ΦΡΟΝΤΙΣΤΗΡΙΑ ΣΤΗ ΣΤΑΤΙΣΤΙΚΗ 1

ΕΚΤΙΜΗΣΗ ΔΙΑΣΤΗΜΑΤΟΣ ΕΜΠΙΣΤΟΣΥΝΗΣ

ΣΥΝΟΨΗ Διαστημάτων Εμπιστοσύνης (που θα ασχοληθούμε)

- Διαστήματα Εμπιστοσύνης για τη Μέση τιμή του πληθυσμού, $\boldsymbol{\mu}$
  - Όταν η Τυπική Απόκλιση του πληθυσμού $\boldsymbol{\sigma}$ είναι Γνωστή
  - Όταν η Τυπική Απόκλιση του πληθυσμού $\boldsymbol{\sigma}$ είναι Άγνωστη
- Διαστήματα Εμπιστοσύνης για το Ποσοστό του Πληθυσμού, $\boldsymbol{\pi}$
- Διαστήματα Εμπιστοσύνης για τη *διαφορά μέσων τιμών* $\boldsymbol{\mu_1-\mu_2}$

Εκτιμήσεις Σημείου και Διαστήματος Εμπιστοσύνης

- Μία εκτίμηση σημείου ή σημειακή εκτίμηση είναι ένας αριθμός
- Ένα διάστημα εμπιστοσύνης παρέχει πρόσθετες πληροφορίες σχετικά με τη μεταβλητότητα της εκτίμησης

Εκτιμήσεις Σημείου

Μπορούμε να εκτιμήσουμε μία Παράμετρο πληθυσμού.		Χρησιμοποιώντας Στατιστική Συνάρτηση (Εκτιμήτρια)
Μέση τιμή	μ	$\bar{X}$
Ποσοστό	π	p

Μια συνάρτηση των παρατηρήσεων ενός <u>δείγματος</u> που εξαρτάται <u>μόνο</u> από τις παρατηρήσεις λέγεται ***στατιστική συνάρτηση*** *(statistic)*.

Η τιμή της εκτιμήτριας συνάρτησης **για ένα συγκεκριμένο δείγμα** είναι η **Σημειακή Εκτίμηση** της παραμέτρου που μας ενδιαφέρει

Η Στατιστική συνάρτηση είναι **τυχαία μεταβλητή**.
Μπορεί να λάβει διαφορετικές τιμές για διαφορετικά δείγματα.

Διαστήματα Εμπιστοσύνης

- Πόση αβεβαιότητα συνδέεται με την σημειακή εκτίμηση μιας παραμέτρου πληθυσμού;
- Μία εκτίμηση διαστήματος παρέχει περισσότερες πληροφορίες από μια σημειακή εκτίμηση σχετικά με ένα χαρακτηριστικό του πληθυσμού
- Αυτές οι εκτιμήσεις διαστήματος αποκαλούνται διαστήματα εμπιστοσύνης

Εκτίμηση Διαστήματος Εμπιστοσύνης

- Ένα διάστημα δίνει ένα **εύρος πιθανών** τιμών για **μια παράμετρο** με βάση τα δεδομένα του δείγματος
  - Λαμβάνει υπόψη **τη μεταβλητότητα** της στατιστικής συνάρτησης/ων από δείγμα σε δείγμα

Με βάση τις παρατηρήσεις από 1 δείγμα

- Παρέχει πληροφορίες σχετικά με την **εγγύτητα** της εκτίμησης στην άγνωστη παράμετρο του πληθυσμού καθορισμένη σύμφωνα με το επίπεδο εμπιστοσύνης
  - π.χ. 95% επίπεδο εμπιστοσύνης, 99% επίπεδο εμπιστοσύνης
  - Ποτέ δεν μπορεί να είναι το επίπεδο εμπιστοσύνης 100%

Παράδειγμα Εκτίμησης Διαστήματος

<u>Παράδειγμα συσκευασίας δημητριακών</u>

- Έστω ότι **γνωρίζουμε** τον πληθυσμό ο οποίος είναι Κανονικός και έχει μ = 368 and σ = 15.

Αν πάρουμε δείγμα μεγέθους n = 25 τότε

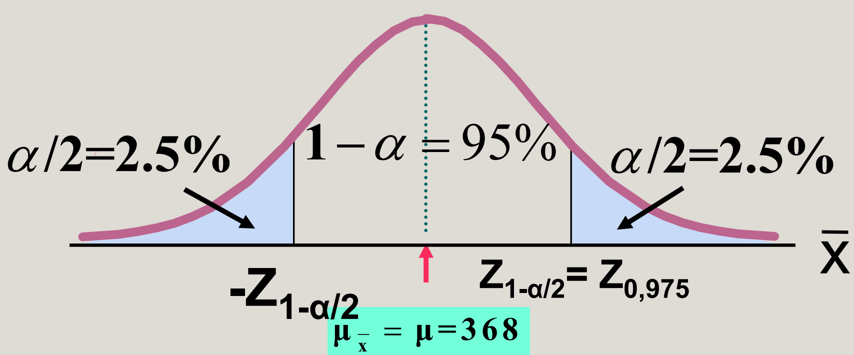
ένα συμμετρικά κατανεμημένο διάστημα γύρω από το μ που περιλαμβάνει το 95% των **δειγματικών μέσων** είναι:

$$368 \pm 1.96 * 15/\sqrt{25} = (362.12, 373.88).$$

Τυπική απόκλιση του $\bar{X}$

Δειγματική κατανομή

$$\bar{X} \sim N(\mu, \frac{\sigma}{\sqrt{n}})$$



Παράδειγμα Εκτίμησης Διαστήματος

<u>Παράδειγμα συσκευασίας δημητριακών</u>

- Όταν **δεν γνωρίζετε** το μ, χρησιμοποιείτε τη **στατιστική συνάρτηση** $\overline{X}$ για να **εκτιμήσετε** το μ
  - Αν η τιμή της εκτιμήτριας για το συγκεκριμένο δείγμα είναι $\overline{X}$ = 362,3 το διάστημα είναι

  $362.3 \pm 1.96 * 15/\sqrt{25} = (356{,}42,\ 368{,}18)$

- Όταν 356,42 ≤ μ ≤ 368,18 το διάστημα που βασίζεται σε αυτό το δείγμα αποτελεί μια σωστή δήλωση για το μ.

Τι γίνεται όμως με τα διαστήματα από άλλα πιθανά δείγματα μεγέθους 25;

Παράδειγμα Εκτίμησης Διαστήματος

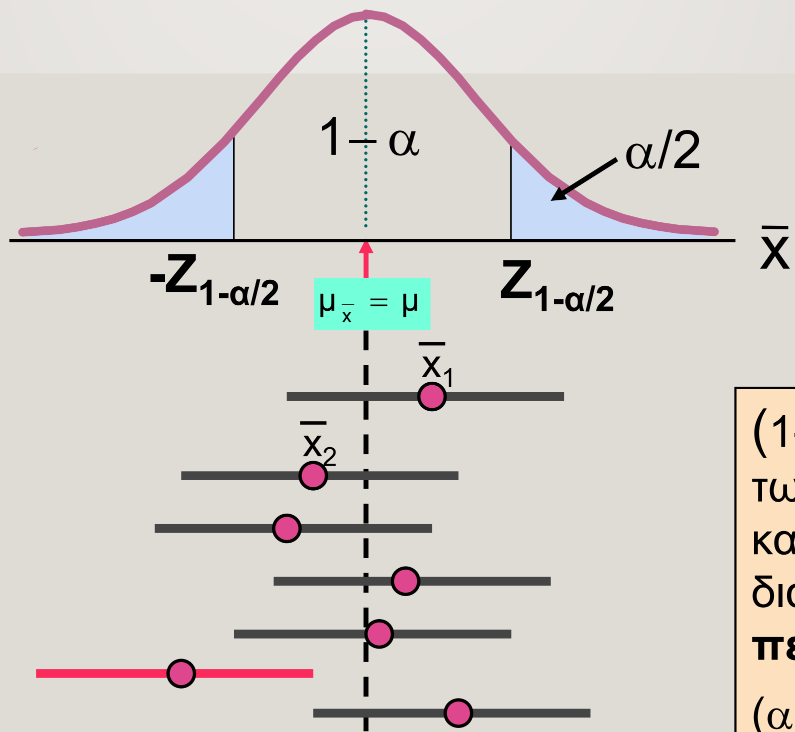
(συνέχεια)

Δείγμα #	$\overline{X}$	Χαμηλότερο Όριο	Υψηλότερο Όριο	**Περιέχει** μ;
1	362.30	356.42	368.18	Ναι
2	369.50	363.62	375.38	Ναι
3	360.00	354.12	365.88	Όχι
4	362.12	356.24	368.00	Ναι
5	373.88	368.00	379.76	Ναι

Διαστήματα και Επίπεδο Εμπιστοσύνης

Κατανομή Δειγματοληψίας της Δειγματικής Μέσης τιμής

z:Pr(Z≤z)=1-α



Διαστήματα Εμπιστοσύνης

Τα διαστήματα κυμαίνονται από

$$\overline{X} - Z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}$$

έως

$$\overline{X} + Z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}$$

$(1-\alpha)\cdot 100\%$ των κατασκευασμένων διαστημάτων **περιέχουν** το μ

$(\alpha)\cdot 100\%$ **δεν** το περιέχουν.

Παράδειγμα Εκτίμησης Διαστήματος

- Στην πράξη λαμβάνετε **μόνο** ένα δείγμα μεγέθους ***n***
- Στην πράξη **δεν γνωρίζετε** το ***μ*** έτσι δεν ξέρετε αν το διάστημα πραγματικά περιέχει το ***μ***
- Ωστόσο, γνωρίζετε ότι το (1-α)% (εδώ 95%) των διαστημάτων που σχηματίζονται με αυτόν τον τρόπο **θα περιέχουν** το ***μ***
- Έτσι, με βάση **ένα** δείγμα, μπορείτε πραγματικά να είστε (1-α)% (εδώ 95%) **βέβαιοι** ότι το διάστημα σας σε μακροχρόνια εφαρμογή **θα περιέχει** το ***μ*** (αυτό είναι ένα 95% διάστημα εμπιστοσύνης)

Σημείωση: Το 95% εμπιστοσύνη βασίζεται στο γεγονός ότι χρησιμοποιήσαμε **Z = 1.96.**

Διαδικασία Εκτίμησης

Γενικός Μαθηματικός Τύπος

- Ο γενικός τύπος για τα διαστήματα εμπιστοσύνης που εξετάζουμε εδώ είναι :

Εκτίμηση Σημείου ± (Κρίσιμη Τιμή)·(Τυπικό Σφάλμα)

Όπου:

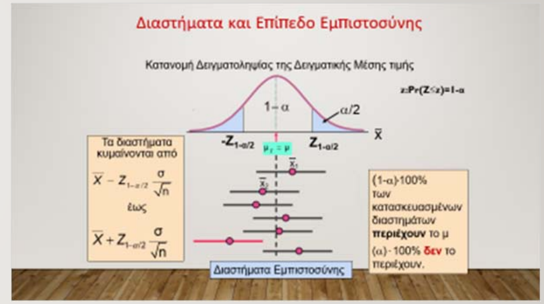
- Εκτίμηση Σημείου είναι **η τιμή** της εκτιμήτρια συνάρτησης που εκτιμά την πληθυσμιακή παράμετρο ενδιαφέροντος **για το συγκεκριμένο δείγμα**
- Κρίσιμη Τιμή είναι μια τιμή πίνακα βάσει της **κατανομής δειγματοληψίας** της εκτιμήτριας συνάρτησης και του **επιθυμητού επιπέδου εμπιστοσύνης**
- Τυπικό Σφάλμα είναι η **τυπική απόκλιση** της κατανομής δειγματοληψίας της εκτιμήτριας συνάρτησης

Επίπεδο Εμπιστοσύνης (1-α)

- Αν επιλέγαμε πολλά τυχαία δείγματα από έναν πληθυσμό και κατασκευάζαμε ένα 1-α% διάστημα εμπιστοσύνης χρησιμοποιώντας κάθε δείγμα, περίπου το **1-α%** των διαστημάτων **θα περιελάμβανε** την παράμετρο ενδιαφέροντος, αν η δειγματοληψία και ο υπολογισμός των διαστημάτων επαναληφθούν πολλές φορές
- Ένα ποσοστό (μικρότερο από 100%)

Επίπεδο Εμπιστοσύνης (1-α)

(συνέχεια)



Ας υποθέσουμε επίπεδο εμπιστοσύνης 95%

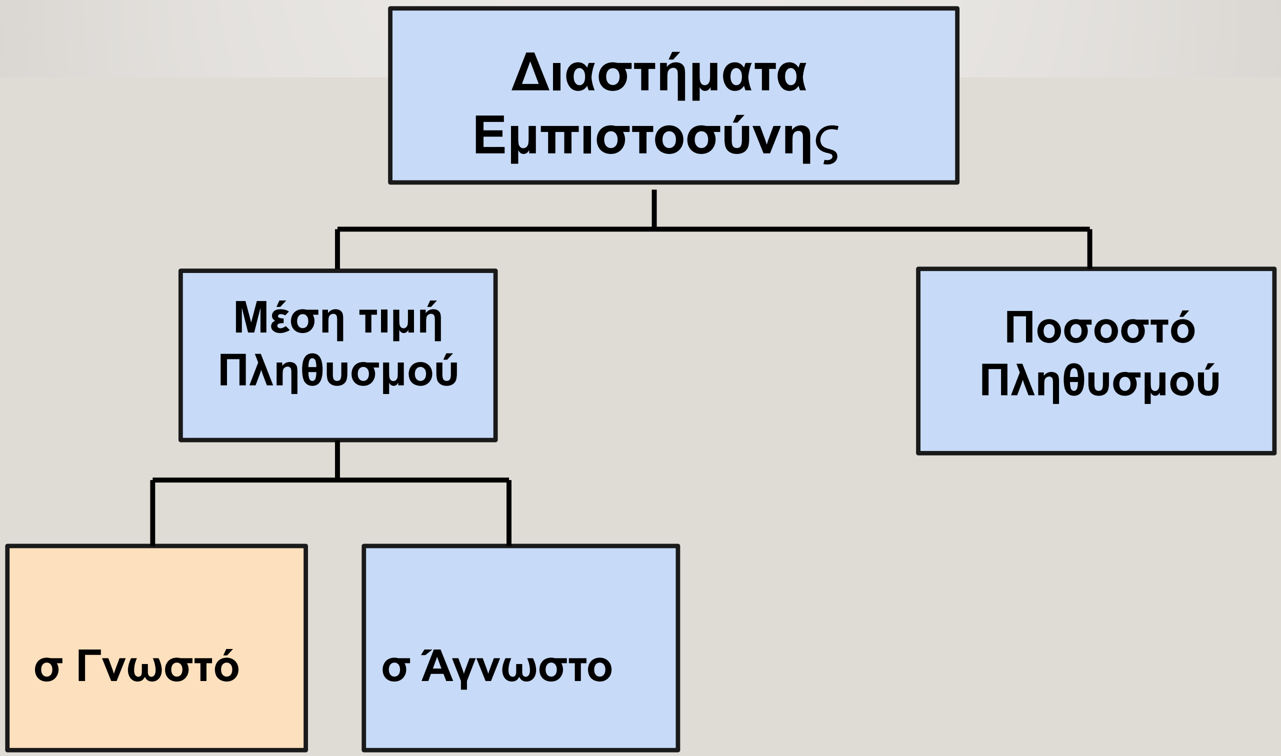
Επίσης γράφεται (1-α)=0.95, (έτσι α=0.05)

Ερμηνεία:

- Το **95% όλων των διαστημάτων εμπιστοσύνης** που μπορούν να κατασκευαστούν **θα περιέχουν** την άγνωστη πραγματική παράμετρο
- Υπάρχει 95% εμπιστοσύνη-βεβαιότητα ότι το διάστημα το οποίο προκύπτει από τη συγκεκριμένη δειγματοληψία, έχει υπολογιστεί με μια μέθοδο <u>που σε βάθος χρόνου</u> παράγει διαστήματα που **θα περιέχουν** την άγνωστη πραγματική παράμετρο.

- **Ένα συγκεκριμένο διάστημα** θα περιέχει ή δεν θα περιέχει την πραγματική παράμετρο
- **Μία παράμετρος** έχει μία σταθερή (άγνωστη) τιμή η οποία θα περιέχεται ή δεν θα περιέχεται σε συγκεκριμένο διάστημα τιμών

Διαστήματα Εμπιστοσύνης



Διαστήματα Εμπιστοσύνης για μ (σ Γνωστό)

- Υποθέσεις
  - Η τυπική απόκλιση πληθυσμού σ είναι **γνωστή**
  - Ο πληθυσμός είναι **κανονικά κατανεμημένος**
    - Αν ο πληθυσμός δεν είναι κανονικός, χρησιμοποιείστε μεγάλο δείγμα (n > 30)

- Εκτίμηση διαστήματος εμπιστοσύνης:

Εκτίμηση Σημείου ± (Κρίσιμη Τιμή)(Τυπικό Σφάλμα)

$$\overline{X} \pm Z_{1-\alpha/2} \frac{\sigma}{\sqrt{n}}$$

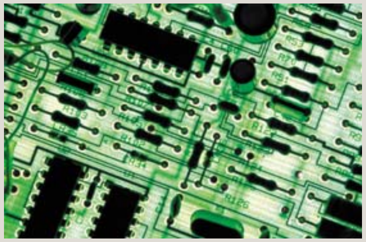
όπου $\overline{X}$ είναι η σημειακή εκτίμηση

$Z_{1-\alpha/2}$ είναι η κρίσιμη τιμή της κανονικής

$\sigma/\sqrt{n}$ είναι το τυπικό σφάλμα της εκτιμήτριας συνάρτησης

Παράδειγμα

- Ένα δείγμα 11 κυκλωμάτων από ένα μεγάλο **κανονικό πληθυσμό** έχει μέση αντίσταση 2.20 Ωμ. Γνωρίζουμε από προηγούμενες δοκιμές ότι η τυπική απόκλιση του πληθυσμού είναι 0.35 Ωμ.
- Προσδιορίστε ένα διάστημα εμπιστοσύνης 95% για την **πραγματική** μέση αντίσταση του **πληθυσμού.**



Παράδειγμα

- **κανονικός** πληθυσμός (η αντίσταση Ωμ)
- **τυπική απόκλιση γνωστή** (0.35 Ωμ)
- $\bar{X} = 2.20$ Ωμ, n=11 α=5%

Εκτίμηση Σημείου ± (Κρίσιμη Τιμή)(Τυπικό Σφάλμα)

Λύση:

$$\bar{X} \pm Z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}} =$$

$$\bar{X} \pm Z_{1-0.025}\frac{\sigma}{\sqrt{n}} = \bar{X} \pm Z_{0.975}\frac{\sigma}{\sqrt{n}}$$

Πίνακας 4. Τυποποιημένη Κανονική Κατανομή
(συνέχεια)

$\Phi(z) = P(Z \le z)$

Αυτός ο Πίνακας δίνεται στις εξετάσεις

	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389
1,0	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177
1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319
1,5	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
1,7	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
1,8	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767
2,0	0,9772	0,9778	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817
2,1	0,9821	0,9826	0,9830	0,9834	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857

Παράδειγμα

(συνέχεια)

$$\overline{X} \pm Z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}} = \overline{X} \pm Z_{0.975}\frac{\sigma}{\sqrt{n}} =$$

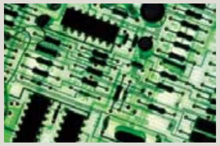
$$= 2.20 \pm 1.96\ (0.35/\sqrt{11})$$

$$= 2.20 \pm 0.2068$$

$$\boxed{1.9932 \le \mu \le 2.4068}$$

Ερμηνεία

- Είμαστε 95% βέβαιοι ότι το διάστημα 1.9932 και 2.4068 Ωμ **θα περιέχει** την πραγματική τιμή της μέσης αντίστασης
- Παρόλο που η τιμή της πραγματικής μέση αντίσταση μπορεί ή δεν μπορεί να βρίσκεται σε αυτό το διάστημα, 95% των διαστημάτων που δημιουργούνται με αυτόν τον τρόπο δειγματοληψίας θα περιέχουν την πραγματική τιμή της μέση τιμή της αντίστασης



<u>1ο Πρόβλημα</u>

Ένα μηχάνημα κατασκευάζει μπαταρίες για ειδικές χρήσεις των οποίων ο χρόνος ζωής ακολουθεί, σύμφωνα με τις τεχνικές προδιαγραφές του μηχανήματος, κανονική κατανομή με μέσο χρόνο ζωής 22 ώρες και τυπική απόκλιση 3 ώρες. Επειδή στις χρήσεις για τις οποίες προορίζονται οι μπαταρίες ο χρόνος ζωής τους είναι μία κρίσιμη παράμετρος, η λειτουργία του μηχανήματος παρακολουθείται συνεχώς με δειγματοληπτικό έλεγχο για το χρόνο ζωής των παραγόμενων μπαταριών. Από τυχαίο δείγμα 9 μπαταριών διαπιστώνεται ότι οι χρόνοι ζωής τους (σε ώρες) είναι:

16	18	22	20	21	20	16	22	25

Με τη βοήθεια της πληροφορίας αυτής να βρεθεί ένα 95% διάστημα εμπιστοσύνης για το μέσο χρόνο ζωής των μπαταριών που παράγει το μηχάνημα

Παράδειγμα

- **κανονικός** πληθυσμός (ο χρόνος ζωής ακολουθεί κανονική κατανομή
- τυπική απόκλιση **γνωστή** (3 ώρες)
- n=9, α=5%

Αρχικά βήματα

Λύση:

Εκτίμηση για το μ

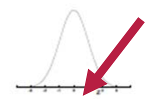
$$\bar{x} = \sum_{i=1}^{9} x_i = \frac{16 + 18 + \cdots + 25}{9} = \frac{180}{9} = 20$$

Εκτίμηση Σημείου ± (Κρίσιμη Τιμή)(Τυπικό Σφάλμα)

$$\overline{\mathrm{X}} \pm Z_{1-\alpha/2} \frac{\sigma}{\sqrt{n}}$$

$$\overline{\mathrm{X}} \pm Z_{1-0.025} \frac{\sigma}{\sqrt{n}} = \overline{\mathrm{X}} \pm Z_{0.975} \frac{\sigma}{\sqrt{n}}$$

Πίνακας 4. Τυποποιημένη Κανονική Κατανομή
(συνέχεια)



$\Phi(z) = P(Z \le z)$

	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389
1,0	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177
1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319
1,5	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
1,7	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
1,8	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	$\boxed{0,9750}$	0,9756	0,9761	0,9767
2,0	0,9772	0,9778	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817
2,1	0,9821	0,9826	0,9830	0,9834	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857

Παράδειγμα

- **κανονικός** πληθυσμός (ο χρόνος ζωής ακολουθεί κανονική κατανομή
- τυπική απόκλιση **γνωστή** (3 ώρες)
- n=9, α=5%

Λύση:

$$\bar{X} \pm Z_{1-0.025}\frac{\sigma}{\sqrt{n}} = \bar{X} \pm Z_{0.975}\frac{\sigma}{\sqrt{n}}$$

$$=20 \pm 1.96\,(3/\sqrt{9}) = 20 \pm 1.96$$

$$18.04 \leq \mu \leq 21.96$$

Είμαστε 95% βέβαιοι ότι το διάστημα 18.04 και 21.96 ώρες **θα περιέχει** την πραγματική τιμή της μέσης διάρκειας ζωής των μπαταριών

Εάν κατασκεύαζα 100 διαστήματα εμπιστοσύνης με μέγεθος δείγματος 9, με παρόμοια χαρακτηριστικά, τα 95 από αυτά θα περιέχουν την πραγματική τιμή της μέσης διάρκειας ζωής των μπαταριών

2ο Πρόβλημα

Μια έρευνα μεταξύ μη κερδοσκοπικών οργανισμών έδειξε ότι οι διαδικτυακές δωρεές προς τους οργανισμούς αυξήθηκαν κατά τη διάρκεια του προηγούμενου έτους. Από ένα τυχαίο δείγμα 55 μη κερδοσκοπικών οργανισμών προέκυψε μέση αξία εφάπαξ δωρεάς κατά το προηγούμενο έτος ύψους $75 με τυπική απόκλιση $9.

Α. Κατασκευάστε ένα 99% διάστημα εμπιστοσύνης για την μέση αξία των εφάπαξ δωρεών του πληθυσμού των δωρητών.

Β. Ερμηνεύστε το διάστημα εμπιστοσύνης.

2ο Πρόβλημα Λύση

κανονικός πληθυσμός (Δεν αναφέρεται αλλά n=55>30)
τυπική απόκλιση **γνωστή** (9$)

Αρχικά βήματα

$\bar{X}$=75 $ n=55, α=1%

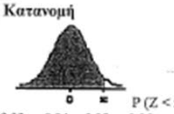
α=0.01 → α/2=0.005 → 1-α/2=0.995

$$\overline{X} \pm Z_{1-0.005}\frac{\sigma}{\sqrt{n}} = \overline{X} \pm Z_{0.995}\frac{\sigma}{\sqrt{n}} =$$

$$75 \pm 2.58\frac{9}{\sqrt{55}} = 75 \pm 3.13$$

$$71.87 \le \mu \le 78.13$$

ΠΙΝΑΚΑΣ 4: Κανονική Κατανομή



P (Z < z)

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0 .08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	$\boxed{0.9951}$	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964

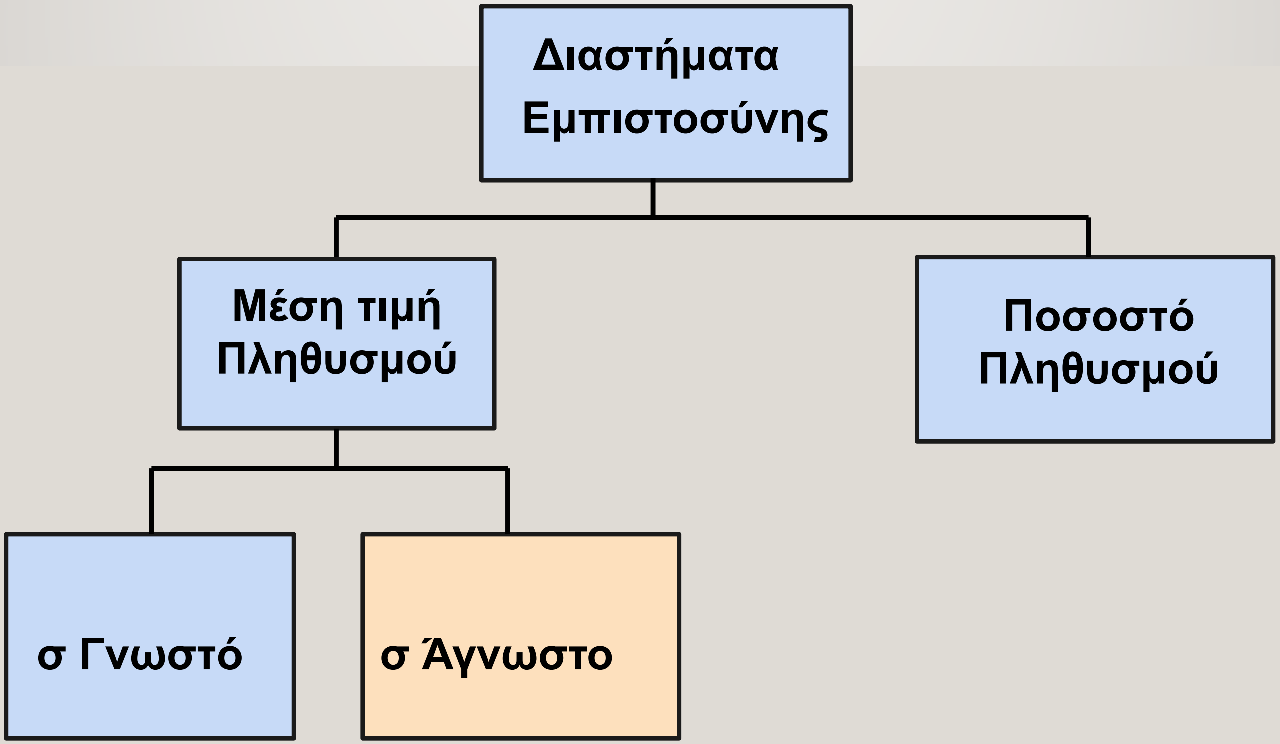
2ο Πρόβλημα Λύση

$$71.87 \le \mu \le 78.13$$

Είμαστε 99% βέβαιοι ότι το διάστημα 71.87 έως 78.13 ώρες **θα περιέχει** την πραγματική τιμή της μέσης αξίας των εφάπαξ δωρεών του πληθυσμού των δωρητών.

Εάν κατασκεύαζα 100 διαστήματα εμπιστοσύνης με μέγεθος δείγματος 55, με παρόμοια χαρακτηριστικά, τα 99 από αυτά θα περιέχουν την πραγματική τιμή της μέσης αξίας των εφάπαξ δωρεών του πληθυσμού των δωρητών.

Διαστήματα Εμπιστοσύνης



Γνωρίζετε πραγματικά το σ;

- Πιθανόν όχι!
- Σε όλες σχεδόν τις πραγματικές επιχειρηματικές καταστάσεις, το $\boldsymbol{\sigma}$ δεν είναι γνωστό.
- Αν υπάρχει μια κατάσταση όπου το $\boldsymbol{\sigma}$ είναι γνωστό τότε το $\boldsymbol{\mu}$ είναι επίσης γνωστό (δεδομένου ότι για να υπολογίσετε $\boldsymbol{\sigma}$ πρέπει να γνωρίζετε το $\boldsymbol{\mu}$).
- Εάν γνωρίζετε πραγματικά το $\boldsymbol{\mu}$, δεν θα χρειαζόταν να συγκεντρωθεί ένα δείγμα για να το εκτιμήσετε.

Διάστημα Εμπιστοσύνης για μ (σ Άγνωστο)

- Εάν η τυπική απόκλιση του πληθυσμού ***σ*** είναι άγνωστη, μπορούμε να την εκτιμήσουμε με την τυπική απόκλιση του δείγματος, ***S***
- Αυτό εισάγει επιπλέον αβεβαιότητα, καθώς το ***S*** μεταβάλλεται από δείγμα σε δείγμα
- Χρησιμοποιούμε την κατανομή ***t*** αντί για την κανονική κατανομή

Κατανομή Student t

- Η *t* είναι μία κατανομή
- Η τιμή $t_{a/2}$ εξαρτάται από τους βαθμούς ελευθερίας (*degrees of freedom* - df.)
  - Ο αριθμός των παρατηρήσεων που είναι ελεύθερες να μεταβληθούν μετά τον υπολογισμό του μέσου του δείγματος

$$\boxed{\text{d.f.} = n - 1}$$

Βαθμοί Ελευθερίας (df)

Ιδέα: Ο αριθμός των παρατηρήσεων που είναι ελεύθερες να μεταβληθούν μετά τον υπολογισμό του μέσου του δείγματος

Παράδειγμα: Υποθέστε ότι ο μέσος όρος 3 αριθμών είναι 8

Έστω $X_1 = 7$

Έστω $X_2 = 8$

X_3;

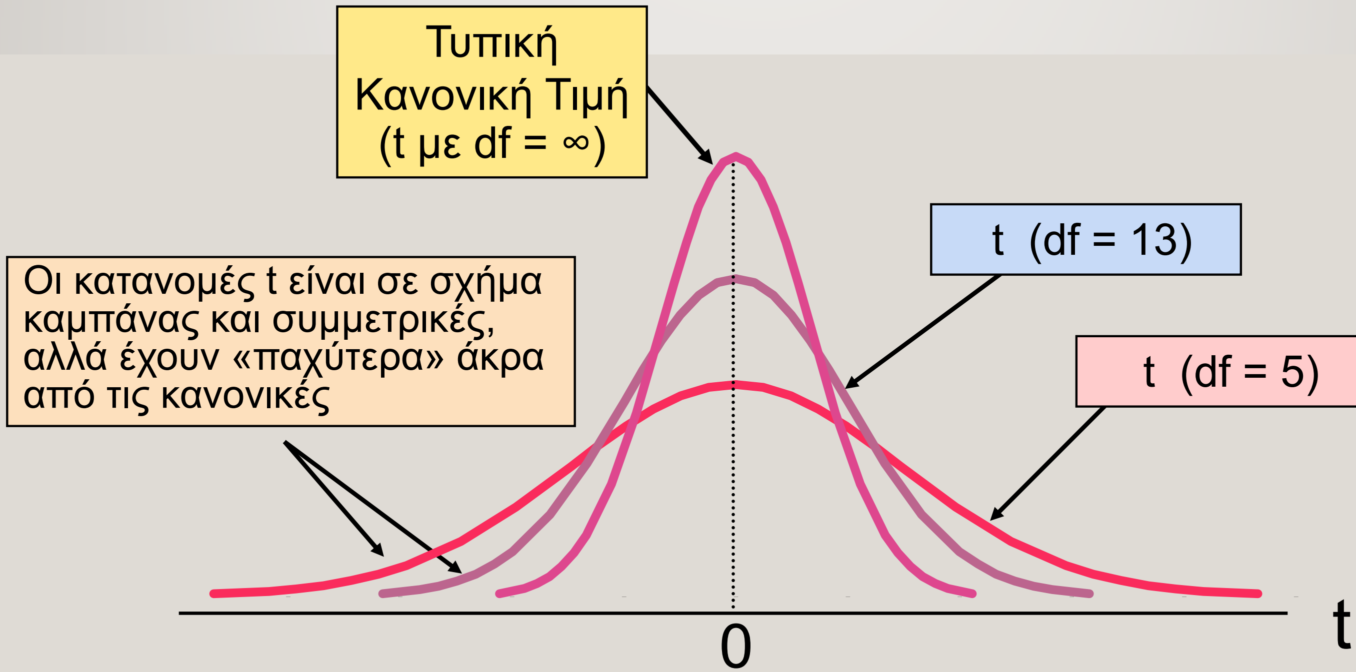
Αν ο μέσος των τριών αριθμών είναι **8**, τότε το X_3 θα είναι **9** (δηλ. το X_3 δεν μπορεί να μεταβάλλεται ελεύθερα)

Εδώ, n = 3, έτσι βαθμοί ελευθερίας = n – 1 = 3 – 1 = 2

(2 τιμές μπορεί να είναι οποιοσδήποτε αριθμός, αλλά ο τρίτος δεν είναι ελεύθερος να μεταβάλλεται για έναν δεδομένο μέσο όρο)

Κατανομή Student t

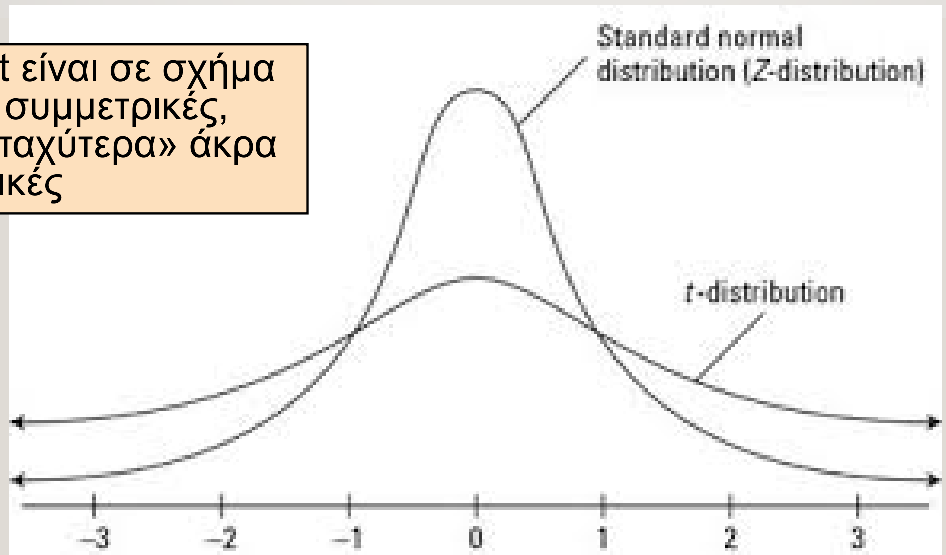
Σημείωση: t $\longrightarrow$ Z όσο το n αυξάνεται



Κατανομή Student t

Σημείωση: t $\longrightarrow$ Z όσο το n αυξάνεται

Οι κατανομές t είναι σε σχήμα καμπάνας και συμμετρικές, αλλά έχουν «παχύτερα» άκρα από τις κανονικές



Διάστημα Εμπιστοσύνης για μ (σ Άγνωστο)

(συνέχεια)

- Υποθέσεις
  - Η τυπική απόκλιση του πληθυσμού είναι **άγνωστη**
  - Ο πληθυσμός κατανέμεται **κανονικά**
    - Εάν ο πληθυσμός δεν είναι κανονικός, χρησιμοποιήστε μεγάλο δείγμα (n> 30)
- Χρησιμοποιήστε την κατανομή Student *t*
- Εκτίμηση Διαστήματος Εμπιστοσύνης:

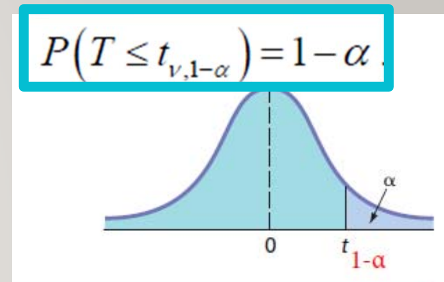
Εκτίμηση Σημείου ± (Κρίσιμη Τιμή)(Τυπικό Σφάλμα)

$$\overline{X} \pm t_{n-1,1-\alpha/2} \frac{S}{\sqrt{n}}$$

(όπου $t_{n-1,a}$ είναι η κρίσιμη τιμή της κατανομής t με n -1 βαθμούς ελευθερίας και μια περιοχή α στη δεξιά ουρά)

Πίνακας Student t

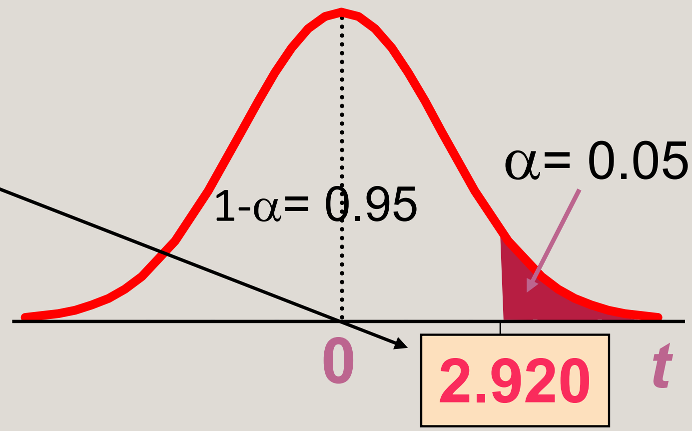
ο Πίνακας που θα δοθεί στις εξετάσεις θα αναφέρεται σε αυτές τις πιθανότητες



	Περιοχή Κάτω Άκρου		
df	.90	**.95**	.975
1	3.078	6.314	12.706
2	1.886	**2.920**	4.303
3	1.638	2.353	3.182

Το σώμα του πίνακα περιέχει τιμές t, **όχι πιθανότητες**

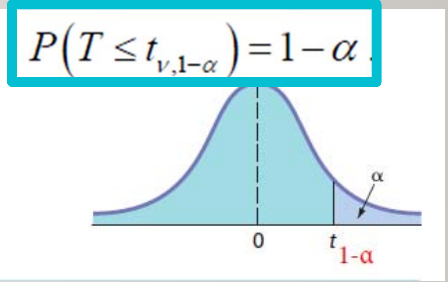
Έστω: n = 3
df = n - 1 = 2
α = 0.05
1-α = 0.95



Πίνακας Student t

Έστω: n = 3
df = *n* - 1 = 2
α = 0.05
1-α = 0.95

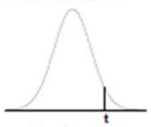
ο Πίνακας που θα δοθεί στις εξετάσεις θα αναφέρεται σε αυτές τις πιθανότητες



	Περιοχή **Κάτω** Άκρου		
df	.90	**.95**	.975
1	3.078	6.314	12.706
2	1.886	**2.920**	4.303
3	1.638	2.353	3.182

Το σώμα του πίνακα περιέχει τιμές t, **όχι πιθανότητες**

Πίνακας 5. Ποσοστιαία Σημεία της Κατανομής *t*



Τα στοιχεία του πίνακα εκφράζουν τα (1-α) ποσοστιαία σημεία της κατανομής t με v βαθμούς ελευθερίας, δηλαδή τις τιμές $t_{v,1-\alpha}$ για τις οποίες $P\left(T \le t_{v,1-\alpha}\right) = 1-\alpha$. Τα κατώτερα ποσοστιαία σημεία προσδιορίζονται από την σχέση $t_{v,\alpha} = -t_{v,1-\alpha}$.

$1-\alpha =$ εμβαδόν

	0,550	0,600	0,650	0,700	0,750	0,800	0,850	0,900	0,950	0,975	0,990	0,995	0,997	0,998	0,999
1	0,158	0,325	0,510	0,727	1,000	1,376	1,963	3,078	6,314	12,706	31,821	63,657	106,100	159,153	318,309
2	0,142	0,289	0,445	0,617	0,816	1,061	1,386	1,886	2,920	4,303	6,965	9,925	12,852	15,764	22,327
3	0,137	0,277	0,424	0,584	0,765	0,978	1,250	1,638	2,353	3,182	4,541	5,841	6,994	8,053	10,215
4	0,134	0,271	0,414	0,569	0,741	0,941	1,190	1,533	2,132	2,776	3,747	4,604	5,321	5,951	7,173
5	0,132	0,267	0,408	0,559	0,727	0,920	1,156	1,476	2,015	2,571	3,365	4,032	4,570	5,030	5,893

Παράδειγμα Διαστήματος Εμπιστοσύνης *t* Κατανομή

Ένα τυχαίο δείγμα n = 25 έχει $\bar{X}$=50 και S = 8.
Δημιουργήστε ένα 95% διάστημα εμπιστοσύνης για το **μ**

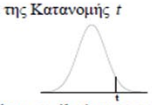
- d.f. = n – 1 = 24
- Κρίσιμη Τιμή $t_{n-1,1-a/2} = t_{24,0.975} = 2.064$

Το διάστημα εμπιστοσύνης είναι

$$\bar{X} \pm t_{n-1,1-{}^{a}/_{2}} \frac{S}{\sqrt{n}} \qquad = 50 \pm 2.064 \frac{8}{\sqrt{25}}$$

$$46.7 \le \mu \le 53.3$$

Πίνακας 5. Ποσοστιαία Σημεία της Κατανομής *t*



Τα στοιχεία του πίνακα εκφράζουν τα (1-α) ποσοστιαία σημεία της κατανομής *t* με *ν* βαθμούς ελευθερίας, δηλαδή τις τιμές $t_{\nu,1-\alpha}$ για τις οποίες $P(T \le t_{\nu,1-\alpha}) = 1-\alpha$. Τα κατώτερα ποσοστιαία σημεία προσδιορίζονται από την σχέση $t_{\nu,\alpha} = -t_{\nu,1-\alpha}$.

$1-\alpha =$ εμβαδόν

	0,550	0,600	0,650	0,700	0,750	0,800	0,850	0,900	0,950	0,975	0,990	0,995	0,997	0,998	0,999
1	0,158	0,325	0,510	0,727	1,000	1,376	1,963	3,078	6,314	12,706	31,821	63,657	106,100	159,153	318,309
2	0,142	0,289	0,445	0,617	0,816	1,061	1,386	1,886	2,920	4,303	6,965	9,925	12,852	15,764	22,327
3	0,137	0,277	0,424	0,584	0,765	0,978	1,250	1,638	2,353	3,182	4,541	5,841	6,994	8,053	10,215
4	0,134	0,271	0,414	0,569	0,741	0,941	1,190	1,533	2,132	2,776	3,747	4,604	5,321	5,951	7,173
5	0,132	0,267	0,408	0,559	0,727	0,920	1,156	1,476	2,015	2,571	3,365	4,032	4,570	5,030	5,893
6	0,131	0,265	0,404	0,553	0,718	0,906	1,134	1,440	1,943	2,447	3,143	3,707	4,152	4,524	5,208
7	0,130	0,263	0,402	0,549	0,711	0,896	1,119	1,415	1,895	2,365	2,998	3,499	3,887	4,207	4,785
8	0,130	0,262	0,399	0,546	0,706	0,889	1,108	1,397	1,860	2,306	2,896	3,355	3,705	3,991	4,501
9	0,129	0,261	0,398	0,543	0,703	0,883	1,100	1,383	1,833	2,262	2,821	3,250	3,573	3,835	4,297
10	0,129	0,260	0,397	0,542	0,700	0,879	1,093	1,372	1,812	2,228	2,764	3,169	3,472	3,716	4,144
11	0,129	0,260	0,396	0,540	0,697	0,876	1,088	1,363	1,796	2,201	2,718	3,106	3,393	3,624	4,025
12	0,128	0,259	0,395	0,539	0,695	0,873	1,083	1,356	1,782	2,179	2,681	3,055	3,330	3,550	3,930
13	0,128	0,259	0,394	0,538	0,694	0,870	1,079	1,350	1,771	2,160	2,650	3,012	3,278	3,489	3,852
14	0,128	0,258	0,393	0,537	0,692	0,868	1,076	1,345	1,761	2,145	2,624	2,977	3,234	3,438	3,787
15	0,128	0,258	0,393	0,536	0,691	0,866	1,074	1,341	1,753	2,131	2,602	2,947	3,197	3,395	3,733
16	0,128	0,258	0,392	0,535	0,690	0,865	1,071	1,337	1,746	2,120	2,583	2,921	3,165	3,358	3,686
17	0,128	0,257	0,392	0,534	0,689	0,863	1,069	1,333	1,740	2,110	2,567	2,898	3,138	3,326	3,646
18	0,127	0,257	0,392	0,534	0,688	0,862	1,067	1,330	1,734	2,101	2,552	2,878	3,113	3,298	3,610
19	0,127	0,257	0,391	0,533	0,688	0,861	1,066	1,328	1,729	2,093	2,539	2,861	3,092	3,273	3,579
20	0,127	0,257	0,391	0,533	0,687	0,860	1,064	1,325	1,725	2,086	2,528	2,845	3,073	3,251	3,552
21	0,127	0,257	0,391	0,532	0,686	0,859	1,063	1,323	1,721	2,080	2,518	2,831	3,056	3,231	3,527
22	0,127	0,256	0,390	0,532	0,686	0,858	1,061	1,321	1,717	2,074	2,508	2,819	3,041	3,214	3,505
23	0,127	0,256	0,390	0,532	0,685	0,858	1,060	1,319	1,714	2,069	2,500	2,807	3,027	3,198	3,485
24	0,127	0,256	0,390	0,531	0,685	0,857	1,059	1,318	1,711	2,064	2,492	2,797	3,014	3,183	3,467
[illegible]	0,127	0,256	0,390	0,531	0,684	0,856	1,058	1,316	1,708	[illegible]	2,485	2,787	3,003	3,170	3,450

Είμαστε 95% βέβαιοι ότι το διάστημα 46.7 έως 53.3 ώρες το οποίο προκύπτει από τη συγκεκριμένη δειγματοληψία, έχει υπολογιστεί με μια μέθοδο που σε βάθος χρόνου παράγει διαστήματα που **θα περιέχουν** την πραγματική τιμή της μέσης τιμής

Παράδειγμα Διαστήματος Εμπιστοσύνης Κατανομής t

$\bar{X}$=50

Ένα τυχαίο δείγμα n = 25 έχει και S = 8.

Δημιουργήστε ένα 95% διάστημα εμπιστοσύνης για το $\boldsymbol{\mu}$

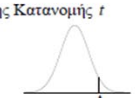
- d.f. = n – 1 = 24
- Κρίσιμη Τιμή $t_{n-1,1-a/2} = t_{24,0.975} = 2.064$

Το διάστημα εμπιστοσύνης είναι

$$\bar{\mathrm{X}} \pm t_{n-1,1-a/2}\frac{S}{\sqrt{n}} = 50 \pm 2.064\frac{8}{\sqrt{25}}$$

$$46.7 \leq \mu \leq 53.3$$

Πίνακας 5. Ποσοστιαία Σημεία της Κατανομής t



Τα στοιχεία του πίνακα εκφράζουν τα (1-α) ποσοστιαία σημεία της κατανομής t με ν βαθμούς ελευθερίας, δηλαδή τις τιμές $t_{\nu,1-\alpha}$ για τις οποίες $P(T \leq t_{\nu,1-\alpha}) = 1-\alpha$. Τα κατώτερα ποσοστιαία σημεία προσδιορίζονται από την σχέση $t_{\nu,\alpha} = -t_{\nu,1-\alpha}$.

$1-\alpha =$ εμβαδόν

	0,550	0,600	0,650	0,700	0,750	0,800	0,850	0,900	0,950	0,975	0,990	0,995	0,997	0,998	0,999
1	0,158	0,325	0,510	0,727	1,000	1,376	1,963	3,078	6,314	12,706	31,821	63,657	106,100	159,153	318,309
2	0,142	0,289	0,445	0,617	0,816	1,061	1,386	1,886	2,920	4,303	6,965	9,925	12,852	15,764	22,327
3	0,137	0,277	0,424	0,584	0,765	0,978	1,250	1,638	2,353	3,182	4,541	5,841	6,994	8,053	10,215
4	0,134	0,271	0,414	0,569	0,741	0,941	1,190	1,533	2,132	2,776	3,747	4,604	5,321	5,951	7,173
5	0,132	0,267	0,408	0,559	0,727	0,920	1,156	1,476	2,015	2,571	3,365	4,032	4,570	5,030	5,893
6	0,131	0,265	0,404	0,553	0,718	0,906	1,134	1,440	1,943	2,447	3,143	3,707	4,152	4,524	5,208
7	0,130	0,263	0,402	0,549	0,711	0,896	1,119	1,415	1,895	2,365	2,998	3,499	3,887	4,207	4,785
8	0,130	0,262	0,399	0,546	0,706	0,889	1,108	1,397	1,860	2,306	2,896	3,355	3,705	3,991	4,501
9	0,129	0,261	0,398	0,543	0,703	0,883	1,100	1,383	1,833	2,262	2,821	3,250	3,573	3,835	4,297
10	0,129	0,260	0,397	0,542	0,700	0,879	1,093	1,372	1,812	2,228	2,764	3,169	3,472	3,716	4,144
11	0,129	0,260	0,396	0,540	0,697	0,876	1,088	1,363	1,796	2,201	2,718	3,106	3,393	3,624	4,025
12	0,128	0,259	0,395	0,539	0,695	0,873	1,083	1,356	1,782	2,179	2,681	3,055	3,330	3,550	3,930
13	0,128	0,259	0,394	0,538	0,694	0,870	1,079	1,350	1,771	2,160	2,650	3,012	3,278	3,489	3,852
14	0,128	0,258	0,393	0,537	0,692	0,868	1,076	1,345	1,761	2,145	2,624	2,977	3,234	3,438	3,787
15	0,128	0,258	0,393	0,536	0,691	0,866	1,074	1,341	1,753	2,131	2,602	2,947	3,197	3,395	3,733
16	0,128	0,258	0,392	0,535	0,690	0,865	1,071	1,337	1,746	2,120	2,583	2,921	3,165	3,358	3,686
17	0,128	0,257	0,392	0,534	0,689	0,863	1,069	1,333	1,740	2,110	2,567	2,898	3,138	3,326	3,646
18	0,127	0,257	0,392	0,534	0,688	0,862	1,067	1,330	1,734	2,101	2,552	2,878	3,113	3,298	3,610
19	0,127	0,257	0,391	0,533	0,688	0,861	1,066	1,328	1,729	2,093	2,539	2,861	3,092	3,273	3,579
20	0,127	0,257	0,391	0,533	0,687	0,860	1,064	1,325	1,725	2,086	2,528	2,845	3,073	3,251	3,552
21	0,127	0,257	0,391	0,532	0,686	0,859	1,063	1,323	1,721	2,080	2,518	2,831	3,056	3,231	3,527
22	0,127	0,256	0,390	0,532	0,686	0,858	1,061	1,321	1,717	2,074	2,508	2,819	3,041	3,214	3,505
23	0,127	0,256	0,390	0,532	0,685	0,858	1,060	1,319	1,714	2,069	2,500	2,807	3,027	3,198	3,485
24	0,127	0,256	0,390	0,531	0,685	0,857	1,059	1,318	1,711	2,064	2,492	2,797	3,014	3,183	3,467
[illegible]	0,127	0,256	0,390	0,531	0,684	0,856	1,058	1,316	1,708	[illegible]	2,485	2,787	3,003	3,170	3,450

Είμαστε **95%** βέβαιοι ότι το διάστημα **46.7** έως **53.3** ώρες το οποίο προκύπτει από τη συγκεκριμένη δειγματοληψία, έχει υπολογιστεί με μια μέθοδο που σε βάθος χρόνου παράγει διαστήματα που **θα περιέχουν** την πραγματική τιμή της μέσης τιμής

Παράδειγμα Διαστήματος Εμπιστοσύνης Κατανομής t

(συνέχεια)

- Η ερμηνεία αυτού του διαστήματος **απαιτεί** την υπόθεση ότι ο πληθυσμός από τον οποίο γίνεται η δειγματοληψία είναι περίπου μια κανονική κατανομή (ειδικά επειδή το ***n*** είναι μόνο 25).
- Η κατάσταση αυτή σε πραγματικά δεδομένα θα πρέπει να ελεγχθεί (πχ με την δημιουργία ενός Διαγράμματος QQ plot)

1^{o} Πρόβλημα

Ένα μηχάνημα κατασκευάζει μπαταρίες για ειδικές χρήσεις των οποίων ο χρόνος ζωής ακολουθεί, σύμφωνα με τις τεχνικές προδιαγραφές του μηχανήματος, κανονική κατανομή. Επειδή στις χρήσεις για τις οποίες προορίζονται οι μπαταρίες ο χρόνος ζωής τους είναι ιδιαίτερα κρίσιμος, η λειτουργία του μηχανήματος παρακολουθείται συνεχώς με δειγματοληπτικό έλεγχο του χρόνου ζωής των παραγόμενων μπαταριών. Από τυχαίο δείγμα 9 μπαταριών διαπιστώνεται ότι οι χρόνοι ζωής τους (σε ώρες) είναι:

16	18	22	20	21	20	16	22	25

Με τη βοήθεια της πληροφορίας αυτής να εκτιμηθεί ένα 95% διάστημα εμπιστοσύνης για το μέσο χρόνο ζωής του συνόλου των παραγόμενων μπαταριών.

Λύση

Για να σχηματίσουμε το 95% διάστημα εμπιστοσύνης για το μέσο χρόνο ζωής των μπαταριών
Βρίσκουμε ότι η τιμή του δειγματικού μέσου είναι

$$\overline{x} = \frac{\sum_{i=1}^{n} X_i}{n} = \frac{\sum_{i=1}^{9} X_i}{9} = \frac{16+18+...+25}{9} = \frac{180}{9} = 20 ώρες$$

Δεδομένου ότι η πληθυσμιακή διακύμανση είναι άγνωστη, θα πρέπει να εκτιμηθεί από το δείγμα.

Η δειγματική τυπική απόκλιση είναι

$$s = \sqrt{\frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{n-1}} = \sqrt{\frac{\sum_{i=1}^{9}(X_i - \overline{X})^2}{9-1}} = \sqrt{\frac{(16-20)^2 + (18-20)^2 + ... + (25-20)^2}{8}} = 2,958$$

και, κατά συνέπεια, το τυπικό σφάλμα εκτίμησης της μέσης τιμής είναι $\sigma_{\overline{X}} = s/\sqrt{n} = 2,958/\sqrt{9} = 0,986$

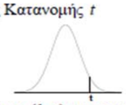
$$\alpha = 0,05 \rightarrow a/2 = 0,025 \rightarrow 1 - a/2 = 0,975$$

Λύση

$$\bar{X} \pm t_{n-1,1-\alpha/2}\frac{S}{\sqrt{n}}$$

$$\alpha = 0{,}05 \rightarrow a/2 = 0{,}025 \rightarrow 1 - a/2 = 0{,}975$$

Πίνακας 5. Ποσοστιαία Σημεία της Κατανομής t



Τα στοιχεία του πίνακα εκφράζουν τα (1-α) ποσοστιαία σημεία της κατανομής t με ν βαθμούς ελευθερίας, δηλαδή τις τιμές $t_{\nu,1-\alpha}$ για τις οποίες $P(T \le t_{\nu,1-\alpha}) = 1-\alpha$. Τα κατώτερα ποσοστιαία σημεία προσδιορίζονται από την σχέση $t_{\nu,\alpha} = -t_{\nu,1-\alpha}$.

$1-\alpha =$ εμβαδόν

	0,550	0,600	0,650	0,700	0,750	0,800	0,850	0,900	0,950	0,975	0,990	0,995	0,997	0,998	0,999
1	0,158	0,325	0,510	0,727	1,000	1,376	1,963	3,078	[illegible]	12,706	31,821	63,657	106,100	159,153	318,309
2	0,142	0,289	0,445	0,617	0,816	1,061	1,386	1,886	2,920	4,303	6,965	9,925	12,852	15,764	22,327
3	0,137	0,277	0,424	0,584	0,765	0,978	1,250	1,638	2,353	3,182	4,541	5,841	6,994	8,053	10,215
4	0,134	0,271	0,414	0,569	0,741	0,941	1,190	1,533	2,132	2,776	3,747	4,604	5,321	5,951	7,173
5	0,132	0,267	0,408	0,559	0,727	0,920	1,156	1,476	2,015	2,571	3,365	4,032	4,570	5,030	5,893
6	0,131	0,265	0,404	0,553	0,718	0,906	1,134	1,440	1,943	2,447	3,143	3,707	4,152	4,524	5,208
7	0,130	0,263	0,402	0,549	0,711	0,896	1,119	1,415	1,895	2,365	2,998	3,499	3,887	4,207	4,785
8	0,130	0,262	0,399	0,546	0,706	0,889	1,108	1,397	1,860	2,306	2,896	3,355	3,705	3,991	4,501
9	0,129	0,261	0,398	0,543	0,703	0,883	1,100	1,383	1,833	2,262	2,821	3,250	3,573	3,835	4,297
10	0,129	0,260	0,397	0,542	0,700	0,879	1,093	1,372	1,812	2,228	2,764	3,169	3,472	3,716	4,144
11	0,129	0,260	0,396	0,540	0,697	0,876	1,088	1,363	1,796	2,201	2,718	3,106	3,393	3,624	4,025
12	0,128	0,259	0,395	0,539	0,695	0,873	1,083	1,356	1,782	2,179	2,681	3,055	3,330	3,550	3,930
13	0,128	0,259	0,394	0,538	0,694	0,870	1,079	1,350	1,771	2,160	2,650	3,012	3,278	3,489	3,852
14	0,128	0,258	0,393	0,537	0,692	0,868	1,076	1,345	1,761	2,145	2,624	2,977	3,234	3,438	3,787
15	0,128	0,258	0,393	0,536	0,691	0,866	1,074	1,341	1,753	2,131	2,602	2,947	3,197	3,395	3,733
16	0,128	0,258	0,392	0,535	0,690	0,865	1,071	1,337	1,746	2,120	2,583	2,921	3,165	3,358	3,686
17	0,128	0,257	0,392	0,534	0,689	0,863	1,069	1,333	1,740	2,110	2,567	2,898	3,138	3,326	3,646
18	0,127	0,257	0,392	0,534	0,688	0,862	1,067	1,330	1,734	2,101	2,552	2,878	3,113	3,298	3,610
19	0,127	0,257	0,391	0,533	0,688	0,861	1,066	1,328	1,729	2,093	2,539	2,861	3,092	3,273	3,579
20	0,127	0,257	0,391	0,533	0,687	0,860	1,064	1,325	1,725	2,086	2,528	2,845	3,073	3,251	3,552
21	0,127	0,257	0,391	0,532	0,686	0,859	1,063	1,323	1,721	2,080	2,518	2,831	3,056	3,231	3,527
22	0,127	0,256	0,390	0,532	0,686	0,858	1,061	1,321	1,717	2,074	2,508	2,819	3,041	3,214	3,505
23	0,127	0,256	0,390	0,532	0,685	0,858	1,060	1,319	1,714	2,069	2,500	2,807	3,027	3,198	3,485
24	0,127	0,256	0,390	0,531	0,685	0,857	1,059	1,318	1,711	2,064	2,492	2,797	3,014	3,183	3,467
25	0,127	0,256	0,390	0,531	0,684	0,856	1,058	1,316	1,708	2,060	2,485	2,787	3,003	3,170	3,450

$$t_{n-1,1-\alpha/2} = t_{8,0,975} = 2{,}306$$

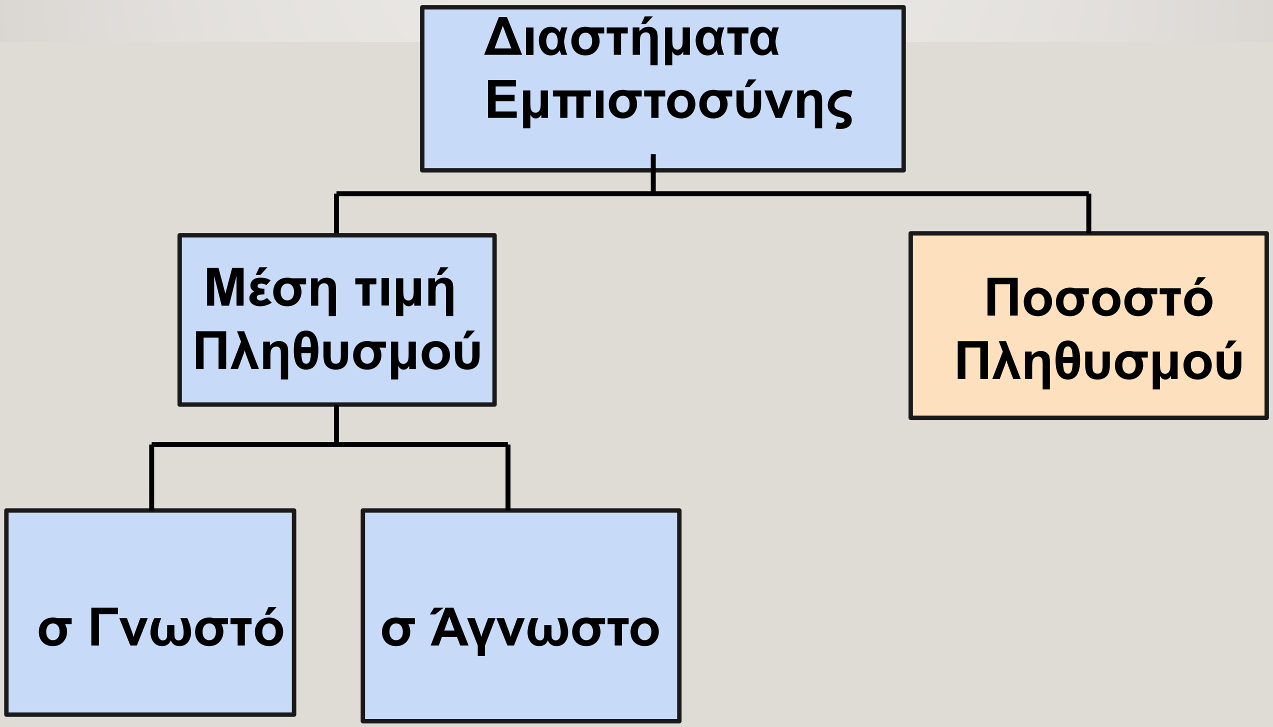
$$\bar{X} \pm t_{n-1,\alpha/2}\frac{S}{\sqrt{n}}$$

$$20 \pm 2{,}306\,\frac{2{,}958}{\sqrt{9}}$$

95% Δ.Ε. για τη μέση τιμή (17.73,22.27)

Είμαστε 95% βέβαιοι ότι το διάστημα 17,73.7 έως 22.27 ώρες το οποίο προκύπτει από τη συγκεκριμένη δειγματοληψία, έχει υπολογιστεί με μια μέθοδο που σε βάθος χρόνου παράγει διαστήματα που **θα περιέχουν** τον πραγματικό μέσο χρόνο ζωής του συνόλου των παραγόμενων μπαταριών της εταιρείας

Διαστήματα Εμπιστοσύνης



Διαστήματα Εμπιστοσύνης για το Ποσοστό Πληθυσμού, π

Μια εκτίμηση διαστήματος για το ποσοστό του πληθυσμού (***π***) μπορεί να υπολογιστεί με την προσθήκη μιας απαίτησης για αβεβαιότητα στο ποσοστό του δείγματος (***p***)

Διαστήματα Εμπιστοσύνης για το Ποσοστό Πληθυσμού, π

(συνέχεια)

- Θυμηθείτε ότι η κατανομή του ποσοστού του δείγματος είναι περίπου κανονική, εάν το μέγεθος του δείγματος είναι μεγάλο ($np > 5$ και $n(1-p) > 5$), με τυπική απόκλιση

$$\sigma_p = \sqrt{\frac{\pi(1-\pi)}{n}}$$

- Θα χρησιμοποιήσουμε δείγματα δεδομένων: $\sqrt{\frac{p(1-p)}{n}}$

Άκρα Διαστήματος Εμπιστοσύνης

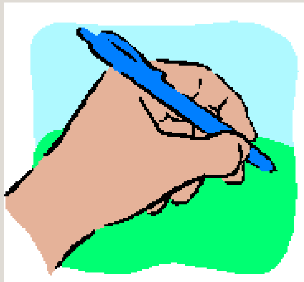
- Τα ανώτερα και κατώτερα όρια εμπιστοσύνης για το ποσοστό του πληθυσμού υπολογίζονται με τον μαθηματικό τύπο

$$p \pm Z_{1-\alpha/2}\sqrt{\frac{p(1-p)}{n}}$$

- όπου
  - $Z_{1-\alpha/2}$ είναι η τυπική κανονική τιμή για το επιθυμητό επίπεδο εμπιστοσύνης
  - p: είναι το ποσοστό του δείγματος
  - n: είναι το μέγεθος του δείγματος
- Σημείωση: πρέπει να ισχύει $np > 5$ και $n(1-p) > 5$

Παράδειγμα

- Ένα τυχαίο δείγμα 100 ατόμων δείχνει ότι τα 25 είναι αριστερόχειρες
- Δημιουργήστε ένα διάστημα εμπιστοσύνης 95% για το πραγματικό ποσοστό των αριστερόχειρων



Παράδειγμα

(συνέχεια)

Ένα τυχαίο δείγμα 100 ατόμων δείχνει ότι τα 25 είναι αριστερόχειρες. Δημιουργήστε ένα διάστημα εμπιστοσύνης 95% για το πραγματικό ποσοστό των αριστερόχειρων.

np=100*25/100=25 > 5

και n(1-p) > 5 (100*0,75=75>5)

$$p \pm Z_{1-\alpha/2}\sqrt{p(1-p)/n}$$

$$=25/100 \pm Z_{0.975}\sqrt{0.25(0.75)/100}$$

$$= 25/100 \pm 1.96\sqrt{0.25(0.75)/100}$$

$$= 0.25 \pm 1.96(0.0433)$$

$$\boxed{= 0.1651 \le \pi \le 0.3349}$$

Παράδειγμα

(συνέχεια)

Ένα τυχαίο δείγμα 100 ατόμων δείχνει ότι τα 25 είναι αριστερόχειρες. Δημιουργήστε ένα διάστημα εμπιστοσύνης 95% για το πραγματικό ποσοστό των αριστερόχειρων.

np=100*25/100=25 > 5
και n(1-p) > 5 (100*0,75=75>5)

$$p \pm Z^*_{\alpha/2}\sqrt{p(1-p)/n} = p \pm Z^*_{0.025}\sqrt{p(1-p)/n} =$$

$$p \pm Z_{1-0.025}\sqrt{p(1-p)/n} = p \pm Z_{0.975}\sqrt{p(1-p)/n}$$

$$=25/100 \pm Z_{0.975}\sqrt{0.25(0.75)/100}$$

$$= 25/100 \pm 1.96\sqrt{0.25(0.75)/100}$$

$$= 0.25 \pm 1.96(0.0433)$$

$$\boxed{= 0.1651 \le \pi \le 0.3349}$$

Ερμηνεία

- Είμαστε 95% βέβαιοι ότι το διάστημα τιμών 16.51% έως 33.49% περιέχει το πραγματικό ποσοστό των αριστερόχειρων στον πληθυσμό

- Αν και το διάστημα από 0.1651 έως 0.3349 μπορεί ή όχι να περιέχει το πραγματικό ποσοστό, το 95% των διαστημάτων που σχηματίζονται από δείγματα μεγέθους 100 με τον τρόπο αυτό θα περιέχουν το πραγματικό ποσοστό.

Πρόβλημα

Ο υπεύθυνος ποιοτικού ελέγχου μιας εταιρείας θέλει να εκτιμήσει το ποσοστό των ελαττωματικών συσκευασιών. Για το λόγο αυτόν επιλέγει ένα τυχαίο δείγμα 200 συσκευασιών και διαπιστώνει ότι 11 από αυτές ήταν ελαττωματικές. Να εκτιμηθεί ένα 95% διάστημα εμπιστοσύνης για το ποσοστό των ελαττωματικών συσκευασιών.

Λύση

Δίνεται ότι n = 200 και X = 11.

Παρατηρούμε ότι $p = 11/200 = 0,055$ και $1 - p = 0.945$ επομένως np=200*0,055=11 > 5 και n(1-p) > 5 (200*0,945=189)

το τυπικό σφάλμα εκτίμησης του ποσοστού είναι:

$$\sigma_p = \sqrt{\frac{p \cdot (1-p)}{200}} = \sqrt{\frac{0,055 \cdot 0,945}{200}} = 0,016$$

Λύση

$1-\alpha = 95\% \Leftrightarrow \alpha = 5\%$

$Z_{a/2} = Z_{0,025} = 1,96 \quad ή \quad Z_{1-a/2} = Z_{0,975} = 1,96$

$p = 11/200 = 0,055$

$q = 1 - \hat{p} = 0,945$

$$p \pm Z_{1-0.025}\sqrt{p(1-p)/n} = p \pm Z_{0.975}\sqrt{p(1-p)/n} =$$

$$0.055 \pm 1.96\sqrt{0.055*0.945/200}$$

$$0.023 \leq \pi \leq 0.086$$

Επομένως, τα άκρα του 95% διαστήματος εμπιστοσύνης για το ποσοστό των ελαττωματικών συσκευασιών είναι 0.023 και 0.086. Είμαστε **95%** βέβαιοι ότι το διάστημα 2,3% και 8,6% περιέχει την πραγματική τιμή του ποσοστού των ελαττωματικών συσκευασιών.

Διαφορά μέσης τιμής μ_1-μ_2 , σ^2 γνωστές Κανονικοί πληθυσμοί Ανεξάρτητα δείγματα

$$\bar{X}_1 - \bar{X}_2 \pm Z_{1-\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

Σε μία έρευνα οι ερευνητές θέλουν να ελέγξουν αν τα δεδομένα που συλλέχθηκαν παρέχουν επαρκή στατιστική ένδειξη διαφοράς στη μέση τιμή του ουρικού οξέος μεταξύ ατόμων χωρίς και με σύνδρομο Down. Τα δεδομένα αποτελούνται από 12 μετρήσεις σε άτομα με σύνδρομο Down και 15 χωρίς σύνδρομο Down. Οι μέσοι όροι για τα δύο δείγματα είναι αντίστοιχα, 3.4 mg/100ml και 4.5 mg/100ml ενώ οι πληθυσμιακές διασπορές είναι 1 και 1.5. Τα δεδομένα αποτελούν ανεξάρτητα δείγματα από κανονικούς πληθυσμούς. Να εκτιμηθεί ένα 95% διάστημα εμπιστοσύνης για τη διαφορά του ουρικού οξέος στους δύο πληθυσμούς

$$\sigma_1^2 = 1.5, \sigma_2^2 = 1.0$$

Πληθυσμοί Κανονικοί, $n_1=15, n_2=12$

$\alpha=5\%$

$$\bar{X}_1 - \bar{X}_2 \pm Z_{1-\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}} = \bar{X}_1 - \bar{X}_2 \pm Z_{0,025}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}} = (4.5-3.4) \pm 1.960.4282$$

95% Διάστημα εμπιστοσύνης για την **διαφορά των πληθυσμιακών μέσων τιμών μ1-μ2** **(χωρίς σύνδρομο- με σύνδρομο Down) είναι (0.26,1.94)**

Είμαστε **95%** βέβαιοι ότι το διάστημα 0.26 έως 1.94 περιέχει την πραγματική μέση τιμή της διαφοράς του ουρικού οξέος στους δύο πληθυσμούς