Chapter 10 Exam Prep: Two-Sample Inference

Statypus Insight: The Universal Translator (Independent Data)

When comparing two entirely distinct groups (like cats vs. dogs), we cannot pair them up. Instead, we look at the gap between their overall averages. The null hypothesis (H_0) usually assumes this gap is zero ($H_0 : \mu_1 - \mu_2 = 0$, or $\mu_1 = \mu_2$).

Raw Exam Question 1 & Output:
A biologist is investigating the impact of dams on platypus health. She collects a random sample of weights from adult platypuses living in a Regulated river and a separate random sample from an Unregulated river. She wants to test if there is a significant difference in the mean weight between the two river types.

```
> t.test(Regulated, Unregulated, alternative = "two.sided")

Welch Two Sample t-test

data:  Regulated and Unregulated
t = -2.314, df = 42.1, p-value = 0.0256
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.68  -0.04
sample estimates:
mean of x mean of y
     1.41      1.77
```

Exam Tasks (Part 1):

A. State the null and alternative hypotheses (H_0 and H_a) using the correct notation.

B. Based on the R output, state your formal conclusion at the $\alpha = 0.05$ significance level.

Bill the Statypus says: Subscript Friction! Make sure you explicitly define which group is "1" (x) and which is "2" (y) (e.g., let μ_1 = Regulated). If you mix up the order, your signs will be backwards!

Your Turn (Record your answer here.)

Statypus Insight: The Power of the Pair (Dependent Data)

If the two sets of data are collected from the *exact same individuals* (like a before-and-after test), the data is **paired**. We don't compare the two groups anymore. We subtract the values for each individual and create a brand new, single list of differences (μ_d).

Raw Exam Question 2 & Output:
A rehabilitation center tracks the health of injured platypuses. They weigh a random sample of $n = 12$ platypuses when they are admitted to the clinic, and then weigh those exact same 12 platypuses 30 days later. They want to estimate the true mean difference in weight (Day 30 minus Day 1).

```
> t.test(Day30, Day1, paired = TRUE, conf.level = 0.95)

Paired t-test

data:  Day30 and Day1
t = 3.12, df = 11, p-value = 0.0097
alternative hypothesis: true mean difference is not equal to 0
95 percent confidence interval:
 0.12  0.64
sample estimates:
mean difference
           0.38
```

Exam Tasks (Part 2):
C. Locate the 95% Confidence Interval in the R output above. Write a formal sentence interpreting this interval in the context of the rehabilitation program.

CHAPTER 10

Sally the Statypus says: The Zero Trap! When interpreting a difference interval, the most important number isn't the bounds... it's zero! Because both numbers in this interval (0.12 and 0.64) are positive, zero is NOT a plausible value.

Bill the Statypus says: Notation check! We are no longer comparing μ_1 and μ_2. Because we are looking at the average of the differences, your interpretation should reference the true **mean difference** (μ_d).

Your Turn (Record your answer here.)