Πιθανότητες και Στατιστική – Λύσεις Θεμάτων

ΠΡΟΣΟΧΗ! ΟΙ ΛΥΣΕΙΣ ΕΝΔΕΧΕΤΑΙ ΝΑ ΜΗΝ ΕΙΝΑΙ 100% ΟΡΘΕΣ

Ασκήσεις Θεωρίας

1. **Αν σε ένα στατιστικό έλεγχο υποθέσεων, έχουμε υπολογίσει το P_{value} και με βάση αυτό η μηδενική υπόθεση απορρίπτεται για επίπεδο σημαντικότητας 5%, τότε απορρίπτεται ή όχι και για επίπεδο σημαντικότητας 10%; Δικαιολογήστε την απάντησή σας (0,5 μονάδα).**

   Αν η μηδενική υπόθεση απορρίπτεται για επίπεδο σημαντικότητας 5%, τότε αυτό σημαίνει ότι: p-value<0,05 Εφόσον 0,05 < 0,10, τότε σίγουρα ισχύει και ότι: p-value<0,10 Άρα, η μηδενική υπόθεση θα απορριφθεί και για επίπεδο σημαντικότητας 10%.

2. **Σε 10 γραμμές το πολύ εξηγήστε ποια τεχνική δειγματοληψίας θα χρησιμοποιούσατε και γιατί, εάν θέλατε να ελέγξετε το βάρος των παιδιών σε ένα δημοτικό σχολείο.**

   Η κατάλληλη τεχνική δειγματοληψίας σε αυτή την περίπτωση είναι η στρωματοποιημένη δειγματοληψία. Σκοπός είναι να διασφαλιστεί ότι όλα τα τμήματα του πληθυσμού (π.χ. τάξεις ή ηλικιακές ομάδες) εκπροσωπούνται σωστά στο δείγμα. Θα χωρίσουμε τα παιδιά σε στρώματα (π.χ. τάξη Α', Β', Γ' έως ΣΤ') και από κάθε στρώμα θα πάρουμε τυχαίο δείγμα. Έτσι αποφεύγουμε την υπερεκπροσώπηση ή υποεκπροσώπηση συγκεκριμένων ηλικιών, οι οποίες επηρεάζουν σημαντικά το βάρος. Με αυτόν τον τρόπο, τα αποτελέσματα θα είναι πιο αξιόπιστα και αντιπροσωπευτικά του συνολικού πληθυσμού του σχολείου.

3. **Αναφέρετε την ερμηνεία του συντελεστή εμπιστοσύνης τόσο σε ένα διάστημα εμπιστοσύνης, όσο και στον έλεγχο υποθέσεων.**

   Ο συντελεστής εμπιστοσύνης (π.χ. 95%) εκφράζει το επίπεδο σιγουριάς που έχουμε για τα συμπεράσματά μας.
   - Σε ένα διάστημα εμπιστοσύνης: Ο συντελεστής εμπιστοσύνης δείχνει ότι αν επαναλαμβάναμε το ίδιο πείραμα πολλές φορές, το 95% των υπολογιζόμενων διαστημάτων εμπιστοσύνης θα περιείχαν την πραγματική τιμή του πληθυσμιακού παραμέτρου (π.χ. μέση τιμή).
   - Στον έλεγχο υποθέσεων: Ο συντελεστής εμπιστοσύνης είναι συμπληρωματικός του επιπέδου σημαντικότητας (α). Δηλαδή, για συντελεστή εμπιστοσύνης 95%, το επίπεδο σημαντικότητας είναι 5% (α = 0,05), και εκφράζει την πιθανότητα να μην απορρίψουμε τη μηδενική υπόθεση όταν αυτή είναι αληθής (μικρότερος κίνδυνος τύπου Ι σφάλματος).

4. **Τα τελευταία 15 χρόνια καταγράφηκαν 20 θαλάσσιοι πνιγμοί σε ένα πολυσύχναστο beach bar της Χαλκιδικής. Ο χρόνος που μεσολαβεί μεταξύ διαδοχικών πνιγμών σε αυτό το beach bar μπορεί να περιγραφεί από την ...**
   (i) εκθετική κατανομή
   (ii) γεωμετρική κατανομή
   (iii) κανονική κατανομή
   (iv) κατανομή Poisson.

   Η **εκθετική κατανομή** χρησιμοποιείται για τη μοντελοποίηση του χρόνου που μεσολαβεί μεταξύ διαδοχικών γεγονότων σε μια διαδικασία που ακολουθεί κατανομή Poisson, όπως στην περίπτωση πνιγμών που συμβαίνουν σποραδικά και ανεξάρτητα. Εφόσον τα πνιγμοί είναι σπάνια γεγονότα που καταγράφονται με βάση το χρόνο, η εκθετική κατανομή είναι κατάλληλη για το μεσοδιάστημα μεταξύ τους.

5. **Δώστε τους ακόλουθους ορισμούς: α) ανεξάρτητα γεγονότα, β) συμπληρωματικά γεγονότα, κάνοντας και σχόλιο για το πλήθος των δοκιμών που αφορούν (ή που σχετίζονται με αυτά)**

   Δύο γεγονότα Α και Β λέγονται ανεξάρτητα αν η εμφάνιση του ενός δεν επηρεάζει την πιθανότητα εμφάνισης του άλλου. Δηλαδή, $\boldsymbol{P(A \cap B) = P(A) \cdot P(B).}$ Σε 2 δοκιμές.

Το γεγονός να μην συμβεί το γεγονός Α λέγεται συμπληρωματικό γεγονός και συμβολίζεται με Α΄, δηλαδή $\boldsymbol{P(A') = 1 - P(A)}$. Σε 1 (στην ίδια δοκιμή)

6. **Δώστε την κατάλληλη απάντηση (ΣΩΣΤΗ ή ΛΑΘΟΣ) στις παρακάτω προτάσεις / φράσεις, αιτιολογώντας όλες τις απαντήσεις σας.**

Σε περίπτωση που μια αεροπορική εταιρία πουλά 75 εισιτήρια για όλες τις πτήσεις της που γίνονται με αεροσκάφη 70 θέσεων (επειδή διαπίστωσε ότι κατά μέσο όρο το 5% των επιβατών με κρατημένες θέσεις δεν εμφανίζονται κατά την αναχώρηση του αεροπλάνου), τότε ο αριθμός των επιβατών που δεν προσέρχονται σε μία πτήση ακολουθεί διωνυμική κατανομή με παραμέτρους p = 0.05 και n = 70. ΛΑΘΟΣ

Η πρόταση λέει ότι η κατανομή είναι διωνυμική με p = 0.05 και n = 70. Αυτό θα σήμαινε ότι μόνο 70 άτομα είχαν εισιτήριο, αλλά στην πραγματικότητα 75 έχουν εισιτήριο και καθένας μπορεί να μην εμφανιστεί. Ο αριθμός «70» αναφέρεται στις θέσεις του αεροσκάφους, όχι στον αριθμό των εισιτηρίων που εκδόθηκαν. Για τη διωνυμική κατανομή μάς ενδιαφέρει ο αριθμός των επιβατών που έκαναν κράτηση (75), όχι οι διαθέσιμες θέσεις.

Όταν η αντίσταση R μιας μεταλλικής ράβδου δίνεται από την ακόλουθη συνάρτηση πυκνότητας πιθανότητας: f(x)=2x για 0 <= x <= 1, τότε η πιθανότητα P(x >= 0,5) ισούται με 0,5. ΛΑΘΟΣ

Αρχικά πρέπει να ισχύει: $\int_0^1 f(x)\,dx = 1 \rightarrow \int_0^1 2x\,dx = [x^2]_0^1 = 1$ **που ισχύει**

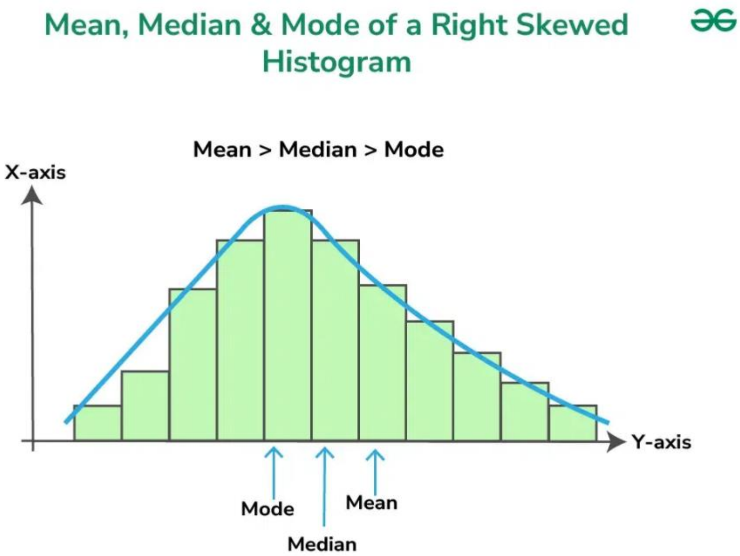
$$\boldsymbol{P(X \geq 0.5) = \int_{0,5}^{1} f(x)\,dx = \int_{0,5}^{1} 2x\,dx}$$

$$\int_{0,5}^{1} \boldsymbol{2x\,dx} = [x^2]_{0.5}^{1} = 1 - 0.25 = 0.75$$

Άρα ισούται με 0.75

Εάν το ιστόγραμμα μιας σειράς δεδομένων είναι λοξό προς τα δεξιά, τότε η μεσαία τιμή τους είναι μεγαλύτερη από τη μέση τιμή τους (δείξτε με κατάλληλο σχήμα). ΛΑΘΟΣ

Όταν ένα ιστόγραμμα είναι λοξό προς τα δεξιά (right–skewed): Η "ουρά" των δεδομένων εκτείνεται προς τις μεγάλες τιμές. Αυτό σημαίνει ότι μερικές μεγάλες τιμές "τραβούν" τη μέση τιμή (mean) προς τα δεξιά. Η διάμεσος (median) δεν επηρεάζεται τόσο από τις ακραίες τιμές, γιατί εξαρτάται από τη θέση των παρατηρήσεων, όχι από τα μεγέθη τους. Η τυπική σχέση είναι: Μέση τιμή (mean) > Διάμεσος (median)> Συχνότερη τιμή (mode)



7. **Σε μια κανονική κατανομή, τι ποσοστό των τιμών βρίσκεται στο διάστημα +1 και -1 τυπικών αποκλίσεων από τη μέση τιμή;**
1. 68,27%

2. 95.45%
3. 97.73%
4. 50%

Σύμφωνα με τον κανόνα των 68-95-99.7 για την κανονική κατανομή:
- ±1σ από το μ: 68,27% των τιμών
- ±2σ από το μ: 95,45%
- ±3σ από το μ: 99,73%

8. **Ποια είναι η παράμετρος «θέσης» (ενός στατικού μεγέθους) που επηρεάζεται λιγότερο από ακραίες τιμές;**
   1. Μέση τιμή
   2. Συχνότερη τιμή
   3. Τυπική απόκλιση
   4. Διάμεσος/μεσαία τιμή

   Μέση τιμή: Επηρεάζεται πολύ από ακραίες τιμές (π.χ. outliers).
   Συχνότερη τιμή (mode): Μπορεί να επηρεαστεί, αλλά δεν είναι αξιόπιστος δείκτης θέσης.
   Τυπική απόκλιση: Δεν είναι μέτρο θέσης, αλλά διασποράς.
   Διάμεσος: Είναι ανθεκτικός δείκτης θέσης, δεν επηρεάζεται σχεδόν καθόλου από ακραίες τιμές γιατί εξαρτάται από τη σειρά των δεδομένων, όχι το μέγεθός τους.

9. **Με τι ισούται η μεταβλητότητα/διακύμανση ενός δείγματος, όταν όλες οι τιμές του είναι ίδιες;**
   1. s^2
   2. 1
   3. ∞
   4. 0

   Όταν όλες οι τιμές είναι ίδιες, τότε δεν υπάρχει καμία απόκλιση από τη μέση τιμή.

10. **Εάν τα γεγονότα A και B είναι ασυμβίβαστα/αμοιβαίως αποκλειόμενα, ποια είναι η πιθανότητα $P(A \cup B)$;**
    1. $P(A) + P(B) - P(A \cap B)$
    2. $P(A) + P(B)$
    3. $P(A) - P(B)$
    4. $P(A) \times P(B)$

    Γενικά, η πιθανότητα της ένωσης δύο γεγονότων είναι: $\boldsymbol{P(A) + P(B) - P(A \cap B)}$**,** αλλά αν τα γεγονότα είναι ασυμβίβαστα (δηλαδή δεν μπορούν να συμβούν ταυτόχρονα), τότε: θα ισχύει ότι επειδή $P(A \cap B) = 0 \rightarrow$ $P(A \cup B) = P(A) + P(B) - P(A \cap B) = P(A) + P(B) - 0 = \boldsymbol{P(A) + P(B)}$

11. **Με τη χρήση ποιας κατανομής θα βρούμε την πιθανότητα να προσέλθουν 10 αυτοκίνητα σε ένα βενζινάδικο το επόμενο 15λεπτο;**
    1. Διωνυμική κατανομή.
    2. Κατανομή Poisson.
    3. Υπεργεωμετρική κατανομή.
    4. Γεωμετρική κατανομή.

    Η κατανομή Poisson χρησιμοποιείται όταν:
    - Μετράμε τον αριθμό γεγονότων (π.χ. αφίξεις αυτοκινήτων) σε συνεχές διάστημα (χρόνος, χώρος).
    - Τα γεγονότα είναι ανεξάρτητα και συμβαίνουν με σταθερό μέσο ρυθμό.

    Αυτό ισχύει στην ερώτηση (προσέλευση 10 αυτοκινήτων σε 15 λεπτά).

Ασκήσεις

Στο πλαίσιο μιας έρευνας σχετικά με τις συνήθειες των φοιτητών και διαπιστώθηκε ότι οι ώρες ανάπαυσής τους ημερησίως ακολουθούν κανονική κατανομή με μέση τιμή 7,389 ώρες και τυπική απόκλιση 1,67 ώρες. (α) Ποια είναι η πιθανότητα οι ώρες ανάπαυσης ενός φοιτητή να μην ξεπερνούν τις 4,5 ώρες; (β) Βρείτε το πλήθος των ωρών πάνω από τις οποίες αναπαύεται μόνο το 5% των φοιτητών. (γ) Ποια είναι η πιθανότητα η δειγματική μέση τιμή των ωρών ανάπαυσης ενός τυχαίου δείγματος 9 φοιτητών να είναι μεταξύ 6 και 8 ωρών;

<u>ΛΥΣΗ</u>

Έχουμε κανονική κατανομή (από εκφώνηση)

Μέση τιμή: μ = 7,389 ώρες
Τυπική απόκλιση: σ = 1,67 ώρες

α) Υπολογίζουμε $\boldsymbol{P(X \leq 4,5)}$

Κάνουμε μετασχηματισμό σε τυποποιημένη κανονική κατανομή Z

$$\boldsymbol{Z} = \frac{\boldsymbol{X-\mu}}{\boldsymbol{\sigma}} = \frac{4{,}5 - 7{,}389}{1{,}67} = \frac{-2{,}889}{1{,}67} \approx -1{,}73$$

Από πίνακα κανονικής κατανομής: $\boldsymbol{P\ (Z \leq -1,73)\ \approx\ 0,0418}$

Η πιθανότητα είναι περίπου **4,18%**

β) Μας ζητείται το 95ο ποσοστημόριο της κατανομής

$$\boldsymbol{P(Z \leq z) = 0,95} \Rightarrow z \approx 1{,}645$$

Μετατρέπουμε σε τιμή Χ: $\boldsymbol{X = \mu + z\sigma} = 7{,}389 + 1{,}645 \cdot 1{,}67 \approx 7{,}389 + 2{,}748 \approx 10{,}137$

Οι φοιτητές αναπαύονται **πάνω από 10,14 ώρες** μόνο στο 5% των περιπτώσεων.

γ) Αντί για Χ, τώρα έχουμε **δειγματική μέση τιμή** $\bar{X}$ με τυπική απόκλιση:

$$\boldsymbol{\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} = \frac{1,67}{\sqrt{9}} = \frac{1,67}{3} \approx 0,5567}$$

Θέλουμε το $\boldsymbol{P(6 <} \bar{X} \boldsymbol{< 8)}$

Για 6 ώρες:

$$\boldsymbol{Z_1} = \frac{\boldsymbol{X-\mu}}{\boldsymbol{\sigma}} = \frac{6 - 7{,}389}{0{,}5567} \approx -2{,}496$$

Για 8 ώρες:

$$\boldsymbol{Z_1} = \frac{\boldsymbol{X-\mu}}{\boldsymbol{\sigma}} = \frac{8 - 7{,}389}{0{,}5567} \approx 1{,}099$$

Από πίνακες κατανομής:

$$\boldsymbol{P(Z \leq -2,50) \approx 0,0062}$$

$$\boldsymbol{P(Z \leq 1,10) \approx 0,8643}$$

Άρα: $\boldsymbol{P(6 <} \bar{X} \boldsymbol{< 8)\ = P(Z \leq 1,10) - P(Z \leq -2,50) \approx 0,8643 - 0,0062 = 0,8581}$

Η πιθανότητα είναι περίπου **85,81%**

Μια μηχανή συσκευάζει καλαμπόκι σε τσουβάλια των 25 Kgr και έχει ρυθμισθεί έτσι ώστε οι ποσότητες καλαμποκιού ανά τσουβάλι να έχουν τυπική απόκλιση 1,5 Kgr. Ο υπεύθυνος παραγωγής γνωρίζει ότι οι ποσότητες καλαμποκιού ανά τσουβάλι ακολουθούν κανονική κατανομή. Επειδή υποψιάζεται ότι η μηχανή έχει απορυθμισθεί τελευταία και θέλει να ελέγξει αν όντως η τυπική απόκλιση των ποσοτήτων καλαμποκιού ανά τσουβάλι είναι 1,5 Kgr επέλεξε τις προάλλες από την παραγωγή 30 τσουβάλια και κατέγραψε τα βάρη τους υπολόγισε δε ότι το μέσο βάρος τους ήταν 24,8 Kgr και η τυπική τους απόκλιση 1,6 Kgr.
(α) Σε επίπεδο σημαντικότητας 5%, διαπιστώνετε ότι είναι βάσιμες οι υποψίες του υπεύθυνου παραγωγής;
(β) Να βρεθεί και να ερμηνευθεί το 98% διάστημα εμπιστοσύνης της μεταβλητότητας των ποσοτήτων καλαμποκιού που συσκευάζονται ανά τσουβάλι

<u>ΛΥΣΗ</u>

Δεδομένα:

- Μέση ποσότητα αναμενόμενη: 25 Kg
- Τυπική απόκλιση σύμφωνα με ρυθμίσεις μηχανής: $\sigma_0 = 1{,}5$ Kg
- Μέσος όρος δείγματος: $\bar{x} = 24{,}8$ Kg
- Τυπική απόκλιση δείγματος: s = 1,6 Kg
- Μέγεθος δείγματος: n = 30

α) Έλεγχος υπόθεσης για τυπική απόκλιση – Χρήση κατανομής x^2
Θέλουμε να εξετάσουμε αν η πραγματική τυπική απόκλιση διαφέρει από 1,5 Kg.

Βήμα 1: Ορισμός υποθέσεων

H_0 : σ = 1,5 (Η μηχανή λειτουργεί σωστά)
H_1 : σ ≠ 1,5 (Η απόκλιση διαφέρει → αμφίπλευρος έλεγχος)

df = 30 – 1 = 29

$$X^2 = \frac{(n-1)s^2}{\sigma^2} = \frac{(30-1)(1{,}6)^2}{(1{,}5)^2} = \frac{29 \cdot 2{,}56}{2{,}25} = \frac{74{,}24}{2{,}25} \approx 32{,}99$$

α = 0,05 → α/2 = 0,025

Κρίσιμες τιμές από τον πίνακα x^2:

Κάτω όριο: $x^2_{0,025} \approx 16{,}047$
Άνω όριο: $x^2_{0,975} \approx 45{,}722$

Η τιμή $x^2 = 32{,}99$ **ανήκει** στο διάστημα αποδοχής: 16,047 < 32,99 < 45,722

Άρα **δεν απορρίπτουμε τη μηδενική υπόθεση**.

Σε επίπεδο σημαντικότητας 5%, **υπάρχουν βάσιμες ενδείξεις ότι η τυπική απόκλιση παραμένει ίση με 1,5 Kg**. Δεν απορρίπτουμε τη H_0.

β) Διάστημα εμπιστοσύνης 98% για την τυπική απόκλιση σ

Χρησιμοποιούμε τον τύπο για διάστημα εμπιστοσύνης τυπικής απόκλισης:

$$P\left(\frac{(n-1)S^2}{X^2_{1-\alpha/2}} \le \sigma^2 \le \frac{(n-1)S^2}{X^2_{\alpha/2}}\right) = 1 - \alpha$$

Επίπεδο εμπιστοσύης 98% → α = 0.02 → α/2 = 0.01

Κρίσιμες τιμές
$x^2_{0,99} \approx 49.59$

$x^2_{0,01} \approx 16.05$

$\left(\frac{(n-1)S^2}{X^2_{1-\alpha/2}} \leq \sigma^2 \leq \frac{(n-1)S^2}{X^2_{\alpha/2}}\right) \rightarrow \left(\frac{29 \cdot 2.56}{49,59} \leq \sigma^2 \leq \frac{29 \cdot 2.56}{16,05}\right) \rightarrow (1,50 \ \leq \sigma^2 \leq 4,63) \rightarrow \left(\sqrt{1,50} \ \leq \sigma \leq \sqrt{4,63}\right) \rightarrow$

$$\mathbf{1,22 \ \leq \boldsymbol{\sigma} \leq 2,15}$$

Το 98% διάστημα εμπιστοσύνης για την τυπική απόκλιση είναι [1.22 , 2.15]

α) Σε μια αποθήκη είναι αποθηκευμένα εξαρτήματα ίδιου τύπου που προέρχονται από τους προμηθευτές Γιάννη, Γιώργο, Κώστα, σε ποσοστά 50%, 40% και 10% αντίστοιχα. Από ιστορικά στοιχεία είναι γνωστό ότι οι 3 αυτοί προμηθευτές παράγουν ελλαττωματικά σε ποσοστά 6%, 10% και 15% αντίστοιχα. Μια μέρα ο γενικός διευθυντής κατέβηκε στην αποθήκη, πήρε ένα εξάρτημα στα χέρια του και άρχισε να το επεξεργάζεται. Ποια είναι η πιθανότητα το εξάρτημα αυτό να είναι ελλαττωματικό;

β) Έστω ότι το εξάρτημα αυτό βρέθηκε ελλαττωματικό. Ποια είναι η πιθανότητα να προέρχεται από τον προμηθευτή Κώστα (όπως άρχισε να ουρλιάζει ο γενικός διευθυντής, που τα «είχε» τελευταία με τον Κώστα);

<u>ΛΥΣΗ</u>

α) $P(\varepsilon\lambda\lambda\alpha\tau\tau\omega\mu\alpha\tau\iota\kappa\acute{o}) = \sum_i P\ (\Pi\rho o\mu\eta\theta\varepsilon\upsilon\tau\acute{\eta}\varsigma\ i) P(\varepsilon\lambda\lambda\alpha\tau\tau\omega\mu\alpha\tau\iota\kappa\acute{o} \mid i)$

Γιάννης: 0.50*0.06 = 0.030
Γιώργος: 0.40*0.10 = 0.040
Κώστας: 0.10*0.15 = 0.015

$P(\varepsilon\lambda\lambda\alpha\tau\tau\omega\mu\alpha\tau\iota\kappa\acute{o}) = 0.030 + 0.040 + 0.015 = 0.085$

Η πιθανότητα να είναι ελλαττωματικό είναι **8.5%.**

β) $P(\ K\acute{\omega}\sigma\tau\alpha\varsigma \mid \varepsilon\lambda\lambda\alpha\tau\tau\omega\mu\alpha\tau\iota\kappa\acute{o}\ ) = \frac{P(K\acute{\omega}\sigma\tau\alpha\varsigma) \cdot P(\varepsilon\lambda\lambda\alpha\tau\tau\omega\mu\alpha\tau\iota\kappa\acute{o} \mid K\acute{\omega}\sigma\tau\alpha\varsigma)}{P(\varepsilon\lambda\lambda\alpha\tau\tau\omega\mu\alpha\tau\iota\kappa\acute{o})}$

P(Κώστας) = 0.10
P(ελλαττωματικό | Κώστας) = 0.015
P(ελλαττωματικό) = 0.085 (από ερώτημα α)

$\frac{0.10 \cdot 015}{0.085} \approx 0.1765$

Η πιθανότητα είναι περίπου **17,65%**

Η ετήσια βροχόπτωση (X) σε μια περιοχή ακολουθεί την κανονική κατανομή με μέση τιμή 70 cm και τυπική απόκλιση 12 cm.
α) Ποια είναι η πιθανότητα να ξεπεράσει τα 92 cm η ετήσια βροχόπτωση;
β) Να βρεθεί η τιμή X_0 της μεταβλητής X, για την οποία ισχύει ότι το 30.5% των τιμών της X είναι μικρότερο από αυτήν.

<u>ΛΥΣΗ</u>

Δεδομένα:
Μέση τιμή μ = 70
Τυπική απόκλιση σ = 12

α) Θέλουμε P(X > 92)

$$\boldsymbol{Z = \frac{X - \mu}{\sigma}} = \frac{92 - 70}{12} = \frac{22}{12} \approx \ 1,83$$

Άρα P(X > 92) = P(Z > 1.83)

Από πίνακα κανονικής κατανομής: $P(Z \leq 1.83) \approx 0.9644$

$P(Z > 1.83) = 1 – 0.9644 = 0.0336$

Η πιθανότητα να ξεπεράσει τα 92 cm είναι περίπου **3,36%**.

β) Ζητάμε X_0 ώστε $P(X < X_0) = 0,305$ (επειδή μικρότερο από 0.5, η Z θα είναι αρνητική)

$$\boldsymbol{Z} = \frac{\boldsymbol{X} - \boldsymbol{\mu}}{\boldsymbol{\sigma}} = \frac{\boldsymbol{X} - 70}{12}$$

Άρα: $P(X < X_0) = P(Z < Z_0) = 0.305$ όπου $Z_0 = \frac{\boldsymbol{X}-70}{12} < 0$

Για αρνητικό Z ισχύει: $P(Z < -z) = 1 - P(Z < z) = P(Z > z)$

$P(0 < Z < |z_0|) = 0,5 - P(Z < z_0) = 0.5 – 0.305 = 0.195$

Ισχύει: $P(Z < z) = P(Z < 0) + P(0 < Z < z) = 0,5 + 0,195 = 0,695$

Από πίνακα κατανομής πιθανότητα αντιστοιχεί στο Z = 0.51, το αρχικό Z ήταν αρνητικό άρα -0.51

$X_0 = \mu + z_0 \cdot \sigma = 70 + (-0,51) \cdot 12 \approx 63,88$

Το 30,5% των ετήσιων βροχοπτώσεων είναι μικρότερο από περίπου **63,88 cm**

Στον έλεγχο ποιότητας των εξαρτημάτων μιας μεγάλης αποστολής, 12 από τα 60 εξαρτήματα που ελέγχθηκαν βρέθηκαν ελαττωματικά.
α) Να βρεθεί και να ερμηνευθεί το 98% διάστημα εμπιστοσύνης για την αναλογία των ελαττωματικών εξαρτημάτων στην αποστολή.
β) Αντίκειται η όχι το εύρημα του ελέγχου ποιότητας στην άποψη του υπευθύνου ποιότητας ότι ποσοστό ελαττωματικών είναι 15% σε επίπεδο σημαντικότητας 5%;

ΛΥΣΗ

Δεδομένα:

Δείγμα: n = 60

Ελαττωματικά: x = 12

α) Εκτιμημένη αναλογία: $\widehat{\boldsymbol{p}} = \frac{\boldsymbol{x}}{\boldsymbol{n}} = \frac{12}{60} = 0.2$

Επίπεδο εμπιστοσύνης 98% → α = 0.02 → α/2 = 0.01 → $Z_{0.01} = 2,33$

$$\boldsymbol{SE} = \sqrt{\frac{\widehat{\boldsymbol{p}}(\mathbf{1} - \widehat{\boldsymbol{p}})}{\boldsymbol{n}}} = \sqrt{\frac{0.2 \cdot 0.8}{60}} = \sqrt{0.002667} \approx 0.0516$$

$\boldsymbol{\delta.\varepsilon} = \widehat{\boldsymbol{p}} \pm \boldsymbol{z_{a/2}} \cdot \boldsymbol{SE} = 0.2 \pm 2.33 \cdot 0.0516 = 0.2 \pm 0.120$

0.2 – 0.12 = 0,08
0.2 + 0.12 = 0.32

Με 98% επίπεδο εμπιστοσύνης, η πραγματική αναλογία των ελαττωματικών εξαρτημάτων στην αποστολή βρίσκεται μεταξύ **8% και 32%**.

β) Υποθετική αναλογία: $p_0 = 0,15$
Επίπεδο σημαντικότητας: α = 0,05

Θέλουμε να ελέγξουμε

$H_0 : p = 0,15$
$H_1 : p \neq 0,15$

Τύπος για μεγάλο δείγμα: $Z = \frac{\hat{P}-p_0}{\sqrt{p_0(1-p_0)/\mathrm{n}}}$

$$\boldsymbol{SE} = \sqrt{\frac{\widehat{\boldsymbol{p}}(\mathbf{1}-\widehat{\boldsymbol{p}})}{\boldsymbol{n}}} = \sqrt{\frac{0.15 \cdot 0.85}{60}} = \sqrt{\frac{0.1275}{60}} = \sqrt{0.002125} \approx 0.0416$$

$$Z = \frac{\hat{P}-p_0}{\sqrt{p_0(1-p_0)/\mathrm{n}}} = \frac{0.2-0.15}{0.0461} = \frac{0.05}{0.0461} \approx 1.085$$

Από πίνακες: α=0.05 → α/2 = 0.025, $Z_{0.025} \approx 1.96$, άρα η H_0 απορρίπτεται εάν $|Z| > 1.96$

Βρήκαμε Z = 1.085 < 1.96 άρα **δεν απορρίπτουμε τη μηδενική υπόθεση**.

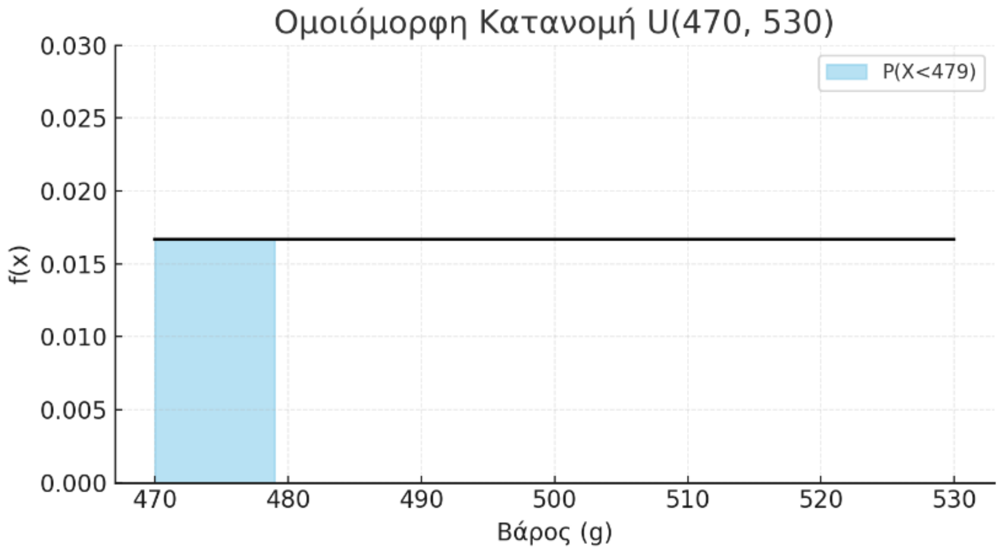
Ένα εργοστάσιο παράγει ζάχαρη σε χάρτινες συσκευασίες. Εάν το βάρος κάθε συσκευασίας ακολουθεί ομοιόμορφη κατανομή (όπως αυτή του σχήματος) στο διάστημα (470 g, 530 g), τότε ποια είναι η πιθανότητα μια συσκευασία ζάχαρης να έχει βάρος μικρότερο από 479 g; Πριν (αλλά και για να) απαντήστε, συμπληρώστε το απαραίτητο σχήμα.

<u>ΛΥΣΗ</u>

Ελάχιστη τιμή: a = 470
Μέγιστη τιμή: b = 530
Μήκος διαστήματος: b – a = 530 – 470 = 60
Πυκνότητα της ομοιόμορφης κατανομής: $\boldsymbol{f}(\boldsymbol{x}) = \frac{\mathbf{1}}{\boldsymbol{b}-\boldsymbol{a}} = \frac{1}{60}$, $470 \leq x \leq 530$

$$\boldsymbol{P}(\boldsymbol{X} < \mathbf{479}) = \frac{\mathbf{479}-\mathbf{470}}{\mathbf{530}-\mathbf{470}} = \frac{9}{60} = 0.15$$

Η πιθανότητα είναι **15%**



Μια τυχαία μεταβλητή t ακολουθεί την κατανομή Student. Λαμβάνεται δείγμα 16 μονάδων. Η πιθανότητα να πάρει η τυχαία μεταβλητή t τιμή μεγαλύτερη από την τιμή ________ είναι προσεγγιστικά ίση με 5%. Πριν (αλλά και για να) απαντήστε, κάντε το απαραίτητο σχήμα.

<u>ΛΥΣΗ</u>

Η τυχαία μεταβλητή ακολουθεί κατανομή t του Student.

Το δείγμα έχει μέγεθος n = 16.

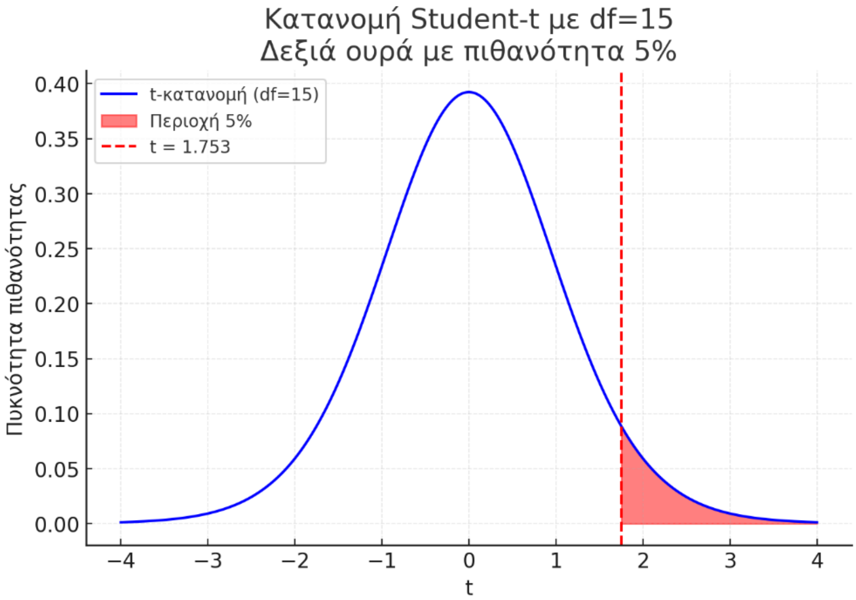
Βαθμοί ελευθερίας: df = n−1 =1 6−1 = 15

Ψάχνουμε $P(T>t_\alpha) = 0.05$

Από πίνακα, για α=0.05 και df = 15 έχουμε $t_{0.05,\,15} = 2.1315$

Άρα $P(T>1.753) \approx 0.05$

Δηλαδή υπάρχει 5% πιθανότητα η τυχαία μεταβλητή T να ξεπεράσει την τιμή 2.1315



Έστω ότι το μέσο βάρος των σωληνάριων κόλλας YUHU - που προέκυψε από ένα δείγμα 25 σωληνάριων - είναι 310,4 gr και η δειγματική τυπική απόκλιση είναι 1 gr. Θεωρώντας ότι το βάρος των σωληνάριων κόλλας ακολουθεί κανονική κατανομή, τότε:
α) μπορείτε να ισχυριστείτε (με συντελεστή εμπιστοσύνης 1%) ότι η παραγωγική διαδικασία παράγει σωληνάρια με το επιθυμητό βάρος 310 gr;
β) να βρεθεί και να ερμηνευθεί το 95% διάστημα εμπιστοσύνης για το μέσο βάρος των σωληνάριων κόλλας YUHU.

<u>ΛΥΣΗ</u>

Δεδομένα

Δείγμα μεγέθους: n = 25
Δειγματικός μέσος: $\bar{x} = 310.4$ gr
Δειγματική τυπική απόκλιση: σ = 1 gr
Ελέγχουμε για $\mu_0 = 310$

α)

H_0: μ = 310
H_1: μ ≠ 310

Διμερής έλεγχος με α = 0.01

$$z = \frac{x - \mu_0}{\sigma/\sqrt{n}} = \frac{310{,}4 - 310}{1/\sqrt{25}} = \frac{0{,}4}{0{,}2} = 2$$

$\alpha = 0.01 \rightarrow \alpha/2 = 0.005 \rightarrow Z_{0.005} = 2.576$

$|Z| = 2 < 2.576 \rightarrow -2.576 < 2 < 2.576$

Δεν απορρίπτουμε την H_0. Δεν υπάρχει στατιστικά σημαντική απόδειξη στο 1% επίπεδο ότι η διαδικασία παράγει σωληνάρια με μέσο βάρος διαφορετικό από 310 gr.

β) $\delta.\varepsilon. = \bar{x} \pm z_{a/2} \cdot \frac{\sigma}{\sqrt{n}}$

Για δ.ε 95% → $\alpha=0.05 \rightarrow \alpha/2 = 0.025 \rightarrow Z_{0.025} = 1.96$

$$\delta.\varepsilon. = \bar{x} \pm z_{a/2} \cdot \frac{\sigma}{\sqrt{n}} = 310.4 \pm 1.96 \cdot \frac{1}{\sqrt{25}} = 310.4 \pm 1.96 \cdot 0.2 = 310.4 \pm 0.392$$

Άρα δ.ε. είναι [310.008 , 310.792]

Με 95% εμπιστοσύνη, το **πραγματικό μέσο βάρος** των σωληναρίων YUHU βρίσκεται περίπου μεταξύ **310,01 gr και 310,79 gr**, δηλαδή κοντά στο επιθυμητό 310 gr.

Ο χρόνος συναρμολόγησης μιας οθόνης Dell ορισμένου τύπου ακολουθεί κανονική κατανομή. Σε τυχαίο δείγμα 25 τέτοιων οθονών η εταιρεία υπολόγισε μέση διάρκεια συναρμολόγησης ίση με 16 min και μεταβλητότητα ίση με 4 min^2.
α) Να προσδιοριστεί (και να ερμηνευτεί) το διάστημα στο οποίο με επίπεδο εμπιστοσύνης 99% θα βρίσκεται η μέση διάρκεια συναρμολόγησης του συνόλου των οθονών τέτοιου τύπου, της Dell.
β) Ποιο θα ήταν το κατάλληλο μέγεθος δείγματος οθονών που θα έπρεπε η Dell να εξετάσει αν ήθελε το πλάτος του διαστήματος εμπιστοσύνης να είναι 1 min;

<u>ΛΥΣΗ</u>

$n = 25$
$\bar{x} = 16$
$s^2 = 4 \rightarrow s = 2$

α) $\delta.\varepsilon. = \bar{x} \pm z_{a/2} \cdot \frac{s}{\sqrt{n}}$

Για 99% εμπιστοσύνη: α=0.01 → α/2 = 0.005 Ψάχνω $Z_{0.995} = 2.576$

$$\bar{x} \pm z_{a/2} \cdot \frac{s}{\sqrt{n}} = 16 \pm 2.576 \cdot \frac{2}{\sqrt{25}} = 16 \pm 1.0304$$

Άρα το δ.ε. είναι [14.97 , 17.03]

Με 99% εμπιστοσύνη, η μέση διάρκεια συναρμολόγησης όλων των Dell οθονών τέτοιου τύπου βρίσκεται μεταξύ **14.97 και 17.03 λεπτών**.

β) Πλάτος δ.ε.

$$\mathbf{2 \cdot z_{\frac{a}{2}} \cdot \frac{s}{\sqrt{n}}} \rightarrow 2 \cdot 2.576 \cdot \frac{2}{\sqrt{n}} = 1 \rightarrow \frac{10.304}{\sqrt{n}} = 1 \rightarrow \sqrt{n} = 10.304 \rightarrow n = (10.304)^2 \approx 106.2$$

Απαιτείται δείγμα **n = 107 οθόνες** (με στρογγυλοποίηση)

Εργοστάσιο παρασκευής γεμιστών μπισκότων αναγράφει σε κάθε συσκευασία τους «καθαρό βάρος 100 gr», ενώ γνωρίζει ακόμη ότι το καθαρό βάρος μιας συσκευασίας, X, ακολουθεί κανονική κατανομή με τυπική απόκλιση 4,5 gr. Ενόψει της αποστολής μιας παρτίδας μπισκότων σε καλό πελάτη της (σούπερ μάρκετ), ελήφθη τυχαίο δείγμα 9 συσκευασιών και υπολογίστηκε μέσο καθαρό βάρος 96 gr. Το αποτέλεσμα αυτό πρέπει να ενεργοποιήσει την εταιρεία ότι τα προϊόντα της συγκεκριμένης παρτίδας γεμίστηκαν ελλιπώς; Θεωρήστε επίπεδο σημαντικότητας 1%.

<u>ΛΥΣΗ</u>

Δηλωμένο βάρος (υποτιθέμενη μέση τιμή): $\mu_0 = 100$ gr
Δείγμα: $n = 9$, $\bar{x}= 96$ gr
Τυπική απόκλιση (πληθυσμιακή): $\sigma = 4.5$ gr
Σημαντικότητα: $\alpha = 0.01$

H_0: $\mu = 100$
H_1: $\mu < 100$

$$z = \frac{x - \mu_0}{\sigma/\sqrt{n}} = \frac{96 - 100}{4.5/\sqrt{9}} = \frac{-4}{1.5} = -2.6667$$

Για α=0.01 (λογω μ<100 έχουμε μονόπλευρο έλεγχο, άρα δεν κάνουμε α/2)
$Z_{0.01} = -2.33$

Έχουμε $-2.6667 < -2.33$

Άρα απορρίπτουμε την H_0

Το αποτέλεσμα δείγματος (μέσος όρος 96 gr) δείχνει σημαντικά χαμηλότερο βάρος από τα 100 gr στο επίπεδο σημαντικότητας 1%. Η εταιρεία πρέπει να ανησυχήσει ότι τα προϊόντα της συγκεκριμένης παρτίδας γεμίστηκαν ελλιπώς.