**Homework 07.**

In this assignment you will have a chance to apply statistics to two published papers that used statistics.

**Assigned: 10 October 2017**

**Due: 5:00PM PST, 17 October 2017**

**Instructions: There are ten multiple choice questions. To receive credit, EMAIL your solution by the deadline to** [**tony\_statman@yahoo.com**](mailto:tony_statman@yahoo.com) **according to the following instructions:**

* The SUBJECT LINE must be “**GSBA545 HW07 for [Last name, First name] –** “ and then the ten letters corresponding to your answers; so, for example, if your name were John Doe, and you believed the answers were CABEDABCCD, then the subject line of the email must be “**GSBA545 HW07** for **Doe, John - CABEDABCCD**”
  + The first seven characters (**GSBA545**) do not have a space between “GSBA” and “545”
  + The ten characters of your answer should have **no spaces in between**
  + If you submit less than 10 letters, it is assumed that the first letter corresponds to your answer to the first question, etc.
* The FIRST LINE of the body of the email should be your last name, your first name, and your student ID
* The SECOND LINE of the body of the email should be five letters, corresponding to the answers to the five questions (make sure your answer consists of five characters)

**For example, a typical email might be**

From: John Doe <john.doe@usc.edu>

To: tony\_statman <tony\_statman@yahoo.com>

Subject: GSBA545 HW07 for Doe, John - CABEDABCCD

DOE, JOHN 123456789  
CABEDABCCD

After the Iraq war ended, no weapons of mass destruction were found, but a large minority of Americans believed that WMD’s had been found. In July to September 2003, U.S. adults in the equivalent of a large simple random sample of U.S. adults were asked two questions:

* “Where do you tend to get most of your news?”
* “Since the war with Iraq ended, is it your impression that the US has or has not found Iraqi weapons of mass destruction?”

A total of 2508 U.S. adults reported one primary source of news. The rate at which various groups believed that weapons of mass destruction were found, based on where they got their news, is given in the table below

Primary source of news n Percent who believed WMD’s were found

Fox 475 33.05%

CNN / NBC / ABC / CBS 1321 20.74%

Print media 633 17.06%

PBS / NPR 79 11.39%

1. What is a 95% confidence interval for the percentage of all Americans who believed that WMD’s had been found?

1. 2.6% to 38.5%
2. 7.8% to 33.3%
3. 19.9% to 27.1%
4. 20.2% to 20.9%
5. 20.2% to 23.5%
6. Is the rate of people who believed that WMD’s were found the same for all groups, regardless of where they get their news? What is the appropriate null hypothesis?
7. All the population rates are different
8. At least one population rate is different from the others
9. All the population rates are the same, but that rate is unknown
10. All population rates are the same, and that rate is equal to 0.0% (i.e., no misinformation)
11. All population rates are the same, and that rate is equal to 21.85%
12. What is the appropriate test of statistical significance?
13. Perform a two-sample T test for difference of population means based on unknown population standard deviations that are assumed to be equal
14. Perform a two-sample T test for difference of population means based on unknown population standard deviations that are not assumed to be equal
15. Perform a one-sample T test based on matched pairs
16. Perform a 2 test of independence
17. Perform a 2 goodness of fit
18. Viewers who received their news from Fox had the highest rate of misperception. Do these data by themselves prove that Fox is the cause of incorrect information?
19. No, because the difference in rates was not statistically significant.
20. No, because although the difference in rates was statistically significant, there could be confounding factors such as political party
21. No, because the differences are under 50 percentage points.
22. Yes, because Fox caused people to believe that WMD’s had been found.
23. Yes, because this was a randomized experiment, and so confounding factors have been randomized away.

In 2016, data were collected from the General Society Survey, the equivalent of a simple random sample of U.S. adults. Among other questions, individuals were asked:

* Their age
* Their sexual orientation
* Number of female sex partners since age 18
* Number of male sex partners since age 18

The data included n=1520 individuals who self-identified as “heterosexual or straight.” (Data are available as a file, or can be downloaded from <https://gssdataexplorer.norc.org/> )

1. From the data, consider women who are between the ages of 18 to 35 (inclusive). Assuming these data are equivalent to a simple random sample of all women age 18 to 35, what is a 95% confidence interval for the average value for “total number of sexual partners (male and female) reported by women age 18 to 35”? (For this analysis, treat answers of “Don’t know”, “No answer”, and “Garbled text” as zero partners.)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| A. 0.00 to 27.57 | B. 4.85 to 7.84 | C. 4.96 to 7.78 | D. 5.02 to 7.87 | E. 5.25 to 7.96 |

1. The 685 men and the 835 women can be divided into three groups: those reporting 0 to 5 partners of the opposite sex; those reporting 6 to 15 partners of the opposite sex; and those reporting 15 or more partners of the opposite sex. The results of this partitioning are given below.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | # of partners of opposite sex | Females | Males |  |
|  | 0 – 5 | 609 | 344 |  |
|  | 6 – 15 | 167 | 195 |  |
|  | 16 or more | 59 | 146 |  |

Based on the above table, does it seem that men and women have the same distribution for “self-reported number of partners of the opposite sex”? What is the most appropriate test of statistical significance?

1. Perform a 2 test of independence
2. Perform a 2 goodness of fit
3. Perform a two-sample T test for difference of population means based on unknown population standard deviations that are assumed to be equal
4. Perform a two-sample T test for difference of population means based on unknown population standard deviations that are not assumed to be equal
5. Take differences, perform a one-sample T test based on matched pairs
6. Based on the above test, what is the appropriate numerical value of the test statistic?
7. C = 44.6, compare to 2 with 2df
8. C = 99.9, compare to 2 with 2df
9. T = 0.28, compare to T with 2 df
10. T = 0.28, compare to T with 4 df
11. T = 0.46, compare to T with 2 df
12. The sample data show a disparity between “average number of females reported by heterosexual males” and “average number of males reported by heterosexual females.” As Michele Alexander noted in the article, “Truth and Consequences” (*Journal of Sex Research*, 2003), “Because a partner is required, it is impossible for men to engage in heterosexual intercourse more often than their female counterparts.” When performing a test of statistical significance about heterosexuals’ self-reported “number of partners of the opposite sex”, what would the null hypothesis be?
13. Heterosexual men exaggerate their self-reported number of female sexual partners
14. Heterosexual women under-report their self-reported number of male sexual partners
15. Both heterosexual men and heterosexual women accurately report their number of partners, and the difference in sample averages is due to random sampling error
16. The average number of female partners reported by heterosexual males is greater than the average number of male partners reported by heterosexual females, but it is not clear why
17. Statements A), B) and D) could all be statements of the null hypothesis
18. Which is the most appropriate mathematical statement of the null hypothesis? Let m be the population average number of female partners reported by men, let f be the population average number of male partners reported by women, and let m and f be the population standard deviations of reported number of partners.
19. Men and women are telling the truth, and m = f and m = f
20. Men and women are telling the truth, and m = f, but m and f might differ
21. Men and/or women lie, so m > f, but m and f are equal
22. Men and/or women lie, so m > f, and m and f might not be equal
23. Men and/or women lie, so m > f, and m and f must not be equal
24. What is the appropriate T-value when testing the null hypothesis? Choose the best answer from the list below.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| A. 0.317 | B. 5.71 | C. 6.14 | D. 8.40 | E. 8.73 |