

University of Guelph  
Department of Mathematics and Statistics

STAT\*2040  
Statistics I

Test 1 (White version)  
March 13 2015

Examiner: Jeremy Balka

**This exam is 80 minutes in duration**

Name:

ID:

Signature:

Please read the instructions:

1. Fill out your name and ID number above.
2. When the examination starts, make sure your question paper is complete. You should have 22 multiple choice questions, along with formula sheets and standard normal and  $t$  tables. The first question is just a bookkeeping question, and does not count for marks, but please fill it in to ensure your exam is properly graded.
3. Do all rough work on this paper.
4. You are allowed to bring in a calculator, and pens and pencils.
5. There is only **one** correct answer for each question. Fill in only one bubble for each question.
6. Fill out the computer answer sheet in pencil as you go. *There will be no extra time given at the end of the exam to fill in the sheet.*
7. The answers given in the exam are often rounded versions of the correct answer. Choose the closest value.

1. The colour of the first page of this examination booklet (the cover sheet) is:
  - (a) White
  - (b) Yellow
  
2. Suppose that green sea urchins are, to a reasonable approximation, randomly and independently distributed on a seabed at an average rate of 6 per 100 square metres. What is the probability that a randomly selected 20 square metre portion of the seabed has at least 2 green sea urchins? (Choose the closest value.)
  - (a) 0.22
  - (b) 0.27
  - (c) 0.34
  - (d) 0.37
  - (e) 0.41
  
3. Only one of the following is an appropriate set of hypotheses for a hypothesis test. Which one?
  - (a)  $H_0: \sigma^2 = 18, H_a: \sigma^2 \neq 18$
  - (b)  $H_0: \bar{X} = 18, H_a: \bar{X} \neq 18$
  - (c)  $H_0: s = 18, H_a: s \neq 18$
  - (d)  $H_0: \bar{X} = 18, H_a: \bar{X} > 18$
  - (e)  $H_0: s = 18, H_a: s > 18$
  
4. Suppose  $P(A) = 0.20$ ,  $P(B) = 0.60$ , and  $P(A \cap B) = 0.12$ . What is  $P(A^c \cup B)$ ?
  - (a) 0.68
  - (b) 0.72
  - (c) 0.80
  - (d) 0.88
  - (e) 0.92

5. Dale (2010) studied birth statistics of elephants born in captivity. In one part of the study, it was found that a sample of 27 newborn male African elephants born in captivity had a mean weight of 112.0 kg and a standard deviation of 18.2 kg. Suppose it was previously claimed that newborn male African elephants born in captivity would have a mean weight of 100 kg, and we wish to carry out a hypothesis test of this claim. The following output from R summarizes the results of the test calculations.

```
data: elephants
t = 3.426, df = 26, p-value = 0.002046
alternative hypothesis: true mean is not equal to 100
95 percent confidence interval:
 104.8003 119.1997
sample estimates:
mean of x
    112
```

Assume for the purposes of this question that the sample of elephants can be thought of as a random sample of newborn male African elephants born in captivity, and assume that the weights are approximately normally distributed.

Of the following options, which one of the most appropriate conclusion to the hypothesis test?

- (a) There is very strong evidence that the true mean weight of newborn male African elephants born in captivity is greater than 100.0 kg.
  - (b) There is very strong evidence that the sample mean weight of the 27 newborn male African elephants born in captivity is greater than 100.0 kg.
  - (c) There is not strong evidence that the sample mean weight of the 27 newborn male African elephants born in captivity differs from 100.0 kg.
  - (d) There is not strong evidence that the true mean weight of newborn male African elephants born in captivity differs from 100.0 kg.
  - (e) There is strong evidence that the true mean weight of newborn male African elephants born in captivity is equal to 100.0 kg.
6. Suppose that the body lengths of adult male crested porcupines in a certain region in Italy are approximately normally distributed with a mean of 73.1 cm and a standard deviation of 5.8 cm. (These values are based on a study by Mori et al. (2014).) If an adult male crested porcupine is randomly selected from this region, what is the probability that its body length lies between 70.0 and 75.0 cm? (Choose the closest value.)
- (a) 0.33
  - (b) 0.43
  - (c) 0.53
  - (d) 0.63
  - (e) 0.73

The next 3 questions refer to the following information.

Suppose that weights of adult female crested porcupines in a certain region in Italy are approximately normally distributed with a mean of 12.1 kg and a standard deviation of 1.9 kg. (These values are based on a study by Mori et al. (2014).)

7. If 10 adult female porcupines are randomly and independently sampled from this region, what is the standard deviation of their total weight?
  - (a) 0.6
  - (b) 1.9
  - (c) 4.4
  - (d) 6.0
  - (e) 19.0
  
8. If 2 adult female porcupines are randomly and independently sampled from this region, what is the probability that both porcupines weigh less than 12.5 kg? (In other words, what is the probability that porcupine #1 weighs less than 12.5 kg and porcupine #2 weighs less than 12.5 kg?)
  - (a) 0.22
  - (b) 0.34
  - (c) 0.58
  - (d) 0.62
  - (e) 0.71
  
9. If 8 adult female porcupines are randomly and independently sampled from this region, what is the 95th percentile of the sampling distribution of their mean weight?
  - (a) 13.2
  - (b) 13.4
  - (c) 13.7
  - (d) 14.2
  - (e) 15.2

10. Suppose we sample 8 observations from a normally distributed population where it is known that  $\sigma = 10$ , and we find a sample mean of 8.2. If we wish to calculate a 43.8% confidence interval for  $\mu$ , what is the appropriate margin of error? (Choose the closest value.)

- (a) 1.95
- (b) 2.05
- (c) 2.15
- (d) 2.25
- (e) 2.35

11. A study investigated the amount of estradiol (an estrogen) found in the blood of young males in the United States. In one part of the study, the mean estradiol concentration in 100 twenty-year-old males in the United States was found to be 38 pg/ml, with an associated confidence interval for  $\mu$  of (37.1, 38.9).

Consider the following statements:

- I. We can be 95% confident that the sample mean estradiol concentration in the blood of these 100 males lies between 37.1 pg/ml and 38.9 pg/ml.
- II. 95% of twenty-year-old males in the United States have a blood estradiol concentration that lies between 37.1 pg/ml and 38.9 pg/ml.
- III. We can be 95% confident that if we measured the blood estradiol concentration of all twenty-year-old males in the United States, the mean would lie between 37.1 pg/ml and 38.9 pg/ml.

Which of these statements is an appropriate interpretation of the 95% confidence interval?

- (a) Just II.
- (b) Just III.
- (c) I and II.
- (d) II and III.
- (e) All of them.

12. Suppose we are testing  $H_0: \mu = 10$  against a two-sided alternative hypothesis at  $\alpha = 0.05$ . Which one of the following statements is false? (If A-D are all true, answer option E.)

- (a) If  $\mu = 10$ , then we cannot make a Type II error.
- (b) If  $\mu = 5$ , then we cannot make a Type I error.
- (c) If  $\mu = 15$ , and the  $p$ -value is 0.01, then we will make a Type II error.
- (d) If  $\mu = 10$ , and the  $p$ -value is 0.01, then we will make a Type I error.
- (e) None of the above.

13. Which one of the following statements is true?

- (a) The variance of the  $t$  distribution is less than the variance of the standard normal distribution.
- (b) The mean of the  $t$  distribution is less than the mean of the standard normal distribution.
- (c) The  $t$  distribution has less area in the tails and a higher peak than the standard normal distribution.
- (d) As the degrees of freedom increase, the  $t$  distribution tends toward a normal distribution with a mean of 1.
- (e) None of the above.

14. Approximately 39% of the Canadian population has blood type O<sup>+</sup>. If Canadians are randomly and independently sampled, what is the probability that the first person with blood type O<sup>+</sup> occurs on or before the third person sampled? (Choose the closest value.)

- (a) 0.15
- (b) 0.63
- (c) 0.71
- (d) 0.77
- (e) 0.86
- (f) 0.89

15. A researcher is about to draw a random sample of 22 observations from a normally distributed population. The mean of the population is unknown, but the standard deviation of the population is known to be 6. The researcher wants to test the null hypothesis that the mean of the population is 11, against the alternative that it is different from 11. The researcher intends to use the test statistic:

$$\frac{\bar{X} - 11}{6/\sqrt{22}}$$

What is the sampling distribution of this test statistic if the null hypothesis is true?

- (a) The standard normal distribution.
- (b) A normal distribution with a mean of 11 and a standard deviation of 6.
- (c) A  $t$  distribution with 22 degrees of freedom.
- (d) A  $t$  distribution with 21 degrees of freedom.
- (e) This statistic does not have a distribution, as it can only take on one possible value.

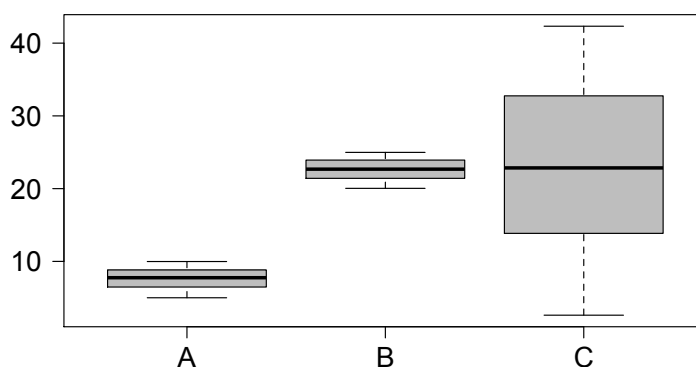
16. A researcher wants to test the null hypothesis that the mean of a certain population is 0, against the alternative hypothesis that the population mean differs from 0. He has taken a statistics course, but he forgets most of the material. He decides to take a random sample of size 50, calculate the  $Z$  statistic (the same one we have used in this course), and he will reject the null hypothesis if the  $Z$  statistic is greater than 1 or less than  $-1$ . If the null hypothesis is true, what is this researcher's probability of making a Type I error? (Choose the closest value.)
- (a) 0.025
  - (b) 0.05
  - (c) 0.10
  - (d) 0.15
  - (e) 0.32
17. Suppose we sample 8 observations from a normally distributed population where it is known that  $\sigma = 10$ , and we find a sample mean of 8.2. If we test  $H_0: \mu = 11$  against a two-sided alternative hypothesis, what is the  $p$ -value of the test? (Choose the closest value.)
- (a) 0.1
  - (b) 0.2
  - (c) 0.3
  - (d) 0.4
  - (e) 0.5
18. Suppose we sample 3 observations from a normally distributed population, and find that the values are:
- 8.1, 12.7, 24.2
- Which one of the following is a 95% confidence interval for the population mean  $\mu$ ?
- (a)  $15 \pm 16.3$
  - (b)  $15 \pm 18.4$
  - (c)  $15 \pm 20.6$
  - (d)  $15 \pm 21.8$
  - (e)  $15 \pm 22.3$

19. Suppose we sample 62 values from a normally distributed population, and using methods learned in STAT\*2040, we find a 95% confidence interval for  $\mu$  of (2.1, 3.1).

Which one of the following statements is false?

- (a) 95% of the area under the sampling distribution of  $\mu$  lies between 2.1 and 3.1.
- (b) If we used the same data to construct a 90% confidence interval for  $\mu$ , the interval would be narrower than the 95% interval.
- (c) If we used the same data to carry out a hypothesis test of  $H_0: \mu = 2.7$  against a two-sided alternative, the  $p$ -value would be greater than 0.05.
- (d) If we used the same data to carry out a hypothesis test of  $H_0: \mu = 5.7$  against a two-sided alternative, the  $p$ -value would be less than 0.05.
- (e) The sample mean was 2.6.

20. The following figure illustrates boxplots for 3 samples (labelled A, B, and C). Each sample has a sample size of 50.



Suppose that for each sample, we carry out a hypothesis test of  $H_0: \mu = 20$  against a two-sided alternative hypothesis. Which one of the 3 tests would have the smallest  $p$ -value, and which one would have the largest  $p$ -value?

- (a) The test for Sample C would have the smallest  $p$ -value, and the test for Sample A would have the largest  $p$ -value.
- (b) The test for Sample A would have the smallest  $p$ -value, and the test for Sample C would have the largest  $p$ -value.
- (c) The test for Sample A would have the smallest  $p$ -value, and the test for Sample B would have the largest  $p$ -value.
- (d) The test for Sample B would have the smallest  $p$ -value, and the test for Sample C would have the largest  $p$ -value.
- (e) The test for Sample B would have the smallest  $p$ -value, and the test for Sample A would have the largest  $p$ -value.



21. Which one of the following statements is false? (If A-D are all true, answer option E.)

- (a) All other factors being equal, the power of a test increases as the sample size increases.
- (b) All other factors being equal, the power of a test increases as the significance level decreases.
- (c) If we are carrying out a hypothesis test at a fixed significance level  $\alpha$ , we reject the null hypothesis if  $p\text{-value} \leq \alpha$ .
- (d) Suppose we are sampling from a normally distributed population and carrying out a  $t$  test on the population mean. If the null hypothesis is true, then on average the  $p$ -value will equal 0.5.
- (e) None of the above.

22. A researcher wants to estimate the average weight of adult skunks in a large town. They trap 16 adult skunks, measure their weights, and find a mean of 4.1 kg and a standard deviation of 1.2 kg. They then calculate a confidence interval using the formula  $4.1 \pm 1.96 \frac{1.2}{\sqrt{16}}$ , which works out to (3.512, 4.688). They use this interval to say that they are 95% confident that the true mean weight of all adult skunks in this large town lies between 3.512 kg and 4.688 kg.

Which of the following 3 statements represent possible problems that may affect the validity of this confidence interval?

- I. The population may not be normally distributed.
  - II. The sampling method may be biased.
  - III. The researcher used a  $z$  value to calculate the interval, when they should have been using a  $t$  value.
- (a) Just I.
  - (b) Just III.
  - (c) I and II.
  - (d) I and III
  - (e) I, II, and III.