As necessary, round all work to two or three decimal places.

1) Is the gender of a nurse related to the specialty he or she chooses?? A hospital administrator counted the number of nurses by gender in three different nursing specialties within her health system. Perform the appropriate hypothesis test to answer this question (at *α* =0.05), showing all necessary parts of the hypothesis test.

|  |  |
| --- | --- |
| **Gender** | **Nursing specialty** |
| ICU | ER | Pediatrics |
| Male | 16 | 20 | 38 |
| Female | 24 | 38 | 116 |

2) A professor used data from last semester’s statistics class to create a regression that predicts how students will do on the final exam. The equation is $\hat{Y}=70.2+3.2\left(X\_{1}\right)-6(X\_{2})$, where *X*1 is the number of hours studied for the exam and *X*2 is the number of sleep interruptions the night before the exam. What is the predicted score of:

a) a student who studied for 10 hours and had 2 sleep interruptions?

b) a student who studied for 6 hours and had no sleep interruptions?

3) An ER nurse notices that when she and her colleagues are asked to work overtime, they seem to get less sleep (rather than just having less free time). She collects data from all of the nurses on her shift on the number of hours of overtime worked on their last shift and the amount of sleep they got that evening.

|  |  |
| --- | --- |
| **Hours of overtime** | **Hours of sleep** |
| 3 | 6 |
| 0 | 9 |
| 1 | 7 |
| 5 | 4 |
| 4 | 6 |

a) Create a scatterplot for these data.

b) Calculate the correlation coefficient.

c) Perform a hypothesis test to determine if there is a significant correlation (*α* =0.05) between number of hours of overtime worked and hours of sleep.

4) An OB office nurse hears about the data collected by the ER nurse in problem #3 and wants to see if the same relationship holds true for her office. She collects data from her coworkers and finds the following data:

|  |  |  |
| --- | --- | --- |
|  | **Hours of overtime** | **Hours of sleep** |
| **Mean** | 1.9 | 8.9 |
| **SD** | 0.7 | 1.4 |

 *r* = -0.734

a) Which correlation shows a stronger relationship, the ER nurse’s correlation in #3 or the OB nurse’s correlation in #4? Why?

b) Using the data from the OB office, generate a regression equation predicting number of hours of sleep from number of hours of overtime worked.

c) Interpret the meaning of the slope from the equation you created in problem #4b.

5) A winery owner wants to determine if one type of her wine is significantly more popular than another. She records the number of bottles of each type of her wine sold during a month.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Merlot** | **Pinot Noir** | **Riesling** | **Chardonnay** |
| **Bottles sold** | 31 | 10 | 47 | 32 |

a) What kind of statistic should you use to test this hypothesis?

 b) Write the set of hypotheses you would use to test this hypothesis.

 c) What are the *df* and critical value for this hypothesis test?

d) Calculate the expected frequencies for these data.

Extra Credit 1: If two variables are correlated, can we assume that one caused the other? Please explain, using an example.

Extra Credit 2: If *r* = 0.6, what is the proportionate reduction in error gained by using regression over the mean of *Y* for the prediction of *Y* values?