Exam 2
Psych 2111, Fall 16

Vocabulary

1. Power: The probability of finding an effect (by rejecting H_0) when there is one

2. Sampling distribution: The probability distribution for values a statistic can take on across different samples

3. Law of large numbers: The fact that the sample mean becomes more reliable, and its standard error gets arbitrarily small, as sample size increases

4. Degrees of freedom: The number of squares in a SS or MS formula that are free to vary

5. Null hypothesis: A conjecture about population parameters, stating that there is no interesting effect, and used as the default hypothesis in hypothesis testing

Conceptual questions

1. Write what kind of hypothesis test would apply for each situation. Be as specific as possible.

a) 20 people are timed holding their breath. We want to know whether the mean is longer than a minute.
One-tailed single-sample t-test: We are testing the mean of a single sample, and we are interested only in deviations from the null hypothesis ($\mu = 1$ min) in one direction.

b) 10 of the people are men and 10 are women. We want to know if there's a difference in either direction.
Two-tailed independent-samples t-test: We are comparing the means of two samples, and we're interested in a difference in either direction.

c) Everyone is tested again, on a treadmill at 10 mph. Two people have longer times than in the first condition, and 18 have shorter times. We predicted that most people would last longer standing still.
One-tailed sign test (a type of binomial test): We're using only the directions (signs) of the differences

2. The size of houseflies follows a uniform distribution: there are flies of all sizes between 5 mm and 12 mm, and all of these sizes occur equally often. Choose what shape each of the following distributions should have: Bimodal, Normal, Uniform, or Other/Unknown.

a) We measure 1000 flies and make a histogram of their sizes.
Uniform: The shape of a large sample should be similar to that of the population

b) Each of 1000 people measures 3 flies and calculates the mean. We make a histogram of the 1000 means.
Other: This is a distribution of sample means, but the sample size (3) is too small for the central limit theorem to apply. (The actual shape will be between a triangle and a Normal with no tails.)

c) Each of 1000 people measures 50 flies and calculates the mean. We make a histogram of the 1000 means.
Normal: The sample size is big enough that the central limit theorem applies

3. An experimental program at the local elementary school gives half the students a cup of black coffee at the start of each day. The other students are given decaf. At the end of the year, all the students are

given a spelling test. A t-test comparing the two groups' spelling scores fails to reject the null hypothesis. Write a sentence summarizing what the researchers conclude, regarding coffee and spelling.

There is insufficient evidence to conclude that caffeine affects how well children learn to spell.

4. We learned about three equivalent ways of doing hypothesis testing, using confidence intervals, test statistics, and p-values. Imagine someone is doing a single-sample two-tailed t-test, and for each method, state the rule for when to reject the null hypothesis.

Confidence Interval: Reject H_0 if μ_0 is outside the confidence interval

Test statistic (t): Reject H_0 if t is outside the critical values

p-value: Reject H_0 if $p < \alpha$

5. Draw a picture of a t distribution being used for a one-tailed hypothesis test. Include the following labels in your diagram: critical value, t (statistic computed from data), α, p-value, critical region, and 0 (zero, on the horizontal axis). Arrange things so that the null hypothesis is retained.
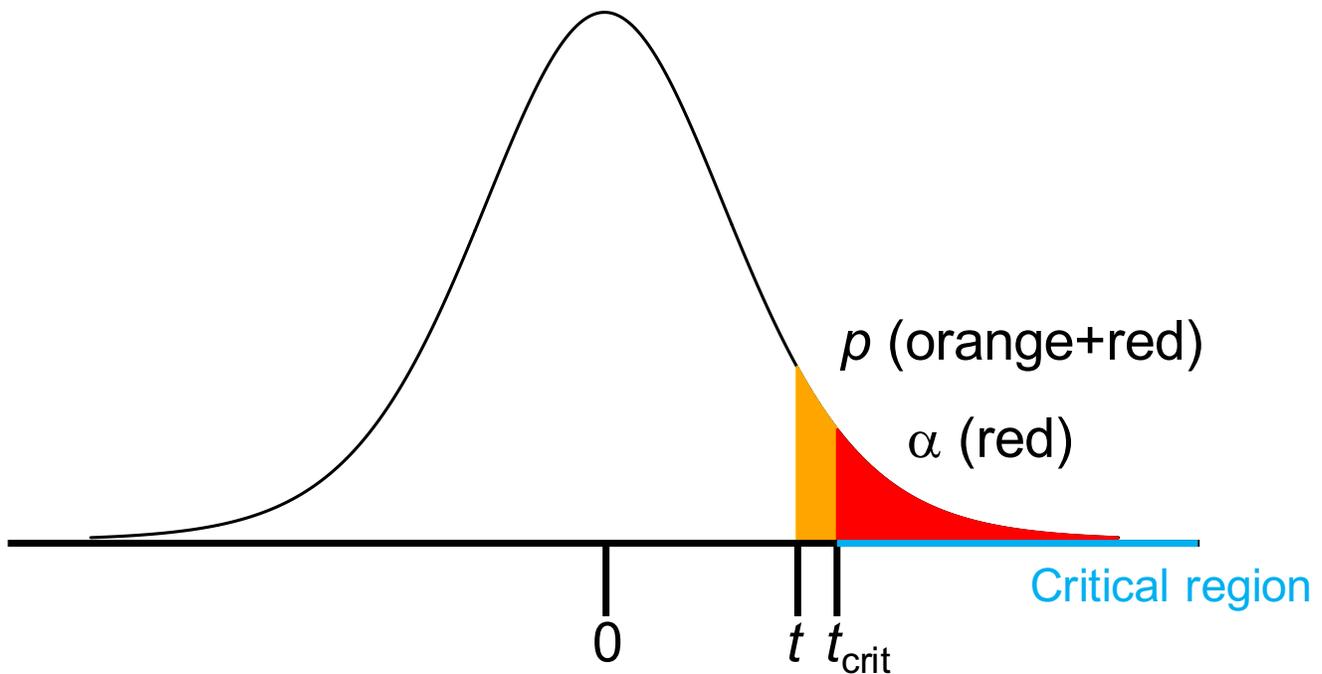
Math questions

You don't need to show your work, but I will give partial credit for partial answers.

1. Infants tend to look at novel objects more than objects to which they've been habituated, but only if they can tell the difference. To test whether 2-month-olds have developed blue-yellow discrimination, you habituate each subject to a blue ball, and then you present them with two blue balls and a yellow ball. If the subjects can't tell the difference, there's a 1/3 chance they'll look at the yellow ball, but if they can then that chance will be greater. Here's the binomial distribution for $q = 1/3$ and $n = 9$:

Frequency:	0	1	2	3	4	5	6	7	8	9
Probability:	.0260	.1171	.2341	.2731	2048	.1024	.0341	.0073	.0009	.0001

a) If 6 of your 9 subjects looked at the yellow ball, what is your p-value for a one-tailed test?

The p-value is the likelihood (according to H_0) of 6 or more subjects looking at the yellow ball.
$.0341 + .0073 + .0009 + .0001 = \underline{.0424}$

b) Write a sentence describing your conclusion (about babies, not about H_0 or H_1), based on an alpha level of .05.

We conclude that babies are more likely than chance to look at the yellow ball, which in turn implies that they can tell the difference between blue and yellow.

2. An old folk remedy for preventing tooth decay is to rinse with diluted wood glue once a day. To scientifically test this treatment, four people tried it for a year and counted the number of new cavities they got. The results for this group were 1, 0, 1, and 2 cavities. Data from a control group of five people (without wood glue) showed 1, 5, 4, 2, and 3 new cavities. Perform a t-test to decide whether the treatment changes people's average cavity rates. Because it seems possible that wood glue is actually bad for people's teeth, you should perform a 2-tailed test. The two-tailed critical t value with $\alpha = 5\%$ is 2.36. Be sure to write which hypothesis you conclude with.

$M_{exp} = 1$, $M_{con} = 3$

$SS_{exp} = 0 + 1 + 0 + 1 = 2$, $SS_{con} = 4 + 4 + 1 + 1 + 0 = 10$

$$MS = \frac{SS_{exp} + SS_{con}}{df} = \frac{10 + 2}{4 + 5 - 2} = \frac{12}{7}$$

$$SE = \sqrt{MS \left(\frac{1}{n_{exp}} + \frac{1}{n_{con}} \right)} = \sqrt{\frac{12}{7} \left(\frac{1}{4} + \frac{1}{5} \right)} \simeq .8783$$

$$t = \frac{3 - 1}{.8783} \simeq 2.28$$

We retain the null hypothesis that wood glue doesn't affect rate of cavities.

3. Calculate a 95% confidence interval for how much the wood glue method reduces the average number of cavities a person will get in a year.

Reduction $= M_{con} - M_{exp} = 2$

$CI = 2 \pm t_{crit} \cdot SE = 2 \pm 2.07 = [-.07, 4.07]$

4. Find the standardized effect size (d) for question 2. Be sure to get the sign right.

$$d = \frac{M_{con} - M_{exp}}{\sqrt{MS}} = \frac{3 - 1}{\sqrt{12/7}} \simeq 1.53$$

5. In a study of social effects on eating, each subject participates in two interviews, one with a male experimenter and one with a female experimenter. There's a bowl of jellybeans on the table throughout, and we count how many each subject eats in each condition. If there are 25 subjects and the standard deviation of the difference scores is $s_{diff} = 2$, how big must the mean difference score (M_{diff}) be in order to conclude the gender of the experimenter affects average jellybean consumption? The critical value (for a two-tailed test, with 24 df and $\alpha = .05$) is 2.06.

To reject the null hypothesis, we must have $t \geq t_{\text{crit}}$

$$\frac{M_{\text{diff}}}{s_{\text{diff}}/\sqrt{n}} \geq t_{\text{crit}} \quad \rightarrow \quad M_{\text{diff}} \geq \frac{t_{\text{crit}} \cdot s_{\text{diff}}}{\sqrt{n}} = \frac{2.06 \cdot 2}{5} = .824$$

R questions

Consider the following command and output for Questions 1-4.

```
> t.test(X,Y)

data:   X and Y
t = -2.5555, df = 88, p-value = 0.01232
95 percent confidence interval:
 -0.8658251 -0.1083051
sample estimates:
 mean of x   mean of y
0.04098254 0.52804766
```

1. What kind of hypothesis test is being run?
Independent-samples t-test

2. What is the null hypothesis?
$M_x = M_y$

3. Which hypothesis is supported by the results (α = 5%)? Point to two separate parts of the output that let you answer this question.
Alternative hypothesis, because $p < .05$ and because zero is not in the 95% confidence interval.

4. If α had been chosen as 1%, which hypothesis would the data support?
Null hypothesis, because $p > .01$

_____

5. What R function can be used to compute a p-value for a binomial test?

pbinom()