



ΕΛΛΗΝΙΚΗ ΔΗΜΟΚΡΑΤΙΑ  
ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ

# Μηχανική Μάθηση Hypothesis Testing

Γιώργος Μπορμπουδάκης  
Τμήμα Επιστήμης Υπολογιστών

# Procedure

---

1. Form the null ( $H_0$ ) and alternative ( $H_1$ ) hypothesis
2. Consider statistical assumptions
3. Find a suitable test statistic  $T$
4. Find the distribution of  $T$  under  $H_0$
5. Select significance level  $\alpha$
6. Gather related data  $D$
7. Compute the observed test statistic  $t = T(D)$
8. Compute the p-value of  $t$
9. Reject  $H_0$  if p-value  $\leq \alpha$ .



## Example: Compare means of two samples

---

- ▶ **Problem:** Given two sets of measurements  $X_1$  and  $X_2$ , decide whether they have equal means.
- ▶ **Example:** Do Greek and Italian adults have the same height?



# Example: Compare means of two samples

---

I. Form the null ( $H_0$ ) and alternative ( $H_1$ ) hypothesis

- ▶ **Null hypothesis**  $H_0: \mu_{x_1} = \mu_{x_2}$
- ▶ **Alternative hypothesis**  $H_1: \mu_{x_1} \neq \mu_{x_2}$



# Example: Compare means of two samples

---

## 2. Consider statistical assumptions

- ▶ **Assumption 1:**  $X_1$  and  $X_2$  follow a normal distribution
- ▶ **Assumption 2:**  $X_1$  and  $X_2$  have the same number of samples
- ▶ **Assumption 3:**  $X_1$  and  $X_2$  have the same variance
- ▶ **Assumption 4:** I.I.D. samples in  $X_1$  and  $X_2$



## Example: Compare means of two samples

---

3. Find a suitable test statistic T

$$t = \frac{\bar{X}_1 - \bar{X}_2}{s_{X_1 X_2} \cdot \sqrt{\frac{1}{n}}} \quad s_{X_1 X_2} = \sqrt{(s_{X_1}^2 + s_{X_2}^2)}$$

**Note:** It is positive if  $X_1$  is greater than  $X_2$  and negative otherwise.



## Example: Compare means of two samples

---

4. Find the distribution of  $T$  under  $H_0$

If all assumptions hold, then  $t$  follows a Student's  $t$  distribution with  $\nu = 2n-2$  degrees of freedom.

$$f(t) = \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\nu\pi} \Gamma(\frac{\nu}{2})} \left(1 + \frac{t^2}{\nu}\right)^{-\frac{\nu+1}{2}},$$



# Example: Compare means of two samples

---

## 5. Select significance level $\alpha$

- ▶ The significance level equals the probability of falsely rejecting the null hypothesis when it is true. This is called a type I error.
- ▶ On the other hand, not rejecting the null hypothesis when it is false is called a type II error.
- ▶ This must be specified before the analysis is done.
- ▶ Common values are 0.01 and 0.05. We use 0.05 for our example.





# Example: Compare means of two samples

---

## 6. Gather related data $D$

- ▶ For our example, we would record the height of  $N$  Greeks and Italians.
- ▶ Ideally, both should be sampled uniformly at random from their respective population, so that the samples are I.I.D.



# Example: Compare means of two samples

---

## 7. Compute the observed test statistic $t = T(D)$

- ▶ Assume that:
  - ▶  $N = 25$
  - ▶  $\text{avg}(X) = 1.75$
  - ▶  $\text{avg}(Y) = 1.70$
  - ▶ both variances equal 100 (standard deviation 10)
- ▶ For the sake of our example, assume that those are the actual averages and variances.
- ▶ Then  $t = 1.7678$



## Example: Compare means of two samples

---

### 8. Compute the p-value of $t$

- ▶  $t$  follows a Student's  $t$  distribution with 48 d.o.f.
- ▶ The p-value is the probability that of getting a more extreme test statistic, given that the null hypothesis holds.
- ▶ In our case, this means that neither is much larger than the other.

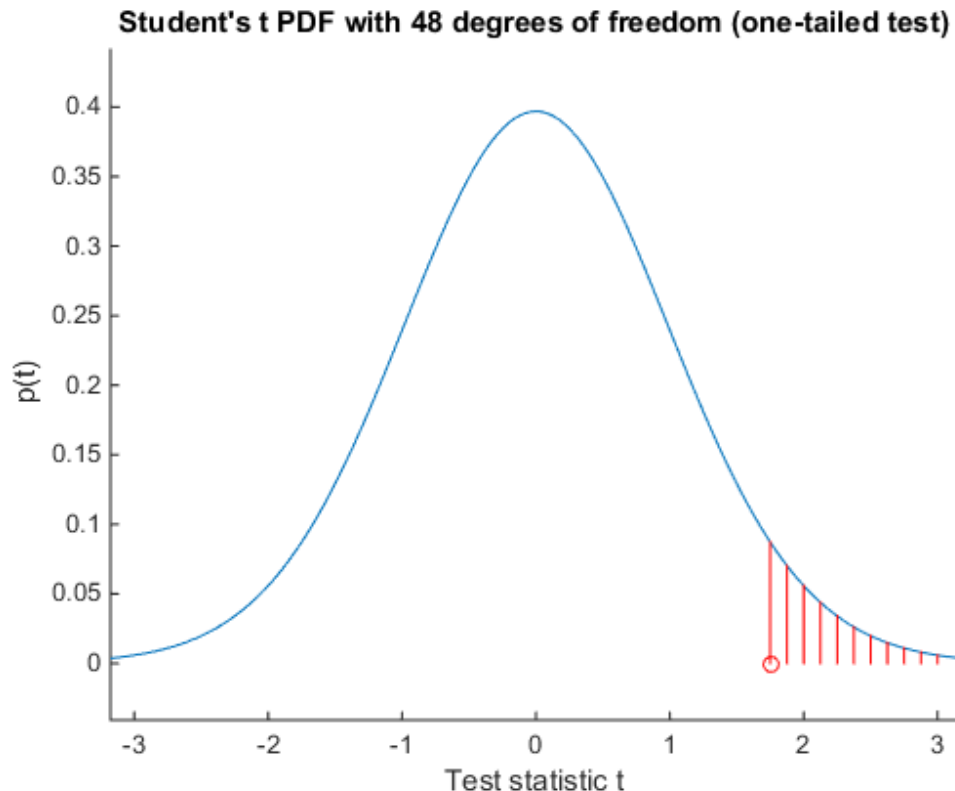


# Example: Compare means of two samples

---

## ► One-tailed test

In this case, we actually only check if  $X_1$  is greater than  $X_2$ !

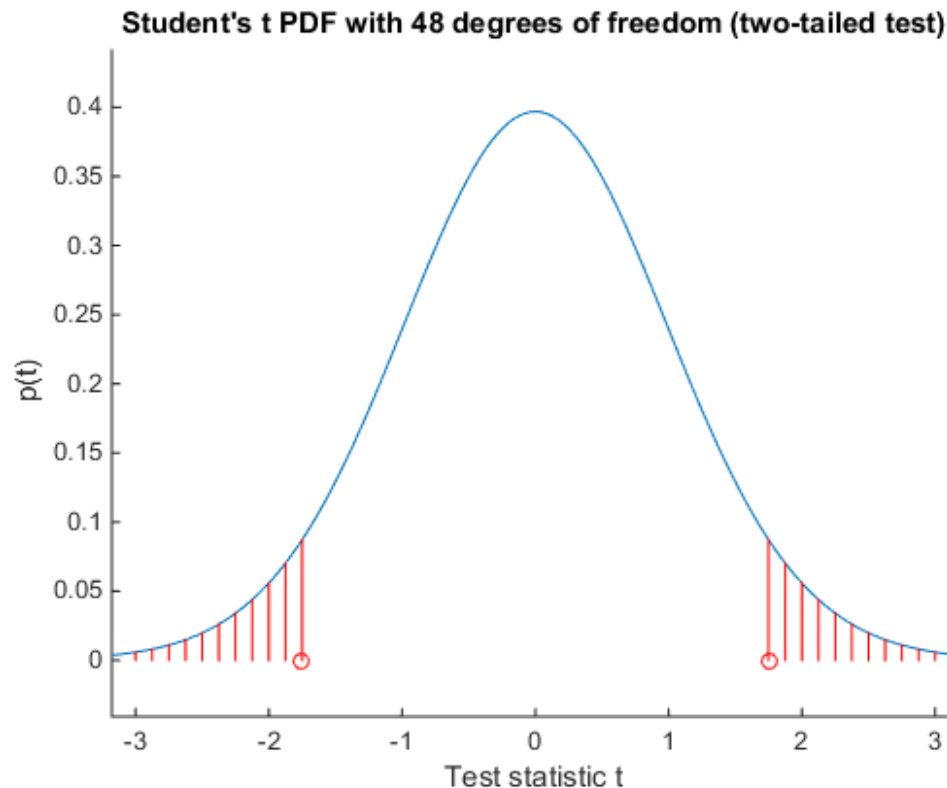


# Example: Compare means of two samples

---

## ▶ Two-tailed test

This checks that neither is larger than the other.



## Example: Compare means of two samples

---

8. Compute the p-value of t

The one-sided p-value is: 0.0417

The two-sided p-value is: 0.0835

For  $\alpha = 0.05$ , we can't reject the null hypothesis!



## Example: Compare means of two samples

---

- ▶ What happened? We know that we correctly estimated the average heights as well as the variances.
- ▶ Let's repeat the experiment and use larger values for N.
  - ▶ Assume for simplicity that we get the same estimates.

Sample Size	Test Statistic	P-value
25	1.7678	0.0835
30	1.9365	0.0577
35	2.0917	<b>0.0402</b>
40	2.2361	<b>0.0282</b>
45	2.3717	<b>0.0199</b>
50	2.5000	<b>0.0141</b>

---



## Example: Compare means of two samples

---

- ▶ We see that, as expected, we can (correctly) reject the null hypothesis with more samples.
- ▶ Of course, this can only happen if the null hypothesis is false. If it is true, we should not be able to reject it even with infinite samples.





# Τέλος Ενότητας



Ευρωπαϊκή Ένωση  
Ευρωπαϊκό Κοινωνικό Ταμείο



Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης



# Χρηματοδότηση

- Το παρόν εκπαιδευτικό υλικό έχει αναπτυχθεί στα πλαίσια του εκπαιδευτικού έργου του διδάσκοντα.
- Το έργο «**Ανοικτά Ακαδημαϊκά Μαθήματα στο Πανεπιστήμιο Κρήτης**» έχει χρηματοδοτήσει μόνο τη αναδιαμόρφωση του εκπαιδευτικού υλικού.
- Το έργο υλοποιείται στο πλαίσιο του Επιχειρησιακού Προγράμματος «Εκπαίδευση και Δια Βίου Μάθηση» και συγχρηματοδοτείται από την Ευρωπαϊκή Ένωση (Ευρωπαϊκό Κοινωνικό Ταμείο) και από εθνικούς πόρους.



Σημειώματα



# Σημείωμα αδειοδότησης (1)

- Το παρόν υλικό διατίθεται με τους όρους της άδειας χρήσης Creative Commons Αναφορά, Μη Εμπορική Χρήση, Όχι Παράγωγο Έργο 4.0 [1] ή μεταγενέστερη, Διεθνής Έκδοση. Εξαιρούνται τα αυτοτελή έργα τρίτων π.χ. φωτογραφίες, διαγράμματα κ.λ.π., τα οποία εμπεριέχονται σε αυτό και τα οποία αναφέρονται μαζί με τους όρους χρήσης τους στο «Σημείωμα Χρήσης Έργων Τρίτων».



[1] <http://creativecommons.org/licenses/by-nc-nd/4.0/>



# Σημείωμα αδειοδότησης (2)

---

- Ως **Μη Εμπορική** ορίζεται η χρήση:
  - που δεν περιλαμβάνει άμεσο ή έμμεσο οικονομικό όφελος από την χρήση του έργου, για το διανομέα του έργου και αδειοδόχο
  - που δεν περιλαμβάνει οικονομική συναλλαγή ως προϋπόθεση για τη χρήση ή πρόσβαση στο έργο
  - που δεν προσπορίζει στο διανομέα του έργου και αδειοδόχο έμμεσο οικονομικό όφελος (π.χ. διαφημίσεις) από την προβολή του έργου σε διαδικτυακό τόπο
- Ο δικαιούχος μπορεί να παρέχει στον αδειοδόχο ξεχωριστή άδεια να χρησιμοποιεί το έργο για εμπορική χρήση, εφόσον αυτό του ζητηθεί.



# Σημείωμα Αναφοράς

Copyright Πανεπιστήμιο Κρήτης, Ιωάννης Τσαμαρδίνος 2015.  
Γιώργος Μπορμπουδάκης «Μηχανική Μάθηση. Hypothesis testing». Έκδοση: 1.0. Ηράκλειο 2015. Διαθέσιμο από τη δικτυακή διεύθυνση:  
<https://opencourses.uoc.gr/courses/course/view.php?id=362>.



# Διατήρηση Σημειωμάτων

Οποιαδήποτε αναπαραγωγή ή διασκευή του υλικού θα πρέπει να συμπεριλαμβάνει:

- το Σημείωμα Αναφοράς
- το Σημείωμα Αδειοδότησης
- τη δήλωση Διατήρησης Σημειωμάτων
- το Σημείωμα Χρήσης Έργων Τρίτων (εφόσον υπάρχει)

μαζί με τους συνοδευόμενους υπερσυνδέσμους.

