

## In-Class Exercise 6

For this in-class exercise, work with your group of 2-3 people, to answer the following questions. These questions are not necessarily easy and sometimes they will not have a clear “correct” answer. The goal is to get you thinking about the material we’ve learned. Some of these questions may require you to discuss and debate with your group members to come up with an answer or can cover topics that we have not yet covered in class.

Be prepared to share your answers with the class and add to the discussion.

After class submit your a do-file with your answers in comments to Moodle for grading. You will be graded as a group on your submission. Only one group member needs to submit the assignment, but make sure add all group member names.

Here’s the dataset you’ll be working with for this exercise: skydrop\_moto\_data.csv

### “SkyDrop” vs. Motorcycle Fleet

You are economic consultants for the **Ministry of Health** in a developing country. The Ministry is deciding between two logistics contracts for delivering vaccines to rural clinics:

1. **SkyDrop (Drones):** High-tech, autonomous drones. They are **expensive** (\$50 per trip) but claim to be weather-independent.
2. **MotoRun (Motorcycles):** Local motorcycle couriers. They are **cheap** (\$8 per trip), but there are concerns about their reliability during the rainy season (floods).

You have been given a dataset of 500 deliveries (250 Drone, 250 Motorcycle) collected over the last rainy season. The dataset includes delivery times, costs, and weather conditions during the trip.

#### **i** Note

You don’t have to write out formulas directly for confidence intervals. Here are some Stata commands that can help you calculate the necessary statistics:

- **Confidence Interval for Mean:** `ci means <var> if <expression>`
- **Two-Sample T-Test:** `ttest <var>, by(<grouping_var>)`
- **Proportion Test:** `prtest <var>, by(<grouping_var>)`

Use `if` expressions and `by` only if needed.

### Question 1

SkyDrop justifies their high price (\$50) by claiming their drones are precision instruments that consistently deliver within 45 minutes, regardless of conditions.

Filter the data to include only **Drone** deliveries. Construct a **95% Confidence Interval** for the true population mean delivery time of drones. Does the interval suggest that SkyDrop meets their “under 45 minutes” promise?

### !Solution

First, load the data and calculate the CI for Drone delivery times.

```
import delimited "skydrop_moto_data.csv", clear
ci means delivery_time if method == "Drone"
```

**Results (approximate based on seed): - Mean:**  $\approx$  40.16 mins - **95% CI:** [39.67, 40.65] mins

**Conclusion:** Since the entire confidence interval is well below 45 minutes (the upper bound is  $\sim$ 40.65), we have strong evidence that SkyDrop **meets** their promise of delivering in under 45 minutes on average.

## Question 2

The Ministry of Health usually employs motorcycles to deliver health supplies. On a normal day, their drivers are just as fast as drones because they know the shortcuts.

Test the hypothesis that **Motorcycles are slower than Drones**.

Run this test twice:

1. Once using the **Full Dataset** (All weather conditions).
2. Once using only data where `weather == "Clear"`.

Compare the p-values. Does the motorcycle disadvantage disappear when the weather is good?

### !Solution

```
* Full Dataset
ttest delivery_time, by(method)

* Clear Weather Only
ttest delivery_time if weather == "Clear", by(method)
```

**Results:** 1. **Full Dataset:** The p-value is extremely small ( $p < 0.001$ ). Drones are significantly faster. 2. **Clear Weather:** The p-value is still extremely small ( $p < 0.001$ ).

**Conclusion:** Even in clear weather, motorcycles are statistically significantly slower than drones. The disadvantage **does not disappear** when the weather is good.

### Question 3

The Minister of Finance wants to cut costs and switch entirely to motorcycles. However, during floods, roads wash out. If a vaccine delivery takes longer than **120 minutes**, the vaccines spoil, costing the Ministry **\$1,000** in wasted medicine value.

Calculate the **Failure Rate** (proportion of trips > 120 mins) for Motorcycles during weather == "Flood".

#### ! Solution

**1. Failure Rate (Probability of > 120 mins):** In the sample, roughly 95.7% of motorcycle trips during floods took longer than 120 minutes.

$$P(\text{Fail}) \approx 0.96$$

Calculate the **Total Expected Cost per Trip** for a Motorcycle during a flood. Compare this to the flat cost of a Drone (\$50).

#### i Hint

What components will go into the expected cost calculation? What are the probabilities of each outcome? Remember that the expectation of an outcome is the sum of each possible outcome multiplied by its probability:

$$E[\text{outcome}] = \sum P(\text{outcome}) \cdot \text{outcome}$$

Based *strictly* on this cost analysis, should the Ministry pay the premium for drones during the rainy season?

#### ! Solution

#### 2. Expected Cost Calculation:

$$E[\text{Cost}] = \text{TripFee} + P(\text{Fail}) \times \text{SpoilageCost}$$

$$E[\text{Cost}] = \$8 + (0.9565 \times \$1000)$$

$$E[\text{Cost}] \approx \$8 + \$956.50 = \$964.50$$

**Comparison:** - **Drone:** \$50 flat (assuming reliable). - **Motorcycle (Flood):** ~\$964.50 expected.

**Decision:** **Yes**, the Ministry should absolutely pay the \$50 premium for drones during floods. The risk of spoilage makes the "cheap" motorcycles astronomically expensive in expected value terms.

### Question 4

Apart from time delays, physical vibration damages some vaccines. In your sample, 8% of motorcycle deliveries arrived with broken vials, compared to 2% for drones.

Test if the proportion of damaged goods is significantly lower for drones.

! Solution

We are testing  $H_0 : p_{drone} = p_{moto}$  vs  $H_A : p_{drone} < p_{moto}$ .

```
prtest broken_vials, by(method)
```

**Results:** -  $\hat{p}_{drone} \approx 0.020$  -  $\hat{p}_{moto} \approx 0.108$  - p-value  $\approx 0.00003$

**Conclusion:** The p-value is well below 0.05. We reject the null hypothesis. The proportion of damaged goods is **significantly lower** for drones compared to motorcycles.

### Question 5

The national analysis shows drones are generally faster. However, a parliament member from the “Highland Zone” argues that in their specific region, the terrain is flat enough that motorcycles are just as fast as drones, so they shouldn’t have to pay for the premium service.

The Ministry has divided the service area into 8 distinct geographic zones (Zones 1 through 8). They ask you to investigate the parliament member’s claim.

- Run an independent two-sample t-test (Drone times vs. Motorcycle times) for each of the 8 separate zones. Use a standard significance level of  $\alpha = 0.05$  for each test. Based only on these results, which zones appear to have a statistically significant difference in delivery times? Does the “Highland Zone” (Zone 5) show a significant difference?
- Explain to the Ministry the danger of running 8 separate statistical tests independently. Calculate the probability of finding at least one false positive (Type I error) among these 8 tests purely by chance, assuming there is actually no difference in any zone. (Use the formula:  $P(\text{at least one error}) = 1 - (1 - \alpha)^m$ ).
- Apply the Bonferroni Correction to address the multiple testing issue. What is your new, adjusted significance level ( $\alpha_{new}$ )? Re-evaluate your findings from Task A using this stricter threshold. Do your conclusions about any of the zones (especially Zone 5) change?

! Solution

**Naive Interpretation ( $\alpha = 0.05$ ):** Running t-tests for each zone, you might find significant differences in most zones.

**Multiple Testing Problem: - Probability of Type I Error:** With 8 tests,  $P(\text{at least one error}) = 1 - (1 - 0.05)^8 \approx 1 - 0.66 = 0.34$ . There is a **34% chance** of finding a “fake” significant result just by luck.

**Bonferroni Correction: - Adjusted Alpha:**  $\alpha_{new} = 0.05/8 = 0.00625$ .

**Re-evaluation:** With the new threshold ( $p < 0.00625$ ): - Zones like 2 and 6 (which had p-values around 0.01-0.02) are **no longer significant**. - Zone 5 remains **not significant** (supporting the MP’s claim that there is no detected speed advantage for drones in that zone). - Stronger zones (like 1, 3, 4, 7, 8) remain significant.

This highlights that while Drones are generally better, we need to be careful about checking many subgroups separately without correcting for the number of “looks” we take at the data.