MULTIPLE CHOICE. Choose the one alternative that best completes the statement or answers the question.

1) A philosophy professor wants to find out whether the mean age of the men in his large lecture class is equal to the mean age of the women in his class. After collecting data from a random sample of his students, the professor tested the hypothesis \( H_0: \mu_M - \mu_W = 0 \) against the alternative \( H_A: \mu_M - \mu_W \neq 0 \). The \( P \)-value for the test was 0.003. Which is true?
   A) There is a 99.7% chance that another sample will give these same results.
   B) It is very unlikely that the professor would see results like these if the mean age of men was equal to the mean age of women.
   C) There is a 0.3% chance that the mean ages for the men and women are different.
   D) There is a 0.3% chance that another sample will give these same results.
   E) There is a 0.3% chance that the mean ages for the men and women are equal.

2) Absorption rates into the body are important considerations when manufacturing a generic version of a brand-name drug. A pharmacist read that the absorption rate into the body of a new generic drug (G) is the same as its brand-name counterpart (B). She has a researcher friend of hers run a small experiment to test \( H_0: \mu_G - \mu_B = 0 \) against the alternative \( H_A: \mu_G - \mu_B \neq 0 \). Which of the following would be a Type I error?
   A) Deciding that the absorption rates are different, when in fact they are not.
   B) Deciding that the absorption rates are the same, when in fact they are not.
   C) Deciding that the absorption rates are different, when in fact they are.
   D) The researcher cannot make a Type I error, since he has run an experiment.
   E) Deciding that the absorption rates are the same, when in fact they are.

3) At one SAT test site students taking the test for a second time volunteered to inhale supplemental oxygen for 10 minutes before the test. In fact, some received oxygen, but others (randomly assigned) were given just normal air. Test results showed that 42 of 66 students who breathed oxygen improved their SAT scores, compared to only 35 of 63 students who did not get the oxygen. Which procedure should we use to see if there is evidence that breathing extra oxygen can help test-takers think more clearly?
   A) 1-sample \( t \)-test
   B) 2-proportion \( z \)-test
   C) matched pairs \( t \)-test
   D) 2-sample \( t \)-test
   E) 1-proportion \( z \)-test

Explain what the \( P \)-value means in the given context.

4) The federal guideline for smog is 12% pollutants per 10,000 volume of air. A metropolitan city is trying to bring its smog level into federal guidelines. The city comes up with a new policy where city employees are to use city transportation to and from work. A local environmental group does not think the city is doing enough and no real decrease will occur. An independent agency, hired by the city, runs its tests and comes up with a \( P \)-value of 0.055. What is reasonable to conclude about the new strategy using \( \alpha = 0.025 \)?
   A) There is a 94.5% chance of the new policy having no effect on smog.
   B) There is a 5.5% chance of the new policy having no effect on smog.
   C) There’s only a 5.5% chance of seeing the new policy having no effect on smog in the results we observed from natural sampling variation. We conclude the new policy is more effective.
   D) We can say there is a 5.5% chance of seeing the new policy having no effect on smog in the results we observed from natural sampling variation. There is no evidence the new policy is more effective, but we cannot conclude the policy has no effect on smog.
   E) We can say there is a 5.5% chance of seeing the new policy having an effect on smog in the results we observed from natural sampling variation. We conclude the new policy is more effective.
5) Two agronomists analyzed the same data, testing the same null hypothesis about the proportion of tomato plants suffering from blight. One rejected the hypothesis but the other did not. Assuming neither made a mistake in calculations, which of these possible explanations could account for this apparent discrepancy?
I. One agronomist wrote a one-tailed alternative hypothesis, but the other used 2 tails.
II. They wrote identical hypotheses, but the one who rejected the null used a higher α-level.
III. They wrote identical hypotheses, but the one who rejected the null used a lower α-level.
A) I or II  B) I only  C) III only  D) I or III  E) II only

Provide an appropriate response.
6) An entomologist writes an article in a scientific journal which claims that fewer than 12% of male fireflies are unable to produce light due to a genetic mutation. Identify the Type I error in this context.
A) The error of failing to accept the claim that the true proportion is at least 12% when it is actually less than 12%.
B) The error of failing to reject the claim that the true proportion is at least 12% when it is actually less than 12%.
C) The error of accepting the claim that the true proportion is at least 12% when it really is at least 12%.
D) The error of rejecting the claim that the true proportion is less than 12% when it really is less than 12%.
E) The error of rejecting the claim that the true proportion is at least 12% when it really is at least 12%.

7) A skeptical paranormal researcher claims that the proportion of Americans that have seen a UFO, p, is less than 5%. Identify the Type II error in this context.
A) The error of rejecting the claim that the true proportion is at least 5% when it really is at least 5%.
B) The error of accepting the claim that the true proportion is at least 5% when it is actually less than 5%.
C) The error of accepting the claim that the true proportion is at least 5% when it really is at least 5%.
D) The error of failing to reject the claim that the true proportion is at least 5% when it is actually less than 5%.
E) The error of rejecting the claim that the true proportion is less than 5% when it really is less than 5%.

Write the null and alternative hypotheses you would use to test the following situation.
8) A skeptical paranormal researcher claims that the proportion of Americans that have seen a UFO is less than 4%.
A) H₀: p > 0.04  B) H₀: p = 0.04  C) H₀: p = 0.04  D) H₀: p < 0.04  E) H₀: p < 0.04
Hₐ: p = 0.04  Hₐ: p < 0.04  Hₐ: p > 0.04  Hₐ: p > 0.04  Hₐ: p = 0.04

Provide the appropriate response.
9) The mayor of a large city will run for governor if he believes that more than 30 percent of the voters in the state already support him. He will have a survey firm ask a random sample of n voters whether or not they support him. He will use a large sample test for proportions to test the null hypothesis that the proportion of all voters who support him is 30 percent. Suppose that 35 percent of all voters in the state actually support him. In which of the following situations would the power of this test be highest.
A) The mayor uses a significance level of 0.01 and n = 250 voters.
B) The mayor uses a significance level of 0.05 and n = 500 voters.
C) The mayor uses a significance level of 0.01 and n = 1,000 voters.
D) The mayor uses a significance level of 0.05 and n = 1,000 voters.
E) The mayor uses a significance level of 0.01 and n = 500 voters.
provide an appropriate response.

10) A weight loss center provided a loss for 72% of its participants. The center’s leader decides to test a new weight loss strategy on a random sample size of 140 and found 109 participants lost weight. Should the center continue its new strategy? Test an appropriate hypothesis using α = 0.02 and state your conclusion. Be sure the appropriate assumptions and conditions are satisfied before you proceed.

A) \( z = 1.54; P\text{-value} = 0.9382 \). The change is statistically significant. A 98% confidence interval is (69.0%, 86.7%). This is clearly higher than 72%. The chance of observing 109 or more participants of 140 is only 94.29% if the weight loss is really 72%.

B) \( z = -1.54; P\text{-value} = 0.0618 \). The change is statistically significant. A 90% confidence interval is (71.6%, 84.1%). This is clearly higher than 72%. The chance of observing 109 or more participants of 140 is only 5.71% if the weight loss is really 72%.

C) \( z = 1.54; P\text{-value} = 0.0618 \). The center should not continue with the new strategy. There is a 6.18% chance of having 109 or more of 140 participants in a random sample weigh less if in fact 72% do. The \( P\text{-value} \) of 0.0618 is greater than the alpha level of 0.02.

D) \( z = -1.54; P\text{-value} = 0.9382 \). The center should continue with the new strategy. There is a 94.29% chance of having 109 or more of 140 participants in a random sample weigh less if in fact 72% do.

E) \( z = 1.54; P\text{-value} = 0.1236 \). The change is statistically significant. A 95% confidence interval is (70.4%, 85.3%). This is clearly lower than 72%. The chance of observing 109 or more participants of 140 is only 11.42% if the weight loss is really 72%.

11) When we accept the null hypothesis, we

A) have committed a Type II error.
B) claim that a significant difference exists between groups.
C) have committed a Type I error.
D) have obtained a t-value greater than our critical t-value.
E) conclude that sampling error is responsible for our obtained difference.

12) A researcher is interested in the academic performance differences between individuals using an optimistic versus a pessimistic approach to their studies. If the researcher fails to find a significant difference, when in fact one exists in the population:

A) the null hypothesis was correctly accepted.
B) the research hypothesis was correctly accepted.
C) a Type 2 error has been made.
D) the null hypothesis was correctly rejected.
E) a Type 1 error has been made.

13) To plan the course offerings for the next year a university department dean needs to estimate what impact the "No Child Left Behind” legislation might have on the teacher credentialing program. Historically, 40% of this university’s pre-service teachers have qualified for paid internship positions each year. The Dean of Education looks at a random sample of internship applications to see what proportion indicate the applicant has achieved the content-mastery that is required for the internship. Based on these data he creates a 90% confidence interval of (33%, 41%). Could this confidence interval be used to test the hypothesis \( H_0: p = 0.40 \) versus \( H_A: p < 0.40 \) at the \( \alpha = 0.05 \) level of significance?

A) Yes, since 40% is in the confidence interval he accepts the null hypothesis, concluding that the percentage of applicants qualified for paid internship positions will stay the same.
B) Yes, since 40% is in the confidence interval he fails to reject the null hypothesis, concluding that there is not strong enough evidence of any change in the percent of qualified applicants.
C) No, because he should have used a 95% confidence interval.
D) Yes, since 40% is not the center of the confidence interval he rejects the null hypothesis, concluding that the percentage of qualified applicants will decrease.
E) No, because the dean only reviewed a sample of the applicants instead of all of them.
14) You could win a $1000 prize by tossing a coin in one of two games. To win Game A, you must get exactly 50% heads. To win Game B, you must get between 45% and 55% heads. Although which game you must play will be chosen randomly, then you may decide whether to toss the coin 20 times or 50 times. How many tosses would you choose to make?
   A) 50 tosses for either.
   B) 20 tosses for either game.
   C) 50 tosses for A, 20 tosses for B.
   D) It does not matter.
   E) 20 tosses for A, 50 tosses for B.

Indicate the correct test procedure and reasoning.
15) A teacher is interested in performing a hypothesis test to compare the mean math score of the girls and the mean math score of the boys. She randomly selects 10 girls from the class and then randomly selects 10 boys. She arranges the girls' names alphabetically and uses this list to assign each girl a number between 1 and 10. She does the same thing for the boys.
   A) Paired t-test. Since the boys and girls are in the same class, and are hence dependent samples, they are can be linked.
   B) Paired t-test. Since there are 10 boys and 10 girls, we can link the two samples.
   C) 1-sample t-test. The teacher should compare the sample mean for the girls against the population mean for the boys.
   D) Either two-sample or paired t-test will work.
   E) Two-sample t-test. There is no natural pairing between the two populations.

Provide an appropriate response.
16) A manufacturer has designed athletic footwear which it hopes will improve the performance of athletes running the 100-meter sprint. It wishes to perform a hypothesis test to compare the times of athletes at the 100 meters with these shoes and with their usual shoes.
   A) Two-sample t-test, since the experiment has two samples.
   B) Pooled t-test, since the standard deviations of the two populations are likely to be the same.
   C) Paired t-test, since a natural pairing exists, and would detect differences between the population means better.
   D) Either two-sample or paired t-test would be equally accurate.
   E) Not enough information is given.

17) Suppose that a manufacturer is testing one of its machines to make sure that the machine is producing more than 97% good parts (H₀: p = 0.97 and Hₐ: p > 0.97). The test results in a P-value of 0.122. Unknown to the manufacturer, the machine is actually producing 99% good parts. What probably happens as a result of the testing?
   A) They correctly fail to reject H₀.
   B) They reject H₀, making a Type I error.
   C) They fail to reject H₀, making a Type I error.
   D) They fail to reject H₀, making a Type II error.
   E) They correctly reject H₀.

18) Not wanting to risk poor sales for a new soda flavor, a company decides to run one more taste test on potential customers, this time requiring a higher approval rating than they had for earlier tests. This higher standard of proof will increase
   I. the risk of Type I error
   II. the risk of Type II error
   III. power
   A) II only
   B) I only
   C) I and II
   D) I and III
   E) III only
SHORT ANSWER. Write the word or phrase that best completes each statement or answers the question.

19) Test identification Suppose you were asked to analyze each of the situations described below. (NOTE: Do not do these problems!) For each, indicate which procedure you would use (pick the appropriate number from the list), the test statistic (z or t), and, if t, the number of degrees of freedom. A procedure may be used more than once.

<table>
<thead>
<tr>
<th>Type</th>
<th>z/t?</th>
<th>df</th>
</tr>
</thead>
<tbody>
<tr>
<td>A.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>E.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>F.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1. proportion – 1 sample
2. difference of proportions, 2 samples
3. mean – 1 sample
4. difference of means – independent samples
5. mean of differences – matched pairs

a. A union organization would like to represent the employees at the local market. A sample of the employees revealed 74 of 120 were in favor of the union. Does the union have the required 3 to 2 majority?

b. An oral surgeon is interested in estimating how long it takes to extract all four wisdom teeth. The doctor records the times for 24 randomly chosen surgeries. Estimate the time it takes to perform the surgery with a 95% confident interval.

c. A microwave manufacturing company receives large shipments of thermal shields from two suppliers. A sample from each supplier’s shipment is selected and tested for the rate of defects. The microwave manufacturing company’s contract with each supplier states the shipment with the smallest rate of defect will be accepted. Do the shipments’ defect rates vary from each other?

d. The owner of a construction company would like to know if his current work teams can build room additions quicker than the time allotted for by the contract. A random sample of 15 room additions completed recently revealed an average completion time of 0.32 days faster than contracted. Is this strong evidence that the teams can complete room additions in less than the contract times?

e. A farmer would like to know if a new fertilizer increases his crop yield. In an effort to decide this, the farmer recorded the yield for 10 different fields prior to adding fertilizer and after adding the fertilizer. The farmer assumes the crop yields are approximately normal. Does the fertilizer work as advertised?

f. In a study to determine whether there is a difference between the average jail time convicted bank robbers and car thieves are sentenced to, the law students randomly selected 20 cases of each type that resulted in jail sentences during the previous year. A 90% confidence interval was created from the results.

20) Improving productivity A packing company considers hiring a national training consultant in hopes of improving productivity on the packing line. The national consultant agrees to work with 18 employees for one week as part of a trial before the packing company makes a decision about the training program. The training program will be implemented if the average product packed increases by more than 10 cases per day per employee. The packing company manager will test a hypothesis using \( \alpha = 0.05 \).

a. Write appropriate hypotheses (in words and in symbols).

b. In this context, which do you consider to be more serious – a Type I or a Type II error? Explain briefly.

c. After this trial produced inconclusive results the manager decided to test the training program again with another group of employees. Describe two changes he could make in the trial to increase the power of the test, and explain the disadvantages of each.
A company manufacturing computer chips finds that 8% of all chips manufactured are defective. Management is concerned that employee inattention is partially responsible for the high defect rate. In an effort to decrease the percentage of defective chips, management decides to offer incentives to employees who have lower defect rates on their shifts. The incentive program is instituted for one month. If successful, the company will continue with the incentive program.

21) Write the company’s null and alternative hypotheses.

22) In this context describe a Type I error and the impact such an error would have on the company.

23) In this context describe a Type II error and the impact such an error would have on the company.

24) Based on the data they collected during the trial program, management found that a 95% confidence interval for the percentage of defective chips was (5.0%, 7.0%). What conclusion should management reach about the new incentive program? Explain.

25) What level of significance did management use?

26) Describe to management an advantage and disadvantage of using a 1% alpha level of significance instead.

27) Management decided to extend the incentive program so that the decision can be made on three months of data instead. Will the power increase, decrease, or remain the same?

28) Over the trial month, 6% of the computer chips manufactured were defective. Management decided that this decrease was significant. Why might management might choose not to permanently institute the employee incentive program?
Answer Key
Testname: UNTITLED1

1) B
2) A
3) B
4) D
5) A
6) E
7) D
8) B
9) D
10) C
11) E
12) C
13) B
14) E
15) E
16) C
17) D
18) A
19)

<table>
<thead>
<tr>
<th>Type</th>
<th>z/t?</th>
<th>df</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>1</td>
<td>( z )</td>
</tr>
<tr>
<td>b.</td>
<td>3</td>
<td>( t )</td>
</tr>
<tr>
<td>c.</td>
<td>2</td>
<td>( z )</td>
</tr>
<tr>
<td>d.</td>
<td>5</td>
<td>( t )</td>
</tr>
<tr>
<td>e</td>
<td>5</td>
<td>( t )</td>
</tr>
<tr>
<td>f</td>
<td>4</td>
<td>( t )</td>
</tr>
</tbody>
</table>

20) a. \( H_0: \mu_d = 0 \) \( H_A: \mu_d > 10 \)

b. A Type I error would be very expensive for the packing company. A Type I error would mean that the manager rejected the null hypothesis when in fact the null hypothesis is true. In this situation, by rejecting the null hypothesis the company thought the training improved productivity, so they paid for the consultant to train all employees. In reality, the training did not improve productivity so the company wasted money on training that did not help.

c. To increase the power of the test, we could increase the level of significance (\( \alpha \)), or increase the sample size. Increasing the level of significance could lead to adopting a training method that actually does not improve productivity. By increasing the sample size, the trial cost would increase and the trial might take more time.

21) The null hypothesis is that the defect rate is 8%. The alternative hypothesis is that the defect rate is lower than 8%. In symbols: \( H_0: p = 0.08 \) and \( H_A: p < 0.08 \)

22) A Type I error would be deciding the percentage of defective chips has decreased, when in fact it has not. The company would waste money on a new incentive program that does not decrease the defect rate of the chips.

23) A Type II error would be deciding the percentage of defective chips has not decreased, when in fact it has. The company would miss an opportunity to decrease the defect rate of the chips.

24) The confidence interval contains values that are all below the hypothesized value of 8%, so the data provide convincing evidence that the incentive program lowers the defect rate of the computer chips.

25) \( \alpha = \frac{1 - 0.95}{2} = 0.025 = 2.5\% \)

26) Advantage: management would be less likely to conclude the incentive program was effective if it really were not. Disadvantage: the test would have less power to detect a positive effect of the new incentive program.

27) The power would increase because of the larger sample size.
28) Although statistically significant, the practical significance (cost of the incentive program compared to the savings due to the decrease in defect rate of the chips) might not be great enough to warrant instituting the program permanently.