Questions 1-5 One of the most controversial and polarizing initiatives of the Obama administration has been the Affordable Care Act, a health care reform that is often referred to as Obamacare. Match each of the five situations below with the parameter of interest from the list. (They should be used only once).

- α) one mean
- β) one proportion
- γ) difference of two dependent means
- δ) difference of two independent means
- η) difference of two independent proportions

1. Will the new health care policy result in actual savings in premiums for the average family, or will it end up costing more or not making a difference?

2. For the program to work, about 40% of enrollees have to be young adults aged between 18 and 34. These generally healthier participants would help subsidize the health care of older enrollees. What percentage of those who have actually enrolled are in this coveted age group?

3. The program’s website was plagued with problems from the moment it opened, eliciting frustration and harsh criticism. What is the average amount of time it takes for a person to complete the enrollment process?

4. More affluent Americans generally have access to health insurance through their place of work, so Obamacare may appeal more to lower income families. How does the average income of those who sign up for the program compare to those who do not?

5. A recent poll found that 53% of Democrats think the health care system will get better as a result of the law, while 83% of Republicans expect the system to get worse.

6. When testing $H_0: \mu_1 - \mu_2 = 0$ vs $H_a: \mu_1 - \mu_2 \neq 0$ the p-value = 0.07. Then we can predict that:
   a) The 90% CI for $\mu_1 - \mu_2$ will include zero but the 95% will not.
   b) The 90% CI for $\mu_1 - \mu_2$ will not include zero but the 95% will.
   c) Neither the 90% CI for $\mu_1 - \mu_2$ nor the 95% CI will include zero.
   d) Both the 90% CI for $\mu_1 - \mu_2$ and the 95% CI will include zero.
   e) It’s impossible to determine whether the 90% CI for $\mu_1 - \mu_2$ or the 95% CI will include zero.

7. We conduct 100 tests, all at $\alpha=0.05$, and in all 100 situations the null hypothesis was true. Then we expect that:
   a) in 5 of the tests we will wrongly reject the null hypothesis, committing a Type I error.
   b) in 5 of the tests we will wrongly reject the alternative hypothesis, committing a Type II error.
   c) in 5 of the tests we will wrongly reject the alternative hypothesis, committing a Type I error.
   d) in 5 of the tests we will wrongly reject the null hypothesis, committing a Type II error.
   e) in none of the tests will we commit an error, as long as the assumption of randomness is satisfied.

8. Which of the following statements are true?
   a) Practical Significance is determined by the p-value, while Statistical Significance is determined by the CI.
   b) Statistical Significance is determined by the p-value, while Practical Significance is determined by the CI.
   c) Both Practical Significance and Statistical Significance are determined by the CI.
   d) Both Practical Significance and Statistical Significance are determined by the p-value.
   e) None of the statements above are true.
Questions 9 – 15 A recent poll conducted by the Pew Research Center asked 1350 randomly selected American adults (ages 18 and over) whether they favor or oppose allowing gays and lesbians to marry legally and also whether this had always been their position or if they had changed their mind on the issue. The results of the poll appear below.

<table>
<thead>
<tr>
<th></th>
<th>Favor</th>
<th>oppose</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Before</strong></td>
<td>495</td>
<td>30</td>
</tr>
<tr>
<td><strong>Oppose</strong></td>
<td>210</td>
<td>615</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>705</td>
<td>645</td>
</tr>
</tbody>
</table>

9. What percentage of survey participants favor gay marriage now?
   a) 52%  b) 70%  c) 94%  d) 39%  e) 74%

10. What percentage of survey participants have changed their mind on this issue?
    a) 18%  b) 30%  c) 6%  d) 25%  e) 2%

11. We should use McNemar’s Test to analyze this data because:
    a) we are comparing two independent proportions.
    b) we are comparing two dependent proportions.
    c) the assumptions necessary for independent proportions procedures are not satisfied.
    d) the assumptions necessary for dependent proportions procedures are not satisfied.
    e) None of the above

12. Compute the Test Statistic for McNemar’s Test.
    a) 13.42
    b) 10.33
    c) 11.62
    d) 17.51
    e) 15.49

13. The p-value for this test would be:
    a) close to 0
    b) close to 1
    c) close to 0.50
    d) close to 0.05
    e) close to 0.25

14. We can conclude from this significance test that there is:
    a) a significant difference in attitudes towards gay marriage in recent years.
    b) a significant number of people who do not support gay marriage now.
    c) a significant number of people who support gay marriage now.
    d) a significant number of people who did not support gay marriage before.
    e) All of the above

15. There were two national polls conducted during the same time period that used slightly different wordings in their questions. The first poll, by the Pew Research Center question asked: “Do you favor or oppose allowing gays and lesbians to marry legally?” The second poll, by the Washington Post/ABC News survey, asked “Do you think it should be legal or illegal for gay and lesbian couples to get married?” A 95% confidence interval to compare the proportion of successes (support for gay marriage) in these two polls was constructed: (-.14, -.06). So we can say that the proportion of Americans who “favors gay marriage” is ____________ the proportion who thinks “gay marriage should be legal.”
    a) not significant different from
    b) significantly lower than
    c) significantly higher than
    d) significantly different from
    e) none of the above
Questions 16–21 Does running a mile burn more calories than walking a mile? And is there a difference between males and females? 15 male and 15 female college students of average fitness volunteered to participate in a study that had each one of them walk a mile at a 20 min/mile pace and also run a mile at a 10 min/mile pace and measured their energy expenditure. Summary statistics and four confidence intervals to analyze the data (all at 95% confidence level) appear below.

<table>
<thead>
<tr>
<th></th>
<th>Walk</th>
<th>Run</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>mean</td>
<td>std dev</td>
</tr>
<tr>
<td>M</td>
<td>100.80</td>
<td>18.20</td>
</tr>
<tr>
<td>F</td>
<td>77.27</td>
<td>9.6</td>
</tr>
</tbody>
</table>

(i) Males Walk vs Run: $(-42.764, -12.84)$
(ii) Females Walk vs Run: $(-28.19, -10.47)$
(iii) Males vs Females Walk: $(12.49, 34.57)$
(iv) Males vs Females Run: $(18.38, 45.62)$

16. In conducting the significant tests associated with each question, according to what the researchers wanted to prove, the signs for the alternative hypotheses for each one (in the same order as the CIs) should be:
   a) (i) >  
   b) (i) ≠  
   c) (i) <  
   d) (i) <  
   e) (i) ≠

17. From the confidence intervals we can predict that, if we conducted two-sided tests for each situation, the p-values for each one (in the same order as the CIs) should be:
   a) (i) negative  
   b) (i) < 0.05  
   c) (i) positive  
   d) (i) < 0.05  
   e) (i) > 0.05

18. Given those summary statistics, and without the original data, all four questions have to be treated as problems for comparing two independent means. But given the original data the best way to treat the questions (in the same order as the CIs) should be:
   a) (i) dependent  
   b) (i) independent  
   c) (i) dependent  
   d) (i) independent  
   e) (i) independent

19. Which of the following conclusions can be correctly derived from the CIs given?
   a) On average, walking burns less calories than running for both men and women. ✓
   b) On average, males burn less calories than females while either walking or running. X
   c) On average, males burn more calories than females while running, but less while walking. X
   d) On average, walking burns less calories than running for males but not for females. X
   e) On average, females burn less calories while running than males do while walking. X

20. Which of the following is true about the randomness assumption for inferences based on this data? The assumption of randomness:
   a) is violated, but participants may still be representative of all other potential exercisers.
   b) is one we still need to make, even when it’s clear the participants were not randomly selected.
   c) is not necessary in this situation because we are making inferences about means.
   d) is not necessary in this situation because we are comparing two groups.
   e) is not necessary in this situation because the samples are large enough.

21. Which of the following is true about the normality assumption for inferences based on this data? Energy expenditure during walking and running, for men and women:
   a) in the sample must be assumed to follow a Normal distribution.
   b) in the population must be assumed to follow a Normal distribution.
   c) in the sample will have Normal distribution because the samples are large.
   d) in the population will have Normal distribution because the populations are large.
   e) None of the above.
22. Concluding there is no significant difference between two population means is the same as saying:
   a) There are no differences in the population means.
   b) There are no differences in the sample means.
   c) The difference in the population means was not large enough to prove a difference in the sample.
   d) The difference in the population means was large enough to prove a difference in the sample.
   e) The difference in the sample means was not large enough to prove a difference in the population.

Questions 23-30 A recent survey asked participants about their favorite computer operating system. Is there a
difference in the proportion of males and females that prefer Mac over Windows? 106 out of 335 Males (group 1) said
they preferred Mac, while 84 out of 249 Females (group 2) said the same. Match the following symbols with the
appropriate values.

23. \( p_1 \)  
   a) 0.3373

24. \( p_2 \)  
   b) 0.3253

25. \( \hat{p}_1 \)  
   c) 0.3164

26. \( \hat{p}_2 \)  
   d) 0.3269

27. \( \hat{p}_{\text{pooled}} \)  
   e) unknown

\[
\hat{p}_{\text{pooled}} = \frac{106 + 84}{335 + 249} = \frac{190}{584} = 0.3253
\]

\[
\hat{M} = \frac{106}{335} = 0.3164
\]

\[
F = \frac{\hat{M}}{0.3253} = 0.3164
\]

28. We want to test Ho: \( p_1 - p_2 = 0 \) vs Ha: \( p_1 - p_2 \neq 0 \). If the test statistic is \( z = -0.34 \) then the p-value will be:
   a) 0.3669
   b) 0.7338
   c) 0.2662
   d) 0.6331
   e) 0.5934

29. Which of the following is a correct interpretation of the relative risk for this problem?
   a) The proportion of women who prefer Mac is 1.07 times higher than the proportion of men.
   b) The proportion of men who prefer Mac is 1.07 times higher than the proportion of women.
   c) The proportion of women who prefer Mac is 2% lower than the proportion of men.
   d) The proportion of men who prefer Mac is 2% lower than the proportion of women.
   e) The proportion of women who prefer Mac is between 2% lower and 1.07% higher than the proportion of men.

30. If the confidence interval for \( p_1 - p_2 \) includes zero then the confidence interval for the relative risk:
   a) will also include zero
   b) will include one
   c) will include the true parameter
   d) will include the statistic
   e) will include the population

Questions 31-32 A recent survey asked participants about their favorite computer operating system. There were 373
Windows users with a mean age of 30.01 and standard deviation of 13.20. There were 190 Mac users with a mean age
of 29.93 and standard deviation of 14.51.

\[
\begin{align*}
N_1 & = 373 \\
N_2 & = 190 \\
\bar{X}_1 & = 30.01 \\
\bar{X}_2 & = 29.93 \\
S_1 & = 13.20 \\
S_2 & = 14.51
\end{align*}
\]

31. What is the standard error of the test statistic to compare the average age of Mac and Windows users?
   a) 1.05
   b) 0.68
   c) 1.58
   d) 1.26
   e) 0.87

32. How many degrees of freedom minimum should be used to perform inferences here?
   a) 189
   b) 190
   c) 372
   d) 373
   e) 561

33. Interpret the p-value for this test: 0.9492.
   a) There is a significant difference in the average age of Windows and Mac users.
   b) There is no significant difference in the average age of Windows and Mac users.
   c) Windows users are, on average, almost one year older than Mac users.
   d) Mac users are, on average, almost one year older than Window users.
   e) Windows users are, on average, between almost one year younger and almost one year older than Mac users.