Έλεγχοι Υποθέσεων

1. Εισαγωγή

- Ο έλεγχος υποθέσεων αναφέρεται στις ιδιότητες μιας άγνωστης παραμέτρους του πληθυσμού:
  - ✓ Ο κατηγορούμενος είναι αθώος
  - ✓ $\mu = 100$
- Κάθε υπόθεση συνοδεύεται από μια εναλλακτική:
  - ✓ Ο κατηγορούμενος είναι ένοχος
  - ✓ $\mu \neq 100$
- Μια υπόθεση μπορεί να είναι αληθινή ή όχι (απορρίπτεται ή δεν απορρίπτεται) και συνεπώς λαμβάνεται απόφαση με βάση των δεδομένων του δείγματος.
- Μια υπόθεση παραμένει να είναι αληθής, μέχρι τη στιγμή που λαμβάνεται η απόφαση να απορριφθεί
- Η απόρριψη ή όχι μιας υπόθεσης μπορεί να είναι σωστή ή όχι
  - ✓ Λάθη τύπου Ι και ΙΙ.
- Η **αρχική υπόθεση**, συμβολίζεται με Η0, είναι μια υπόθεση για μία ή περισσότερες παραμέτρους του πληθυσμού. Η υπόθεση αυτή θεωρούμε ότι ισχύει μέχρις ότου έχουμε επαρκή στατιστικά ευρήματα για να αποφασίσουμε να την απορρίψουμε.
  - ✓ Η0: $\mu = 100$
- Η **εναλλακτική υπόθεση**, συμβολίζεται με Η1, είναι η υπόθεση που καλύπτει όλες τις άλλες περιπτώσεις που δεν συμπεριλαμβάνονται στην αρχική.
  - ✓ Η1: $\mu \neq 100$
- Η0 και Η1 είναι:
  - ✓ Αμοιβαία αποκλειόμενες
    - Μόνο μία μπορεί να είναι ¨αληθινή¨.
  - ✓ Συμπληρωματικές
    - Ο συνδυασμός τους καλύπτουν <u>όλα</u> τα πιθανά ενδεχόμενα και επομένως είτε η μία είτε η άλλη θα είναι ¨αληθινή¨.
- Η **αρχική υπόθεση**:
  - ✓ Αντιπροσωπεύει την αντίληψη μας για το τι ¨επικρατεί¨ στον πληθυσμό.

- ✓ Θεωρείται ότι είναι ¨αληθινή¨ μέχρι να αποδείξουμε το αντίθετο χρησιμοποιώντας μια **στατιστική συνάρτηση (συνάρτηση ελέγχου)** που θα μας οδηγήσει να απορρίψουμε την αρχική.

- Η **στατιστική ελέγχου (test statistic)** είναι απλώς μια συνάρτηση που υπολογίζεται από τα δεδομέναis. Με βάση τη τιμή που λαμβάνει, αποφασίζουμε αν θα απορρίψουμε ή όχι την αρχική υπόθεση.
- Ο **κανόνας απόφασης (decision rule)** μιας στατιστικής υπόθεσης είναι ένας κανόνας που καθορίζει τις συνθήκες κάτω από τις οποίες η αρχική υπόθεση μπορεί να απορριφθεί.
- Για παράδειγμα, έστω ότι H0: μ = 100. Θα πρέπει να δημιουργήσουμε ένα κανόνα που να βασίζεται στην εξής λογικής:¨Απέρριψε την H0 αν ο <u>δειγματικός</u> μέσος είναι μικρότερος από την τιμή 95 ή μεγαλύτερος από την τιμή 105¨
- Υπάρχουν δύο εναλλακτικές:
  - ✓ H0 είναι *αληθινή*
  - ✓ H0 είναι *λανθασμένη*
- Υπάρχουν δύο πιθανές αποφάσεις:
  - ✓ *Απορρίπτω την H0*
  - ✓ *Δεν απορρίπτω H0*
- Μια απόφαση μπορεί να είναι σωστή με δύο τρόπους:
  - ✓ Δεν απορρίπτω μια αληθινή H0
  - ✓ Απορρίπτω μια λανθασμένη H0
- Μια απόφαση μπορεί να είναι λανθασμένη με δύο τρόπους :
  - ✓ Λάθος Τύπου I: Απορρίπτω μια αληθινή H0
    - Η πιθανότητα του σφάλματος I συμβολίζεται με α.
  - ✓ Λάθος Τύπου II: Δεν απορρίπτω μια λανθασμένη H0
    - Η πιθανότητα του σφάλματος II συμβολίζεται με β
- Η τελική απόφαση μπορεί να είναι λανθασμένη με δύο τρόπους:
  - ✓ Λάθος τύπου I: Απορρίπτω την H0 όταν είναι αληθινή
    - Η πιθανότητα λάθος τύπου I ορίζεται με το α.
    - α ονομάζεται **επίπεδο σημαντικότητας (level of significance)**
  - ✓ Λάθος τύπου II: δέχομαι μια μη αληθινή H0

- Η πιθανότητα λάθος τύπου ΙΙ ορίζεται με το β.
- 1 - β ονομάζεται δύναμη (power) του τεστ.

	Κατάσταση	
Απόφαση	**Ho: Αληθινή**	**H1: Αληθινή**
Δεν απορρίπτω Ho	Σωστή: Πιθανότητα =1-α	Λάθος τύπου ΙΙ (β)
Απορρίπτω Ho	Λάθος τύπου Ι (α) Πιθανότητα = α	Σωστή Πιθανότητα =1-β

Η **δύναμη (power)** ενός στατιστικού ελέγχου είναι η πιθανότητα να απορρίψουμε την αρχική υπόθεση όταν αυτή είναι λανθασμένη.

Power = (1 - β)

- **Επηρεάζεται από την απόσταση μεταξύ της τιμής της παραμέτρου κάτω από την αρχική υπόθεση και της πραγματικής τιμής του πληθυσμού:** ***όσο μεγαλύτερη είναι η απόσταση, τόσο μεγαλύτερη είναι η δύναμη του τεστ.***
- **Επηρεάζεται από την τυπική απόκλιση του πληθυσμού:** ***όσο μικρότερη είναι, τόσο μεγαλύτερη είναι η δύναμη του τεστ.***
- **Επηρεάζεται από το μέγεθος του δείγματος:** ***όσο μεγαλύτερο είναι, τόσο μεγαλύτερη είναι η δύναμη του τεστ.***

2. Η τιμή *p*-Value

Η τιμή ***p-value*** είναι η πιθανότητα η στατιστική συνάρτηση να λάβει μια ακραία τιμή, ή μεγαλύτερη από αυτήν, όταν η αρχική υπόθεση είναι αληθινή.

Η τιμή p-value είναι το μικρότερο επίπεδο σημαντικότητας, α, στο οποίο η αρχική υπόθεση μπορεί να απορριφθεί.

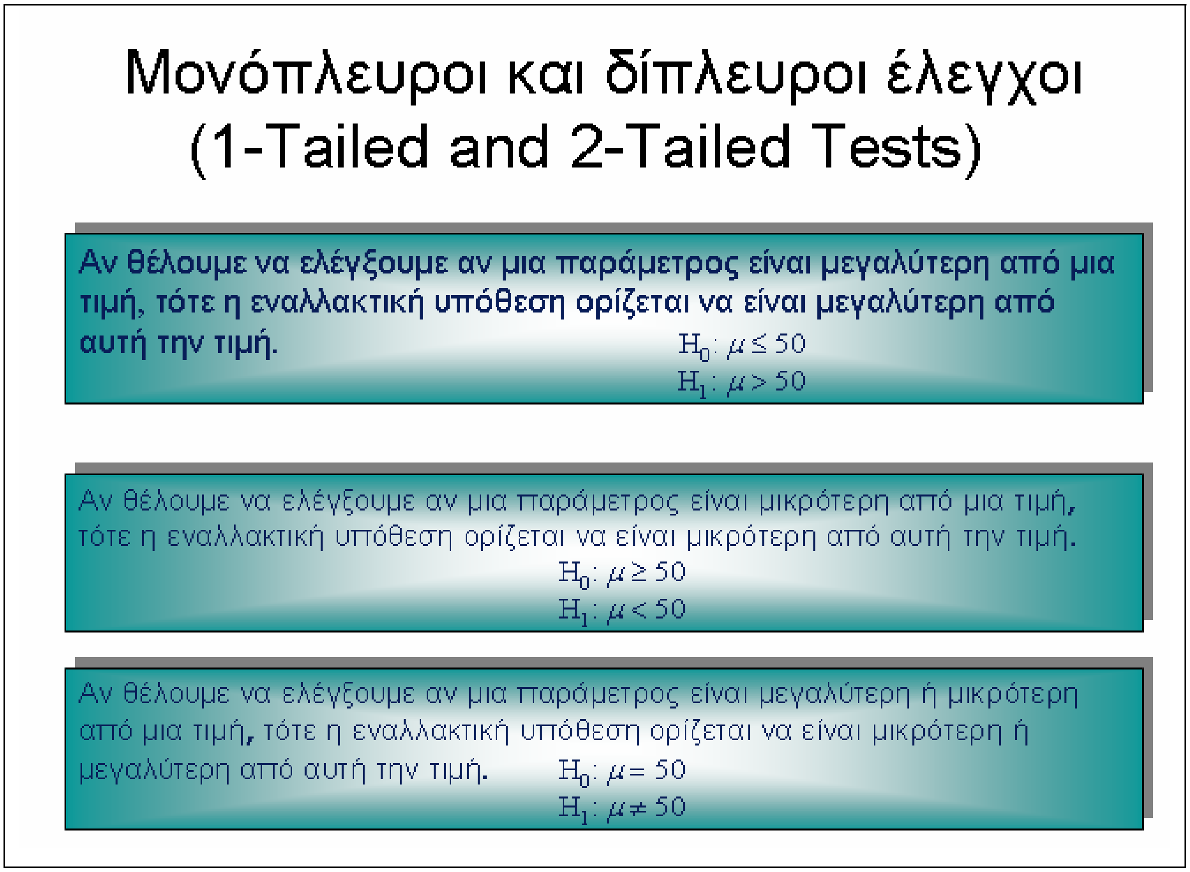
Κανόνας: Όταν η τιμή *p*-value είναι μικρότερη από a , απέρριψε την H0.

- Δεν υπάρχει γενικά αποδεκτός κανόνας που να οδηγεί στην επιλογή του επιπέδου σημαντικότητας στα προβλήματα ελέγχου στατιστικών υποθέσεων.
  - Οι δυσκολίες αυτές οδήγησαν τους Στατιστικούς να δημιουργήσουν την *p*-Value τιμή.
  - Η τιμή αυτή είναι το λεγόμενο παρατηρούμενο επίπεδο σημαντικότητας.
- **Ορισμός**: Ορίσουμε ως παρατηρούμενο επίπεδο σημαντικότητας την πιθανότητα η στατιστική συνάρτηση ελέγχου να πάρει μια τιμή τόσο ακραία ή περισσότερο ακραία από αυτήν που πήρε για το συγκεκριμένο δείγμα, κάτω από την μηδενική υπόθεση.
- **Η χρησιμοποίηση της** *p*-Value τιμής αντί του επιπέδου σημαντικότητας α στην αντιμετώπιση ενός προβλήματος στατιστικού ελέγχου υποθέσεως δεν μεταβάλλει την κλασική στατιστική μεθοδολογία. Απλά εκείνο που συμβαίνει είναι ότι ο ερευνητής αναφέρει στον ενδιαφερόμενο την τιμή *p*-Value και αφήνει την επιλογή του κατά πόσο θα πρέπει να απορριφθεί ή όχι η μηδενική υπόθεση στον ενδιαφερόμενο.
  - Για την αποφυγή παρανοήσεων, θα πρέπει να τονισθεί ότι η τιμή *p*-Value **δεν είναι** η πιθανότητα ότι η μηδενική υπόθεση είναι σωστή. Στην κλασική στατιστική θεωρία δεν υπάρχει τρόπος να προσδιορίσουμε την πιθανότητα να είναι σωστή η μηδενική υπόθεση. Η μηδενική υπόθεση θα είναι πάντα σωστή ή λάθος.
  - Αυτό που παρέχει η τιμή *p*-Value είναι η πιθανότητα να βρούμε ενδείξεις αντίθετες με την μηδενική υπόθεση τόσο ισχυρές από αυτές που έχουμε διαθέσιμες αν η μηδενική υπόθεση ίσχυε.

3. Μονόπλευρες και δίπλευρες εναλλακτικές υποθέσεις

- Όταν η μηδενική υπόθεση έχει τη μορφή Ho:$\theta \leq \theta_0$, τότε η εναλλακτική έχει τη μορφή $H_1 :\ \theta > \theta_0$. Η εναλλακτική αυτή ονομάζεται μονόπλευρή προς τα δεξιά.
- Όταν η μηδενική υπόθεση έχει τη μορφή Ho:$\theta \geq \theta_0$, τότε η εναλλακτική έχει τη μορφή $H_1 :\ \theta < \theta_0$. Η εναλλακτική αυτή ονομάζεται μονόπλευρή προς τα αριστερά.

- Όταν η μηδενική υπόθεση έχει τη μορφή Ho:$\theta = \theta_0$, τότε η εναλλακτική έχει τη μορφή $H_1 : \theta \neq \theta_0$. Η εναλλακτική αυτή ονομάζεται μονόπλευρή προς τα αριστερά.



4. Γενική μορφή της στατιστικής ελέγχου και ο τρόπος ελέγχου

Τα βήματα που πρέπει να ακολουθήσουμε όταν ελέγχουμε μία υπόθεση είναι τα εξής:

1. Διατύπωση των υποθέσεων Ho και H1
2. Επιλογή του επιπέδου σημαντικότητας α και της κατάλληλης στατιστικής ελέγχου
3. Προσδιορισμός του κανόνα αποφάσεως και συνεπώς των περιοχών απόρριψης ή μη απορρίψεως της αρχικής υποθέσεως με βάση την τιμή του α και της μορφής της H1
4. Υπολογισμός της τιμής της στατιστικής συνάρτησης ελέγχου από το δείγμα. Αν η τιμή αυτή βρίσκεται στην περιοχή απορρίψεως, τότε η H0 απορρίπτεται.

Εκτός από τις περιπτώσεις ελέγχου υποθέσεων για τις διακυμάνσεις, η στατιστική συνάρτηση ελέγχου έχει την εξής μορφή:

$$TS = \frac{\hat{\theta} - \theta_0}{se\left(\hat{\theta}\right)}$$

Ανάλογα με τα δεδομένα, όπως και στις περιπτώσεις των διαστημάτων εμπιστοσύνης, η TS θα έχει N(0,1) ή t κατανομή.

1. Έλεγχοι Υποθέσεων για το μ ενός πληθυσμού

Όταν θέλουμε να κατασκευάσουμε ένα έλεγχο υποθέσεων για το μέσο μ ενός πληθυσμού, θέτουμε $\hat{\theta} = \overline{x}, se\left(\hat{\theta}\right) = \sigma_{\overline{x}}$ ή $se\left(\hat{\theta}\right) = s_{\overline{x}}$ αν η $\sigma_{\overline{x}}$ είναι άγνωστη.

1.1 Έλεγχοι υποθέσεων για το μ όταν $n < 30$

- Η τιμή του σ^2 είναι γνωστή και ο πληθυσμός που προέρχεται ο πληθυσμός είναι κανονικός
- Η τιμή του σ^2 είναι άγνωστη και ο πληθυσμός που προέρχεται ο πληθυσμός είναι κανονικός

1.1.1 Έλεγχοι υποθέσεων για το μ όταν $n < 30$, η τιμή του σ^2 είναι γνωστή και ο πληθυσμός που προέρχεται ο πληθυσμός είναι κανονικός

Εφόσον η κατανομή του πληθυσμού υποτίθεται ότι είναι $N\left(\mu, \sigma^2\right)$, έπεται ότι $\overline{X} \sim N\left(\mu, \sigma_{\overline{X}^2}\right)$, $\sigma_{\overline{X}}^2 = \frac{\sigma^2}{n}$. Επομένως:

$$TS = \frac{\overline{X} - \mu_o}{\sigma/\sqrt{n}} \sim N\left(0,1\right)$$

Με βάση αυτή τη μορφή της H1, διακρίνουμε τις εξής τρεις περιπτώσεις:

1. Αν H1: μ>μο, τότε απορρίπτουμε την Ho αν TS > Zα
2. Αν H1: μ<μο, τότε απορρίπτουμε την Ho αν TS <- Zα
3. Αν H1: μ$\neq$μο, τότε απορρίπτουμε την Ho αν TS > Zα/2 ή TS < -Zα/2

Παράδειγμα

Μία βιομηχανία γνωρίζει ότι η ποσότητα του τυριού που παράγεται από 1000 κιλά γάλα είναι 300 κιλά με τυπική απόκλιση 11 κιλά τυρί. Προτείνεται μια νέα μέθοδος. Η μέση ποσότητα παραγόμενου τυριού είναι 305 με τυπική απόκλιση 11. Τα στοιχεία αυτά επιλέχθηκαν από ένα δείγμα μεγέθους 10.

Α)Υπήρξε βελτίωση? Χρησιμοποιείστε επίπεδο σημαντικότητας α=5% και α=10%

Β)Ποιο είναι το μικρότερο επίπεδο σημαντικότητας στο οποίο θα μπορούσε να απορριφθεί η αρχική υπόθεση (p-value)

Γ)Να υπολογισθεί και να ερμηνευτεί η δύναμη του ελέγχου για α=5% και μ=310.

Λύση

Α)

$H_o: \ \mu = 300$

$H_1: \ \mu > 300$

$$TS = \frac{\overline{X} - \mu_o}{\sigma/\sqrt{n}} = \frac{305 - 300}{11/\sqrt{10}} = 1.44$$

Για α=5%, Ζα=1.645, επομένως δεν απορρίπτω την αρχική υπόθεση

Για α=10%, Ζα=1.28, επομένως απορρίπτω την αρχική υπόθεση

Β)

Το μικρότερο επίπεδο εμπιστοσύνης στο οποίο θα μπορούσαμε να απορρίψουμε την Ηο είναι η τιμή P=P(Z>1.44) = 0.0749. Επομένως, αν μας δοθεί α=8%, τότε απορρίπτουμε την Ηο. Αν μας δοθεί α=7% τότε δεν την απορρίπτουμε

Γ)

Για να υπολογίσουμε τη δύναμη του ελέγχου όταν α=5% και μ=μ1=310, ξεκινάμε από τον κανόνα αποφάσεως, ο οποίος είναι

Απορρίπτω την αρχική υπόθεση αν $TS = \frac{\overline{X} - \mu_o}{\sigma/\sqrt{n}} = \frac{\overline{X} - 300}{11/\sqrt{10}} > 1.645$

Δηλαδή απέρριψε την αρχική αν $\overline{X} + 1.645 \times 3.4785 \Rightarrow \overline{X} > 305.7221$ και επομένως μην απορρίπτεις όταν $\overline{X} \leq 305.7221$. Επομένως:

$$\beta\left(\mu = 310\right) = P\left(\overline{X} \leq 305.7221 | \mu = 310\right) = P\left(Z < \left(305.7221 - 310\right)/3.4785\right)$$

$$= P\left(Z < -1.23\right) = 0.1093$$

οπότε δύναμη =1-β(μ=310) = 1-0.1093=0.8907

Ερμηνεία

Αν η υπόθεση Ηο:μ=300 είναι εσφαλμένη και στην πραγματικότητα ισχύει ότι μ=310, τότε ο έλεγχος θα απορρίψει την Ηο με πιθανότητα 89.07%.

Παράδειγμα

Για το προηγούμενο παράδειγμα, ας υποθέσουμε ότι η νέα μέθοδος απλώς δοκιμάζεται, χωρίς να υπάρχει κανένας ισχυρισμός ότι είναι αποτελεσματικότερη από την παραδοσιακή. Υπάρχει, με άλλα λόγια, περισσότερη αβεβαιότητα για την τιμή του μέσου μ. Να διατυπωθούν οι υποθέσεις Ho και H1 και να ελεγχθεί στα επίπεδα α=5% και α=10%, όπως και να απαντηθούν και τα β και γ ερωτήματα του προηγούμενου παραδείγματος.

Λύση

Α)Ο έλεγχος είναι δίπλευρος

$H_o : \mu = 300$

$H_1 : \mu \neq 300$

Για α=5%, οι κριτικές τιμές είναι -1.96 και 1.96, ενώ για α=10% αυτές είναι -1.645 και 1.645, ενώ TS = 1.44. Επομένως επειδή η TS βρίσκεται στην περιοχή μη απορρίψεως δεν απορρίπτουμε την αρχική υπόθεση.

Β)Επειδή έχουμε δίπλευρο έλεγχο και μας ενδιαφέρει τόσο η περίπτωση μ>300 όσο και η περίπτωση μ<300, η τιμή P=P(Z>1.44 ή z<-1.44)=2*P(Z>1.44) = 2*0.0749 = 0.1498

Γ)

Ο κανόνας αποφάσεως απόρριψης είναι

$$\frac{\overline{X} - 300}{3.4785} < -1.96$$

ή

$$\frac{\overline{X} - 300}{3.4785} > 1.96$$

δηλαδή απέρριψε την Ho αν $\overline{X} < 293.1821$ και $\overline{X} > 306.8179$.

Άρα η πιθανότητα σφάλματος ΙΙ για μ=310 είναι

$$\beta\left(\mu = 310\right) = P\left(293.1821 \leq \overline{X} \leq 306.8179 | \mu = 310\right)$$

$$P\left(\left(293.1821 - 310\right)/3.4785 < Z < \left(306.8179 - 310\right)3.4785\right) = P\left(-4.83 < Z < -0.91\right) = 0.1814$$

Επομένως η δύναμη του ελέγχου για μ=310 είναι ισή με 1-β=0.8186

1.1.2 Έλεγχοι υποθέσεων για το μ όταν $n < 30$, η τιμή του σ^2 είναι άγνωστή και ο πληθυσμός που προέρχεται ο πληθυσμός είναι κανονικός

Εφόσον η διακύμανση είναι άγνωστη, έπεται ότι

$$TS = \frac{\overline{X} - \mu_o}{S/\sqrt{n}} \sim t_{n-1}$$

Παράδειγμα

Για το προηγούμενο παράδειγμα, αν η διακύμανση του πληθυσμού ήταν άγνωστη, τότε θα έπρεπε να χρησιμοποιήσουμε τη δειγματική εκτίμηση. Έστω s=11.8. να κάνετε τον μονόπλευρο έλεγχο για επίπεδο σημαντικότητας α=5%

Λυση

A)

$H_o: \ \mu = 300$

$H_1: \ \mu > 300$

$$TS = \frac{\overline{X} - \mu_o}{\sigma/\sqrt{n}} = \frac{305-300}{11.8/\sqrt{10}} = 1.34$$

Επειδή $t_{9,\,0.05} = 1.833$ δεν απορρίπτω την αρχική υπόθεση.

1.2 Μεγάλα δείγματα

Με βάση το κεντρικό οριακό θεώρημα μπορώ να χρησιμοποιώ την κανονική κατανομή.

Παράδειγμα

Στο προηγούμενο παράδειγμα, θεωρούμε ότι το μέγεθος του δείγματος είναι n=36, ο δειγματικός μέσος ισούται με 305 με τυπική απόκλιση s= 11. Να πραγματοποιήσετε τον

$H_o: \ \mu = 300$

έλεγχο $H_1: \ \mu > 300$ με βάση το p value.

Λύση

$$TS = \frac{\overline{X} - \mu_o}{S/\sqrt{n}} = \frac{305-300}{11/\sqrt{36}} = 2.73$$

Η τιμή P είναι η πιθανότητα P(Z>2.73)=0.0032. Για οποιοδήποτε επίπεδο σημαντικότητα α>0.0032 απορρίπτω την αρχική υπόθεση.

Παράδειγμα

Μέχρι το έτος 1987 στις εξετάσεις στο μάθημα της Στατιστικής στη σχολή Ικάρων έχει διαπιστωθεί από παρατηρήσεις πολλών ετών, ότι η μέση βαθμολογία των σπουδαστών ήταν 87 μονάδες με τυπική απόκλιση σ=9 μονάδες. Στις εξετάσεις της περιόδου Ιανουαρίου του τρέχοντας έτους από τυχαίο δείγμα 36 σπουδαστών διαπιστώνεται μέση βαθμολογία 90 μονάδων. Με την προϋπόθεση ότι η τυπική απόκλιση σ=9 ισχύει, μπορούμε να ισχυριστούμε ότι η μέση βαθμολογία δεν έχει αυξηθεί; Επίπεδο σημαντικότητας 5%.

Λύση

Η υπόθεση θα είναι $\begin{array}{l} H_0 : \mu = 87 \\ H_1 : \mu > 87 \end{array}$

Επειδή σ είναι γνωστή και έχουμε μεγάλο δείγμα θα χρησιμοποιήσουμε την κανονική κατανομή.

Στατιστικό στοιχείο: $\dfrac{\overline{X} - \mu_0}{\sigma / \sqrt{n}} = \dfrac{90 - 87}{9/6} = 2$

Κριτική τιμή (έλεγχο μιας ουράς): $z_\alpha = z_{5\%} = 1{,}645$

Επειδή 2>1,645 απορρίπτεται ότι η βαθμολογία έχει παραμένει σταθερή (απορ. H_0)

Παράδειγμα

Μια αυτόματη μηχανή γεμίζει με δύο λίτρα (2000 cc) μπουκάλια κοκα κόλας. Ένας καταναλωτής θέλει να ελέγχει την υπόθεση ότι τα λίτρα είναι τουλάχιστον 2. επιλέχθηκε ένα τυχαίο δείγμα 40 μπουκαλιών με μέση τιμή 1999.6 cc. Η τυπική απόκλιση του πληθυσμού ισούται με 1.30 cc. Υπολογίστε το p-value για αυτό το τεστ.

H_0: $\mu \geq 2000$
H_1: $\mu < 2000$
$n = 40$, $\mu_0 = 2000$, x-bar = 1999.6,
$\sigma = 1.3$

Η στατιστική ελέγχου: $z = \dfrac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}$

$$z = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{1999.6 - 2000}{\frac{1.3}{\sqrt{40}}}$$

$$= -1.95$$

$$p\text{-value} = P(Z < -1.95)$$

$$= 0.5000 - 0.4744$$

$$= 0.0256$$

Παράδειγμα

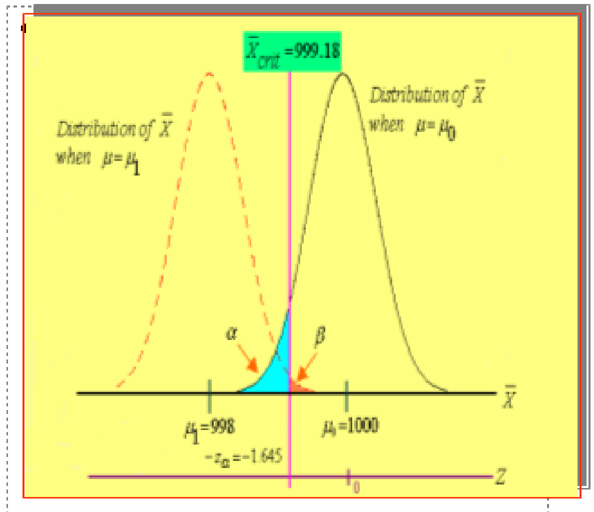
Θεωρείστε τον παρακάτω έλεγχο υποθέσεων:

$$H_0: \mu \geq 1000$$
$$H_1: \mu < 1000$$

έστω $\sigma = 5$, $\alpha = 5\%$, και $n = 100$. Θέλουμε να υπολογίσουμε το β όταν $\mu = \mu_1 = 998$.

Λύση

- Το γράφημα δείχνει την κατανομή του δειγματικού μέσου όταν μ = μ0 = 1000, και όταν μ = μ1 = 998.
- Η Ho απορρίπτεται όταν ο δειγματικός μέσος είναι μικρότερος από την κριτική τιμή, ήτοι, μ0 -zα * σ/√n = 1000 – 1.645*5/ √100 = 999.18.
- Συνεπώς, η Η0 δεν θα απορρίπτεται όταν ο δειγματικός μέσος είναι μεγαλύτερος από την κριτική τιμή.



- Όταν $\mu = \mu_1 = 998$, β είναι η πιθανότητα μη απόρριψης της H_0 το οποίο σημαίνει P{(δειγματικός μέσος > κριτική τιμή του δειγματικού μέσου}.
- Όταν $\mu = \mu_1$, ο δειγματικός μέσος ακολουθεί την κανονική κατανομή με μέσο μ_1 και τυπική απόκλιση= $\sigma/\sqrt{n}$. Επομένως,

$$\beta = P\left[Z > \frac{\bar{X}_{crit} - \mu_1}{\sigma/\sqrt{n}}\right] = P(Z > 1.18/0.5) = P(Z > 2.360)$$
$$= 0.0091$$

- Η δύναμη του τεστ ισούται με 1 – 0.0091 = 0.9909.

Παράδειγμα

Ένας μύλος για να είναι λειτουργικός, θα πρέπει η ταχύτητα του αέρα να ξεπερνάει τα 20 μίλια την ώρα. Σε 50 διαφορετικές στιγμές καταγράφηκε η ταχύτητα του αέρα (=25, σ=12). Η περιοχή που επιλέχθηκε είναι κατάλληλη? Δηλαδή, η ταχύτητα του αέρα είναι μεγαλύτερη από 20?

- Στην περίπτωση του σφάλματος τύπου Ι (α), θα απορρίπταμε την αρχική υπόθεση όταν αυτή είναι σωστή, δηλαδή, θα είχαμε λανθασμένα συμπεράνει ότι η ταχύτητα του αέρα ξεπερνάει τα 20 μίλια. Η συνέπεια αυτής της απόφασης θα ήταν να κτίζαμε σε λάθος μέρος τον ανεμόμυλο. Το κόστος θα ήταν σημαντικό και γι΄ αυτό θέτουμε α=0,01.

- Στην περίπτωση του σφάλματος τύπου ΙΙ (β), δεν θα απορρίπταμε την αρχική υπόθεση όταν αυτή είναι λάθος, δηλαδή, θα είχαμε λανθασμένα συμπεράνει ότι η ταχύτητα του αέρα δεν ξεπερνάει τα 20 μίλια. Η συνέπεια αυτής της απόφασης είναι ότι δεν θα κτίζαμε τον ανεμόμυλο σε ένα μέρος που είναι κατάλληλο. Το κόστος αυτού του σφάλματος δεν είναι μεγάλο.
- Ωστόσο, θεωρείστε ότι όταν η ταχύτητα του αέρα ξεπερνάει τα 25 μίλια, τα κέρδη είναι τεράστια. Άρα θα πρέπει να υπολογίσουμε την πιθανότητα αυτού του σφάλματος.

Λύση

Θεωρείστε τον παρακάτω έλεγχο υποθέσεων:

$$H_0: \mu < 20$$
$$H_1: \mu > 20$$

έστω $\sigma = 12$, $\alpha = 1\%$, και $n = 50$. Θέλουμε να υπολογίσουμε το β όταν $\mu = \mu_1 = 25$.

•Το πρώτο βήμα είναι να υπολογίσουμε την περιοχή απόρριψης για το δειγματικό μέσο. $z = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}} > z_\alpha = 2.33 \Rightarrow \frac{\bar{x} - 20}{\frac{12}{\sqrt{50}}} > 2.33 \Rightarrow$

$\bar{x} > 23.95$

•Το δεύτερο βήμα είναι να υπολογίσουμε την περιοχή μη απόρριψης της Ho (<23,95). Άρα $\beta = P(\bar{x} < 23.95 / \mu = 25) = P\left(\frac{\bar{x} - \mu}{\sigma / \sqrt{n}} < \frac{23.95 - 25}{12 / \sqrt{50}}\right) = P(z < -0.62) = 0.2676$

Σχολιασμός

- Η πιθανότητα να μην απορρίψω την αρχική υπόθεση όταν μ=25 ισούται με 0,2676, το οποίο σημαίνει ότι όταν η ταχύτητα του αέρα είναι 25 μίλια, υπάρχει πιθανότητα (26,76%) να μην απορρίψω την αρχική, ενώ θα έπρεπε.
- Πως μειώνω το συγκεκριμένο πρόβλημα?
    - Αύξηση του δείγματος,

Παράδειγμα

Μια αυτόματη μηχανή γεμίζει με δύο λίτρα **(2000 cc)** μπουκάλια κοκα κόλας. Ένας καταναλωτής θέλει να ελέγχει την υπόθεση ότι τα λίτρα είναι τουλάχιστον **2.** επιλέχθηκε ένα τυχαίο δείγμα **40** μπουκαλιών με μέση τιμή **1999.6 cc.** Η τυπική απόκλιση του πληθυσμού ισούται με **1.30 cc.**

H_0: $\mu \geq 2000$
H_1: $\mu < 2000$
$n = 40$
Για $\alpha = 0.05$, η κριτική τιμή της z είναι -1.645

Η στατιστική ελέγχου: $z = \dfrac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$

Δεν απορρίπτω H_0 εάν: [$z \geq -1.645$]
Απορρίπτω H_0 αν: [$z < -1.645$]

$n = 40$
$\bar{x} = 1999.6$
$\sigma = 1.3$

$$z = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{1999.6 - 2000}{\frac{1.3}{\sqrt{40}}}$$

$$= -1.95 \Rightarrow \text{Reject } H_0$$

Υπολογισμός P-value

Μια αυτόματη μηχανή γεμίζει με δύο λίτρα **(2000 cc)** μπουκάλια κοκα κόλας. Ένας καταναλωτής θέλει να ελέγχει την υπόθεση ότι τα λίτρα είναι τουλάχιστον **2.** επιλέχθηκε ένα τυχαίο δείγμα **40** μπουκαλιών με μέση τιμή **1999.6 cc.** Η τυπική απόκλιση του πληθυσμού ισούται με **1.30 cc.**

H_0: $\mu \geq 2000$
H_1: $\mu < 2000$
$n = 40$
Για $\alpha = 0.05$, η κριτική τιμή της z είναι -1.645

Η στατιστική: $z = \dfrac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$

Δεν απορρίπτω H_0 αν: [$p\text{-value} \geq 0.05$]
Απορρίπτω H_0 αν: [$p\text{-value} < 0.05$]

$$z = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{1999.6 - 2000}{\frac{1.3}{\sqrt{40}}}$$

$$= -1.95$$

$$p\text{-value} = P(Z < -1.95)$$

$$= 0.5000 - 0.4744$$

$$= 0.0256 \Rightarrow \text{Reject } H_0 \text{ since } 0.0256 < 0.05$$

Παράδειγμα

Μια αεροπορική εταιρεία επιθυμεί να εξετάσει το μέσο βάρος των βαλιτσών που μεταφέρουν οι επιβάτες στις καμπίνες των αεροσκαφών. Ο ερευνητής θα εξετάσει αν το μέσο βάρος είναι $\mu_0 = 12$ pounds, σε σχέση με την εναλλακτική ότι το μέσο βάρος δεν είναι 12 pounds. Το επίπεδο εμπιστοσύνης είναι 5%, $\alpha = 0.05$.

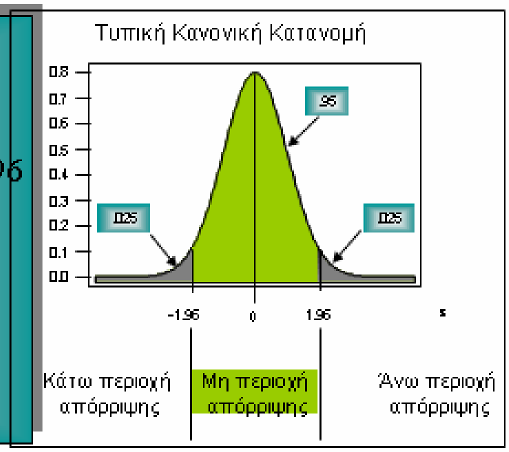
H_0: $\mu = 12$
H_1: $\mu \neq 12$

Για $\alpha = 0.05$, η κριτική τιμή της z είναι ± 1.96

Η στατ. συνάρτηση ισούται με $z = \dfrac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}$

Μην απόρριψης H_0 εάν: $[-1.96 \leq z \leq 1.96]$

Απέρριψε H_0 εάν: $[z < -1.96]$ η $[z > 1.96]$



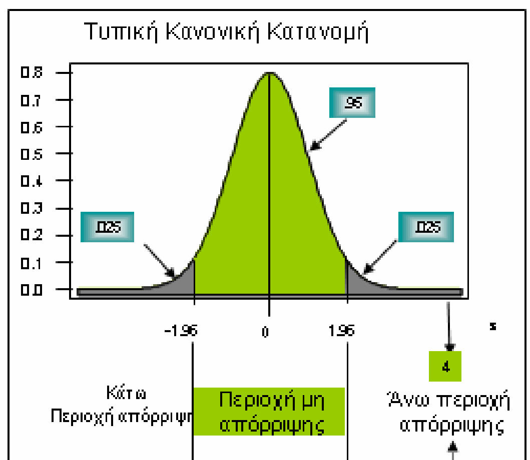
n = 144

x = 14.6

s = 7.8

$$z = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}} = \frac{14.6 - 12}{\frac{7.8}{\sqrt{144}}}$$

$$= \frac{2.6}{0.65} = 4$$



Δεδομένου ότι η τιμή της στατιστικής ανήκει στο άνω διάστημα απόρριψης η H_0 απορρίπτεται και επομένως μπορούμε να καταλήξουμε στο συμπέρασμα ότι το μέσο βάρος των αποσκευών είναι διαφορετικό από τα 12 pounds.

Παράδειγμα

Μια ασφαλιστική εταιρεία πιστεύει ότι το μέσο ύψος των ασφαλειών είναι $2000. Χρησιμοποιώντας $\alpha = 0.01$, ελέγξτε αν αυτή η υπόθεση είναι αληθινή.

$H_0: \mu = 2000$

$H_1: \mu \neq 2000$

Για $\alpha = 0.01$, οι κριτικές τιμές της z είναι ± 2.576

Η στατιστική ελέγχου ισούται $= \dfrac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}$

Δεν απορρίπτω H_0 εάν: $[-2.576 \le z \le 2.576]$

Απορρίπτω H_0 εάν: $[z < -2.576]$ ή $[z > 2.576]$

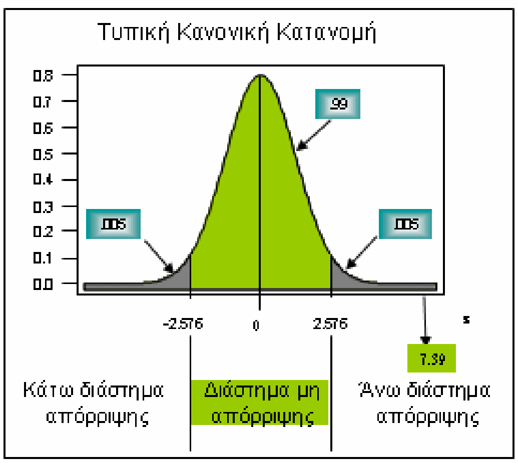
$n = 100$

$\bar{x} = 2700$

$s = 947$

$$z = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}} = \frac{2700 - 2000}{\frac{947}{\sqrt{100}}}$$

$$= \frac{700}{94.7} = 7.39 \quad \Rightarrow \text{Reject } H_0$$



Δεδομένου ότι η τιμή της στατιστικής συνάρτησης ανήκει στο άνω διάστημα απόρριψης, η $H0$ απορρίπτεται, και μπορούμε να καταλήξουμε στο συμπέρασμα ότι το μέσο ύψος των ασφαλειών είναι διαφορετικό των $2000

Παράδειγμα

Ο μέσος χρόνος που χρειάζεται ένας υπολογιστής για να επεξεργαστεί τα δεδομένα είναι 3.24 δευτερόλεπτα. Εφαρμόζεται ένας καινούργιος αλγόριθμός. Χρησιμοποιώντας επίπεδο εμπιστοσύνης 5%, ελέγξτε την υπόθεση.

H_0: $\mu = 3.24$
H_1: $\mu \neq 3.24$

Για $\alpha = 0.05$, οι τιμές τις z είναι ± 1.96

Η στατιστική ισούται: $z = \dfrac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}$

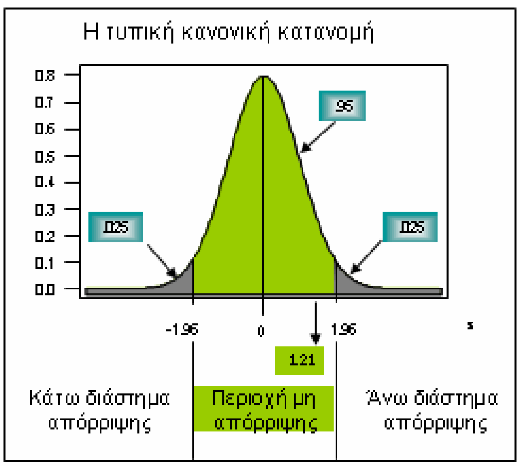
Δεν απορρίπτω H_0 εάν: $[-1.96 \le z \le 1.96]$

Απορρίπτω H_0 εάν: $[z < -1.96]$ or $[z > 1.96]$

n = 200
$\bar{x}$ = 3.48
s = 2.8

$$z = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}} = \frac{3.48 - 3.24}{\frac{2.8}{\sqrt{200}}}$$

$$= \frac{0.24}{0.20} = 1.21 \quad \Rightarrow \text{Do not reject } H_0$$



Δεδομένου ότι η τιμή της στατιστικής ανήκει στο διάστημα μη απόρριψης, η H_0 δεν απορρίπτεται, και μπορούμε να συνάγουμε ότι ο μέσος χρόνος επεξεργασίας δεν έχει αλλάξει.

Παράδειγμα

Η αξία της γης στο Τόκιο αυξήθηκε κατά 49% τους τελευταίους 6 μήνες του 1995. Ένας διεθνής οργανισμός θέλει να ελέγξει την υπόθεση χρησιμοποιώντας επίπεδο εμπιστοσύνης 1%.

H_0: $\mu = 49$
H_1: $\mu \neq 49$
$n = 18$
Για $\alpha = 0.01$ και $(18\text{-}1) = 17$ βαθμούς ελευθερίας (df) ,
Οι κριτικές τιμές της t είναι ± 2.898

Η στατιστική ελέγχου ισούται: $t = \dfrac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}$

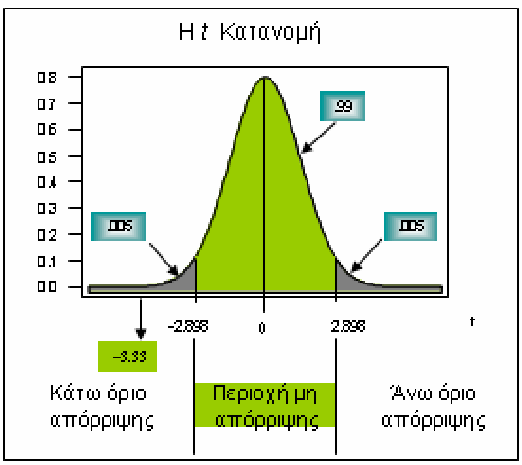
Δεν απορρίπτω H_0 εάν: $[-2.898 \leq t \leq 2.898]$

Απορρίπτω H_0 if: $[t < -2.898]$ ή $[t > 2.898]$

$n = 18$
$\bar{x} = 38$
$s = 14$

$$t = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}} = \frac{38 - 49}{\frac{14}{\sqrt{18}}}$$

$$= \frac{-11}{3.3} = -3.33 \Rightarrow \text{Reject } H_0$$



Δεδομένου ότι η στατιστική ελέγχου ανήκει στο άνω διάστημα, η H_0 απορρίπτεται, και μπορούμε να συμπεράνουμε ότι η αξία της γης δεν αυξήθηκε κατά49%.

Παράδειγμα

Η Canon, Inc,. Κατασκεύασε έναν νέο εκτυπωτή που αναμένεται να εκτυπώνει 27 αντίγραφα το λεπτό. Χρησιμοποιώντας $\alpha = 0.05$ ελέγξτε την υπόθεση.

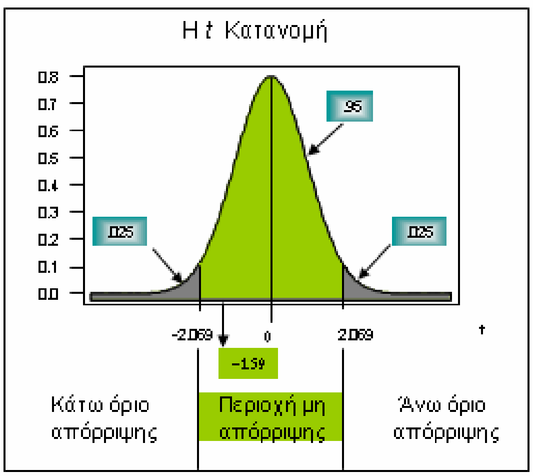
H_0: $\mu = 27$
H_1: $\mu \neq 27$
n = 24
Για $\alpha = 0.05$ και (24-1) = 23 df ,
Οι κριτικές τιμές της t είναι ±2.069

Η στατιστική ελέγχου ισούται: $t = \dfrac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}$

Δεν απορρίπτω H_0 εάν: $[-2.069 \le t \le 2.069]$
Απορρίπτω H_0 εάν: $[t < -2.069]$ or $[t > 2.069]$

n = 24
$\bar{x}$ = 24.6
s = 7.4

$$t = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}} = \frac{24.6 - 27}{\frac{7.4}{\sqrt{24}}} = \frac{-2.4}{1.51} = -1.59 \Rightarrow \text{Do not reject } H_0$$



Δεδομένου ότι η στατιστική συνάρτηση ανήκει στο διάστημα μη απόρριψης, η H_0 δεν απορρίπτεται, και επομένως μπορούμε να συμπεράνουμε ότι ο μέσος χρόνος εκτύπωσης είναι 27 αντίγραφα το λεπτό.

Παράδειγμα

Ένας αναλυτής της Goldman Sachs επιθυμεί να ελέγξει την υπόθεση το 70% των ξένων επενδυτών στην χρηματαγορά της Αγγλίας είναι Αμερικάνοι. Ο αναλυτής συνέλεξε δεδομένα από 210 λογαριασμούς, από τους οποίους διαπίστωσε ότι 130 ανήκουν σε Αμερικάνους. Πραγματοποιήστε τον έλεγχο χρησιμοποιώντας $\alpha = 0.05$.

$H_0: p = 0.70$

$H_1: p \neq 0.70$

$n = 210$

Για $\alpha = 0.05$ οι κριτικές τιμές της z είναι ± 1.96

Η στατιστική ελέγχου ισούται: $z = \dfrac{\hat{p} - p_0}{\sqrt{\dfrac{p_0 q_0}{n}}}$

Δεν απορρίπτω H_0 εάν: $[-1.96 \leq z \leq 1.96]$

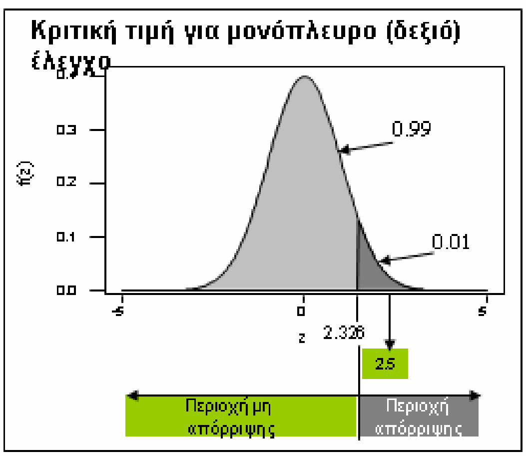
Απορρίπτω H_0 εάν: $[z < -1.96]$ or $[z > 1.96]$

$n = 210$

$\hat{p} = \dfrac{130}{210} = 0.619$

$$z = \frac{\hat{p} - p_0}{\sqrt{\dfrac{p_0 q_0}{n}}} = \frac{0.619 - 0.70}{\sqrt{\dfrac{(0.70)(0.30)}{210}}}$$

$$= \frac{-0.081}{0.0316} = -2.5614 \Rightarrow \text{Reject } H_0$$



Δεδομένου ότι η τιμή της στατιστικής ανήκει στην περιοχή απόρριψης, η H_0 απορρίπτεται και επομένως μπορούμε να συμπεράνουμε ότι οι βιομηχανίες παράγουν περισσότερα αέρια από ότι επιτρέπεται.

Παράδειγμα

Ένας κατασκευαστής ενός προϊόντος ισχυρίζεται ότι το μέσο μέρος της κάθε συσκευασίας είναι 12. Ωστόσο, καταναλωτές διαμαρτυρήθηκαν ότι το βάρος του προϊόντος είναι μικρότερο. Πραγματοποιήθηκε δειγματοληψία σε 144 πακέτα στα οποία το μέσο βάρος ήταν11,8 με τυπική απόκλιση 6. Πραγματοποιήστε τον έλεγχο

H_0: $\mu \geq 12$

H_1: $\mu < 12$

$n = 144$

Για $\alpha = 0.05$, η κριτική τιμή της Z ισούται με z is -1.645

Η στατιστική συνάρτηση ελέγχου ισούται με: $z = \dfrac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}$

Δεν απορρίπτω H_0 εάν: [$z \geq -1.645$]

Απορρίπτω H_0 εάν: [$z < -1.645$]

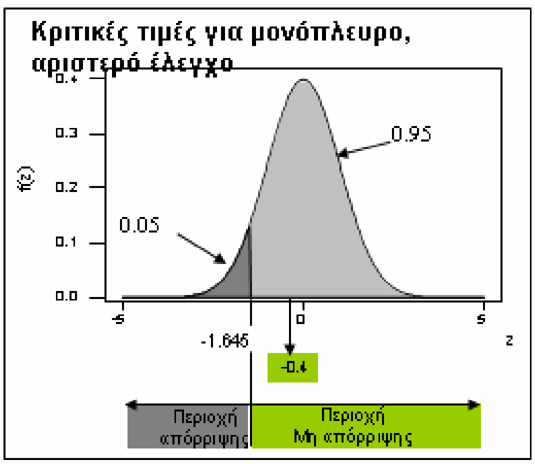
$n = 144$

$\bar{x} = 11.8$

$s = 6$

$$z = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}} = \frac{11.8 - 12}{\frac{6}{\sqrt{144}}}$$

$$= \frac{-.2}{.5} = -0.4 \quad \Rightarrow \text{Do not reject } H_0$$



Δεδομένου ότι η τιμή της στατιστικής συνάρτησης ελέγχου ανήκει στο διάστημα μη απόρριψης, δεν απορρίπτω Ho. Επομένως το μέσο βάρος είναι στατιστικά ίσο με 12

Παράδειγμα

Μια μπαταρία λέγεται ότι διαρκεί 65 ώρες. Ένας ανταγωνιστής πιστεύει ότι διαρκεί λιγότερο. Επιλέχθηκε ένα τυχαίο δείγμα μεγέθους **21** στο οποίο η μέση διάρκεια ήταν 62,5 ώρες με τυπική απόκλιση 3. Για α=0,01, πραγματοποιήστε τον έλεγχο.

H_0: $\mu \geq 65$

H_1: $\mu < 65$

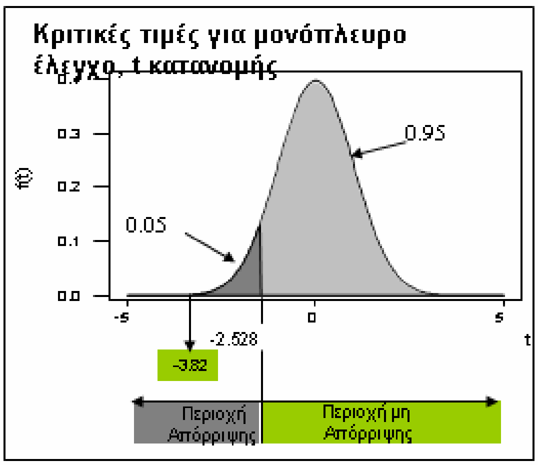
n = 21

Για $\alpha = 0.01$ και (21-1) = 20 df, η κριτική τιμή ισούται με -2.528

Η στατιστική ελέγχου ισούται με: $t = \dfrac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}} = -3.82$

Δεν απορρίπτω H_0 εάν: [$t \geq -2.528$]

Απορρίπτω H_0 εάν: [$t < -2.528$]



Δεδομένου ότι η στατιστική συνάρτηση ανήκει στο διάστημα απόρριψης, η H_0 απορρίπτεται και επομένως μπορούμε να ισχυριστούμε ότι η μέση διάρκεια είναι μικρότερη από 65 ώρες.

2. Έλεγχοι υποθέσεων για την αναλογία (p), υποθέτοντας «μεγάλο» δείγμα

Έστω ότι έχουμε ένα δείγμα n ανεξάρτητων Bernoulli δοκιμών και θέλουμε να κατασκευάσουμε ένα έλεγχο υπόθεσης για την (p) στον πληθυσμό. Υποθέτουμε ότι το μέγεθος του δείγματος είναι μεγάλο και ισχύουν οι παρακάτω προϋποθέσεις:

$$n\hat{P} \geq 5$$
$$n\left(1-\hat{P}\right) \geq 5$$

όπου $\hat{P}$ είναι η αναλογία στο δείγμα. Η εκτιμήτρια της πιθανότητας p της διωνυμικής κατανομής είναι η αναλογία των επιτυχιών που παρατηρούμε στις n ανεξάρτητες διαδοχικές δοκιμές, δηλαδή $\hat{P} = X/n$. Η εκτιμήτρια $\hat{P}$ είναι τυχαία μεταβλητή. Η μέση τιμή και η μεταβλητότητα της $\hat{P}$ είναι $\mu_{\hat{P}} = p,\ \sigma^2_{\hat{P}} = p\left(1-p\right)/n$ αντίστοιχα.

Παράδειγμα

Ένα νόμισμα εξετάζεται για το αν είναι «δίκαιο». Το «στρίβετε» 25 φορές και μόνο 8 «κεφαλές» εμφανίζονται. Εξετάστε αν είναι «δίκαιο» α=5%

Έστω p συμβολίζει την πιθανότητα εμφάνισης «κεφαλής»
H_0: p = 0.5
H_1: p ≠ 0.5
Επειδή ο έλεγχος είναι δίπλευρος, p-value = 2*P(X ≤ 8)
Από τους διωνυμικούς πίνακες, με n = 25, p = 0.5, p-value = 2*0.054 = 0.108.
Αφού 0.108 > α = 0.05, δεν απορρίπτω H_0

Παράδειγμα

Διενεργείται μία έρευνα για να διαπιστωθεί αν η αναλογία του εκλογικού σώματος πιστεύει ότι θα αυξηθεί ο πληθωρισμός. Παραδοσιακά η αναλογία υποτίθεται ότι ισούται με 1/3. Γίνεται μια έρευνα σε 100 άτομα και βρίσκεται ότι 25 από αυτούς πιστεύουν ότι α αυξηθεί.

Α) Να διατυπωθούν οι υποθέσεις και να ελεγχθεί η αρχική με α =5%

Β)Υπολογίστε το p-value

Γ)Να υπολογισθεί η δύναμη του ελέγχου για p=0.2

Λύση

$$H_o:\ p = 1/3$$
$$H_1:\ p \neq 1/3$$

Το δείγμα είναι «μεγάλο» γιατί

$$np = 100/3 > 5\ n(1-p) = 200 \times 2/3 > 5$$

Το τυπικό σφάλμα ισούται με

$$\sigma_{\hat{p}} = \sqrt{1/3 \times 2/3}/\sqrt{100} = 0.0471$$

άρα

$$TS = \frac{0.25 - 0.33333}{0.0471} = -1.77$$

Δεν απορρίπτω την αρχική, αφού η κριτική τιμή ισούται με Ζα/2=1.96

Β)P=P(Z<1.77 ή Z>1.77) = 2P(Z>1.77)=0.0768. Άρα μπορούμε να απορρίψουμε την Ho σε επίπεδο σημαντικότητας 7.7% ή μεγαλύτερο

Γ)Ο κανόνας αποφάσεως για απόρριψη της Ho είναι:

$$\frac{\hat{P} - 1/3}{0.0471} < -1.96$$

ή

$$\frac{\hat{P} - 1/3}{0.0471} > 1.96$$

Από τις ανισότητες αυτές προκύπτει ότι η περιοχή απορρίψεως πάνω στην κατανομή δειγματοληψίας της στατιστικής $\hat{P}$ ορίζεται ως εξής: Απορρίψτε την Ho αν $\hat{P} < 0.2410$ ή $\hat{P} > 0.4256$. Η πιθανότητα σφάλματος τύπου ΙΙ για την p=0.2 είναι

$$\beta\left(p = 0.2\right) = P\left(0.2410 \leq \hat{P} \leq 0.4256 | p = 0.2\right)$$

Επειδή όμως σε αυτή την πιθανότητα υποθέτουμε ότι p=0.2 και όχι p=1/3 για να υπολογίσουμε κατά τα γνωστά πρέπει να υπολογίσουμε το τυπικό σφάλμα της αναλογίας στο δείγμα. Επομένως $\sigma_{\hat{P}} = \sqrt{0.2 \times 0.8/100} = 0.04$. Άρα

$$\begin{aligned}\beta\left(p = 0.2\right) &= P\left(\left(0.2410 - 0.2\right)/0.04 \leq Z \leq \left(0.4256 - 0.2\right)/0.04\right) \\ &= P\left(1.025 < Z < 5.64\right) = 0.153\end{aligned}$$

Παράδειγμα

Ένας αναλυτής της Goldman Sachs επιθυμεί να ελέγξει την υπόθεση το 70% των ξένων επενδυτών στην χρηματαγορά της Αγγλίας είναι Αμερικάνοι. Ο αναλυτής συνέλεξε δεδομένα από 210 λογαριασμούς, από τους οποίους διαπίστωσε ότι 130 ανήκουν σε Αμερικάνους. Πραγματοποιήστε τον έλεγχο χρησιμοποιώντας $\alpha = 0.05$.

$H_0: p = 0.70$

$H_1: p \neq 0.70$

$n = 210$

Για $\alpha = 0.05$ οι κριτικές τιμές της z είναι ± 1.96

Η στατιστική ελέγχου ισούται: $z = \dfrac{\hat{p} - p_0}{\sqrt{\dfrac{p_0 q_0}{n}}}$

Δεν απορρίπτω H_0 εάν: $[-1.96 \le z \le 1.96]$

Απορρίπτω H_0 εάν: $[z < -1.96]$ or $[z > 1.96]$

$n = 210$

$$\hat{p} = \frac{130}{210} = 0.619$$

$$z = \frac{\hat{p} - p_0}{\sqrt{\dfrac{p_0 q_0}{n}}} = \frac{0.619 - 0.70}{\sqrt{\dfrac{(0.70)(0.30)}{210}}}$$

$$= \frac{-0.081}{0.0316} = -2.5614 \Rightarrow \text{Reject } H_0$$

3. Έλεγχος υποθέσεων για τη διαφορά δύο μέσων $\mu_X - \mu_Y$

Έστω ότι έχουμε δύο τυχαίες μεταβλητές, X και Y, και θέλουμε να κατασκευάσουμε ένα διάστημα εμπιστοσύνης για τη διαφορά των μέσων του, $\mu_X - \mu_Y$. Στη διάθεσή μας, έχουμε ένα δείγμα από κάθε πληθυσμό μεγέθους n_X και n_Y, αντίστοιχα. Το πρώτο πράγμα που μας ενδιαφέρει να διαπιστώσουμε εδώ είναι αν δείγματα είναι εξαρτημένα ή ανεξάρτητα.

3.1 Εξαρτημένα δείγματα

Η εκτίμηση της διαφοράς μεταξύ δύο διαδοχικών τιμών των ίδιων παρατηρήσεων είναι μια συνηθισμένη ανάλυση στις επιχειρηματικές δραστηριότητες. Για παράδειγμα, διαφορές στις καταμετρήσεις των απογραφών, μέτρηση ικανοποίησης ενός καταναλωτή πριν και μετά την εφαρμογή νέου τρόπου εξυπηρέτησης, μέτρηση έντασης γνώμης πριν και μετά την εκστρατεία ενημέρωσης κλπ. Έτσι το δείγμα αποτελείται από 2 ομάδες παρατηρήσεων. Την πρώτη πριν από την εφαρμογή και τη δεύτερη μετά την εφαρμογή της μεθόδου.

Στην περίπτωση που τα δύο δείγματα έχουν το ίδιο μέγεθος $n_X = n_Y = n$ και είναι εξαρτημένα μεταξύ τους, προχωρούμε στην κατασκευή μιας νέας μεταβλητής

$D_i = X_i - Y_i$ και στην συνέχεια υπολογίζουμε το διάστημα εμπιστοσύνης με βάση την προηγούμενη ανάλυση. Η στατιστική συνάρτηση ελέγχου:

$$TS = \frac{\overline{D} - D_o}{S_{\overline{D}}}$$

ακολουθεί είτε την κανονική κατανομή είτε την t ανάλογα με το μέγεθος του δείγματος και αν γνωρίζουμε το διακύμανση του πληθυσμού

Παράδειγμα

Ένα τυχαίο δείγμα 16 τηλεθεατών επιλέχθηκε. Το ποσό των χρημάτων που είχαν ξοδέψει οι τηλεθεατές τις περασμένες διακοπές έχει καταγραφεί. Το ίδιο συνέβηκε την επόμενη χρονιά. Εξετάστε αν αυξήθηκε το ποσό που ξοδεύουν οι τηλεθεατές

Καταναλωτής	Προηγούμενη	Τωρινή	Διαφ.
1	334	405	71
2	150	125	-25
3	520	540	20
4	95	100	5
5	212	200	-12
6	30	30	0
7	1055	1200	145
8	300	265	-35
9	85	90	5
10	129	206	77
11	40	18	-22
12	440	489	49
13	610	590	-20
14	208	310	102
15	880	995	115
16	25	75	50

H_0: $\mu_D \leq 0$
H_1: $\mu_D > 0$

$df = (n-1) = (16-1) = 15$

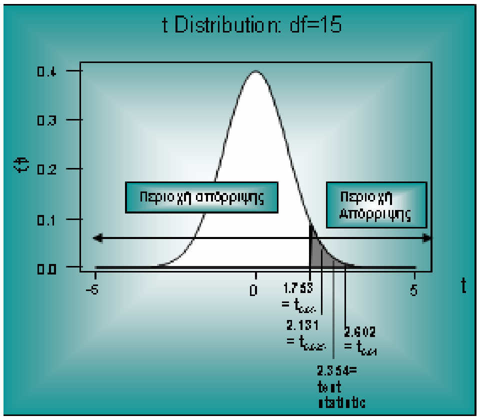
$$t = \frac{\overline{D} - \mu_{D_0}}{\frac{s_D}{\sqrt{n}}}$$

Κριτική τιμή: $t_{0.05} = 1.753$

Δεν απορρίπτω H_0 if : $t \leq 1.753$
Απορρίπτω H_0 if: $t > 1.753$

$$t = \frac{\bar{D} - \mu_{D_0}}{\frac{s_D}{\sqrt{n}}} = \frac{32.81 - 0}{\frac{55.75}{\sqrt{16}}} = 2.354$$

$t = 2.354 > 1.753$, επομένως απορρίπτω H_0 και μπορούμε να συμπεράνουμε ότι υπάρχουν ενδείξεις ότι έχει αυξηθεί



Παράδειγμα

Θέλετε να ελέγξετε αν η μέση μηνιαία απόδοση κάποιων μετοχών αλλάζει από το γεγονός ότι εμφανίζονται στην στήλη της *The Wall Street Journal* "Heard on the Street."

H_0: $\mu_D \leq 0$
H_1: $\mu_D > 0$

n = 50
D = 0.1%
s_D = 0.05%

Η στατιστική: $z = \frac{\bar{D} - \mu_{D_0}}{\frac{s_D}{\sqrt{n}}}$

$$z = \frac{\bar{D} - \mu_{D_0}}{\frac{s_D}{\sqrt{n}}} = \frac{0.1 - 0}{\frac{0.05}{\sqrt{50}}} = 14.14$$

p - value: $p(z > 14.14) \approx 0$

Ανεξάρτητα δείγματα

Ας υποθέσουμε τώρα ότι τα δύο δείγματά μας μπορούν να θεωρηθούν ανεξάρτητα. Στην περίπτωση αυτή, δεν είναι απαραίτητο να υποθέσουμε ότι $n_X = n_Y$. Θα διακρίνουμε δύο περιπτώσεις:

1. και τα δύο δείγματα είναι μικρά
2. και τα δύο δείγματα είναι μεγάλα

3.2.1 Μικρά δείγματα $(n_X < 30,\ n_Y < 30)$

Σε αυτή την περίπτωση, είναι απαραίτητο να υποθέσουμε ότι και οι δύο πληθυσμοί είναι κανονικοί. Θα διακρίνουμε 2 υποπεριπτώσεις:

1. Οι διακυμάνσεις των πληθυσμών $(\sigma_X^2,\ \sigma_Y^2)$
2. Οι διακυμάνσεις είναι άγνωστες.

3.2.1.1 Οι διακυμάνσεις θεωρούνται γνωστές

Έστω $X_1, X_2, ...X_{n_X}$ και $Y_1, Y_2, ..., Y_{n_Y}$ δύο τυχαία δείγματα μεγέθους $n_X,\ n_Y$ αντίστοιχα από δύο ανεξάρτητους κανονικούς πληθυσμούς $N\left(\mu_X, \sigma_X^2\right)$ και $N\left(\mu_Y, \sigma_Y^2\right)$. Έστω ότι $\sigma_X^2,\ \sigma_Y^2$ είναι γνωστά μεγέθη. Ενδιαφερόμαστε να κατασκευάσουμε ένα διάστημα εμπιστοσύνης για τη διαφορά $\mu_X - \mu_Y$.

Όπως είναι γνωστό, μια σημειακή εκτιμήτρια για τη διαφορά $\mu_X - \mu_Y$ είναι το $\overline{X} - \overline{Y}$, όπου $\overline{X}$ και $\overline{Y}$ είναι οι δειγματικοί μέσοι των δειγμάτων.

Δεδομένου ότι:

$$\overline{X} \sim N\left(\mu_X, \sigma_X^2/n_X\right)$$

$$\overline{Y} \sim N\left(\mu_Y, \sigma_Y^2/n_Y\right)$$

$$\overline{X} - \overline{Y} \sim N\left(\mu_X - \mu_Y,\ \frac{\sigma_X^2}{n_X} + \frac{\sigma_Y^2}{n_Y}\right) \Rightarrow$$

$$Z = \frac{\overline{X} - \overline{Y} - \left(\mu_X - \mu_Y\right)}{\sqrt{\frac{\sigma_X^2}{n_X} + \frac{\sigma_Y^2}{n_Y}}} \sim N(0,1)$$

Η στατιστική συνάρτηση ελέγχου είναι:

$$TS = \frac{\overline{X} - \overline{Y} - D_o}{\sigma_{\overline{X}-\overline{Y}}} \sim N(0,1)$$

3.2.1.2 Οι διακυμάνσεις θεωρούνται άγνωστες, αλλά ίσες

Δουλεύουμε όπως και στο προηγούμενο παράδειγμα, αλλά μόνο χρησιμοποιούμε την t κατανομή και την από κοινού διακύμανση

$$S_p^2 = \frac{n_X S_X^2 + n_Y S_Y^2}{n_X + n_Y - 2} = \frac{(n_X - 1)\, S_X^{*2} + (n_Y - 1)\, S_Y^{*2}}{n_X + n_Y - 2}$$

Παράδειγμα

Ο CPI αυξάνεται με τον ίδιο ρυθμό ή όχι όταν η τιμή του πετρελαίου είναι 27.50$ ή 20$?

Population 1: Oil price = \$27.50

$n_1 = 14$

$\bar{x}_1 = 0.317\%$

$s_1 = 0.12\%$

Population 2: Oil price = \$20.00

$n_2 = 9$

$\bar{x}_2 = 0.21\%$

$s_2 = 0.11\%$

$df = (n_1 + n_2 - 2) = (14 + 9 - 2) = 21$

$H_0: \mu_1 - \mu_2 = 0$

$H_1: \mu_1 - \mu_2 \neq 0$

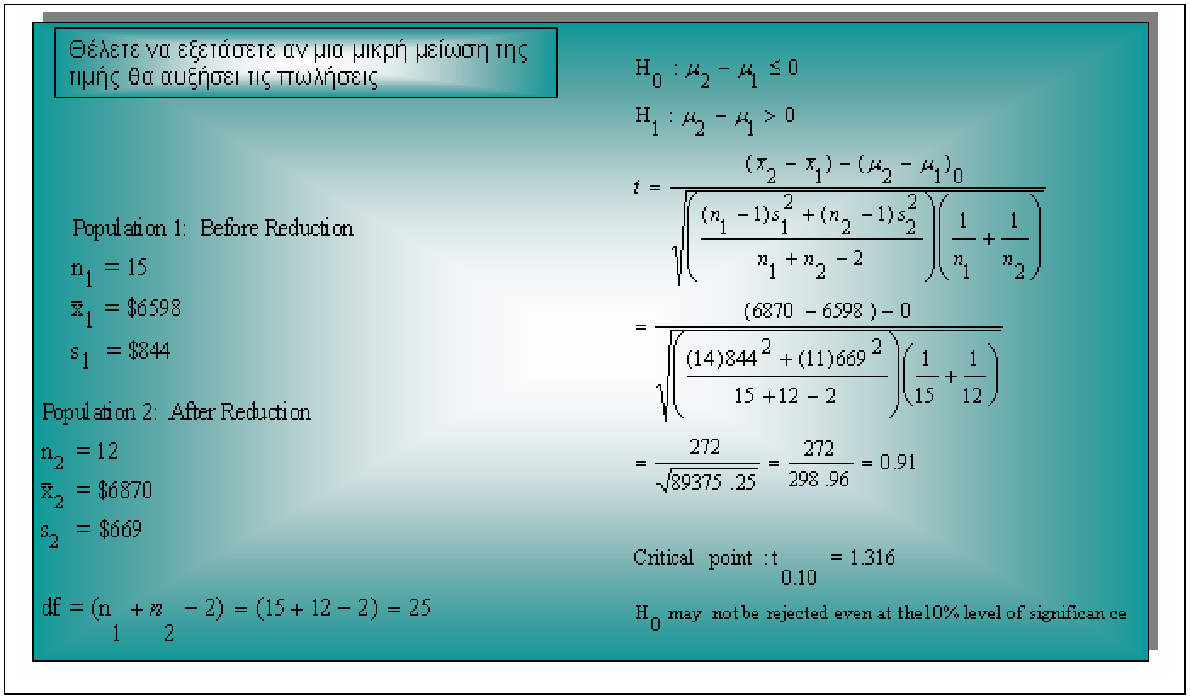
$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)_0}{\sqrt{\left(\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}\right)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$$= \frac{0.107}{\sqrt{0.00247}} = \frac{0.107}{0.0497} = 2.154$$

Critical point: $t_{0.025} = 2.080$

H_0 may be rejected at the 5% level of significance

Παράδειγμα



Από δύο κανονικούς πληθυσμούς πήραμε δυο δείγματα Α και Β, τα οποία μας έδωσαν τις παρακάτω τιμές:

X_i: 0,114 0,127 0,143 0,132

Y_i: 0,131 0,107 0,104 0,111 0,108 0,110

Αν υποθέσουμε ότι $\sigma_1^2 = \sigma_2^2$ να ελεγχθεί η υπόθεση , δηλαδή και τα δυο δείγματα προέρχονται από δύο πληθυσμούς με τον ίδιο μέσο ως προς την εναλλακτική υπόθεση $\mu_1 \neq \mu_2$.

Λύση

$\overline{X} = 0{,}129$

$\overline{Y} = 0{,}1118$

$n_1 = 4$

$n_2 = 6$

$$S_1^2 = \frac{\sum (X_i - \overline{X})^2}{n_1 - 1} = 0{,}000220$$

$$S_2^2 = \frac{\sum (Y_i - \overline{Y})^2}{n_2 - 1} = 0{,}000094$$

$$S^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2} = 0{,}0001412$$

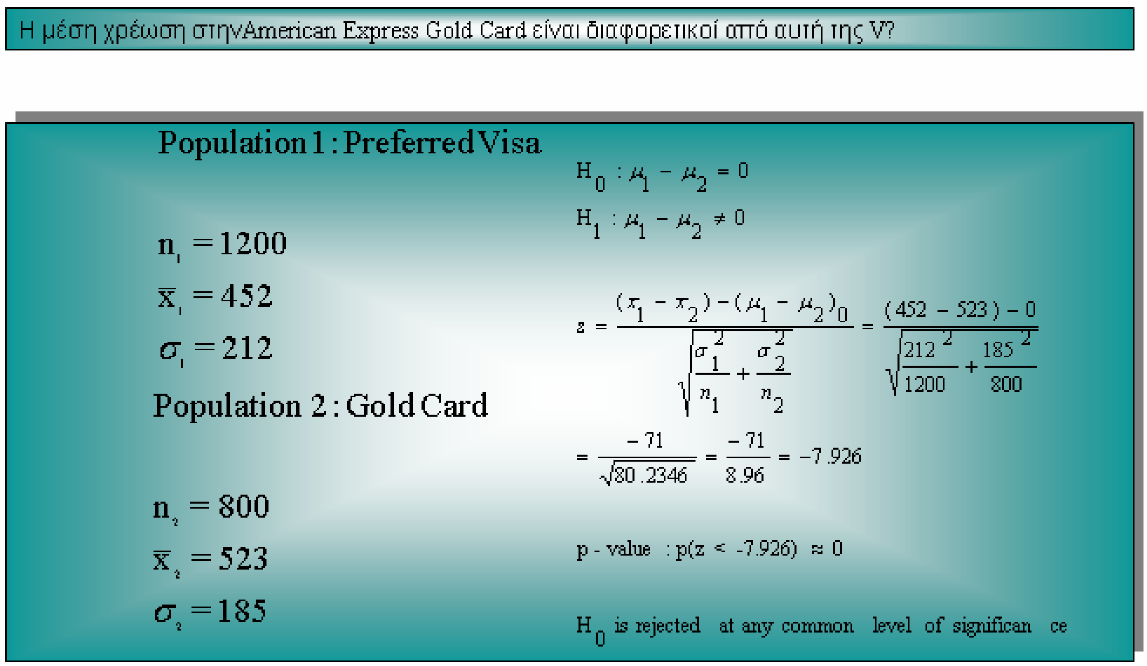
$$t = \frac{0{,}129 - 0{,}1118}{\sqrt{0{,}0001412\left(\frac{1}{4} + \frac{1}{6}\right)}} = 2{,}24$$

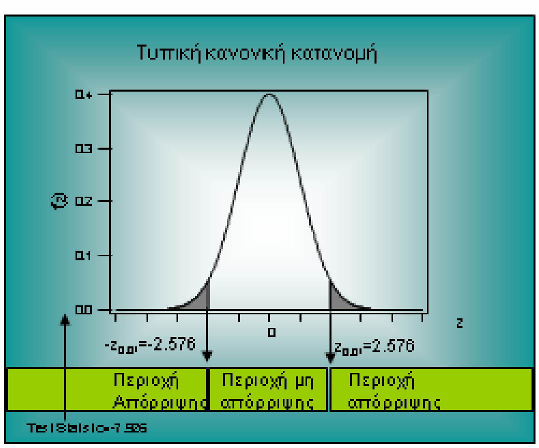
Οι βαθμοί ελευθερίας είναι v=n$_1$+n$_2$-2=8, από τους πίνακες $t_{8,5\%} = 2{,}306$. Επειδή 2,24<2,306 δεν απορρίπτουμε την H$_o$ σε επίπεδο 5%.

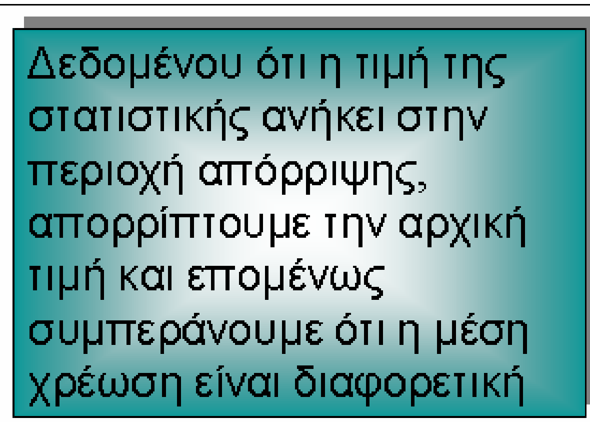
3.2.2 Μεγάλα δείγματα $(n_X \geq 30,\ n_Y \geq 30)$

Όταν τα δείγματα είναι μεγάλα, δεν είναι απαραίτητο οι κατανομές των δύο πληθυσμών να είναι κανονικές. Με βάση το Κεντρικό Οριακό Θεώρημα, θα μπορούσαμε να χρησιμοποιήσουμε τη $N(0,1)$ κατανομή.

Παράδειγμα







Έλεγχοι υποθέσεων για αναλογίες

Παραδείγματα

Δίπλευρος έλεγχος για το ποσοστό δανείων που αφορούν τα αυτοκίνητα (1980 και 1995).

Population 1: 1980

$n_1 = 100$

$x_1 = 53$

$\hat{p}_1 = 0.53$

Population 2: 1995

$n_2 = 100$

$x_2 = 43$

$\hat{p}_2 = 0.43$

$$\hat{p} = \frac{x_1 + x_2}{n_1 + n_2} = \frac{53 + 43}{100 + 100} = 0.48$$

$H_0: p_1 - p_2 = 0$

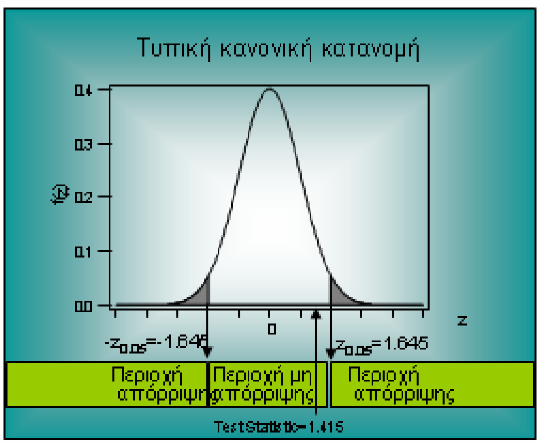
$H_1: p_1 - p_2 \neq 0$

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - 0}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{0.53 - 0.43}{\sqrt{(.48)(.52)\left(\frac{1}{100} + \frac{1}{100}\right)}}$$

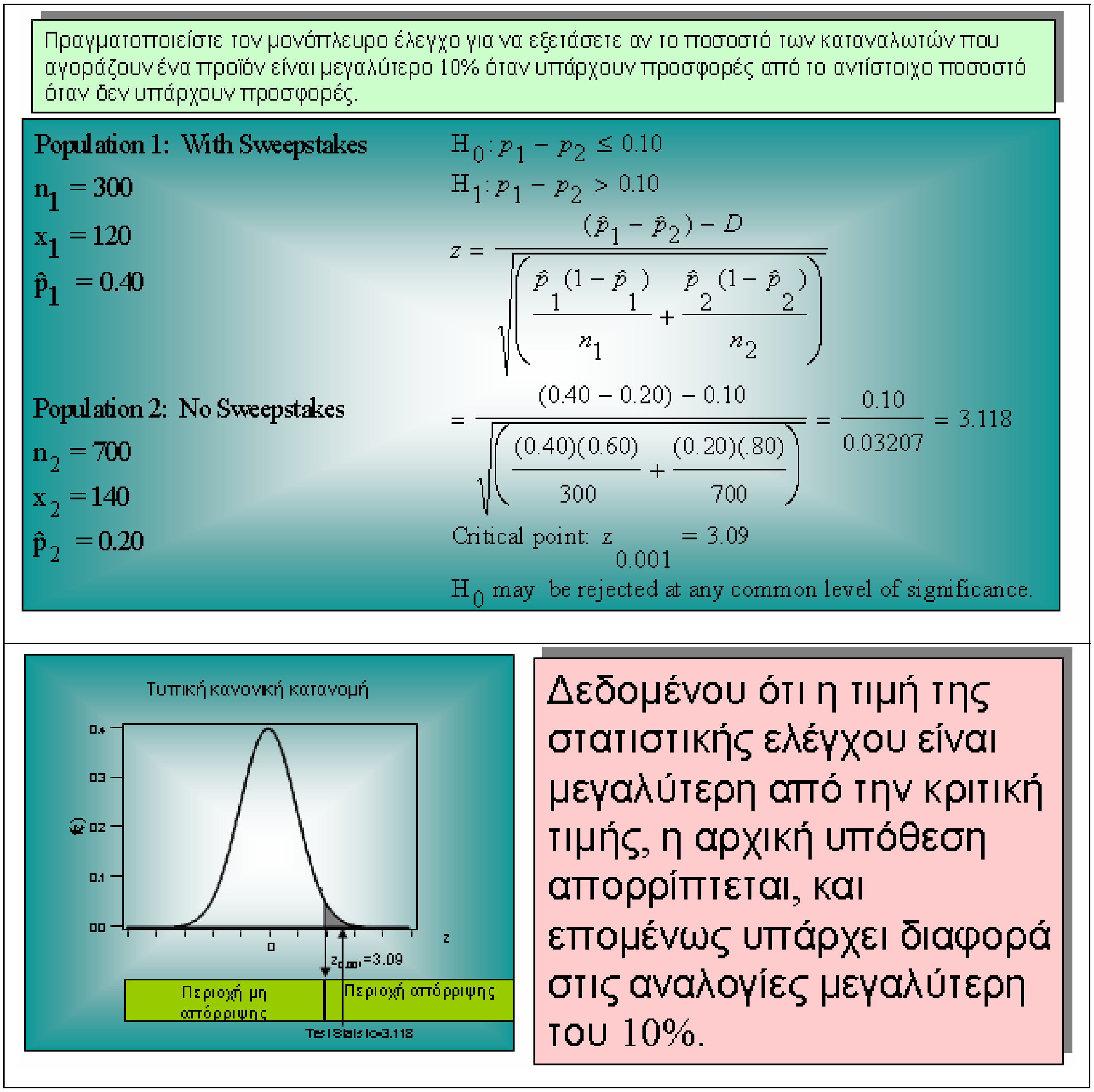
$$= \frac{0.10}{\sqrt{0.004992}} = \frac{0.10}{0.07065} = 1.415$$

Critical point: $z_{0.05} = 1.645$

H_0 may not be rejected even at a 10% level of significance.



Δεδομένου ότι η στατιστική συνάρτηση ανήκει στην περιοχή μη απόρριψης (επίπεδο σημαντικότητας 10%), συμπεράνουμε ότι δεν υπάρχουν διαφορές μεταξύ των δύο αναλογιών (ποσοστά δανείων που αφορούν αυτοκίνητα το 1980 και το 1995)



4. Έλεγχοι υποθέσεων για τη διακύμανση σ^2 ενός κανονικού πληθυσμού.

Ο έλεγχος υποθέσεων για τη διακύμανση ενός πληθυσμού δεν έχει τη γενική μορφή που έχουμε ήδη περιγράψει. Απαραίτητη προϋπόθεση είναι ότι ο πληθυσμός κατανέμεται κανονικά για να χρησιμοποιήσουμε το θεώρημα:

$$\frac{(n-1)\,S^2}{\sigma^2} \sim X^2_{n-1}$$

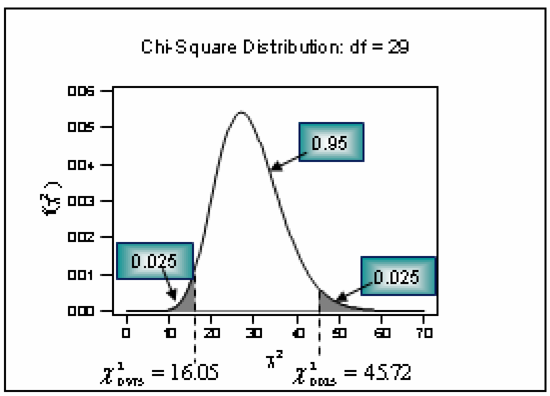
, όπου S^2 είναι η μεταβλητότητα του δείγματος. Η κατανομή X^2 είναι ασυμμετρική και γι΄ αυτό λέμε ότι, για παράδειγμα, τα 95% όρια εμπιστοσύνης καθορίζεται από δύο τιμές $X^2_{0.025}$ και $X^2_{0.975}$, όπως φαίνεται από το επόμενο γράφημα. Οι τιμές $X^2_{0.025}$ και $X^2_{0.975}$

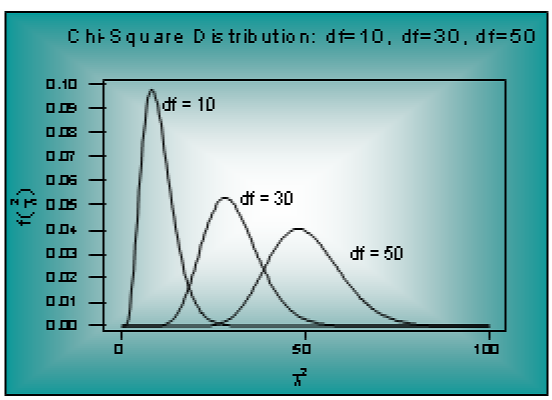
εξαρτώνται από τους βαθμούς ελευθερίας από τους οποίους εξαρτάται οπωσδήποτε και η μορφή της κατανομής.

Χαρακτηριστικά της X^2 κατανομής

- Η χ2 μεταβλητή δεν μπορεί να λάβει αρνητικές τιμές, επομένως το κάτω όριο της είναι το μηδέν.
- Η χ2 μεταβλητή είναι δεξιά ασύμμετρη.
- Η χ2 μεταβλητή προσεγγίζει την κανονική όταν αυξάνονται οι βαθμοί ελευθερίας.

	Area in Right Tail									
df	.995	.990	.975	.950	.900	.100	.050	.025	.010	.005
.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.
28	12.46	13.56	15.31	16.93	18.94	37.92	41.34	44.46	48.28	50.99
29	13.12	14.26	16.05	17.71	19.77	39.09	42.56	45.72	49.59	52.34
30	13.79	14.95	16.79	18.49	20.60	40.26	43.77	46.98	50.89	53.67





Για τον έλεγχο υποθέσεων χρησιμοποιούμε τη στατιστική ελέγχου:

$$TS = \frac{(n-1)\,S^2}{\sigma_o^2} \sim \chi^2_{n-1}$$

Παράδειγμα

Από έναν κανονικό πληθυσμό πήραμε ένα δείγμα με τιμές 30, 40, 28, 54. Ζητείται να ελεγχθεί η υπόθεση:
H_0: σ^2 =30
H_1: σ^2 >30
σε επίπεδο σημαντικότητας 5%.

Λύση

Στατιστικό στοιχείο: $\frac{(n-1)S^2}{\sigma_0^2} = \frac{3(141{,}3)}{30} = 14{,}13$ όπου $S^2 = \frac{\sum (X_i - \overline{X})^2}{n-1} = 141{,}3$

Κριτική τιμή: οι βαθμοί ελευθερίας είναι ν=n-1=4-1=3, α=5%, οπότε $\chi^2_{\nu,\alpha} = \chi^2_{3,5\%} = 7{,}81$

Επειδή 14,13>7,81 απορρίπτεται η H_o.

Παράδειγμα

Σε μία μεγάλη πόλη η μηνιαία δαπάνη για κρέας μίας τετραμελούς οικογένειας κατανέμεται κανονικά με διακύμανση σ^2 Σε επίπεδο σημαντικότητας α =5% να ελέγξετε την υπόθεση Ho: σ_0^2 =20 έναντι της εναλλακτικής $H_1 : \sigma^2 \neq 20$ αν σε ένα δείγμα 15 οικογενειών η διακύμανση είναι $S^2 = 7.5$. Επίσης να υπολογίσετε τη δύναμη του ελέγχου για σ^2 =10

Λύση

Υπολογίζουμε την τιμή της στατιστικής ελέγχου

$$TS = \frac{(n-1)\,S^2}{\sigma_o^2} = (15-1) \times 7.5/20 = 5.25$$

Η τιμή αυτή βρίσκεται στην περιοχή απόρριψης, αφού $\chi_{14,\,0.975} = 5.63$, ενώ $\chi_{14,\,0.025} = 26.12$

Υπολογισμός της δύναμης του ελέγχου για σ^2 =10

Επειδή $\chi_{14,\,0.975} = 5.63$, $\chi_{14,\,0.025} = 26.12$, $\frac{\sigma_0^2}{\sigma^2} = 20/10 = 2$η δύναμη ισούται με

$$P\left(\sigma^2=10\right)=P\left(\frac{(n-1)\,S^2}{\sigma_0^2}<\chi^2_{n-1,\,1-a/2}|H_1\right)+P\left(\frac{(n-1)\,S^2}{\sigma_0^2}>\chi^2_{n-1,\,a/2}|H_1\right)=$$

$$P\left(\frac{(n-1)\,S^2}{\sigma^2}<\frac{\sigma_0^2}{\sigma^2}\chi^2_{n-1,\,1-a/2}|H_1\right)+P\left(\frac{(n-1)\,S^2}{\sigma^2}>\frac{\sigma_0^2}{\sigma^2}\chi^2_{n-1,\,a/2}|H_1\right)=$$

$$=P\left(\chi^2_{n-1}<\frac{\sigma_0^2}{\sigma^2}\chi_{n-1,\,1-a/2}\right)+P\left(\chi^2_{n-1}>\frac{\sigma_0^2}{\sigma^2}\chi_{n-1,\,a/2}\right)$$

$$=P\left(\chi^2_{14}<2\times 5.63\right)+P\left(\chi^2_{14}>2\times 26.12\right)=0.3345+0.00000256=0.335$$

Η δύναμη αυτή είναι πολύ μικρή, αφού στα 2/3 των περιπτώσεων δεν θα απορρίπταμε την αρχική υπόθεση ότι $\sigma^2=20$, ενώ στην πραγματικότητα η τιμή της θα ήταν : $\sigma^2=10$