Homework 5

Statistical Inference, Fall 2021

1- According to a 2012 survey of 2,254 American adults, 17% of cell phone owners prefer to browse on their phone rather than on a computer or other device.
   a. According to an online article, a report from a mobile research company indicates that 38 percent of Chinese mobile web users only access the internet through their cell phones. Conduct a hypothesis test to determine if these data provide strong evidence that the proportion of Americans who only use their cell phones to access the internet differs from the Chinese proportion of 38%.
   b. Interpret the p-value in this context.
   c. Calculate a 95% confidence interval for the proportion of Americans who access the internet on their cell phones and interpret the interval in this context.

2- Some people report that just one sip can detect the difference between a diet and a regular soda. A researcher wanted to test this claim, so he chose 80 persons randomly. He then filled 80 simple white cups with Soda, half diet and half regular, and asked each person to take one taste from their cup and identify the Soda as diet or regular based on a random assignment. The drink was accurately identified by 53 of the participants.
   a. Do these data show that these participants can tell the difference between diet and regular Soda, or, in other words, are the results significantly better than random guessing?
   b. Interpret the p-value in this context.

3- The US National Institutes of Health announced in July 2008 that it was terminating clinical research early due to unexpected results. The trial included HIV-positive women in Sub-Saharan Africa who were given a single dosage of Nevaripine (an HIV therapy) after childbirth to avoid HIV transmission to the baby. The trial compared the effects of continuing therapy with Nevaripine vs. Lopinavir, a second HIV medicine after a woman had a successful delivery. The trial included 240 women, 120 of whom were randomly assigned to one of the two therapies. Each woman was checked twenty-four weeks after starting the study therapy to see if her HIV infection was getting worse (an outcome called virologic failure). Twenty-six of the 120 women treated with Nevaripine developed virologic failure, compared to ten of the 120 women treated with the other medicine who experienced virologic failure.
   a. Create a two-way table presenting the results of this study.
   b. State appropriate hypotheses to test for difference in virologic failure rates between treatment groups.
   c. Complete the hypothesis test and state an appropriate conclusion. (Reminder: verify any necessary conditions for the test).

Homework 5

Statistical Inference, Fall 2021

4- The National Center for Health Statistics (NCHS) provided data on the weight distribution (category) of Americans in 2002. This distribution is based on a specific value of Body Mass Index (BMI) calculated as the square of body weight (kilograms) and height (meters). Underweight was defined as BMI <18.5, normal weight was defined as BMI from 18.5 to 24.9, overweight was defined as BMI from 25 to 29.9, and obesity was defined as BMI over 30. Americans in 2002 were distributed as follows: 2% underweight, 39% normal weight, 36% overweight, and 23% obesity. Suppose you want to evaluate whether the BMI distribution is different in the Framingham Offspring sample. Using data from n = 3,326 participants who participated in the 7th examination of the descendants of the Framingham Heart Study, we created a defined BMI category and observed the following:

	Underweight BMI<18.5	Normal Weight BMI 18.5-24.9	Overweight BMI 25.0-29.9	Obese BMI $\geq$ 30	Total
# of Participants	20	932	1374	1000	3326

a. Set up hypotheses and determine the level of significance.
b. Select the appropriate test statistic.
c. Set up decision rule.
d. Compute the test statistic.
e. Conclusion.

5- In Iran, an insurance company asked motorcycle and car drivers whether they prefer not to use their insurance to get a discount for next year or use the insurance to get their damage cost. A person can be a motorcycle driver or car driver in this poll, not both. Below is the distribution of responses by the driver's type.

	Motor cycle	car
Yes	23	43
No	27	27
Total	50	70

Is the proportion of these two different groups identical or not?

Homework 5

Statistical Inference, Fall 2021

6- **(R)** The experimenters investigated whether a mother can recognize the smell of her child. 26 practitioners of mothers were kept in an isolated room. The children were given to their mothers in order to be recognized. The mother should be able to identify her child's smell. There were a total of 320 trials, of which the mother correctly recognized her child's smell 110 times.

   a. What is the null hypothesis, i.e., how often would we expect the participants to be correct by chance (in raw number and in percentage)?

   b. Using a chi-square test, what do you conclude about whether mothers recognize their child's smell?

7- **(R)** Consider the dataset named "caith" in the "MASS" library. Caith consists of 4 rows of eye colors: blue, light, medium, and dark and 5 rows of hair colors: fair, red, medium, dark and black. Test the hypothesis of whether the eye color is independent of the hair color at a .05 significance level. calculate the expected value.

8- **(R)** An experiment was done to check the temperature required to damage a computer system. This experiment is important for computer servers to keep them in a good environment. Because of the price of testing, twenty computers were tested to see what temperature would damage the computer. In this test, all of the computers were damaged in more than 110 degrees. Perform a hypothesis test for the variable's success rate by means of the simulation method.