**STAT 405 Fall 2014 - Homework 2 (59 points)**

**Due Friday, September 12th**

Instructions: You must show ALL of your work in order to receive full credit. For example, if you use SAS to calculate a p-value, you must include your SAS command lines in your written solutions. If you use R or JMP the same thing applies.

**Problems
1)** Suppose 10 gonorrhea cases are reported over a 3-month period among 10,000 people living in an urban county. The statewide incidence of gonorrhea is 50 cases per 100,000 over a 3-month period. Is the number of gonorrhea cases in this county unusual for this time period? (2 pts.)

**2)** An article was published concerning the incidence of cardiac death attributable to the earthquake in Los Angeles County on January 17, 1994. In the week before the earthquake there were an average of 15.6 cardiac deaths per day in Los Angeles County. On the day of the earthquake, there were 51 cardiac deaths.

1. What is the exact probability of 51 deaths occurring on one day if the cardiac death rate in the previous week continued to hold on the day of the earthquake? (2 pts.)
2. Is the occurrence of 51 deaths unusual? Explain. (2 pts.)
3. What is the maximum number of cardiac deaths that could have occurred on the day of the earthquake to be consistent with the rate of cardiac deaths in the past week? Use a cutoff probability of .05 to determine the maximum number. (2 pts.)

**3)** Studies have been undertaken to assess the potential relationship between abortion and the development of breast cancer. In one study among nurses (the Nurses’ Health Study II), there were 16,539 abortions among 2,169,321 person-years of follow-up for women of reproductive age.

1. What is a *person-year*? (1 pt.)
2. What is the expected number of abortions among nurses over this time period if the incidence of abortion is 25 per 1000 women per year and no woman has more than 1 abortion? (2 pts.)
3. Does the abortion rate among nurses differ significantly from the national experience? Why or why not? Use a Poisson distribution to justify your answer, i.e. a simple yes/no answer is NOT acceptable. (3 pts.)

**4) Bone density study**

Consider the data set (BONEDEN.JMP on the course1 site). Calculate the difference in bone density of the lumbar spine (g/cm2) between the heavier-smoking twin and lighter-smoking twin (bone density for the heavier-smoking twin minus the bone density for the lighter-smoking twin) for each of the 41 twin pairs.

1. Suppose smoking has no relationship to bone density. What would be the expected number of twin pairs with negative difference scores? What is the actual number of twin pairs with negative difference scores? Do you feel smoking is related to bone density of the lumbar spine, given the observed results? Why or why not? A yes/no answer is NOT acceptable. (Hint: Use the binomial distribution.) (4 pts.)
2. Sort the differences in smoking between members of a twin pair (expressed in pack-years). Identify the subgroup of 20 twin pairs with the largest differences in smoking. Answer the previous question (a) based on this subgroup of 20 twin pairs. (3 pts.)
3. Answer part (a) for bone density of the femoral neck. (2 pts.)
4. Answer part (b) for bone density of the femoral neck. (2 pts.)
5. Answer part (a) for bone density of femoral shaft. (2 pts.)
6. Answer part (b) for bone density of femoral shaft. (2 pts.)

**5) Environmental Health**

Much discussion has taken place concerning possible health hazards from exposure to anesthetic gases. In one study conducted in 1972, 525 Michigan nurse anesthetists were surveyed by mail questionnaires and telephone interviews to determine the incidence rates of cancer. Of this group, 7 individuals reported having a new malignancy other than skin cancer during 1971.

1. What is the best estimate of the 1971 incidence rate from these data? (1 pt.)
2. Provide a 95% CI for the true incidence rate. (3 pts.)

A comparison was made between the Michigan report and the 1969 cancer-incidence rates from the Connecticut tumor-registry, where the expected incidence, based on the age distribution of the Michigan nurse anesthetists, was determined to be 402.8 per 100,000 person-years.

1. Comment on the comparison between the observed incidence rate in the Michigan study and the Connecticut tumor-registry. (2 pts.)

**6) Hypertension**

Suppose 100 hypertensive people are given an antihypertensive drug and the drug is effective in 20 of them. By effective, we mean that the diastolic blood pressure is lowered by at least 10 mmHg as judged from a repeat blood pressure measurement 1 month after taking the drug.

1. What is the best point estimate of the probability *p* of the drug being effective? (1 pt.)
2. Suppose we know that 10% of all hypertensive patients who are given a placebo will have their diastolic blood pressure lowered by 10 mmHg after 1 month. Can we carry out some procedure to be sure we are not simply observing the placebo effect? If so (and you can) carry out this procedure and summarize your conclusions. (3 pts.)

**7) Sports Medicine**

Injuries are common in football and may be related to a number of factors, including the type of playing surface, the number of years of playing experience, and whether any previous injury exists. A study of factors affecting injury among Canadian football players was recently reported.

The rate of injury to the upper extremity (that is, shoulder to hand) on a dry field consisting of natural grass was 2.73 injuries per 1000 games. Assume this rate is known without error.

1. The study reported 45 injuries to upper extremity on a dry field consisting of artificial turf over the course of 10,112 games. What procedure can be used to assess whether the risk of injury is different on artificial turf versus natural grass? Carry out this procedure and summarize your findings. (3 pts.)
2. Provide a 95% CI for the rate of injury to the upper extremity on artificial turf. (Hint: Use the Poisson distribution) Express each rate (LCL and UCL) as the number of injuries per 1000 games. (4 pts.)