Name ________________

1 It is of interest to test the hypotheses H_0: $p = 0.8$ versus H_a: $p \neq 0.8$. Using a sample of size $n = 40$ observations, the test statistic is $\hat{p} = 0.65$, The p-value for this test was found to be 0.03.

If the test was performed correctly, where should the randomization distribution be centered?

a) 0 b) 0.03 c) 0.65 d) 40 e) 0.8

2 Fill in the blanks using one of the options provided.

a) If a test is statistically significant at the $\alpha = 0.05$ then it is ________ also significant at the $\alpha = 0.01$ level.

a) never b) sometimes c) always

b) If a test is statistically significant at the $\alpha = 0.01$ then it is ________ also significant at the $\alpha = 0.05$ level.

a) never b) sometimes c) always

3 Alice and Bob are each testing the same hypothesis using the same study design. Each of them uses a significance level of $\alpha = 0.05$ as the threshold for statistical significance. Alice has a sample of size 65 and Bob has a sample of size 30.

Circle the correct option in each item below.

a) Assuming the null hypothesis is true, who is more likely to make type I error?

Alice Bob both the same impossible to tell

b) Assuming the null hypothesis is false, who is more likely to make type II error?

Alice Bob both the same impossible to tell

Explain your reasoning:

4 You Want to Know

In each of the following situations, pretend you want to know some information and you are designing a statistical study to find out about it. For each scenario

(i) say **what variables you would need** to have in your data set,

(ii) for each variable, say whether it is **categorical or quantitative**,

(iii) state the **null hypothesis** you would use in a hypohtesis test to answer your question. Do this **twice**: once using words and once using proper statistical notation.

Record your answers in the table below. Part a) has been done as an example.

a) You want to know whether the mean body temperature of college students is really 98.6 degrees.

b) You want to know whether more Americans consider themselves to be a "morning person" or a "night person".

c) You want to know whether Calvin students and Hope students are equally likely to study abroad.

d) You want to know whether more hours of sleep is associated with a better GPA.

e) You want to know if students perform any better or worse when their tests are printed on blue paper instead of white paper.

	Variable(s)	Null Hypothesis
a)	body temperature (Quant)	$H_0 : \mu = 98.6$ The mean body temperature of all college students is 98.6 degrees F.
b)		
c)		
d)		
e)		

5 The 95% confidence interval for the difference in mean weights for two varieties (A and B) of pumpkins is 2.2 ± 1.8 pounds. What can you say about the p-value of the hypothesis test with $H_0 : \mu_A - \mu_B = 0$, and $H_a : \mu_A - \mu_B > 0$?

a) The p-value is larger than 0.10,

b) The p-value is between 0.05 and 0.10.

c) The p-value is between 0.025 and 0.05.

d) The p-value is less than 0.025

e) We can't tell which of these three is correct from this information.

6 According to recent US growth charts, heights of 10-year old boys are approximately normally distributed with a mean of 139 cm and a standard deviation of 6.5 cm.

Accompany your answers to each of these questions with a picture that shows how you are thinking.

a) Approximately what percent of 10 year-old boys are taller than 139 cm?

b) Approximately what percent of 10 year-old boys are taller than 152 cm?

c) Approximately what percent of 10 year-old boys are between 135 and 145 cm tall?

d) Johnny is in the 80th percentile for height. How tall is he?

7 Male radiologists have long suspected that they tend to have fewer sons than daughters. In one sample of 87 children of "highly irradiated" radiologist fathers, 27 were male and 60 were female.

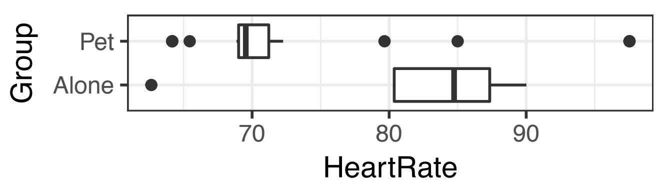
a) What null hypothesis should we test in this situation?

b) The standard error for this test is $SE = 0.05$. Use this to compute the p-value.

c) What conclusion should we draw from this study? Explain.

8 In a study to determine whether the presence of a pet can reduce stress, 30 women, all self-proclaimed dog-lovers, were randomly divided into two groups of subjects. Each subject performed a stressful task. Women in the first group did so alone. Women in the other group did the task with their pet dog present. The average heart rate during the task was used as a measure of stress.
Here is a summary of the data.

```
  Group    min      Q1 median      Q3    max     mean        sd  n missing
1 Alone 62.646 80.3690 84.738 87.3385 90.015 82.58582  7.633606 11       0
2   Pet 64.169 69.0155 69.538 71.2155 97.538 72.94326 10.005395 19       0
```



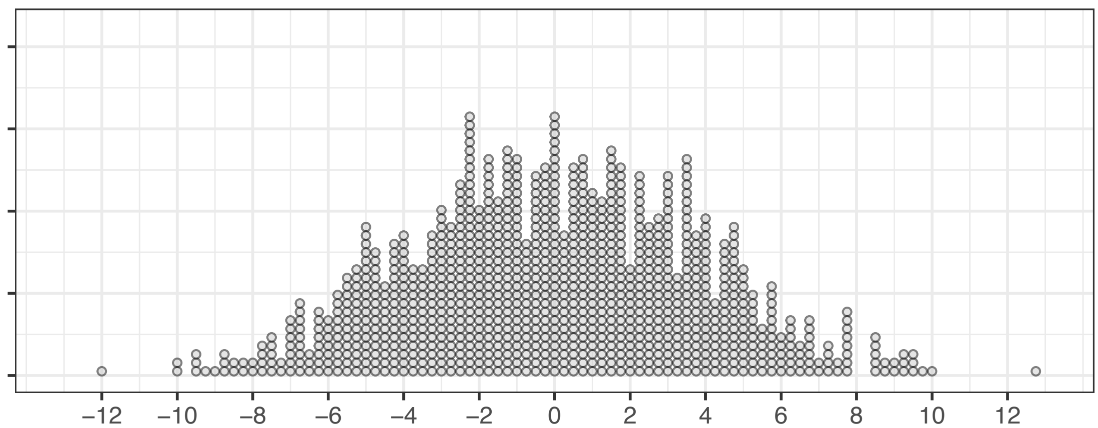
a) Write an R command that can be used to generate the summary table above.

b) Write an R command that can be used to generate the plot above.

c) State the null and alternative hypotheses the researchers should use to answer their question. Do this using **both symbols and words**.

d) Compute the test statistic for this test.

Below is a dotplot showing 1000 randomizations.



e) The x- and y-axes of the dotplot are not labeled. Add good labels.

f) Estimate the p-value. Make sure it is clear how you are getting your value.

g) Based on your p-value, what conclusion should we draw?

h) Give an R command that could be used to generate the randomization distribution for this test.

Space for additional work. Please label.

Solutions

4 a) reading score (quant), gender(cat); 2-sample t.
b) weight (quant); 1-sample t
c) been to a foreign country? (cat); 1-proportion
d) pulse rate (quant), sex (cat); 2-sample t
e) pulse rate during lecture (quant), pulse rate during test (quant); paired t
f) program (cat), success at quitting smoking (cat); Chi-squared for 2-way table

5 C. Since 0 is not in the 95% confidence interval, we will not reject $H_0 : \mu = 0$ at the $\alpha = 0.05$ level. Thus the p-value must be less than 0.05.

6

```
1 - pnorm(b, m, s)

[1] 0.02275013

pnorm(c2, m, s) - pnorm(c1, m, s)

[1] 0.5528661

qnorm(percentile, m, s)

Warning in qnorm(percentile, m, s):  NaNs produced

[1] NaN NaN NaN NaN
```

7

```
data_frame(
  x = c(26, 30, 32, 27),
  p.hat = x / 87,
  SE = round(sqrt(p.hat * (1-p.hat) / 87), 3),     # wrong SE
  z = (p.hat - 0.5) / SE,
  pval1 = pnorm(x / 87, 0.5, SE),
  pval2 =  2 * pnorm(x / 87, 0.5, SE)
)

# A tibble: 4 x 6
      x p.hat     SE     z     pval1     pval2
  <dbl> <dbl>  <dbl> <dbl>     <dbl>     <dbl>
1  26.0 0.299 0.0490 -4.11 0.0000202 0.0000404
2  30.0 0.345 0.0510 -3.04 0.00117   0.00235
3  32.0 0.368 0.0520 -2.54 0.00551   0.0110
4  27.0 0.310 0.0500 -3.79 0.0000744 0.000149
```

Unused Problems

9
Identify distributions

10 What is **blinding** and why is it used in many experimental designs?

11 When we generated a randomization distributions in class, we always made a histogram or a dotplot of the distribution. Why is it important to do this?

12 The Carolina Biological Supply Company markets something called the Wisconsin FastPlant (`http://www.fastplants.org/life_cycle/`) that is advertized to grow from seed to fully mature plant with harvestable seeds in approximately 28 days. Some biology students want to determine the average height (in cm) of 10-day old Fast-Plant plants fed with a new fertilizer they have developed, so they planted the 20 seeds and measured their heights 10 days after planting.

Here is the top bit of their data

plant	height
1	18
2	21
⋮	⋮

What R command could you use to generate the bootstrap distribution in this situation?

13 Test Preparation. An article reports its findings about the difference between two groups of students on a standardized test. Group A had a special one-day preparatory course. Group B did not have any special preparation. The article includes the following chart:

Group	mean score
A	70.4
B	61.5

and then concludes "Our data strongly suggest that the preparatory class improves test scores. (t = 2.19, df= 16, p-value = 0.0218)"

a) Is this result significant at the 5% level? Is it significant at the 1% level? [Be sure your answer indicates how you are getting your answers.]

b) P-values are probabilities, probabilities of what? Carefully explain what probability is being measured by the p-value *in this specific situation*.

14 How Much is that Doggie's Cholesterol? High cholesterol levels in the blood are not good for humans or dogs. Cholesterol levels are affected by diet. A study compared cholesterol levels of healthy dogs owned by a canine research clinic (and fed carefully prescribed diets) with healthy dogs owned as pets that were brought to the clinic to be neutered. If dogs owned as pets differ from dogs owned by the clinic, then researchers feared that studies done on clinic dogs might not be generalizable to dogs owned as pets.
The summary statistics for a sample of these dogs is given below:

Group	n	mean	standard deviation
pets	26	193	68
clinic dogs	23	174	44

a) How strong is the evidence that pets have a different mean cholesterol level than clinic dogs?
Carry out an appropriate test to answer this question. Be sure to

- State the null and alternative hypotheses *in both symbols and words*, and
- Give the p-value and state your conclusion.

b) Give a 95% confidence interval for the difference in mean cholesterol levels bewteen pets and clinic dogs.

c) Give a 95% confidence interval for the mean cholesterol in pets.

15 Put check marks in the appropriate boxes.

	...makes the confidence interval ...		
	wider	narrower	about the same width
Increasing the sample size ...			
Decreasing the sample size ...			
Increasing the level of confidence ...			
Decreasing the level of confidence ...			

16 CORN YIELD.
The mean yield of corn in the U.S. is approximately 120 bushels per acre. A researcher is gathering data on corn yields in a certain region to see if this region does significantly better or worse than the national average. The researcher selects an SRS of 42 farmers from the population of farmers in the region and determines that the mean yield for this sample is 123.8 bushels per acre with a standard deviation of 12.1 bushels per acre.

a) What are the null and alternative hypotheses involved in this study? [Be sure to comment on your choice of alternative hypothesis.]

b) Carry out the test, give the P-value, and state your conclusions.

17 PREDICTING A VOTE. A local paper wants to predict the outcome of a vote on a local referendum, so they use a random-digit dialer to call 200 voters in their city. Each is asked whether they favor or oppose the referendum that will be on the upcoming ballot. Of the 200 respondents, 112 said they were in favor, 88 were opposed.

a) Find a 95% confidence interval for the proportion of voters who were in favor.

b) Can the outcome of the vote be predicted with 95% confidence? Or should the newspaper say that the race is "too close to call." Explain.

c) The answers above rely on the assumption that those called are an SRS from the appropriate population. What is the population of interest? What factors might cause the method used **not** be an SRS from this population?

d) How many people must be called to predict the percentage of people in favor of the referendum to within ±5% with 95% confidence regardless of the split? [Remember, the largest standard deviation occurs when there is a 50-50 split.]

18 A Simple Game. Consider the following game: The player rolls two six-sided dice. If the player rolls double 1's, she wins \$1. If she rolls double 6's, she wins \$6. Otherwise she loses.

a) Let X be the random variable that measures the amount of money won by a player in one try. Fill in the table below to describe the distribution of X:

X	0	1	6
probability			

If you have trouble filling in the table, I will give you the results so that you can continue.

a) What is the mean value of X?

b) What is the standard deviation of X?

c) How much must the house charge per play if is going to make money at this game?

19 A statistically significant difference between two sample means can fail to have practical importance.

20 Bowling with Don and Celeste. Celeste and Don like to bowl. Assume that their scores are (approximately) normally distributed with the mean and standard deviation listed below:

	mean	standard deviation
Celeste	215	15
Don	202	12

a) Which player is more *consistent*? How do you know?

b) What is Celeste's average margin of victory?

c) What is the probability that Don's score is higher than Celeste's in a single head-to-head match?

21 t procedures should never be used with strongly skewed data.

22 An increased sample size means

a) a reduction in the sampling variability (the variability from sample to sample).

b) a reduction in the cost of the study.

c) a reduction in type I error.

d) we're feeding the sample too much.

True/False. Place your answer to each item in the margin to the left of the question. You are not required to give any explanation, but you may if you like. Explanations may result in the awarding of partial credit.

23 Every hypothesis test that is significant at the $\alpha = 0.05$ level is also significant at the $\alpha = 0.01$ level.

24 Every hypothesis test that is significant at the $\alpha = 0.01$ level is also significant at the $\alpha = 0.05$ level.

25 If the degrees of freedom are the same, then a larger value of X^2 (chi-square) produces a larger P-value.

26 Age of Brides. You want to test whether the mean age of a bride in Cumberland County is greater than 30 years. You gather sample data by randomly selecting 24 marriage licenses issued in the last 6 months. The mean bride's age in this sample is 33.83 years with a standard deviation of 13.56.

a) Are the numbers 33.83 and 13.56 parameters or statistics?

b) What are the null and alternative hypotheses for our question?

c) Carry out the hypothesis test. That is, compute the appropriate test statistic and determine the p-value.

d) Write a sentence or two about how to interpret the results of your test.

27 In order to study the effectiveness of a vaccine, 120 experimental animals were given the vaccine and 180 were not. Then all 300 animals were infected with the disease. The results are in the 2-way table below. We want to know if the vaccination reduces the rate of death.

	Unvaccinated	Vaccinated	Total
Lived	162	114	276
Died	18	6	24
Total	180	120	300

a) What is the null hypotheses for this test? (State this in words within the particular context of this problem.)

b) What percentage of the vaccinated animals died?

c) What percentage of the unvaccinated animals died?

d) How many vaccinated animals would we expect to die if the null hypothesis is true? (That is, compute the expected count for that cell of the table. Be sure to show how you calculate this value from the information above.)

Here is some R output:

```
chisq.test(rbind(c(162, 114), c(18, 6)), correct = FALSE)


Pearson's Chi-squared test

data:  rbind(c(162, 114), c(18, 6))
X-squared = 2.4457, df = 1, p-value = 0.1179
```

e) The p-value for this test is 0.1179. The p-value is a probability. Carefully say what probability it is *in this particular situation.*

f) What conclusion do we draw from this test?

Space for extra work.

10 Blinding is withholding information from either experimenters or subjects. It is used to reduce or eliminate bias caused by people interpreting results differently based on such knowledge. Placebo effects are one example of this.

14 For 2-sample t test:

```
SE = sqrt( 68^2/26 + 44^2/23 ); SE

[1] 16.18703

t = ( 193 - 174) / SE; t

[1] 1.173779

2 * pt( -abs(t), df=22 )  # conservative p-value

[1] 0.2530345
```

For 2-sample t interval:

```
t.star <- qt(.975, df=22); t  # conservative t

[1] 1.173779

marginOfE <- t.star * SE; marginOfE

[1] 33.56985

(193 - 174) - marginOfE

[1] -14.56985

(193 - 174) + marginOfE

[1] 52.56985
```

For 1-sample t:

```
SE = 68/sqrt(26); SE

[1] 13.3359

t.star <- qt(.975,df=25); t.star

[1] 2.059539

marginOfE <- t.star * SE; marginOfE

[1] 27.46579

193 - marginOfE

[1] 165.5342

193 + marginOfE

[1] 220.4658
```

16

a) H_0: $\mu = 120$; H_a: $\mu \neq 120$. Since we have not reason (before collecting our data) to know that the local region cannot be worse, we must use a 2-sided test, even if we hope it is better and even if the data comes out showing it is better.

b) $t = \frac{123.8-120}{12.1/\sqrt{42}} = 2.035$. $P(|T| > 2.035) = 2 \cdot P(T > 2.035)$ is between 4% and 5%. This gives evidence sufficient to reject H_0 at the 5% level. We have moderately strong evidence that there is a difference between the local corn production and the national corn production.

18 probabilities: $\frac{3}{4}, \frac{1}{16}, \frac{1}{16}, \frac{1}{16}, \frac{1}{16}$
$\mu_W = (0)(\frac{3}{4}) + 1(\frac{1}{16}) + 2(\frac{1}{16}) + 3(\frac{1}{16}) + 4(\frac{1}{16}) = .625$ dollars.
At least the mean payout, 62.5 cents per play.
B is approximately $N(.625, .1218)$ – use $SE = \frac{1.218}{\sqrt{100}}$.
The probability of coming out ahead after 1 play is $\frac{1}{16} + \frac{1}{16} + \frac{1}{16} + \frac{1}{16} = \frac{1}{4}$. The probability of coming out ahead after 100 plays is $P(\bar{x} > 0.75) \approx P(Z > \frac{.75-.625}{.128} = 1.026) \approx 15.15\%$.
$\sigma_W^2 = (0-.625)^2(\frac{3}{4}) + (1-.625)^2(\frac{1}{16}) + +(2-.625)^2(\frac{1}{16}) + +(3-.625)^2(\frac{1}{16}) + +(4-.625)^2(\frac{1}{16}) = 1.4843$ dollars, so $\sigma_W = 1.218$ dollars.

19 True, especially if the difference is "small". Small, of course, depends on context. With large sample sizes, even very small differences may be detected.

20 the distribution of the differences in their scores is approximately normal with a mean of 13 and a standard deviation of $\sqrt{16^2 + 12^2} = 20$. The probability of Celeste's score is more than Don's $= Pr(X > 0) = Pr(Z > \frac{13}{20}) = Pr(Z > 0.65) = 0.2578$.

21 False. The t procedures are robust against skewness – more so as the sample size increases. For large samples ($n \geq 40$) even strong skewness can be handled.

23 False. If the p-value is 0.03, for example, then the result is significant at the 0.05 level but not at the 0.01 level.

24 True. If the P-value is less than 0.01, then it is certainly less than 0.05.

27

a) Null Hypothesis: Death rates are the same for vaccinated and unvaccinated animals. Alternative Hypothesis: Death rates are different for vaccinated animals than for unvaccinated.

b) $6/120 = 5\%$

c) $18/180 = 10\%$

d) $\frac{120}{300} \cdot 24 = 9.6$

e) Even if there really is no difference in death rates between the vaccinated and unvaccinated populations, we would expect to see a difference at least this large in a sample nearly 12% of the time just by chance.

f) The χ^2 value for this test is 2.446. p-value is 0.1179, so this is not significant at the 5% level. This data does not provide enough evidence to be convinced that the vaccination is doing anything to reduce the death rate.

28 True/False. If an item is false, explain why it is false. (Your answer should not simply be a restatement of the fact that the item is false but should get at the heart of what makes it false.)

a) Half of all data values in a distribution fall below the mean of the distribution. (The other half are above the mean.)

b) When evaluating the results of a study, it is important to know how the data were collected.

c) If the p-value is 0.65, this means the null hypothesis is true.

29 Compute the **mean**, **variance**, and **standard deviation** of the following small data set:

$$1 \quad 4 \quad 7 \quad 8$$

Show all of your work. The point is to show me that you know how these numbers are calculated. Do not use any statistical functions on your calculator or computer. (You may use your calculator or RStudio for basic arithmetic like addition, subtraction, multiplication, division, etc.)

30 SAT SCORES. The distribution of SAT scores for students in a particular school district are approximately a normal distribution with a mean of 500 and a standard deviation of 100.

a) Approximately how many SAT-takers score below 400?

b) Approximately what percentage of SAT-takers score between 400 and 575?

c) If Joe was told that his score placed him in the 28 percentile, what was his SAT score (round to the nearest point)?

d) Approximately what percentage of SAT-takers received a higher SAT score than Joe?

31 A RANDOM VARIABLE. The table below describes the distribution of the random variable X.

value of X	0	1	5	10
probability	0.69	0.20	0.10	0.01

a) What is $P(X > 0)$?

b) Determine the **mean** of X.

32 The Blitz Battery Company claims that the lifetime of their batteries (under continuous use) is normally distributed with a mean of 11 hours and a standard deviation of 2 hours. Suppose you buy a 4-pack of Blitz batteries. You are hoping that you will get at least 40 hours of battery life from the 4 batteries. That's an average of 10 per battery, so it seems like you should have a reasonably good chance.

a) Use the Central Limit Theorem to calculate the probability that the four batteries have a mean lifetime of at least 10 hours.

b) What assumption(s) are you making when you apply the Central Limit Theorem in this case?

28

a) False. The median splits the distribution into two equal pieces.

b) True.

c) False. A large p-value says our data are not unusual when the null hypothesis is true, but that doesn't mean the null hypothesis is true – especially if the sample size is small.

29

```
x <- c(1, 4, 7, 8)
(1 + 4 + 7 + 8)/4  # mean

[1] 5

mean(x)

[1] 5

v <- ((-4)^2 + (-1)^2 + 2^2 + 3^2)/3
v  # variance

[1] 10

var(x)

[1] 10

sqrt(v)  # st dev

[1] 3.162278

sd(x)

[1] 3.162278
```

30

a) A score of 400 corresponds to a Z-score of -1. Using the 68-95-99.7 rule gives $P(Z < -1) \approx 16\%$. Using Table A: $P(Z < -1) \approx 0.1587 = 15.87\%$.

b) The Z-score for 575 is 0.75. $P(-1 < Z < 0.75) = P(Z < 0.75) - P(Z < -1) =$ `pnorm(.75) - pnorm(-1) = 0.6147174`

c) Of course, you should have drawn sketches to accompany your work on this problem, but since I am doing it by computer, I am omitting the sketches in these solutions. Locate 0.2800 (or something close) in Table A. This indicates that Joe's Z-score is between -0.58 and -0.59. Those Z-scores correspond to SAT scores of $500 - (0.58)(100) = 500 - 58 = 442$ and $500 - (0.59)(100) = 500 - 59 = 441$, respectively. So his score was probably one of those two.

d) Since approximately 28% scored below Joe, approximately 72% scored above him.