SOLUTION

1. Suppose a simple random sample of size n = 16 from some population of interest has a sample mean of 3.5 and sample variance of 4.

   (a) What is the value of the t-statistic for testing the hypothesis that the population mean is 2? (3)

   $H_0 : \mu = 2$

   $$t = \frac{3.5 - 2}{\sqrt{4/16}} = \frac{1.5}{2/4} = 3$$

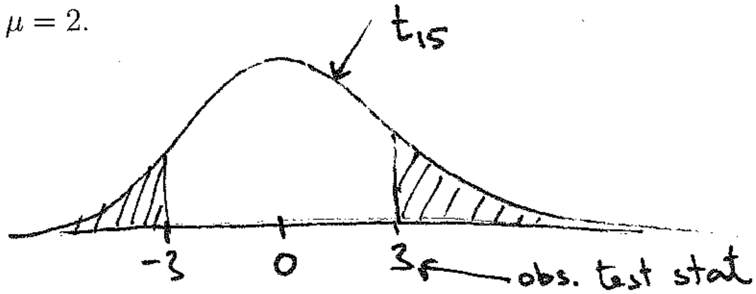
   (b) What critical value should you use to perform a **one-sided** level $\alpha = 0.05$ test of $H_0 : \mu = 2$ vs. a greater alternative $H_A : \mu > 2$? You can write your answer in R code or notation. (3)

   - $t_{(n-1), 1-\alpha} = t_{(15), 0.95}$ (or notation like: $t_{15}(0.95)$)

   OR

   - qt(0.95, df = 15)

   (c) Sketch a diagram of the appropriate reference distribution, including the observed test statistic, and shading in the area required for a **two-sided** p-value for the hypothesis $H_0 : \mu = 2$. (3)



   (d) Construct a 95% confidence interval for the population mean. (3)

   Some possibly useful quantiles of the t-distribution are:

   $t_{(16),0.95}$ = qnorm(0.95, df = 16) = 1.75
   $t_{(16),0.975}$ = qnorm(0.975, df = 16) = 2.12
   $t_{(15),0.95}$ = qnorm(0.95, df = 15) = 1.75
   $t_{(15),0.975}$ = qnorm(0.975, df = 15) = 2.13 ←
   (Make sure you clearly indicate which one you use.)

   $$\bar{Y} \pm t_{(n-1), 1-\alpha/2} \sqrt{\frac{s^2}{n}}$$

   $$3.5 \pm 2.13 \left(\sqrt{\frac{4}{16}}\right)$$

   $$3.5 \pm 2.13 (0.5)$$

   95% CI: $(2.435, 4.565)$

2. **Short Answer** Give a 1-2 sentence answer to each question.

   (a) When are causal inferences statistically justified? (3)

   - In a randomized experiment
   - If treatments are randomly assigned

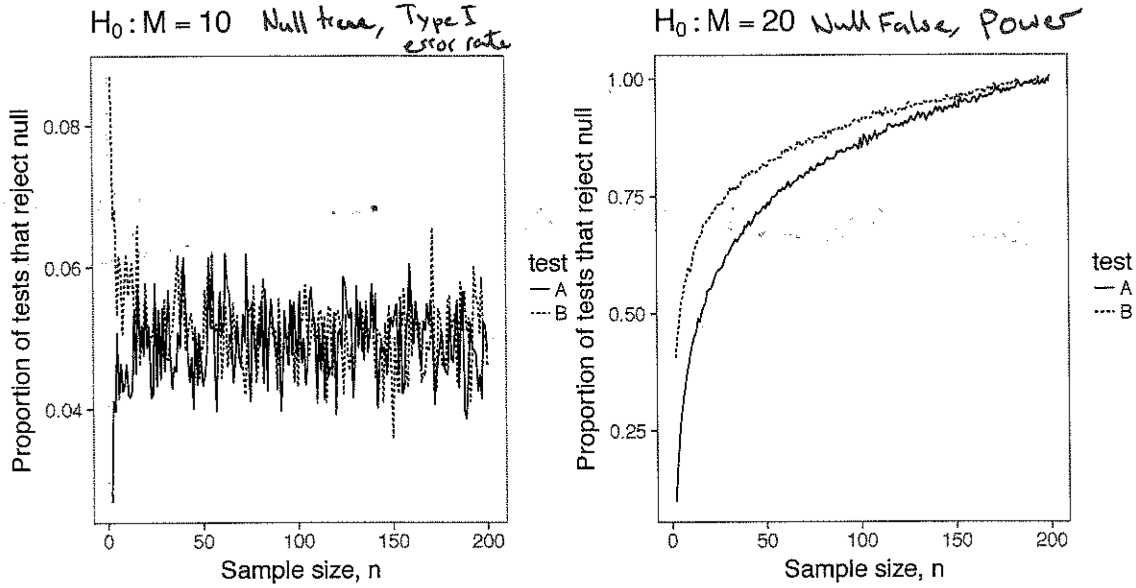
   (b) Describe what is meant by the sampling distribution of a statistic. (3)

   The distribution of the statistic under all possible samples from the population.

   (c) Give the interpretation/definition of a p-value. (3)

   The probability, if the null is true, of seeing a ~~test~~ statistic as or more extreme than that observed

(d) Consider two possible test procedures for testing a hypothesis about the population median: test A, and test B. Below you see plots of the rejection rate for a level $\alpha = 0.05$ test for various sample sizes, based on simulation, for a particular setting where the population median is 10. In this setting, which test would you prefer to use and why? (3)



Prefer B: Both A & B seem to have ~~quite goo~~ - Type I error close to $\alpha = 0.05$.

B has higher power at all sample sizes.

(e) Is the Wilcoxon Signed Rank test a good test for a population median? Why or why not? (3)

Only if you add the assumption that the population is symmetric. Otherwise the ~~Type I error~~ it is not exact nor consistent.

3. I bought a bag of fun size candy bars containing Mars and Snickers. I'm suspicious that the manufacturer doesn't put in equal proportions of the two types of candy bar. In my bag there were 20 bars total, and 8 were Mars.

   Assume the contents of my bag are like a random sample from the population of candy bars that get put in bags.

   You'll find two Binomial distributions that may be helpful on the opposite page.

   (a) Find the exact two-sided p-value for the null hypothesis $H_0 : p = 0.5$. (4)

   Use Binomial (20, 0.5) ← dist for X = # Mars when null is true

   $p = 2\,(0.12 + 0.074 + 0.037 + 0.015 + 0.005 + 0.001)$ OR $(1 - (0.16 + 0.176 + 0.16))$

   $= 0.504$

   (b) What conclusion would you make at level $\alpha = 0.05$? (2)

   $p \gg 0.05$

   Fail to reject H_0

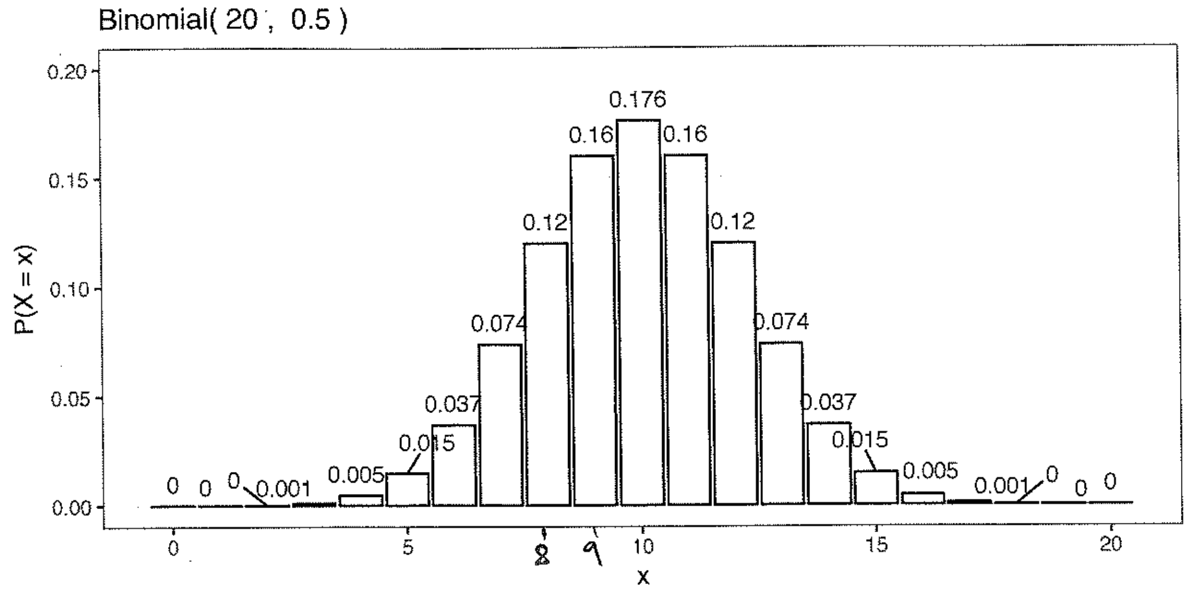
   (c) Would you recommend using the *Approximate Binomial Test* here? Why or why not? (2)

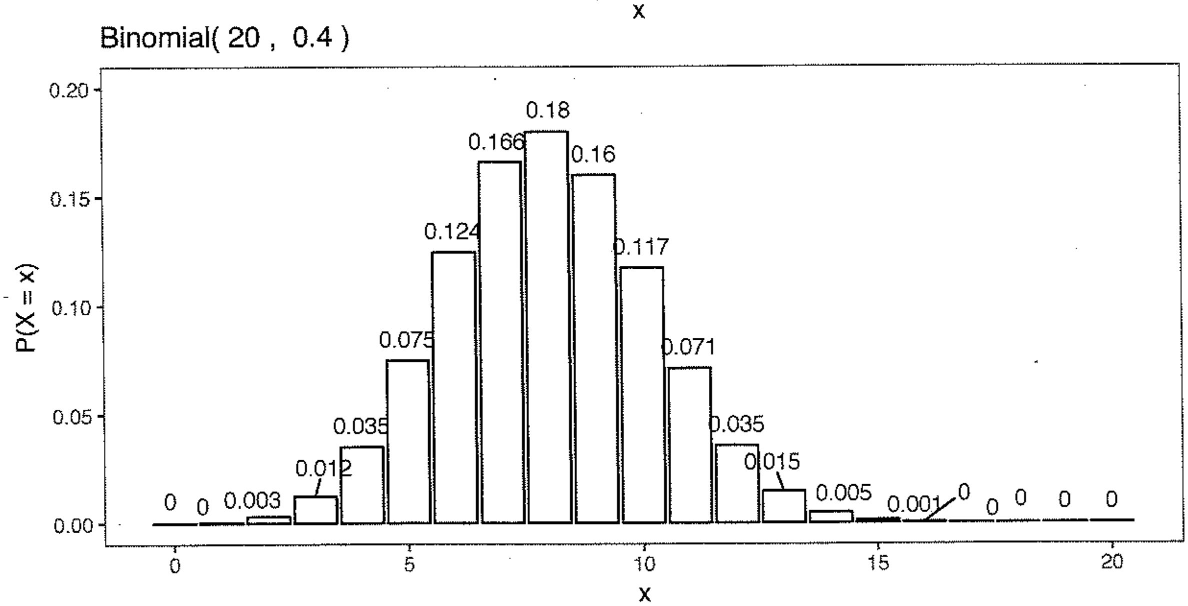
   $np_0 = 10 > 5$ & $n(1-p_0) = 10 > 5$

   Based on guideline the Approximate Binomial Test should be OK.

   (Credit also for talking about small n, assymtotic exactness etc.)

 Question 3 continues on the next page...


Binomial(20 , 0.5)
P(X = x)
0.20
0.15
0.10
0.05
0.00
0.176
0.16
0.16
0.12
0.12
0.074
0.074
0.037
0.037
0.015
0.015
0.005
0.005
0.001
0.001
0 0 0
0 0 0
0
5
8 9
10
15
20
x


Binomial(20 , 0.4)
P(X = x)
0.20
0.15
0.10
0.05
0.00
0.18
0.166
0.16
0.124
0.117
0.075
0.071
0.035
0.035
0.012
0.015
0.003
0.005
0.001
0 0
0 0 0 0 0
0
5
10
15
20
x

4. Researchers are interested in understanding the relationship between the analytic skills of young gifted children and their parents' IQ.

   For 36 gifted children they find the difference in IQ between their mother and father (father's IQ - mother's IQ).

   The researchers conduct a t-test of the hypothesis the mean difference in IQ between the father and mother is zero, against a two-sided alternative. They obtain the following results from their statistical package.

   p-value = 0.009898
   Sample mean = -3.3888889
   95% confidence interval = (-5.9108828 , -0.866895)

   (a) Write a statistical summary based on these results. (5)

   There is convincing evidence the mean IQ difference between father's and mother's of gifted children is not zero (one sample t-test of differences, $p = 0.01$).

   It is estimated that mothers have, on average, an IQ 3.4 points higher than the fathers IQ.

   With 95% confidence, the mothers have, on average an IQ between 0.87 and 5.91 points higher than the fathers.

   (b) What additional information do you need to decide whether the inferences in your summary are justified? (2)

   To make inferences to all gifted children we need to know how these 36 children were sampled. Population inferences are only justified is these 36 children were a <u>random</u> sample from a larger population of gifted children.

Question 4 continues on the next page...

(c) What procedure would you recommend to the researchers if they were more interested in the median difference in IQ between the father and mother? (1)

Sign test.

(d) What is wrong with conducting two hypothesis tests, one for the mean and one for the median, and then reporting only the result with the smaller p-value? (2)

Conducting many tests and only reporting some (i.e. the more "significant") is an example of p-hacking. It is unethical and leads to invalidating the desired performance of the tests.

5. **True/False** Circle one, you do not need to give justification.

   (a) With a large sample size, the t-test will produce approximately valid inference for the population mean even if the population distribution is not normal. (1)

   A. TRUE B. FALSE

   (b) A p-value of 0.20 provides convincing evidence the null hypothesis is true. (1)

   A. TRUE B. FALSE

   (c) If a two-sided p-value in a t-test for testing $H_0 : \mu = 0$ is greater than 0.05, then 0 will be in the 95% confidence interval for μ. (1)

   A. TRUE B. FALSE

   (d) The null hypothesis in the one-sample Z test is that the sample average is some value μ_0. (1)

   A. TRUE B. FALSE

   (e) As the sample size increases the power of the Z test for a fixed alternative decreases. (1)

   A. TRUE B. FALSE