ECON 0150 | Spring 2025 | Homework 3.3

Homework is designed to both test your knowledge and challenge you to apply familiar concepts in new applications. Answer clearly and completely. You are welcomed and encouraged to work in groups so long as your work is your own. Submit your figures and answers to Gradescope.

Study Setup

A university runs a campus wellness program designed to reduce student anxiety. Anxiety is measured on a 0–21 scale for the same students before and after the program. For each student, we compute the paired difference:

$$x_i = \text{after}_i - \text{before}_i \tag{1}$$

A negative x_i means the student's anxiety decreased. Our aim is to study whether the program worked to reduce student anxiety.

Q1. Known μ and σ

Suppose we know the population of paired differences has mean $\mu = -3$ and standard deviation $\sigma = 5$ and we have a sample of $n = 25$ students.

a) Compute the standard error (SE) of $\bar{x}$. Then compute the probability that $\bar{x}$ falls within 1 SE of μ. *(Hint: use `stats.norm.cdf()`)*.

b) Construct a 1 SE confidence interval centered on μ. Then simulate 1,000 samples of size $n = 25$ from this population. What fraction of the 1,000 sample means fall inside the confidence interval?

```
ci_lower = mu - SE
ci_upper = mu + SE

samples = np.random.normal(loc=mu, scale=sigma, size=(1000, n)) # Samples
sample_means = samples.mean(axis=1) # Sample Mean

inside = (sample_means >= ci_lower) & (sample_means <=ci_upper) # Check in CI
mean_inside = np.mean(inside) # Compute the proportion inside
```

Q2. Testing a Model

In practice we don't know σ or μ. Instead of constructing confidence intervals centered on μ using σ, we center them on $\bar{x}$ and swap in the sample standard deviation S for σ. This swap introduces extra uncertainty which we account for using the t-distribution. Let's use this information to test whether the wellness program actually works.

a) A university runs the program on $n = 25$ students and finds $\bar{x} = -2.8$ and $S = 4.6$. State the null hypotheses for testing whether the program has any effect on anxiety.

b) Compute the probability of seeing a sample as extreme as $\bar{x}$ given the null hypothesis (*two-sided*).

```
p_value = 2 * stats.t.cdf(-abs(x_bar), df=n-1, loc=0, scale=S/np.sqrt(n))
```

c) A large national study of a similar program surveys $n = 2,000$ students and finds $\bar{x} = -0.3$ and $S = 5$. Compute the p-value.

d) A clinician considers a drop of at least 2 points on the anxiety scale to be meaningful. The first study found a 2.8-point drop. The second found a 0.3-point drop. Which result is *practically* significant? What does this tell you about the difference between statistical significance and practical significance?

Q3. Interpreting p-Values

A researcher studying the effect of a new teaching method on exam scores conducts a hypothesis test and reports a p-value of 0.032.

a) In your own words, what does this p-value tell us?

b) If the researcher used a significance level of $\alpha = 0.05$, would they reject the null hypothesis? What if they used $\alpha = 0.01$?

c) For each of the following statements, say whether it is **true or false** and explain your reasoning.

1. A p-value of 0.10 means there is a 10% chance that the null hypothesis is true.
2. If we fail to reject the null hypothesis, we have proven that it is true.
3. A small p-value indicates that our observed result would be unlikely if the null hypothesis were true.