Probability and Hypothesis Testing: Vartanian

1. Assume you are tossing a coin 11 times. The following distribution gives the likelihoods of getting a particular number of heads:

Heads	Probability
0	.0005
1	.0054
2	.0269
3	.0806
4	.1611
5	.2256
6	.2256
7	.1611
8	.0806
9	.0269
10	.0054
11	.0005

The null hypothesis is that you are no more likely to get a head than a tail. The alternative hypothesis is that the coin is not a fair coin and that you are more likely to get a head than a tail.

A. Is this a one-tailed test or a two-tailed test?

B. If you are using a 95% confidence interval or a 5% rejection region and you got 9 heads out of 11 tosses of the coin, what would you conclude?

C. If you were using a 5% rejection region and got 8 heads, what would you conclude?

D. If you were using a 15% rejection region and got 8 heads, what would you conclude?

E. If you rejected the null hypothesis in question D but found out that in the population (or in reality) that the coin was a fair coin, what type of error would you have made?

F. What is the type of error you could have made in part C of this question?

G. If you increased the likelihood of a type I error, would you increase or decrease the rejection region?

H. If you decreased the likelihood of a type II error, what would happen to the likelihood of a type I error. Give an example of increasing the likelihood of a type I error.

(9

2. You are examining the difference in income levels between women and men. Your null hypothesis is that there is no difference in income levels between the two groups. Your research hypothesis states that there is a difference in income levels between the groups, but you're not sure which group has a greater income. In your sample you find that men have a mean income level of $12,500 and women have a mean income level of $11,000. You find that the grouped standard error is $500. You have set up a 95% confidence interval. Assume that the samples were drawn independently, a very large sample size and normality.

A. Have you found a significant difference in income levels?

B. What will you do with your null hypothesis?

C. What type of error could you have made?

D. Assuming that you have a random sample from the population, can you generalize your results?

E. What are your independent and dependent variables?

F. Have you controlled for alternative explanations to gender differences?

3. Assume that we've conducted another sample of men and women, examining their levels of income. In this sample we find that the mean level of income for men is $11,000 and the mean level of income for women is $13,000. The grouped standard error is 2,000. Assume that the samples were drawn independently, a very large sample size and normality.

Answer questions A-F from question 2.

Answers to the Worksheet Questions: Vartanian

1.

A. 1 tailed test.

B. Reject the null since 0.0269+0.0054+0.0005=.0328<.05

C. Fail to reject the null since 0.0806+0.0269+0.0054+0.0005=0.1134>0.05

D. Reject the null since 0.1134<0.1500

E. Type I error. You rejected the null when there is no relationship between the independent and dependent variable. Whenever you reject the null, the only type of error you could have possibly made is a type I error. If you fail to reject the null, the only type of error you could have made is a type II error.

F. You could have made a type II error since you fail to rejected the null hypothesis.

G. To increase the likelihood of a type I error, you must increase the rejection region. If you increase the rejection region, your rejection region may go from .05 to .10 and the likelihood of a type I error increases from 5% to 10%.

H. The likelihood of a type I error would increase, from say 5% to 10%.

2.

A. Since this is a two tailed test, you will have a 2 1/2 % rejection region on each tail of the normal distribution. In order to reject the null hypothesis you must land in the rejection region. The area between the central point of the normal distribution and the rejection region is 47.50.

Compute a Z score: Z=(12,500-11,000)/500=3. You are thus 3 standard error units away from the hypothesized difference between the two groups (which is 0). Looking this value up in the Z table gives you a value of 49.87. You are therefore at the 99.87th percentile (50+49.87). Hence, you are within the 2 1/2% rejection region and you will therefore reject your null hypothesis. Any Z value that would have taken you to the 97 1/2 percentile would have allowed you to reject the null hypothesis.

B. Reject it.

C. You may have made a type I error if, in the population there truly is not a relationship between gender and level of income. There is only a 5% likelihood that there is no relationship between gender and income.

D. Yes.

E. The independent variable is gender and the dependent variable is income.

F. No. You may wish to control for alternative explanations and thus use control variables within your statistical model. Such factors as education, family background, level of work experience may be important factors in the analysis of the determinants of income.

3.

A. Z=(11,000-13,000)/2,000 = -1. Thus, you are 1 standard error unit away from the hypothesized difference between the groups. The area between the hypothesized difference, and the actual value of this difference, is 34.13. We are thus at the 15.87th percentile (50-34.13=15.87). Since we are not at either the top 2 1/2 percentile (at the 97 1/2 percentile) or at the bottom 2 1/2 percentile, we are not in the rejection region.

B. Fail to reject the null hypothesis. Even if this was a one-tailed test you would not have rejected the null hypothesis since the value of the difference divided by the grouped SE did not fall in the lowest 5% of the normal distribution.

C. You may have made a type II error. Whenever you fail to reject the null hypothesis, there is a chance you have made a type II error.

D. You can say that in all likelihood there is no relationship between gender and income in the population.

E. IV:Gender. DV: Income

F. No.