Math 323

Lab #4

Review order statistics.

Please note: Answers may vary depending on rounding and whether the table or computer is used.

Minitab instructions
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.

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

(a) $P(Z \le -1.44)$	(.0749)
(b) $P(Z \ge 0.52)$	(.3015)
(c) $P(Z \ge -1.52)$	(.9357)
(d) $P(-0.33 \le Z \le 0)$	(.1293)
(e) $P(-0.62 \le Z \le 0.62)$	(.4648)
(f) $P(-0.50 \le Z \le 1.98)$	(.6676)
(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 the value of $\mathbf{Z_o}$ if it is known that:

(a) $P(-Z_o \leq Z \leq Z_o) = 0.90$ (1.645)
(b) $P(-Z_o \leq Z \leq Z_o) = 0.10$ (.1257)
(c) $P(Z \geq Z_o) = 0.20$ (.842)
(d) $P(-1.66 \leq Z \leq Z_o) = 0.25$ (-.529)
(e) $P(Z \leq Z_o) = 0.40$ (-.253)
(f) $P(Z_o \leq Z \leq 1.80) = 0.20$ (.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 between the interval in (b)? (75.1)

4. 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)
   (c) If we convert Celsius to Fahrenheit, do the answers change?

5. The time it takes a runner to run a certain course follows a normal distribution. The probability that it takes him less than 5 minutes is 0.1515. The probability that it takes him/her more than 7 minutes is 0.05. Find the average time and variance (μ and σ^2) of this normal distribution. (~5.7747, ~.5548)

6. 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)

7. SAT verbal scores are normally distributed with a mean of 430 and a standard deviation of 120. Randomly selected SAT verbal scores are obtained from the population of students who took a test preparatory course from a training school.
(a) If one of the students is randomly selected, find the probability that he or she obtained a score greater than 440.
(b) If 100 students of an SAT preparation course achieve a sample mean of 440, does it seem reasonable to conclude that the course is effective because the students perform better on the SAT?

8. Lengths of pregnancies are normally distributed with a mean of 268 days and a standard deviation of 15 days.
   (a) If we stipulate that a baby is premature if born at least three weeks early, what percentage of babies are born prematurely? (.0808)
   (b) A wife claimed to have given birth 308 days after a brief visit from her husband. Find the probability of a pregnancy lasting 308 days or longer. What does the result suggest? (.0038)
   (c) If 25 randomly selected females are put on a special diet just before they become pregnant, find the probability that their lengths of pregnancy have a mean that is less than 260 days.9. (.0038)

9. The total weights of garbage that households discard weekly is approximately normally distributed, with a mean of 12.5 kg and a standard deviation of 5.7 kg.
   (a) If 120 households are randomly selected, find the probability that the mean weight of their discarded garbage is over 13.5 kg. (.0274)
   (b) If the town's waste transfer stations allocated capacity of 1690kg of temporary garbage storage per 120 households per week, and the typical garbage route is based on 120 households, what percentage of garbage routes will exceed the allocated capacity for their garbage? Is this an acceptable level or should the town council take corrective action? (.0012)

The directions for the t, chi-square, and F distribution for MINITAB are the same as the directions for the standard normal. The only difference is that you have to also plug in degrees of freedom. Practice going between the tables and the computer.

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.

Some of these are best found using the computer.

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

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

(g) $P(-T_o \leq T \leq T_o) = 0.90$ df =5 (2.015)
(h) $P(-T_o \leq T \leq T_o) = 0.10$ df = 10 (.1289)
(i) $P(T \geq T_o) = 0.20$ 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)

12. A firm establishes a committee to investigate the amount each contract costs over and above the amount quoted in the original contract (overruns). The committee has determined that the standard deviation of overruns is $17,500.
(a) The average overrun for a random sample of 50 contracts is $12,000. Determine a 98% confidence interval estimate of the true mean overrun based on this sample. {$6233.54,$17,766.45}
(b) A random sample of 36 contracts is selected to estimate the average overrun. What is the probability that the sample mean will over-estimate the population mean overrun by at least $5000 {.0436}

13. 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%]

14. The earnings per share for a random sample of technology stocks listed on the NYSE were (in $'s):

1.90	2.15	2.01	0.89	1.53	1.89	2.12	2.05	1.75	2.22	3.44

(a) Assuming that earnings per share are normally distributed, determine a 95% confidence interval estimate of the average earnings per share of the NYSE technology stocks. { $\bar{y} = 1.9955$, s = 0.608} {$1.58, $2.41}
(b) A broker stated that the NYSE technical average earning was $1.25 per share. Do the data confirm this or not. Use the results of (a) only
(c) How can we decrease/increase the error? Assume that the variability does not change from the data given above.

15. 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.) [17.5111, 19.0889]

16. 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]

17. A random sample of 41 quarters has a mean weight of 5.622g and a standard deviation of 0.068g.
    (a) Construct a 98% confidence interval estimate of the population mean of all quarters in circulation. [5.5963g, 5.6477]
    (b) The U.S. Department of the Treasury claims that it mints quarters to yield a mean weight of 5.640g. Is this claim consistent with the confidence interval? Explain why.

18. The 95% confidence interval for the true mean distance by male students in one year is 11,290 to 12,466. This was based on a sample of 121 randomly selected male students. Find the sample standard deviation that was used. [s=3266.6667]

19. 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]

20. Because a proposed survey is time-consuming, an enterprising pollster posts it on the Internet and promises free software to everyone who responds by completing the survey. Results include 2250 responses, and 80% of them indicate that a fax machine is owned. Construct a 95% confidence

interval for the percentage of all people who have a fax machine. Are the results valid? Why or why not? [78.3%, 81.7%] Results are not valid because the sample is self-selected (not a random sample)

21. A random survey of 85 CEOs in British Columbia showed 70 respondents have a computer on their desk. Based on those results, construct a 98% confidence interval for the percentage of all CEOs in British Columbia who **do not** have a computer on their desk. (.0802, .2728)

22. The drug Ziac is used to treat hypertension. In a clinical test, 3.2% of 221 Ziac users experienced dizziness.
    (a) Construct a 99% confidence interval estimate of the percentage of all Ziac users who experience dizziness.(.0015, .0625)
    (b) In the same clinical test, people in the placebo group didn't take Ziac but 1.8% of them reported dizziness. Based on the results in parts (a) and (b), what can we conclude about dizziness as an adverse reaction to Ziac?

23. **Some of these are best found using the computer.**

(a) $P(\chi^2 \leq 19.02)$	df = 9	(.975)
(b) $P(\chi^2 \leq 1.44)$	df = 6	(0.0366)
(c) $P(\chi^2 \geq 10.52)$	df = 20	(.9577)
(d) $P(4.33 \leq \chi^2 \leq 1\ .33)$	df = 15	(.1979)

24. Determine the value of $\chi^2_{\boldsymbol{o}}$ if it is known that:

(a) $P(\chi^2 \geq \chi^2\ _o) = 0.20$	df = 25	(.30.6752)
(b) $P(5.66 \leq \chi^2 \leq \chi^2\ _o) = 0.25$	df = 12	(9.2412)
(c) $P(\chi^2 \leq \chi^2\ _o) = 0.40$	df = 15	(13.0297)

23. Suppose that a machine dispenses sand into bags. A random sample of 100 bags is taken from a new machine and the standard deviation is 14kg. The population standard deviation is known to be 18kg. Is 14 significantly lower that 18 kg?[χ^2 =59.8889, P(χ^2<59.8889, df = 99)= .0007, yes]

24. Last year, the mean number of books borrowed per cardholder at a major university was 18.2 books per semester with a st.deviation of 4.2. A random sample of 25 cardholders showed the following results for this semester: s^2=6.17.
    a. The library administration would like to know whether this semester's variance is smaller than last year's variance. What assumptions were made? [χ^2=8.39, P(χ^2<8.39, df = 24)= .0013, yes]
    b. Construct a 95% confidence interval for σ^2. Comment on this. [3.7618,11.9409]
    c. Construct a 95% confidence interval for σ. Comment on this. [1.9395, 3.4556]

1