Department of Epidemiology and Biostatistics
McGill University

EPI 513-607 (Inferential Statistics)
Midterm Examination
November 4, 1993

INSTRUCTIONS

The answers are to be written in the spaces provided.

When writing, be brief and W R I T E C L E A R L Y .

Unless specifically asked for, complete calculations are not needed. To avoid writing out formulae, just indicate which table or formula would be appropriate and give a reference; explain where one obtains each of the components of the formula.

The points add up to 120 so one point deserves one minute of effort.

To allow choices, the mark you get out of 120 will count as if it is out of 100. Thus, for example, if you get 75/120, it will count as 75%.

_____
your ID number or *nom-de-plume*

2 points each

___ In a study of a large number of persons over the age of 75, the mean age was 79.9 (sd=4.2). What can you say about the shape of the age distribution?

___ In a study of the effects of acid rain, a random sample of 100 trees from a particular forest are examined. Forty percent of these show some signs of damage. Which of the following statements are correct?
(a) 40% is a parameter
(b) 40% is a statistic
(c) 40% of all trees in this forest show signs of damage
(d) more than 40% of the trees in this forest show some signs of damage
(e) less than 40% of the trees in this forest show some signs of damage

___ You have measured the systolic blood pressure (SBP) of a random sample of 25 employees of a company. A 95% confidence interval for the mean systolic blood pressure for the employees of this company is (122, 138). Which of the following statements gives a valid interpretation of this interval?
(a) 95% of the sample of employees have a systolic blood pressure between 122 and 138.
(b) 95% of the population of employees have a SBP between 122 and 138.
(c) If the procedure were repeated many times, 95% of the resulting confidence

intervals would contain the population mean SBP
(d) The probability that the population mean SBP is between 122 and 138 is .95.
(e) If the procedure were repeated many times, 95% of the sample means would be between 122 and 138.

___ A significance test was performed to test the null hypothesis H_0: $\mu = 2$ versus the alternative H_a: $\mu \neq 2$. The test statistic is $z = 1.40$. The P-value for this test is approximately

(a) .16 (b) .08 (c) .003 (d) .92 (e) .70

___ A significance test gives a P-value of .04. From this we can
(a) reject H_0 at the $\alpha = .01$ level
(b) reject H_0 at the $\alpha = .05$ level
(c) say that the probability that H_0 is false is .04.
(d) say that the probability that H_0 is true is .04.
(e) none of the above

___ A 95% confidence interval for the mean reading score for a population of third grade students is (44.2, 54.2). Suppose you compute a 99% confidence interval. Which of the following statements is correct?
(a) both intervals have the same width
(b) the 99% interval is wider
(c) the 95% interval is wider
(d) you cannot determine which interval is wider unless you know n and s

___ You are reading an article in your field. The article says that the P-value for a significance test is 0.045. Is this result significant at the 5% significance level?

___ You want to compute a 90% confidence interval for the mean of a population with unknown population standard deviation. The sample size is 30. The value of t you would use for this interval is
(a) 1.96 (b) 1.645 (c) 1.699 (d) .90 (e) 1.311

___ *True or False ?* A 90% confidence interval will have a larger margin of error than a 95% confidence interval using the same data.

___ *True or False ?* In a test of significance, a P-value of 0.03 means that there is only probability 0.03 that the null hypothesis is true.

___ *True or False ?* The margin of error in a confidence interval does not take into account errors from such sources as failing to contact subjects or incorrect responses due to the subject's bad memory.

4 points each

___ The Standard Deviation (SD) of the measurements {1, 2, 3, 4, 5} is 1.6.

What is the SD of the measurements {10, 20, 30, 40, 50}? _____

What is the SD of the measurements {11, 21, 31, 41, 51}? _____

___ Scores on a standard IQ test are normally distributed with mean 100 and standard deviation 15. How high must a person's IQ be in order to fall in the highest 25%?

___ A dataset consists of 20 zeros (0's) and 80 ones (1's). What is the approximate SD of these 100 integers ?

___ Construct a dataset with 100 integers so that their mean is 5 and their SD is approximately 2.

___ The diameter of red blood cells in healthy adults has a Gaussian distribution with mean 7.5 microns and standard deviation 0.3 microns. What fraction of red blood cells have diameter between 7.1 and 7.6 microns?

Would the standard deviation of red blood cells be about the same in a bigger person (with more red blood cells) as in a smaller person (with fewer red cells)? Why/why not?

6 points each

___ A study of emergency medical care looks at medical calls to the 911 emergency number in a large city. Among the variables studied are the time X (in minutes) required for the operator to gather information and alert an ambulance service, the time Y required for the ambulance to reach the patient, and the additional time Z required to bring the patient to a hospital. The study finds that X, Y, and Z are independent (and therefore uncorrelated) with the following means and standard deviations.

$\mu_X = 2.3 \quad \sigma_X = 1.6$
$\mu_Y = 25.2 \quad \sigma_Y = 14.4$
$\mu_Z = 10.9 \quad \sigma_Z = 6.3$

From this information, could you find the mean of the total time $T = X + Y + Z$ from the time a call reaches 911 until the patient reaches a hospital? If so, explain how *but don't finish the calculation*. If not, say why not.

Can you find the standard deviation of the total time $T = X + Y + Z$ from this information? If so, explain how, *without doing the calculation*. If not, why not?

If X, Y and Z are non-Gaussian, what can you say about the shape of the distribution of T ?

___ An athlete suspected of having used steroids is given two tests that operate independently of each other. Test A has probability 0.9 of being positive if steroids

have been used. Test B has probability 0.8 of being positive if steroids have been used. What is the probability that neither test is positive if steroids have been used? *A tree may help*

(a) 0.02 (b) 0.72 (c) 0.30 (d) 0.28 (e) none of the above

Suppose the two tests are indeed negative. What can now be said about whether the athlete has used steroids? *[one sentence]*

Use this example to explain why the alpha used in a statistical test cannot be used in the same way following statistical tests as positive and negative predictive values are following medical tests.*[one sentence]*

___ Suppose that the concentration of a toxic chemical in waste water output from a factory varies over time according to a normal (Gaussian) distribution with mean 0.004 and standard deviation 0.0012. Samples on 9 dates are taken and the sample mean is calculated. If the sample mean is .005 or above, the factory owner will be required to pay a large penalty. How does one calculate the probability that the owner will have to pay the penalty? *no need to do actual calculation.. you have a research assistant*

___ A scale used in analytical chemistry gives weights in repeated weighings of the same object that have a normal distribution, with mean equal to the true weight of the object. (That is, the scale has no bias.) The standard deviation of repeated weighings of the same object is $\sigma = 0.010$ gram. The student wants to estimate the true weight within ± 0.01 gram with 95% confidence. How does she determine how many weighings should she average to achieve this? *[calculations not needed]*

Refer to item 1 (supplied at last class) It gives data on the age at which persons first have full sexual activity.

4 pts Consider a more general statistical question , about the onset of any behaviour or characteristic: is it statistically possible that the <u>average</u> age of onset of a behaviour or stage of development (e.g. puberty or menarche) is say x.x years but that <u>fewer than 50%</u> have onset by this age? Use a histogram to explain. If you prefer, use the random variable length of life (equivalently, the age at death). Could it be that a majority (e.g. 60%) of people live longer than average?

Refer to item 2. The letter writer asks the Daily «why in your first sentence do you use the phrase "rare compared to their colleagues" when the percentages are 36 and 42? »[1]
We could ask the letter writer «why do you use the phrase "deviation of only 6%" when a simpler "difference of only 6%" would do equally well» and «why complicate things by mentioning Gallop polls and margins of error and the square root of two?»

2 pts In *one sentence*, explain why one doesn't need inferential statistics here.

4 pts Also, explain to this writer that if (s)he is going to bring statistical inference about proportions into this, (s)he should get his margin of error correct.[2]

For the next series of questions, refer to item 4 (ultrasound imaging in pregnancy).

2 pts Suppose, in the **Summary**, you were given the 95% confidence interval 1.09 to 1.67, but not the p-value. From the confidence interval alone, what could you say about the p-value associated with a test of Relative Risk=1 vs. Relative Risk ≠1?

4 pts The section on Statistical Analysis gives the sample size calculations they did. For the one dealing with duration of neonatal stay, which formula or table from our course is relevant? What values would you substitute for each of the components of the formula or table *(you don't have to evaluate the formula)*.

3 pts For the outcome variable "length of neonatal stay", state, (i) in symbols, and (ii) in carefully chosen words, the null hypotheses that they tested by the "t test for

---

[1]The writer should also complain about the use of "compared to" ; the correct usage is "compared with".

[2]Restrict yourself to professors who have either a Canadian or a U.S. PhD (ie base your calculations on the revised data "in the faculty of Arts, of the professors who have either a Canadian or a U.S. PhD, some 54% (42%/(42%+36%) have American PhDs whereas only 46% (36%/78%) have Canadian".

independent parameters" mentioned in the Statistical Analysis section.

2 pts Which statistical reference distribution governs the random variation in the number, out of a total of 2834, who could have been assigned to the intensive group?

2 pts Considering the shape of the distribution duration of neonatal stay (Figure), are you worried about the validity of the t-test used by the authors?

2 pts Why then did the authors use median(interquartile range) in the 2nd last row of Table 1?

3 pts "The mean birthweight in the intensive group was 25 g less than in the regular group, although this difference was not statistically significant" (last sentence or **Results**). Outline the calculations and/or tables used to justify this statement, filling in plausible values for any pieces that you lack. *You are not asked to complete the calculation.. imagine you are directing your research assistant to do it*

2 pts Explain in less technical words what is meant by the sentence "this difference was not statistically significant".

2 pts Explain in less technical words what is meant by the sentence "intrauterine growth restriction was significantly higher in the intensive group; p=0.006".

Refer to item 5 (Oestradiol-17ß).

3 pts From what we have covered in the course thus far, how could you formally test the 569 vs. 658 in the second last row of the Table? *Calculations not needed.* Do you have enough data to do the test?

2 pts What do the two intervals in the rightmost half of Fig 1 (total exercise time) represent?

3 pts How do you think the authors calculated them?

3 pts Do they help you to assess the statistical significance of the differences?

2 pts (Approximately, since you have to read it from the diagram) what is the median of the n=11 total exercise times in the placebo condition?

2 pts Is it more accurate in this type of study to use the term "difference of medians" or "median of differences". Why?

Of the 11 differences in total exercise time, 9 are increases, 1 is a decrease and 1 is practically zero. Thus if we leave out the zero, we have 9/10 changes in the positive direction.

2 pts Under the null hypothesis, what proportion would one expect to be positive?

3 pts What is the probability, under the null hypothesis, of obtaining

(a) 9 or more positives? _____

(b) 9 or more negatives?_____

(c) 9 or more positives or 9 or more negatives?_____

2 pts Given you answer to (c), how do you interpret the 9/10 positive changes?