WESTMINSTER
International University in Tashkent

# Week 7
# Hypothesis testing

*By*

Olmos Isakov
Office hours: Tuesday, 09:00 – 11:00 (ATB 216)



WESTMINSTER
International University in Tashkent

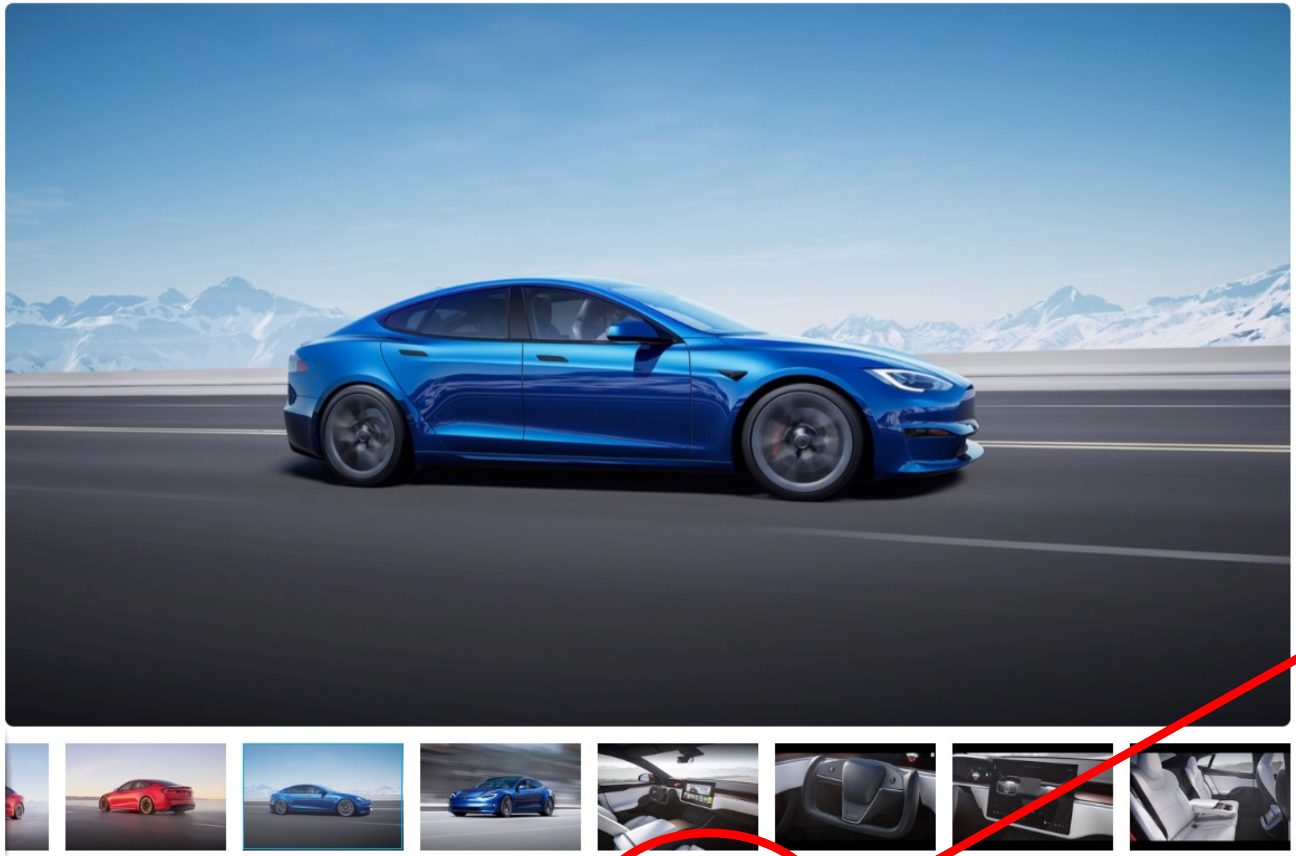
# AGENDA

- Define hypothesis testing;
- Conduct hypothesis testing for μ (population mean) and π (population proportion);
- Understand Type I and Type II errors;
- Compute and interpret the p-value in hypothesis testing.

# Tesla Model S Range

Tesla Model S Dual Motor

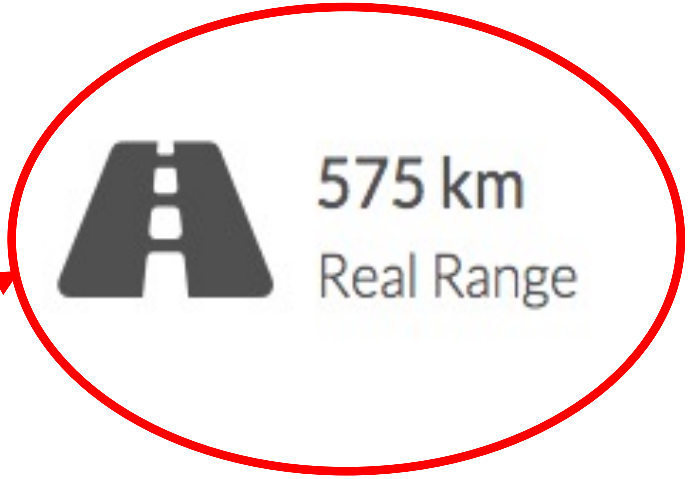
Available since January 2023



95.0 kWh *
Useable Battery

575 km
Real Range

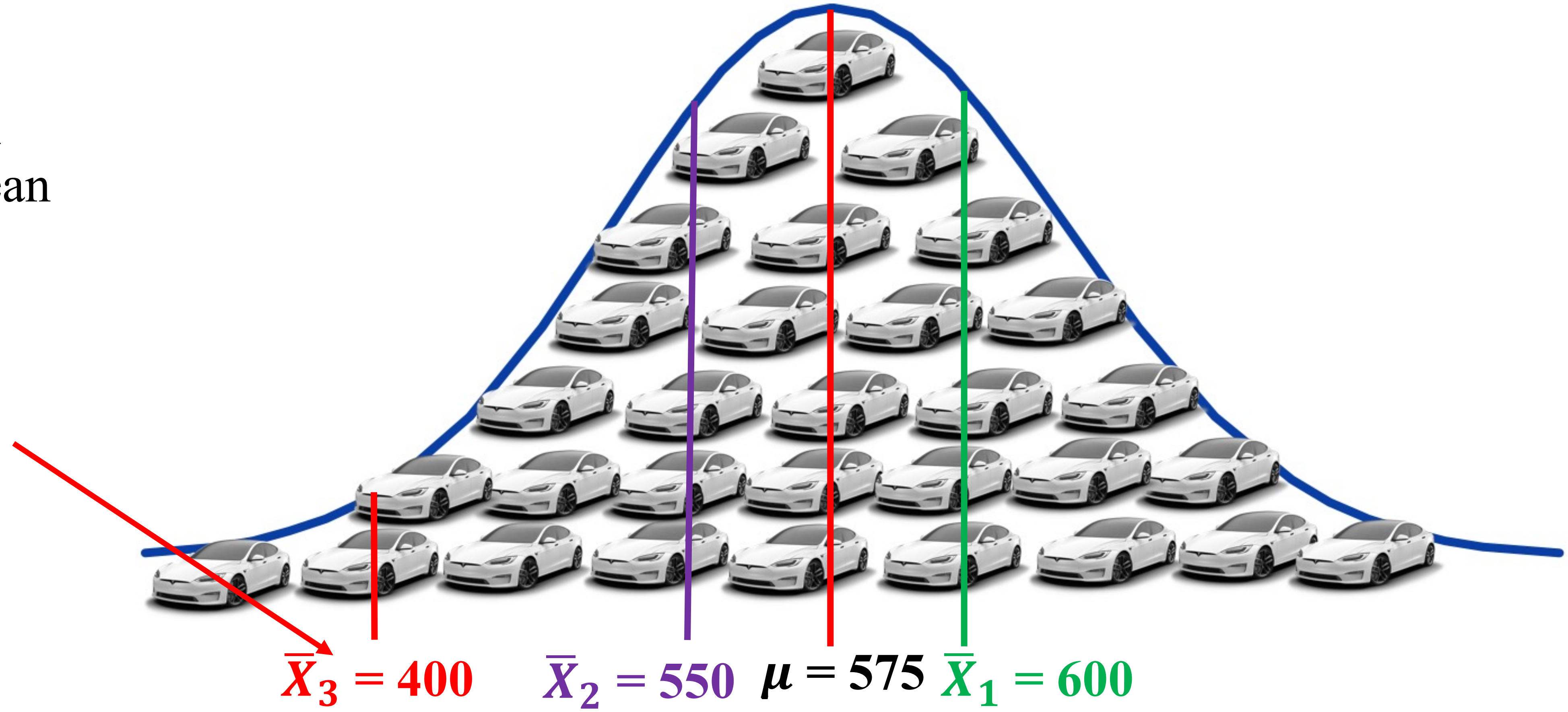
165 Wh/km
Efficiency



# Is actual range lower than 575 km?

You randomly selected 50 Tesla model S cars and found sample mean range (in km):

Unreasonable



# Hypothesis Testing

A **hypothesis** is a claim about a population parameter (such as a population proportion $\pi$ or population mean $\mu$) or some other characteristic of a population.

**Hypothesis testing** is a statistical method used to determine if there is enough evidence in a sample data to draw conclusions about a population parameter.

**The Null Hypothesis, $H_0$** (read "H-naught")**:** General statement or default position that there is nothing new happening. Refers to the Status Quo. Always contains the "=" sign.

**The Alternative Hypothesis, $H_1$ or $H_a$ :** Opposite of the null hypothesis. Challenges the Status Quo. Never contains the sign "=".

# Hypothesis testing for μ

Setting Hypotheses → Calculating test statistic → Formulating decision rule → Making decision

# Step 1: Setting hypotheses

**Case 1:** Is the average age in Uzbekistan different than 30 years?

**Two-sided (Two-tailed):** $H_0: \mu = \mu_0$ $\quad$ $H_0: \mu = 30$

$H_a: \mu \neq \mu_0$ $\quad$ $H_a: \mu \neq 30$

**Case 2:** Is the mean range of Tesla Model S less than 575 kilometers?

**Left-sided (Left-tailed):** $H_0: \mu \geq \mu_0$ $\quad$ $H_0: \mu \geq 575$

$H_a: \mu < \mu_0$ $\quad$ $H_a: \mu < 575$

**Case 3:** Is average income in Uzbekistan more than $400?

**Right-sided (Right-tailed):** $H_0: \mu \leq \mu_0$ $\quad$ $H_0: \mu \leq 400$

$H_a: \mu > \mu_0$ $\quad$ $H_a: \mu > 400$

# Step 2: calculating test statistic

**When $\sigma$ is known**

$$z_{stat} = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$$

**When $\sigma$ is unknown ($n \geq 30$)**

$$z_{stat} = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

***Example***: www.olx.uz claims mean rental cost for two-bedroom apartments is \$400 in Tashkent. A sample mean and standard deviation of 36 randomly sampled apartments are \$415 and \$55 respectively, can we claim the mean cost is higher than \$400. Calculate the test statistic.

***Solution***: **$\mu_0$ = \$400; $\bar{x}$ = \$415; n =36 > 30; s = \$55**

The population standard deviation is unknown, yet sample size is greater than 30, hence we can use sample standard deviation.

$$z_{stat} = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{415 - 400}{55/\sqrt{36}} = 1.636$$

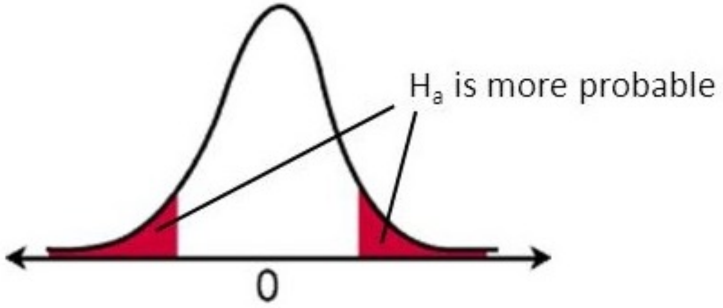
# Step 3: Formulate the decision rule (Rejection region)

**Case 1:** Is the average age in Uzbekistan different than 30 years?

**Two-sided:** ***Reject $H_0$ if:*** $|z_{stat}| > z_{\alpha/2}$ **(z critical)**

$H_0$: $\mu = \mu_0$

$H_a$: $\mu \neq \mu_0$

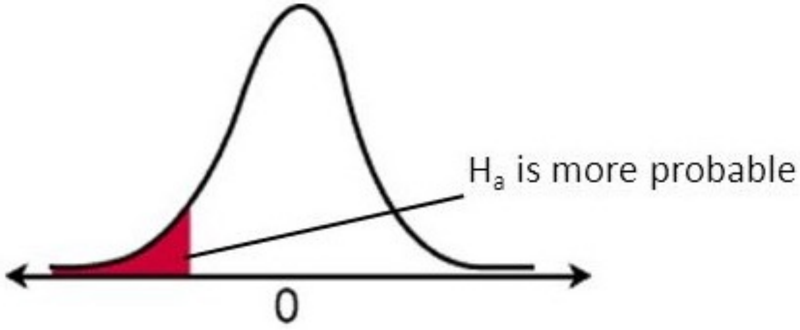



**Case 2:** Is the mean range of Tesla Model S less than 575 kilometers?

**Left-sided:** ***Reject $H_0$ if:*** $z_{stat} < -z_{\alpha}$ **(z critical)**

$H_0$: $\mu \geq \mu_0$

$H_a$: $\mu < \mu_0$

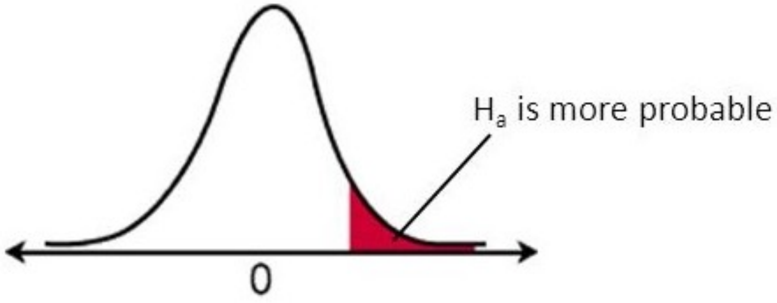



**Case 3:** Is average income in Uzbekistan is more than $400?

**Right-sided:** ***Reject $H_0$ if:*** $z_{stat} > z_{\alpha}$ **(z critical)**

$H_0$: $\mu \leq \mu_0$

$H_a$: $\mu > \mu_0$



# Step 4: Making a decision

Either ***reject Null hypothesis ($H_o$)*** or

***fail to (do not) reject Null hypothesis ($H_o$)***

based on the comparison on step 3.

**Interpret the result:** Provide your final conclusion in the context of the research question.

**Rejecting $H_0$:** There is sufficient evidence to support the claim …

**Failing to reject $H_0$:** There is **not** sufficient evidence to support the claim…

# Example:

Suppose you are looking for a two-bedroom apartment for rent. Local website www.olx.uz claims average cost of rent for two-bedroom apartments is \$400 in Tashkent. But you suspect the real number is higher than \$400. So you randomly selected 36 two-bedroom apartments from that website and calculated sample mean to be \$415 and sample standard deviation to be \$55. Use $\alpha = 0.10$.

***Conduct a hypothesis test to see whether rent rate is higher than \$400.***

# Solution: Step 1 & 2

Suppose you are looking for a two-bedroom apartment for rent. Local website www.olx.uz claims average cost of rent for two-bedroom apartments is $400 in Tashkent. But you suspect the real number is higher than $400. So you randomly selected 36 two-bedroom apartments from that website and calculated sample mean to be $415 and sample standard deviation to be $55. Use $\alpha = 0.10$. *Conduct a hypothesis test to see whether average rent rate is higher than $400.*

***Solution.***

***Step 1.*** *State null and alternative hypotheses:*

$$H_0: \mu \leq 400$$

$$H_a: \mu > 400$$

***Step 2.*** *Calculate the test statistic.*

$$z_{stat} = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{415 - 400}{55/\sqrt{36}} = 1.636$$

# Solution: Step 3 & 4

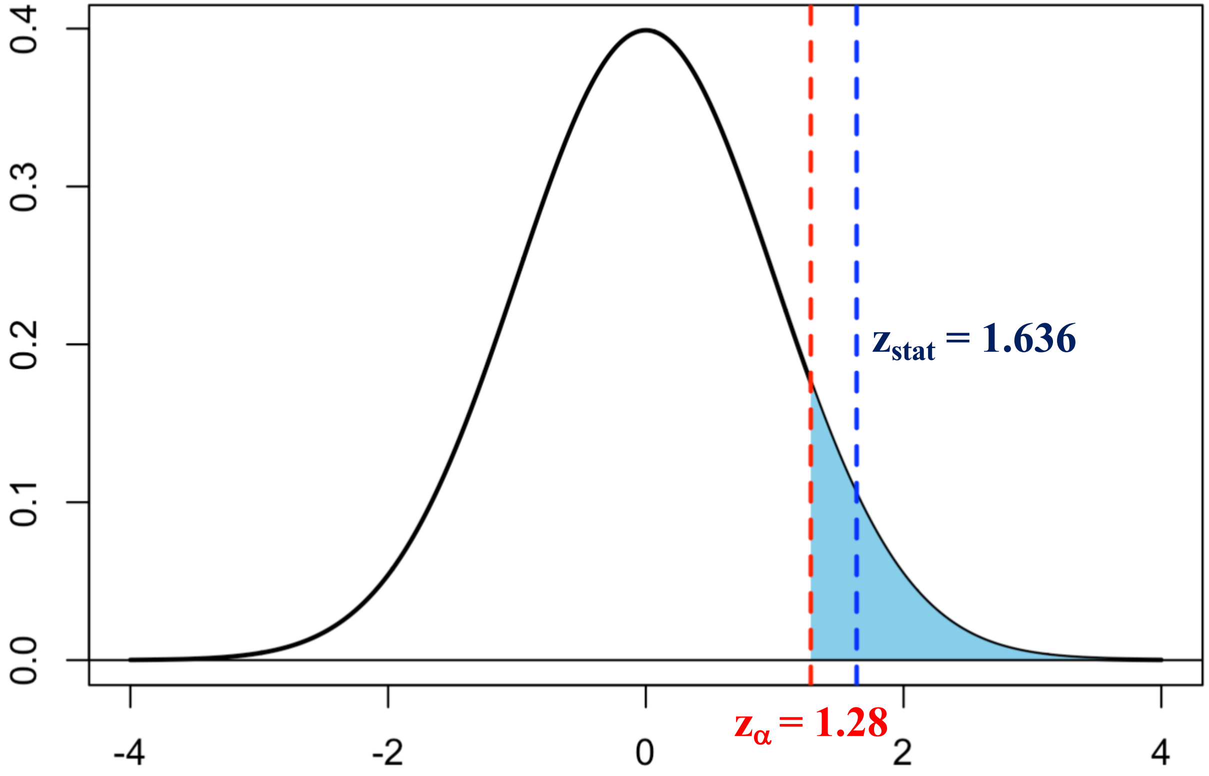
***Step 3. Formulate the decision rule.***

We reject $H_o$ if $z_{stat} > z_\alpha$ ➔ $z_\alpha = z_{0.10} =$ **1.28 (from z-table)**

***Step 4. Make a decision and interpret the result.***

Since $z_{stat}$ = **1.636** > **1.28** = $z_\alpha$, we **reject** $\mathbf{H_0}$.

**There is a sufficient evidence to claim that average rent in Tashkent is greater than $400.**



# Type I and Type II errors

| ***Null hypothesis is …*** | ***True*** | ***False*** |
|---|---|---|
| ***Not rejected*** | Correct decision<br>True negative<br>Probability = 1 - $\alpha$ | Type II error<br>False negative<br>Probability = β |
| ***Rejected*** | Type I error<br>False positive<br>Probability = α | Correct decision<br>True positive<br>Probability = 1-β |

**Type I error (*false-positive*)**: rejecting a null hypothesis that is actually true in the population;
**Type II error (*false-negative*)**: failing to reject a null hypothesis that is actually false in the population.

# Example: Fraud Detection System

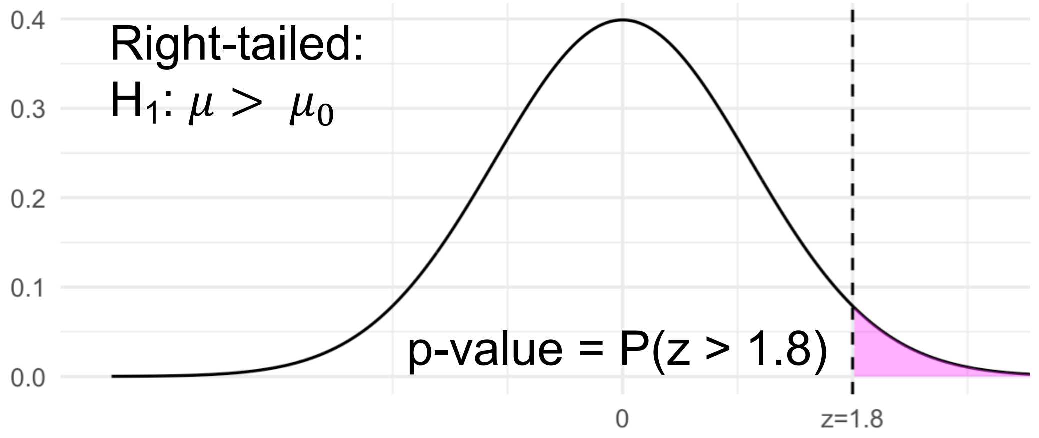

$H_0$: The transaction is *legitimate*.

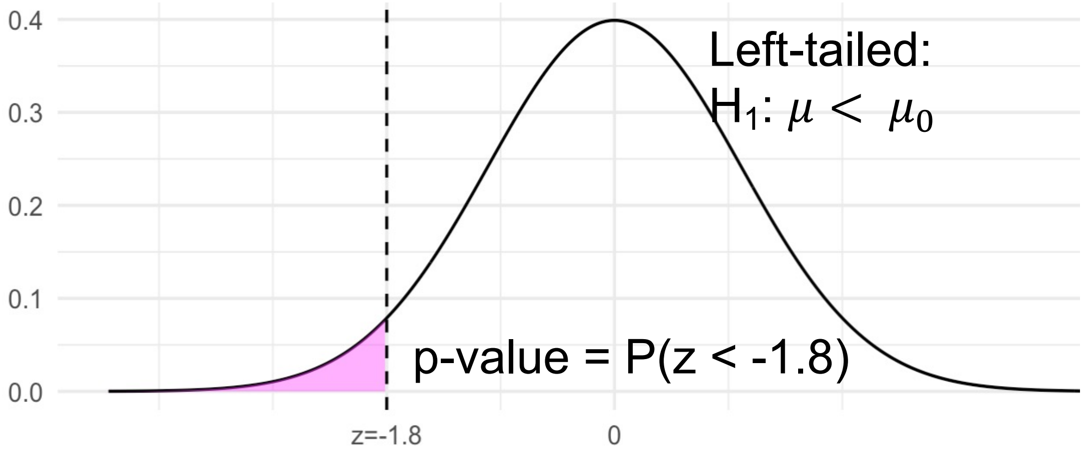
$H_a$: The transaction is *fraudulent*.

| Error Type | What It Means | Real-World Consequence |
|---|---|---|
| **Type I** | The system flags a legitimate transaction as fraud (rejects $H_0$ when it's true) | The customer's card is blocked unnecessarily, leading to inconvenience and loss of trust. |
| **Type II** | The system fails to detect an actual fraudulent transaction (fails to reject $H_0$ when it's false). | The fraud goes undetected, causing financial loss to the bank and customer. |

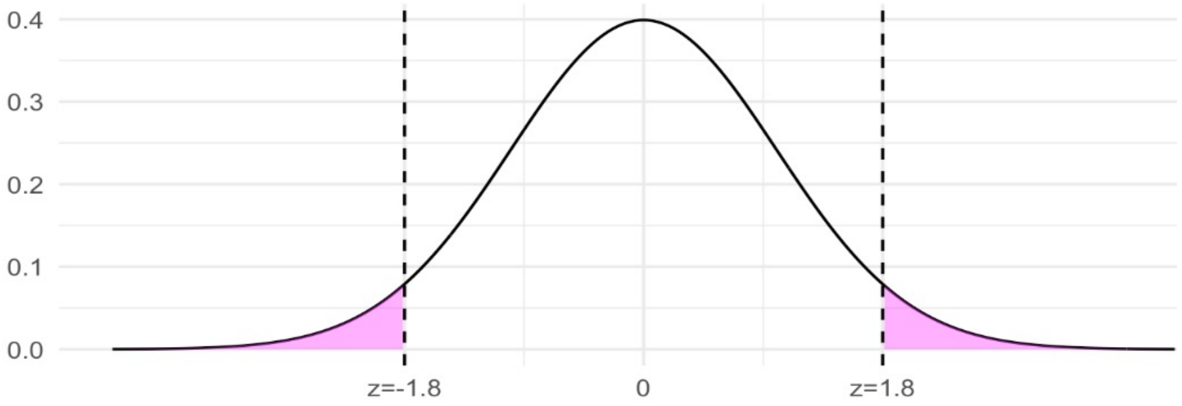
# P-value

The **p-value** is the probability of observing a test statistic as extreme as (or more extreme than) the one you got if the null hypothesis is true.




Two-tailed:
$H_1$: $\mu \neq \mu_0$




p-value = 2* P(z < -1.8) =
= 2* P(z > 1.8)

**INTERPRETING THE WEIGHT OF EVIDENCE AGAINST $H_0$**

If the *p*-value is less than

(a) .10, we have *some* evidence that $H_0$ is not true.

(b) .05, we have *strong* evidence that $H_0$ is not true.

(c) .01, we have *very strong* evidence that $H_0$ is not true.

(d) .001, we have *extremely strong* evidence that $H_0$ is not true.

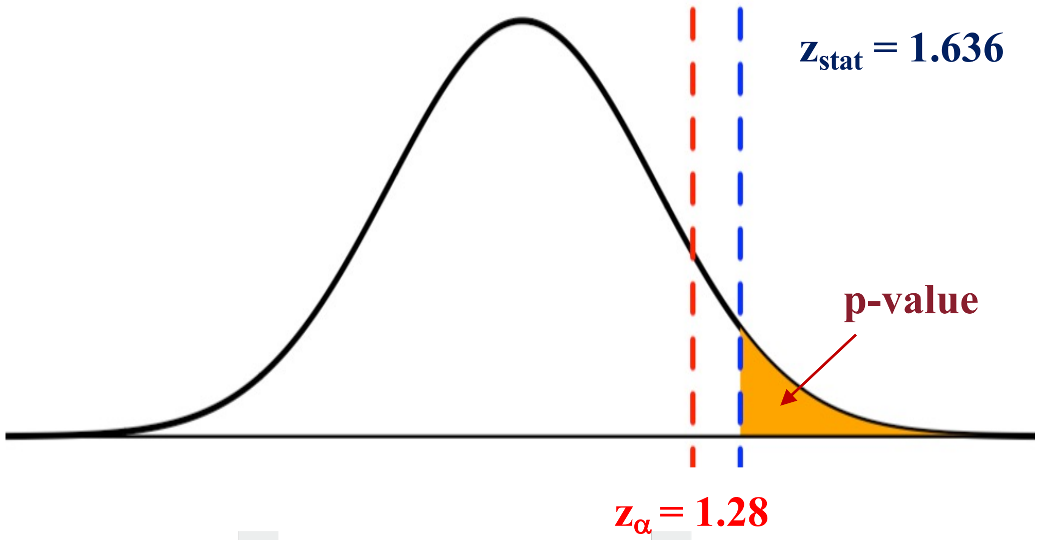
**Decision rule: Reject Ho if p-value < $\alpha$**
**This is a universal rule for many statistical tests.**

Refer to the previous example about the rental cost of two-bedroom apartments in Tashkent.

$z_{stat} = 1.636$

**p-value = P(z > 1.636) = 1 - P(z < 1.636) = 0.0505 < $\alpha$ = 0.10**

We reject the null hypothesis. The conclusion remains the same.



# Hypothesis testing for $\pi$

Setting Hypotheses → Calculating test statistic → Formulating decision rule → Making decision

# Step 1: Setting hypotheses


**Case 1:** Is Messi's penalty performance different than that of Maradona? Note, Maradona has 82.5% accuracy

**Two-tailed**

$H_0: \pi = \pi_0$ $H_0: \pi = 0.825$

$H_a: \pi \neq \pi_0$ $H_a: \pi \neq 0.825$

**Case 2:** Is Messi's penalty performance worse than that of Maradona? Note, Maradona has 82.5% accuracy

**Left-tailed**

$H_0: \pi \geq \pi_0$ $H_0: \pi \geq 0.825$

$H_a: \pi < \pi_0$ $H_a: \pi < 0.825$

**Case 3:** Is Messi's penalty performance better than that of Maradona? Note, Maradona has 82.5% accuracy

**Right-tailed** $H_0: \pi \leq \pi_0$ $H_0: \pi \leq 0.825$

$H_a: \pi > \pi_0$ $H_a: \pi > 0.825$

# Step 2: Calculating test statistic value

$$Z_{stat} = \frac{p - \pi_0}{\sqrt{\frac{\pi_0 * (1 - \pi_0)}{n}}}$$

***p*** = sample proportion

$\boldsymbol{\pi_0}$ = hypothesized population proportion under the Null hypothesis

***n*** = total sample size

***Example***: Many football fans think that Lionel Messi has worse scoring records of penalties compared to Diego Maradona. As of now, Messi has scored 111 out of 142 penalty kicks. Maradona's penalty success rate was 82.5%. Let's test whether Messi's performance in penalty taking is statistically worse than that of Maradona or just based on bad luck/chance. Use $\alpha = 0.05$.

***Solution***: **$\pi_0 = 0.825$;** **$p = 111/142 = 0.782$;** **np = 111 > 15 & nq = 31 > 15**

**since np = 111 & nq = 31 and both are greater than 15, we may assume normality**

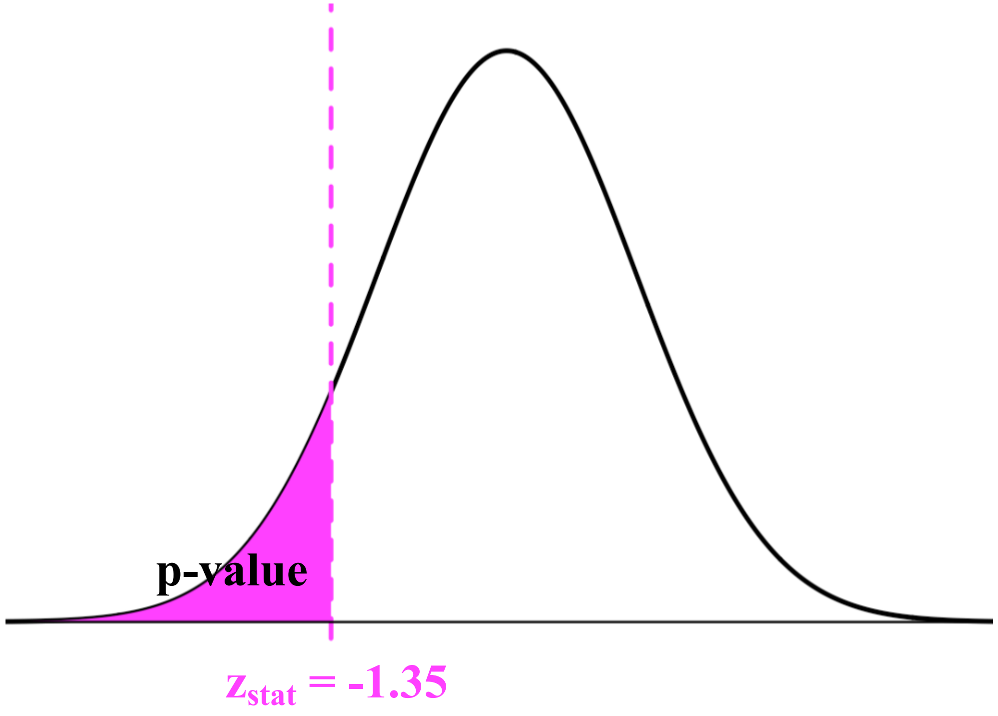
$H_0: \pi \geq 0.825$

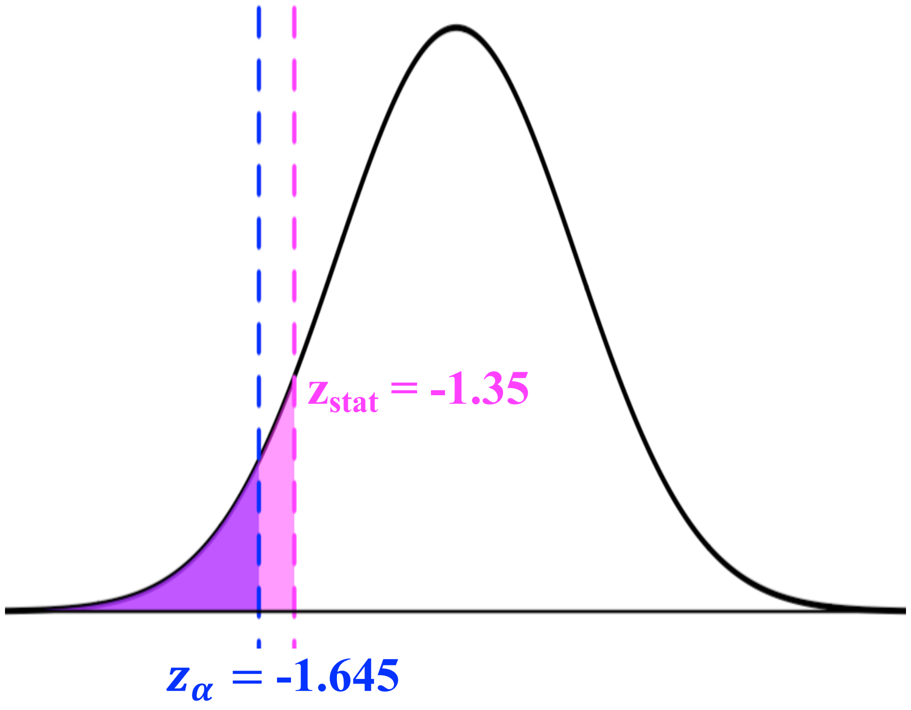
$H_a: \pi < 0.825$

$$z_{stat} = \frac{0.782 - 0.825}{\sqrt{\frac{0.825 * (1 - 0.825)}{142}}} = -1.35$$

# Step 3. Decision rule (Rejection region)

**Reject $H_0$ if: p-value < 0.05** **or** $z_{stat} < -z_{\alpha}$



# Step 4: Making a decision

- p-value = $P(z < -1.35) = P(z > 1.35) = 1 – 0.9115 = 0.0885$
- Since p-value is greater than 0.05, we fail to reject $H_0$.
- There is no evidence to conclude that Messi's penalty performance is worse than that of Maradona's.

# REFERENCES


1. Lind et al. (ISBN 978-1-260-18750-2), Chapter 9 & 10.
2. McClave & Sincich (ISBN 978-0-321-75593-3), Chapter 7 & 8.
3. Ott & Longnecker (ISBN 978-0-495-01758-5), Chapter 5 & 10.

# Thank You!