Z \le 1.80) = 0.20$ (.720)

3. Assume that the height of male college students are normally distributed with a mean of 178.05 cm and a standard deviation of 6.86 cm.
   (a) Find the percentage of male students who fall between the Toronto Maple Leaf's average height of 183.9 cm and the Philadelphia Flyers' average height of 189.48 (15.02%)
   (b) If the question in (a) was changed to inches (conversion is to divide cm by 2.54), would the answer change?
   (c) Among 500 randomly selected male college students, how many would fall within the interval in (b)? (75.1)

4. Find the percentage of years in which Canadian imports from the Middle East and Africa are between \$339 million and \$8004 million. Assume that the annual import values exhibit no trend, and are normally distributed with a mean of \$2502 million and a standard deviation of \$905 million (99.15%)

5. Assume that human body temperatures are normally distributed with a mean of 36.4°C and a standard deviation of 0.62°C.
   (a) If we define a fever to be a body temperature above 37.8°C, what percentage of normal and healthy persons would be considered to have a fever? Does this percentage suggest that a cutoff of 37.8°C is appropriate? (1.19%)
   (b) If we defined a fever so that only 0.5% of the population would have a body temperature above a certain temperature, what is this temperature? (37.997°C)

6. The time it takes Bob to get from home to his office follows a normal distribution. The probability that it takes him less than 3 minutes is 0.345. The probability that it takes him more than 10 minutes is 0.01. Find the average time and variance (μ and σ^2) of this normal distribution. (~4.0256, ~6.5746) Hint: substitute one equation into the other.

7. It takes on average 12.3 minutes to run a race with a standard deviation of 0.4 minutes.
(a) What is the probability that the runner will take between 12.1 and 12.6 minutes to finish the race? (.4649)
(b) What is the maximum time (in minutes) the runner must have for the time to be classified "among the **fastest** 5% of his times"? (11.642 min)

(c) The times for a random sample of 6 of runners is considered. What is the probability that the average time for this sample is more than 12.75 minutes? (0.0039)

8. In human engineering and product design, it is often important to consider the weights of people so that airplanes or elevators aren't overloaded, chairs don't break, and other such dangerous or embarrassing mishaps do not occur. Given that the population of players in offensive back positions (including quarterback) in the CFL have weights that are approximately normally distributed, with a mean of 197.5 lb and a standard deviation of 14.2 lb

(a) Find the probability if one player is randomly selected, his weight is greater than 200 lb. (.4286)

(b) If 36 different players are randomly selected, find the probability that their mean weight is greater than 200 lb.? (0.1469)

9. Electrical connectors last on average 18.2 months with a standard deviation of 1.7 months. Assume that the life of the connectors is normally distributed.

(a) The manufacturer agrees to replace, free of charge, any connectors that fail within 17 months of installation (warranty). What percentage of the connectors can he expect to have replaced free of charge? (.2404)

(b) The manufacturer does not want to have to replace more than 2.5% of the connectors free of charge. What should he set the warranty for (number of months for free replacement)? (14.868)

The directions for the t distribution for MINITAB is the same as the directions for the standard normal. The only difference is that you have to also plug in degrees of freedom.

10. (a) $P(T \le -1.44)$ df = 6 (0.1)
(a) $P(T \ge 0.52)$ df = 20 (.3044)
(b) $P(T \ge -1.52)$ df = 10 (.9203)
(c) $P(-0.33 \le T \le 0)$ df = 15 (.1270)
(d) $P(-0.62 \le T \le 0.62)$ df = 10 (.4508)
(e) $P(-0.50 \le T \le 1.98)$ df = 12 (.6513)
(f) $P(0.34 \le T \le 2.33)$ df = 16 (.3525)

11. Determine the value of $\mathbf{T}_o$ if it is known that:

(g) $P(-T_o \le T \le T_o) = 0.90$ df =5 (2.015)
(h) $P(-T_o \le T \le T_o) = 0.10$ df = 10 (.1289)
(i) $P(T \ge T_o) = 0.20$ df = 25 (.8562)
(j) $P(-1.66 \le T \le T_o) = 0.25$ df = 12 (-.5049)
(k) $P(T \le T_o) = 0.40$ df = 15 (-.2579)

Note: If you want to calculate the mean and standard deviation of a data set,

1. input all the data into one column.
2. Click on the header Calc>Column Statistics.
3. Click on the statistic that you are interested in (mean, st.dev etc)
4. Type the column in which the data is in, in the input variable box (or click on the input variable box and then double click on the column where the data is located.
5. Hit enter or click on OK

Note: You should familiarize yourself with some of the other functions of MINITAB. Check them out. You may find some time saving techniques.

Confidence intervals

1. In developing patient appointment schedules, a medical center desires to estimate the mean time a staff member spends with each patient. How large a sample should be taken if the precision of the estimate is to be ± 2 minutes at a 95% level of confidence? How large a sample is needed for a 99% level of confidence? Use a planning value for the population standard deviation of 8 minutes. [62, 107]

2. A simple sample of five people provided the following data on ages: 21, 25, 20, 18, and 21. Develop a 95%confidence interval for the mean age of the population being sampled. State any assumptions you must make in you method. [(17.8349, 24.1651), use t-distribution because σ is not known, small random sample, assume normal population]

3. The time (in minutes) taken by a biological cell to divide into two cells has a normal distribution. From past experience, the standard deviation can be assumed to be 3.5 minutes. When 16 cells were observed, the mean time taken by them to divide was 31.2 minutes. Estimate the true mean time for a cell division using a 98 percent confidence interval. [29.1645, 33.2355 use z distribution (σ=3.5)]

4. Based on a random sample of 100 cows of a certain breed, a confidence interval for estimating the true mean yield of milk is given by $41.6 < \mu < 44.0$. If the yield of milk of a cow may be assumed to be normally distributed with a standard deviation of 5, what was the level of confidence used? [98.36%]

5. When 16 cigarettes of a particular brand were tested in a laboratory for the amount of tar content, it was found that their mean content was 18.3 milligrams with a standard deviation of 1.8 milligrams. Set a 90 percent confidence interval for the mean tar content in cigarettes of this brand. (Assume that the amount of tar in a cigarette is normally distributed and sample is random.) [(17.5111, 19.0889) t-distribution]

6. In 10 half-hour programs on a TV channel, Mary found that the number of minutes devoted to commercials were 6, 5, 5, 7, 5, 4, 6, 7, 5, and 5. Set a 95% confidence interval for the true mean time devoted to commercials during a half-hour program. Assume that the amount of time devoted to commercials is normally distributed. [(4.8049, 6.1951), t-distribution]

7. A random sample of 16 servings of canned pineapple has a mean carbohydrate content of 49 grams. If it can be assumed that population is normally distributed with a variance of 4 grams, find a 98 percent confidence interval for the true mean carbohydrate content of a serving. [(47.835, 50.165) z-dist]

8. An archaeologist found that the mean cranial width of 17 skulls was 5.3 inches with a standard deviation of 0.5 inches. Using a 90% confidence level, set a confidence interval for the true mean cranial width. Assume that the cranial width is normally distributed. [(5.0883,5.5117) t-dist]

9. It is suspected that a substance called actin is linked to various movement phenomena of non-muscle cells. In a laboratory experiment when eight fertilized eggs were incubated for 14 days the following amounts (mg) of total brain actin were obtained: 1.2, 1.4, 1.5, 1.2, 1.4, 1.7, 1.5, 1.7. Assuming that brain-actin amount after 14 days of incubation is normally distributed,
   (a) Find a 95 percent confidence interval for the true mean brain-actin amount. [(1.2890, 1.6111) t-dist]
   (b) How can we decrease/increase the error? Assume that the variability does not change from the data given above. (Hint: there are two things that we can change)

10. An economist wants to estimate the mean income for the first year of work for a college graduate who has had the profound wisdom to take a statistics course. How many such incomes must be found if we want to be 95% confident that the sample mean is within $500 of the true population mean? Assume that a previous study has revealed that for such incomes, σ = 6250. [601]

11. If we want to estimate the mean weight of plastic discarded by households in one week, how many households must we randomly select if we want to be 99% confident that the sample mean is within 0.250 lb of the true population means when preliminary results show that the standard deviation is 1.067? [121]

12. Wawanesa Mutual Insurance Company wants to estimate the percentage of drivers who change tapes or CDs while driving. A random sample of 850 drivers results in 544 who change tapes or CDs while driving.
   (a) Find the point estimate of the percentage of all drivers who change tapes or CDs while driving. [64.0%]

(b) Find a 90% interval estimate of the percentage of all drivers who change tapes or CDs while driving. [61.29%<p<66.71%]

13. In a study of store checkout scanners, 1234 items were checked and 20 of them were found to be overcharges.
    (a) Using the sample data, a confidence interval for the proportion of all such scanned items that are overcharges was found to be from 0.00915 to 0.02325. What was the level of confidence that was used? [~95% level of confidence]
    (b) Find the sample size necessary to estimate the proportion of scanned items that are overcharges. Assume that you want 99% confidence that the estimate is in error by no more than 0.005.
        (i) Use the sample data as a pilot study [4228, Note: the sample size may be off due to rounding]
        (ii) Assume, instead, that we do not have prior information on which to estimate the value of $\hat{p}$. [66307, Note: sample size may be off due to rounding]

14. A hotel chain gives an aptitude test to job applicants and considers a multiple-choice test question to be easy if at least 80% of the responses are correct. A random sample of 6503 responses to one particular question includes 84% correct responses. Construct the 99% confidence interval for the true percentage of correct responses. Is it likely that the question is really easy? Why? [82.83%<p <85.17%, yes]

15. In a survey, 1039 adults were asked " How much respect and confidence do you have in the public school system?" The results, reported in the Toronto Star (Sept. 26, 1988), are shown below:

Responses	A great deal	Quite a lot	Some	Very little	No opinion
Percentages	12%	30%	35%	13%	10%

    Estimate with a 90% confidence the proportion of all adults who had "a great deal" or quite a lot" of respect for the public school system. Interpret this interval [.3948, .4452]

16. Of the 200 individuals interviewed, 80 said they were concerned about fluorocarbon emissions in the atmosphere. Obtain a 99% confidence interval estimate for the true proportion of individuals who are concerned. [.3108, .4892] Interpret this interval.

17. As part of a study of post-secondary education, a random sample of the graduating classes of colleges and universities is to be selected to estimate their expected success in finding employment. It is desired to estimate the success rate to within ± 0.01, with a confidence of 95%. No reliable planning value for the success rate is available.
    (a) What is a conservatively large sample size to meet the precision requirements? [9604, Note; sample size may be off due to rounding]
    (b) It was finally decided to select 2500 graduating students for the sample. Of these, 1141 were successful in finding immediate employment. Estimate the true success rate of this graduating class in a 99% confidence interval. Interpret the meaning of the interval. [.4307, .4821]

18. A marketing research organization wishes to estimate the proportion of television viewers who watch a particular prime-time situation comedy on December 14. The proportion is expected to be approximately 0.30. At a minimum, how many viewers should be randomly selected to ensure that a 95% confidence interval for the true proportion of viewers will have a width of at most 0.01? [32270, Note: sample size may be off due to rounding]