Στατιστικός έλεγχος υποθέσεων (Μέρος 1ο)

Η γενική ιδέα της διαδικασίας στατιστικού ελέγχου υποθέσεων

- Πρόκειται για μια διαδικασία απόφασης μεταξύ δύο υποθέσεων
- Η μια υπόθεση ονομάζεται μηδενική (**H_0**) και η άλλη εναλλακτική (**H_1**)
- Θέτουμε ως μηδενική αυτή που αμφισβητείται και εξετάζουμε αν ένα τυχαίο δείγμα που παίρνουμε από τον πληθυσμό συνηγορεί (δίνει αποδείξεις) υπέρ της απόρριψής της έναντι της εναλλακτικής
- Έτσι, υποθέτοντας ότι η **H_0** είναι αληθής, αν αυτό που «παρατηρείται στο δείγμα» είναι ακραίο, δηλαδή, αν έχει πολύ μικρή πιθανότητα να συμβεί, τότε **απορρίπτουμε την H_0**
- Σε αντίθετη περίπτωση, δηλαδή, αν αυτό που παρατηρείται στο δείγμα δεν είναι ακραίο – σπάνιο (όταν δηλαδή είναι αληθής η **H_0**), τότε το δείγμα που πήραμε **δε μας δίνει αρκετές ενδείξεις για την απόρριψη της H_0** και «αποτυγχάνουμε να την απορρίψουμε»

Είδη σφαλμάτων

- Σφάλμα ***Τύπου I***: Η λανθασμένη απόρριψη της $\mathbf{H_0}$
- Σφάλμα ***Τύπου II***: Η λανθασμένη μη απόρριψη της $\mathbf{H_0}$

Παράδειγμα

- Το όριο αντοχής ενός τύπου καλωδίων είναι τυχαία μεταβλητή *X* με μέση τιμή μ=1500 Kg και τυπική απόκλιση σ=175 Kg
- Το εργοστάσιο που κατασκευάζει αυτόν τον τύπο καλωδίων ισχυρίζεται ότι βελτίωσε τα υλικά που χρησιμοποιεί και πλέον το όριο αντοχής των καλωδίων έχει αυξηθεί
- Έλεγχος υποθέσεων
  - **H_0: μ = 1500 Kg**
  - **H_1: μ > 1500 Kg**

Κανόνες καθορισμού H_0

1. Ως $\mathbf{H_0}$ διαλέγουμε την υπόθεση που δηλώνει ότι στον πληθυσμό η κατάσταση παραμένει αμετάβλητη (δεν υπάρχει αλλαγή / διαφορά)
2. Ως $\mathbf{H_0}$ θέτουμε την υπόθεση της οποίας η λανθασμένη απόρριψη εγκυμονεί τους περισσότερους κινδύνους

Είδη ελέγχων

- Μονόπλευροι έλεγχοι:
  1. $H_0: \mu \leq \mu_0$, $H_1: \mu > \mu_0$
  2. $H_0: \mu \geq \mu_0$, $H_1: \mu < \mu_0$
- Αμφίπλευρος έλεγχος:
  1. $H_0: \mu = \mu_0$, $H_1: \mu \neq \mu_0$

Στατιστική συνάρτηση ελέγχου

- Η στατιστική συνάρτηση την οποία επιλέγουμε για να εκφράσει «αυτό που παρατηρείται στο δείγμα»
- Συναρτήσεις που επιλέγονται ως στατιστικές συναρτήσεις ελέγχου:
  1. Ο δειγματικός μέσος, $\bar{\mathrm{X}}$
  2. Η συνάρτηση $\mathrm{Z} = \frac{\bar{\mathrm{X}}-\mu_0}{\sigma/\sqrt{n}} \sim \mathrm{N}(0,1)$
  3. Η συνάρτηση $\mathrm{T} = \frac{(\bar{\mathrm{X}}-\mu_0)\sqrt{n}}{\mathrm{S}} \sim t_{n-1}$

Επίπεδο σημαντικότητας, κρίσιμη τιμή και περιοχή απόρριψης

- Το μέγιστο αποδεκτό επίπεδο της *πιθανότητας σφάλματος Τύπου Ι* (λανθασμένης απόρριψης της $\mathbf{H_0}$), συμβολίζεται με **α** και ονομάζεται **επίπεδο σημαντικότητας** του ελέγχου
- Προκαθορίζεται, και με βάση αυτό ορίζεται η **κρίσιμη τιμή** του ελέγχου, δηλαδή, η τιμή της στατιστικής συνάρτησης ελέγχου με βάση την οποία κρίνεται αν αυτό που παρατηρείται στο δείγμα είναι ακραίο ή όχι, και επομένως αν δίνει στατιστικά σημαντικές αποδείξεις εναντίον της $\mathbf{H_0}$
- Οι τιμές της στατιστικής συνάρτησης ελέγχου για τις οποίες, σε προκαθορισμένο επίπεδο σημαντικότητας **α**, απορρίπτεται η $\mathbf{H_0}$, ορίζουν την **περιοχή απόρριψης** της $\mathbf{H_0}$
- Συνήθως το επίπεδο σημαντικότητας ορίζεται ίσο με 0,01 ή 0,05

P-τιμή

- Είναι η ελάχιστη τιμή του επιπέδου σημαντικότητας (πιθανότητα) για την οποία απορρίπτεται η $\mathbf{H_0}$
  - Αν **a ≥ P-τιμή**, τότε σε επίπεδο σημαντικότητας **a**, η $\mathbf{H_0}$ απορρίπτεται
  - Αν **a < P-τιμή**, τότε σε επίπεδο σημαντικότητας **a**, η $\mathbf{H_0}$ δεν απορρίπτεται
- Η **P-τιμή** είναι ένα μέτρο το οποίο εκφράζει πόσο ισχυρές είναι οι αποδείξεις που προκύπτουν εναντίον της $\mathbf{H_0}$ από το συγκεκριμένο δείγμα που εξετάζουμε

Στατιστικά σημαντικό δείγμα

- Όταν σε επίπεδο σημαντικότητας **α** απορρίπτεται η $\mathbf{H_0}$, το δείγμα χαρακτηρίζεται στατιστικά σημαντικό (σε επίπεδο σημαντικότητας **α**) και έχει την έννοια ότι δίνει σημαντικές αποδείξεις εναντίον της $\mathbf{H_0}$
- Όσο πιο μικρό είναι το επίπεδο σημαντικότητας στο οποίο απορρίπτεται η $\mathbf{H_0}$ και επομένως όσο πιο μικρή είναι η **P-τιμή**, τόσο πιο σημαντικό είναι το δείγμα

Ισχύς του ελέγχου

- Είναι η πιθανότητα να μην αποτύχει ο έλεγχος να απορρίψει την $\mathbf{H_0}$ όταν αληθής είναι η $\mathbf{H_1}$
- Δηλαδή, η ισχύς του ελέγχου εκφράζει την ικανότητα του ελέγχου να απορρίπτει σωστά την $\mathbf{H_0}$
- Συμβολίζεται με **1-β**, όπου **β** η *πιθανότητα σφάλματος Τύπου ΙΙ*
- Δηλαδή
  - **β = P(μη απόρριψη της H_0 | αληθής η H_1)**
  - **1-β = P(απόρριψη της H_0 | αληθής η H_1)**

Στατιστικός έλεγχος υποθέσεων σε επίπεδο σημαντικότητας **α** για τη μέση τιμή **μ** ενός πληθυσμού με ένα τυχαίο δείγμα μεγέθους **n**

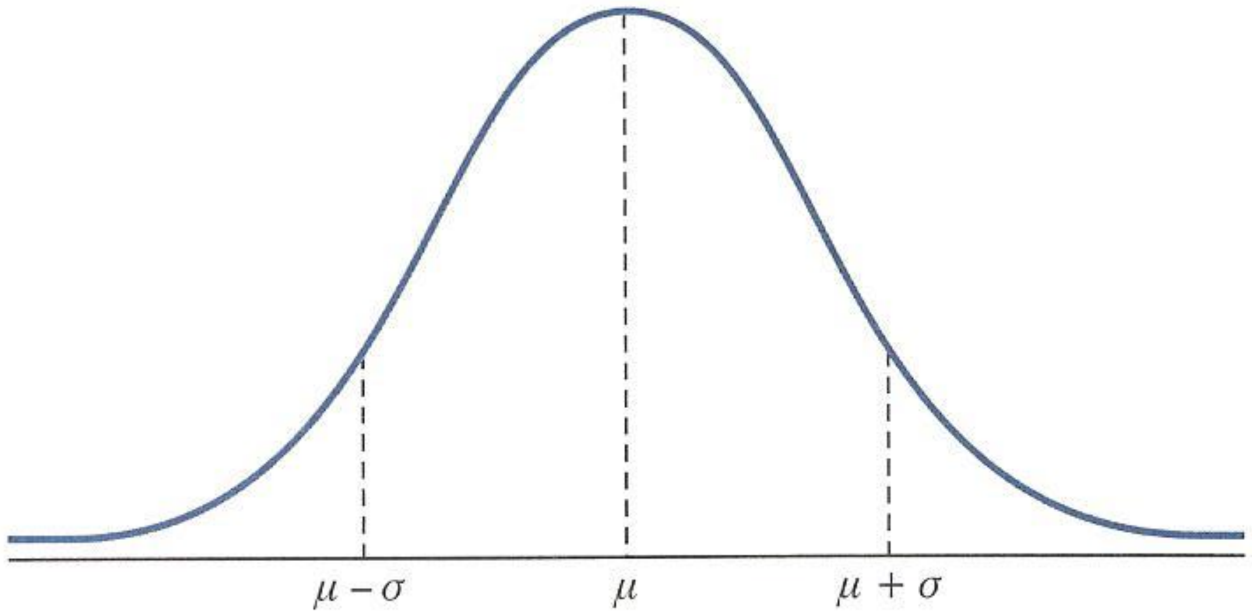
H_0: μ = μ_0

- Η διακύμανση **σ^2** είναι γνωστή και ο πληθυσμός είναι κανονικός (οτιδήποτε μέγεθος πληθυσμού)
- Η διακύμανση **σ^2** είναι γνωστή και το μέγεθος του πληθυσμού **n** μεγάλο (>30) (πληθυσμός οτιδήποτε) [*έλεγχος επιπέδου σημαντικότητας **α** κατά προσέγγιση*]

Περιοχή απόρριψης της H_0

H_1: μ ≠ μ_0	H_1: μ > μ_0	H_1: μ < μ_0
$\lvert Z\rvert = \dfrac{\lvert\bar{X} - \mu_0\rvert}{\sigma/\sqrt{n}} \geq z_{a/2}$	$Z = \dfrac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \geq z_a$	$Z = \dfrac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \leq -z_a$

Γραφική παράσταση της κανονικής κατανομής



Οι θέσεις των σημείων καμπής και της κορυφής της κανονικής καμπύλης

Στατιστικός έλεγχος υποθέσεων σε επίπεδο σημαντικότητας **α** για τη μέση τιμή **μ** ενός πληθυσμού με ένα τυχαίο δείγμα μεγέθους **n**

H_0: μ = μ_0

- Η διακύμανση **σ^2** είναι άγνωστη και το μέγεθος του πληθυσμού **n** μεγάλο (οτιδήποτε πληθυσμός) [έλεγχος *επιπέδου σημαντικότητας **α** κατά προσέγγιση*]

Περιοχή απόρριψης της H_0

H_1: μ ≠ μ_0	H_1: μ > μ_0	H_1: μ < μ_0
$\lvert Z \rvert = \frac{\lvert \bar{X} - \mu_0 \rvert}{S/\sqrt{n}} \geq z_{a/2}$	$Z = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} \geq z_a$	$Z = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} \leq -z_a$

Η τ.μ. $\frac{(\bar{X}-\mu)\sqrt{n}}{S/\sqrt{n}}$, για n ≥ 30, προσεγγίζεται ικανοποιητικά από την τυποποιημένη κανονική κατανομή, N(0,1)

Στατιστικός έλεγχος υποθέσεων σε επίπεδο σημαντικότητας **α** για τη μέση τιμή **μ** ενός πληθυσμού με ένα τυχαίο δείγμα μεγέθους **n**

H_0: μ = μ_0

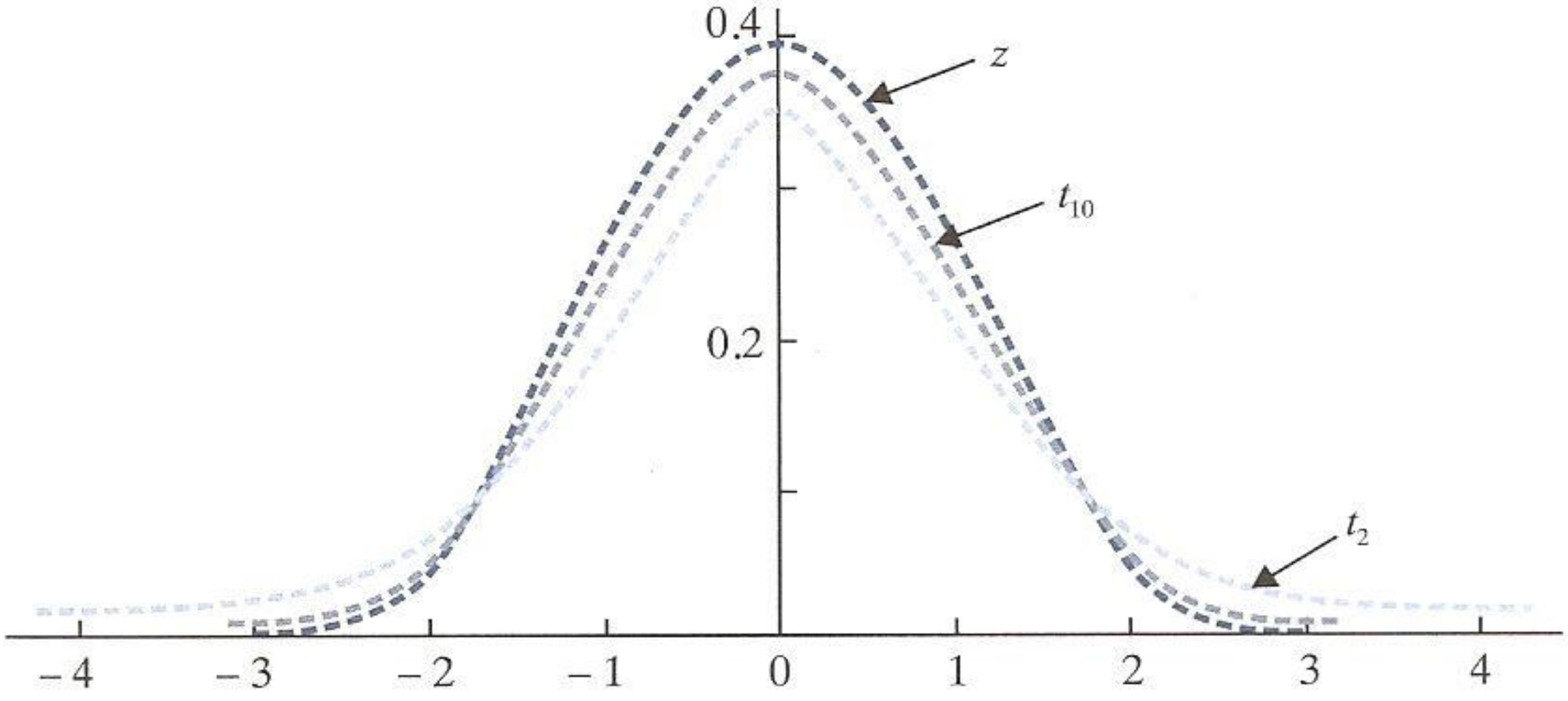
- Η διακύμανση **σ^2** είναι άγνωστη και ο πληθυσμός είναι κανονικός (οτιδήποτε μέγεθος πληθυσμού)

Περιοχή απόρριψης της H_0

H_1: $\mu \neq \mu_0$	H_1: $\mu > \mu_0$	H_1: $\mu < \mu_0$
$\lvert \mathrm{T} \rvert = \dfrac{\lvert \bar{\mathrm{X}} - \mu_0 \rvert}{S/\sqrt{n}} \geq t_{n-1;a/2}$	$\mathrm{T} = \dfrac{\bar{\mathrm{X}} - \mu_0}{S/\sqrt{n}} \geq t_{n-1;a}$	$\mathrm{T} = \dfrac{\bar{\mathrm{X}} - \mu_0}{S/\sqrt{n}} \leq -t_{n-1;a}$

Η τ.μ. $\frac{(\bar{\mathrm{X}}-\mu)\sqrt{n}}{S/\sqrt{n}}$, για κανονικό πληθυσμό, προσεγγίζεται ικανοποιητικά από την t-κατανομή με n-1 βαθμούς ελευθερίας

Η κατανομή *t* (ή Student) με *n* βαθμούς ελευθερίας – καμπύλη



Η συνάρτηση πυκνότητας της t_v για $v = 2$ και $v = 10$ και της $N(0, 1)$.

Στατιστικός έλεγχος υποθέσεων σε επίπεδο σημαντικότητας **α** για τη μέση τιμή **μ** ενός πληθυσμού με ένα τυχαίο δείγμα μεγέθους **n**

H_0: μ = μ_0

- Το t-test προϋποθέτει ότι ο πληθυσμός είναι κανονικός
- Στην πράξη εντούτοις αποδεικνύεται «ανθεκτικό» σε αυτή την υπόθεση
- Δηλαδή, το επίπεδο σημαντικότητας του ελέγχου είναι κοντά στο **α** ακόμα και αν η υπόθεση της κανονικότητας του πληθυσμού δεν ικανοποιείται όταν:
  1. Η κατανομή του πληθυσμού δεν απέχει δραματικά από την κανονική κατανομή
  2. Το μέγεθος του δείγματος δεν είναι πολύ μικρό

Στατιστικός έλεγχος υποθέσεων σε επίπεδο σημαντικότητας **α** για τη μέση τιμή **μ** ενός πληθυσμού με ένα τυχαίο δείγμα μεγέθους **n**

H_0: μ = μ_0

- Όταν το **n** είναι μικρό, ο πληθυσμός όχι κανονικός και η διακύμανση **σ^2** είτε γνωστή είτε άγνωστη δεν μπορούμε να πραγματοποιήσουμε στατιστικό έλεγχο υποθέσεων σε επίπεδο σημαντικότητας **α**

Στατιστικός έλεγχος υποθέσεων σε επίπεδο σημαντικότητας **α** για το διωνυμικό ποσοστό **p** με ένα τυχαίο δείγμα μεγέθους **n**

H_0: p = p_0

- $np_0 \geq 5$ και $n(1-p_0) \geq 5$
- $\hat{p}$ είναι το ποσοστό επιτυχιών στο δείγμα

Περιοχή απόρριψης της H_0

H_1: p ≠ p_0	H_1: p > p_0	H_1: p < p_0
$\lvert z \rvert = \frac{\lvert \hat{p}-p_0 \rvert/\sqrt{n}}{\sqrt{p_0(1-p_0)}} \geq z_{a/2}$	$z = \frac{(\hat{p}-p_0)/\sqrt{n}}{\sqrt{p_0(1-p_0)}} \geq z_a$	$z = \frac{(\hat{p}-p_0)/\sqrt{n}}{\sqrt{p_0(1-p_0)}} \leq -z_a$

Παράδειγμα 1

- Στον παρακάτω Πίνακα φαίνεται για κάθε μια από 50 τυχαία επιλεγμένες γαλακτοπαραγωγές αγελάδες ο χρόνος (σε μήνες) από την πρώτη εκδήλωση μιας συγκεκριμένης ασθένειας που προσβάλλει τις αγελάδες μέχρι την επανεμφάνισή της

2,1	**1,7**	**0,8**	**0,8**	**4,1**	**8,7**	**1,4**	**2,9**	**1,9**	**2,7**
4,4	**2,2**	**5,5**	**7,0**	**1,8**	**0,2**	**1,0**	**0,9**	**4,0**	**0,7**
2,0	**6,5**	**0,7**	**4,3**	**0,2**	**5,6**	**2,4**	**1,4**	**1,3**	**1,2**
0,5	**3,9**	**7,4**	**3,3**	**8,8**	**0,3**	**2,0**	**5,7**	**0,8**	**2,6**
9,9	**1,6**	**2,8**	**1,0**	**0,6**	**1,3**	**0,8**	**5,9**	**0,9**	**0,4**

- Από τη σχετική βιβλιογραφία είναι γνωστό ότι ο μέσος χρόνος μέχρι την επανεμφάνιση της συγκεκριμένης ασθένειας είναι 3 μήνες
- Μήπως τα συγκεκριμένα δεδομένα δίνουν σε επίπεδο σημαντικότητας 5%, στατιστικά σημαντικές αποδείξεις ότι ο μέσος χρόνος επανεμφάνισης της ασθένειας δεν είναι 3 μήνες αλλά λιγότερο;

Παράδειγμα 2

- Το μέσο ύψος των κατοίκων μιας χώρας (ενήλικων αρρένων) είναι μ_0 **= 177cm**
- Σε μια περιοχή της χώρας αυτής όμως οι κάτοικοι ισχυρίζονται ότι έχουν διαφορετικό μέσο ύψος **μ** από τους υπόλοιπους κατοίκους της χώρας
- Προκειμένου να ελεγχθεί αν ο ισχυρισμός αυτός αληθεύει επιλέχθηκε τυχαίο δείγμα **20** ανδρών από την περιοχή αυτή και μετρήθηκαν τα ύψη:

176,3	**172,5**	**179,2**	**185,7**	**173,9**
179,8	**172,9**	**178,7**	**176,5**	**184,0**
178,8	**178,7**	**178,6**	**178,2**	**182,9**
178,5	**186,5**	**175,3**	**172,8**	**182,3**

- Χρησιμοποιώντας το SPSS να ελέγξετε σε επίπεδο σημαντικότητας **α = 0.05** αν ισχύει ο παραπάνω ισχυρισμός