ΦΡΟΝΤΙΣΤΗΡΙΟ #1

Εφαρμογές Κεντρικού Οριακού Θεωρήματος (Κ.Ο.Θ.)

Το Κεντρικό Οριακό Θεώρημα κατέχει εξέχουσα θέση στη Θεωρία Πιθανοτήτων και στη Στατιστική. Κάτω από αρκετά γενικές συνθήκες, η κατανομή του αθροίσματος τυχαίων μεταβλητών μπορεί να προσεγγιστεί ικανοποιητικά από μία κανονική κατανομή. Ως συνέπεια του Κ.Ο.Θ. η κανονική κατανομή, επιπλέον, προσφέρει "εξαιρετικές" προσεγγίσεις για διάφορες κατανομές όπως π.χ. η διωνυμική, η Poisson κλπ. δικαιολογώντας τον κεντρικό ρόλο που **κατέχει στη Στατιστική.**

Κ.Ο.Θ. Έστω $X_1, X_2, \ldots, X_n$ ανεξάρτητες παρατηρήσεις πάνω σε μία τυχαία μεταβλητή X που περιγράφει ένα υπό μελέτη χαρακτηριστικό ενός πληθυσμού, με πεπερασμένη μέση τιμή μ και πεπερασμένη διακύμανση σ^2. Έστω $\bar{X}_n$ ο δειγματικός μέσος των παρατηρήσεων δηλ.

$$\bar{X}_n = \frac{\sum_{i=1}^{n} X_i}{n}$$

Έστω η τυποποιημένη τυχαία μεταβλητή

$$Z_n = \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} = \frac{\sqrt{n}\,(\bar{X}_n - \mu)}{\sigma}$$

και $F_{Z_n}(z)$ η συνάρτηση κατανομής της τ.μ. Z_n.

Τότε, $\lim_{n\to\infty} F_{Z_n}(z) = \Phi(z),\ \forall z \in \mathbb{R}$, όπου $\Phi(z)$ είναι η συνάρτηση κατανομής της τυποποιημένης κανονικής κατανομής $N(0,1)$.

Ισοδύναμα, το Κ.Ο.Θ. μπορεί να διατυπωθεί ως εξής:

Ο δειγματικός μέσος $\overline{X}_n$ ακολουθεί ασυμπτωτικά την κανονική κατανομή $N(\mu, \frac{\sigma^2}{n})$ ή ισοδύναμα το άθροισμα $S_n = \sum_{i=1}^{n} X_i$, όπου X_i, $i=1,\dots,n$ ανεξάρτητες τ.μ. ακολουθεί ασυμπτωτικά την κανονική κατανομή $N(n\mu, n\sigma^2)$.

Μπορούμε να γράψουμε:

$$Z_n = \frac{\dfrac{\sum_{i=1}^{n} X_i - n\mu}{n}}{\dfrac{\sigma\sqrt{n}}{n}} = \frac{\dfrac{S_n}{n} - \mu}{\sqrt{\dfrac{\sigma^2}{n}}} = \frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}} =$$

$$= \frac{\sqrt{n}(\overline{X}_n - \mu)}{\sigma} \underset{n\to\infty}{\sim} N(0,1).$$

Το σημαντικότερο στοιχείο του Κ.Ο.Θ. είναι ότι το συμπέρασμά του δεν εξαρτάται από τη μορφή του πληθυσμού τον οποίον περιγράφει η τυχαία μεταβλητή X (δηλαδή, δεν εξαρτάται από την κατανομή του πληθυσμού).

Με βάση το Κ.Ο.Θ. μπορούμε να εξηγήσουμε γιατί οι τ.μ. που εμφανίζονται σε ένα μεγάλο πλήθος εφαρμογών ακολουθούν (κατά προσέγγιση) την κανονική κατανομή αναδεικνύοντας την κεντρική θέση που αυτή κατέχει στη Θεωρία Πιθανοτήτων και στη Στατιστική.

Η ακρίβεια της προσέγγισης της κατανομής της τ.μ. Z_n από την κανονική κατανομή μέσω του Κ.Ο.Θ. μεταβάλλεται με την τιμή του n. Επειδή η κανονική κατανομή είναι συμμετρική γύρω από τη μέση τιμή της, όσο πιο ασύμμετρη είναι η κοινή κατανομή των τ.μ. $X_1, \ldots, X_n$ τόσο μεγαλύτερο μέγεθος δείγματος απαιτείται για να επιτευχθεί μία καλή προσέγγιση της $F_{Z_n}(z)$.

Στις φυσικές, κοινωνικές και οικονομικές επιστήμες οι τ.μ. οι οποίες χρησιμοποιούνται για να περιγράψουν τα χαρακτηριστικά ενός υπό μελέτη πληθυσμού που μας ενδιαφέρουν, μπορούν να θεωρηθούν ως αποτέλεσμα άθροισης πολλών μικρών τυχαίων παραγόντων. Αν και κάθε συγκεκριμένος παράγοντας έχει ίσως μικρή ή και αμελητέα συμβολή, η συσσώρευση πολλών τέτοιων παραγόντων οδηγεί σε μία μη αμελητέα τυχαία ποσότητα, η οποία σύμφωνα με το Κ.Ο.Θ. θα ακολουθεί την κανονική κατανομή (κατά προσέγγιση). Ορισμένα παραδείγματα στα οποία, με βάση το παραπάνω σκεπτικό, μπορούμε να χρησιμοποιήσουμε την κανονική κατανομή είναι τα εξής: βάρος (ή ύψος) ενός ατόμου, μηνιαίο ή ετήσιο εισόδημα ενός εργαζομένου, ύψος αποζημίωσης που απαιτεί ένας ασφαλισμένος, η θέση και η ταχύτητα των μορίων ενός αερίου.

#1 <u>Εφαρμογές</u>

Το σφάλμα μέτρησης ενός οργάνου ακολουθεί την Ομοιόμορφη κατανομή στο διάστημα $(-0.05, 0.05)$. Ποια είναι η πιθανότητα (κατά προσέγγιση)

(α) Το σφάλμα μέτρησης για το άθροισμα 100 μετρήσεων να είναι κατ' απόλυτη τιμή μικρότερο του 0.25;

(β) Το άθροισμα των τετραγώνων 100 μετρήσεων

(i) να μην υπερβαίνει το $1/12$;

(ii) να βρίσκεται μεταξύ 0.15 και 0.20;

<u>Λύση</u> (α) Έστω η τ.μ. X_i που αναπαριστά το σφάλμα που γίνεται κατά την i-οστή μέτρηση, $i=1,2,\dots,100$.

Τότε: $\mu = E[X_i] = \dfrac{0.05+(-0.05)}{2} = 0$

και $\sigma^2 = Var(X_i) = \dfrac{(-0.05-0.05)^2}{12} = \dfrac{1}{1200}$

Έστω η τ.μ. $S_{100} = X_1 + X_2 + \dots + X_{100}$

Η ζητούμενη πιθανότητα θα είναι:

$$P(|S_{100}|<0.25) = P(-0.25 < S_{100} < 0.25)$$

$$\underset{\text{ΚΟΘ}}{=} P\left(\frac{-0.25-0}{\sqrt{1/1200}\sqrt{100}} < \frac{S_{100}-\mu 100}{\sigma\sqrt{100}} < \frac{0.25-0}{\sqrt{1/1200}\sqrt{100}}\right)$$

$$= P(-0.87 < Z_{100} < 0.87) \cong \Phi(0.87) - \Phi(-0.87)$$

$$\overset{(*)}{=} 2\Phi(0.87) - 1 \underset{\substack{\text{από}\\\text{πίνακες}}}{=} 2(0.8078) - 1 \cong 62\%.$$

(*) γιατί $\forall z \in \mathbb{R}$, $\Phi(z) + \Phi(-z) = 1$.

(β) Για το άθροισμα $S^*_{100} = X_1^2 + \dots + X_{100}^2$ μπορούμε να χρησιμοποιήσουμε την κανονική προσέγγιση του Κ.Ο.Θ.

Είναι $E[X_i^2] = Var(X_i) + (E[X_i])^2$

$$= \frac{1}{1200} + 0 = \frac{1}{1200}$$

Ενώ, $Var(X_i^2) = E[X_i^4] - (E[X_i^2])^2$,

με
$$E(X_i^4) = \int_{-\infty}^{\infty} x^4 f_X(x)\,dx,$$ όπου $f_X(x)$ η συνάρτηση πυκνότητας της τ.μ. $X \sim$ Ομοιόμορφη στο διάστημα $(-0.05, 0.05)$.

Είναι
$$E[X_i^4] = \frac{1}{0.1}\int_{-0.05}^{0.05} x^4\,dx = \frac{1}{0.1}\left[\frac{x^5}{5}\right]_{-0.05}^{0.05} = \frac{1}{8\cdot 10^5}$$

Άρα
$$Var(X_i^2) = \frac{1}{8\cdot 10^5} - \left(\frac{1}{1200}\right)^2 = \frac{1}{22500}$$

Η τ.μ.
$$Z^*_{100} = \frac{S'^*_{100} - 100\,\frac{1}{1200}}{\sqrt{100\,\frac{1}{22500}}} = \frac{S'^*_{100} - \frac{1}{12}}{\frac{1}{15}}$$
$$= 15\left(S'^*_{100} - \frac{1}{12}\right)$$

ακολουθεί κατά προσέγγιση την τυποποιημένη κανονική κατανομή. Οι ζητούμενες πιθανότητες θα είναι, κατά προσέγγιση, ίσες με:

(i) $P\left(S'^*_{100} \le \frac{1}{12}\right) = P\left(15\left(S'^*_{100} - \frac{1}{12}\right) \le 15\left(\frac{1}{12} - \frac{1}{12}\right)\right)$

$= P(Z^*_{100} \le 0) = \Phi(0) \overset{\text{(από πίνακες)}}{=} 0.5$, όπου $\Phi(z) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{z} e^{-\frac{y^2}{2}}\,dy$ η σ.κ. της $N(0,1)$.

(ii) $P(0.15 < S'^*_{100} < 0.2) = P\left[15\left(\frac{3}{20} - \frac{1}{12}\right) < 15\left(S'^*_{100} - \frac{1}{12}\right) < 15\left(\frac{1}{5} - \frac{1}{12}\right)\right]$

$= P\left(1 < Z^*_{100} < \frac{7}{4}\right) = \Phi(1.75) - \Phi(1) = 0.9599 - 0.8413$
$= 11.86\%$.
(από πίνακες)

#2 (Υπολογισμός μεγέθους δείγματος για εκτίμηση μέσου).

Μία ασφαλιστική εταιρεία θέλει να εκτιμήσει το μέσο ύψος μ των ετήσιων απαιτήσεων ανά ασφαλισμένο (σε μία συγκεκριμένη κατηγορία ασφάλισης). Για το λόγο αυτό παίρνει ένα τ.δ. n ασφαλισμένων, εξετάζει το ύψος των απαιτήσεων που έχει το κάθε άτομο σ' αυτό το οικονομικό έτος και υπολογίζει το μέσο όρο n παρατηρήσεων. Αν δεχτούμε ότι τα ύψη των απαιτήσεων είναι ανεξάρτητες και ισόνομες τ.μ. με διακύμανση 256 ευρώ, τι μέγεθος δείγματος θα πρέπει να χρησιμοποιηθεί ώστε με **πιθανότητα** 99% τουλάχιστον, η εκτίμησή μας να μην απέχει από τον πραγματικό (θεωρητικό) μέσο μ περισσότερο από 3 ευρώ;

<u>Λύση</u> Έστω X_i το ύψος των ετήσιων **απαιτήσεων του i** ασφαλισμ**ένου**. Οι τ.μ. $X_1, X_2, \ldots$ θα είναι ανεξάρτητες και ισόνομες με $E[X_i] = \mu$, $Var(X_i) = \sigma^2 = 16^2 = 256$, $i = 1, \ldots, n$.

Το συνολικό ύψος των ετήσιων απαιτήσεων για τους n ασφαλισμένους περιγράφεται από την τυχαία μεταβλητή

$$S_n = X_1 + X_2 + \ldots + X_n$$

Θέλουμε να προσδιορίσουμε την τιμή του n για την οποία ισχύει: $P\left(-3 < \frac{S_n}{n} - \mu < 3\right) \geq 0.99$

Για αρκετά μεγάλο n, μπορούμε να εφαρμόσουμε το Κ.Ο.Θ. (το n δεν το γνωρίζουμε!) (Εικάζουμε ότι θα είναι μεγάλο!), έχουμε:

$$P\left(-3 < \frac{S_n}{n} - \mu < 3\right) = P\left(-3n < S_n - n\mu < 3n\right)$$

$$= P\left(-\frac{3n}{16\sqrt{n}} < \frac{S_n - n\mu}{\sigma\sqrt{n}} < \frac{3n}{16\sqrt{n}}\right)$$

$$= P\left(-\frac{3\sqrt{n}}{16} < Z_n < \frac{3\sqrt{n}}{16}\right)$$

$$= \Phi\left(\frac{3\sqrt{n}}{16}\right) - \Phi\left(-\frac{3\sqrt{n}}{16}\right) \qquad \Phi(z) + \Phi(-z) = 1 \ \forall z$$

$$= 2\Phi\left(\frac{3\sqrt{n}}{16}\right) - 1$$

Θα πρέπει: $2\Phi\left(\frac{3\sqrt{n}}{16}\right) - 1 \geq 0.99$

ή $\Phi\left(\frac{3\sqrt{n}}{16}\right) \geq 0.995$

Από πίνακες της τυπικής κανονικής κατανομής $N(0,1)$, έχουμε: $\Phi(2.57) = 0.9949$, $\Phi(2.58) = 0.9951$

Άρα, αν ζητήσουμε να ισχύει $\Phi\left(\frac{3\sqrt{n}}{16}\right) \geq \Phi(2.58)$ εξασφαλίζεται η ισχύς της προηγούμενης ανισότητας.

Αφού η $\Phi(z)$ είναι γνησίως αύξουσα, $\frac{3\sqrt{n}}{16} \geq 2.58$

$$\Rightarrow n \geq \left[\frac{16\cdot(2.58)}{3}\right]^2 = 189.3$$

Το ελάχιστο μέγεθος δείγματος είναι $n = 190$. Η πολύ μεγάλη τιμή του n, δικαιολογεί (εκ των υστέρων) τη χρήση του Κ.Ο.Θ.

Διόρθωση Συνέχειας

Αν το Κ.Ο.Θ. χρησιμοποιείται για την προσέγγιση της κατανομής του αθροίσματος διακριτών τ.μ. τότε στον προσεγγιστικό τύπο ενσωματώνεται η διόρθωση συνέχειας. Αν οι $X_1, X_2, \dots$ είναι ανεξάρτητες και ισόνομες διακριτές τ.μ. με μέση τιμή μ και διασπορά σ^2 και i, j δύο ακέραιοι αριθμοί με $i \le j$ τότε ο τύπος

$$P(i \le S_n \le j) = \Phi\left(\frac{j+0.5-n\mu}{\sigma\sqrt{n}}\right) - \Phi\left(\frac{i-0.5-n\mu}{\sigma\sqrt{n}}\right)$$

(με διόρθωση συνέχειας)

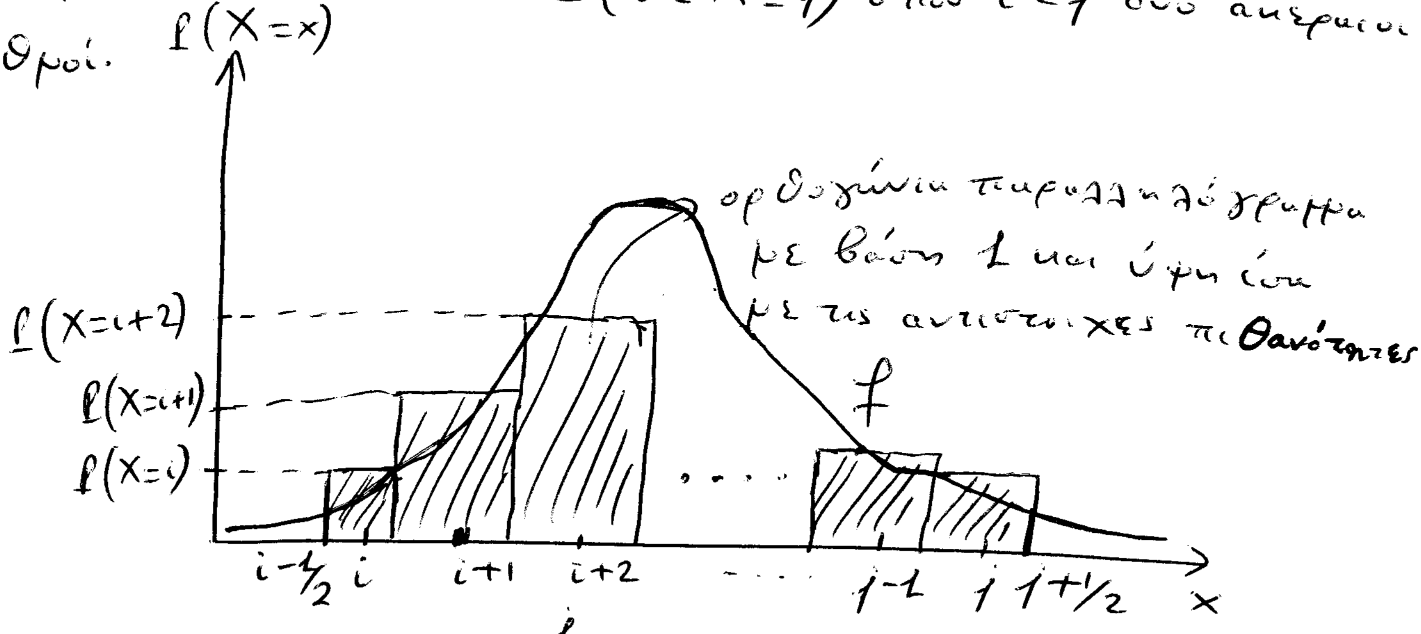
δίνει ακριβέστερα αποτελέσματα από τον τύπο:

$$P(i \le S_n \le j) = \Phi\left(\frac{j-n\mu}{\sigma\sqrt{n}}\right) - \Phi\left(\frac{i-n\mu}{\sigma\sqrt{n}}\right)$$

(χωρίς διόρθωση συνέχειας).

Εξήγηση:

Έστω X διακριτή τ.μ. με ακέραιες τιμές. Θέλουμε να προσεγγίσουμε την πιθανότητα $P(i \le X \le j)$ όπου $i < j$ δύο ακέραιοι αριθμοί.



Τότε: $P(i \le X \le j) = \sum_{x=i}^{j} P(X=x)$

Προσεγγίζουμε την παραπάνω πιθανότητα με μία συνάρτηση πυκνότητας f (συνεχή κατανομή).

Ως προσεγγιστική τιμή της πιθανότητας $P(i \le X \le j)$ μπορούμε να θεωρήσουμε το εμβαδόν που περικλείεται μεταξύ του οριζόντιου άξονα, της γραφικής παράστασης της f και των κατακόρυφων ευθειών $x = i - 1/2$, $x = j + 1/2$.

Προσεγγιστικά, μπορούμε να γράψουμε

$$P(i \le X \le j) \cong \int_{i-0.5}^{j+0.5} f(x)\,dx \qquad (*)$$

Ανάλογοι τύποι μπορούν να γραφούν για τις πιθανότητες $P(X=i)$, $P(X\ge i)$ και $P(X \le j)$.

Ο τύπος $(*)$ (διόρθωση συνέχειας) χρησιμοποιείται όταν κάποια διακριτή τ.μ. προσεγγίζεται από μία συνεχή.

Για διακριτή τ.μ. X με $E(X)=\mu$, $Var(X)=\sigma^2$ από το Κ.Ο.Θ. έχουμε:

$$P(i \le X \le j) \cong \int_{i-0.5}^{j+0.5} \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{(x-\mu)^2}{2\sigma^2}}\,dx$$

$$= P\left(\frac{(i-0.5)-\mu}{\sigma} \le Z \le \frac{(j+0.5)-\mu}{\sigma}\right)$$

όπου $Z = \dfrac{X-\mu}{\sigma}$

#3 Ο χρόνος χρήσης (σε ώρες) ενός οργάνου στη διάρκεια μιας μέρας περιγράφεται από μία συνεχή τ.μ. X με συνάρτηση κατανομής $F(x) = 1 - e^{-\frac{x}{4}}$, $x > 0$.

(α) Να υπολογιστεί η πιθανότητα σε μία ημέρα να χρησιμοποιηθεί το όργανο το πολύ 2 ώρες.

(β) Να βρεθεί ο αναμενόμενος αριθμός ημερών σε ένα έτος (365 ημέρες) κατά τις οποίες το όργανο χρησιμοποιείται λιγότερο από 2 ώρες.

(γ) Να υπολογιστεί προσεγγιστικά η πιθανότητα σε ένα έτος να υπάρχουν τουλάχιστον 150 ημέρες κατά τις οποίες το όργανο χρησιμοποιήθηκε το πολύ 2 ώρες. Να γίνει υπολογισμός με διόρθωση συνέχειας, χωρίς διόρθωση συνέχειας.

<u>Λύση</u> (α) $P(X \leq 2) = F(2) = 1 - e^{-2/4} \cong 0.3934$

(β) Έστω $Y :=$ # ημερών σε ένα έτος κατά τις οποίες χρησιμοποιείται το όργανο λιγότερο από 2 ώρες

Τότε η τ.μ. $Y \sim B(n = 365, p = P(X \leq 2) \cong 0.39)$

Άρα $E[Y] = np = 365 \cdot (0.39) = 142.35$ ημέρες

(γ) Ζητάμε την $P(Y \geq 150) = \sum_{y=150}^{365} P(Y = y)$

$$= \sum_{y=150}^{365} \binom{365}{y} (0.39)^y (1 - 0.39)^{365-y} = 0.220991$$

(υπολογισμός με χρήση υπολογιστή)

Επειδή $n = 365$ μπορούμε να χρησιμοποιήσουμε το Κ.Ο.Θ.

$$Z = \frac{Y - np}{\sqrt{np(1-p)}} = \frac{Y - 142.35}{\sqrt{86.8}} \sim N(0,1)$$

κατά προσέγγιση

Άρα

$$P(Y \geq 150) = P\left(\frac{Y - 142.35}{\sqrt{86.8}} \geq \frac{150 - 142.35}{\sqrt{86.8}}\right)$$

$$= P(Z \geq 0.82) \cong 1 - \Phi(0.82) = 0.2061$$

Η ίδια πιθανότητα, με χρήση διόρθωσης συνέχειας, δίνεται από τον προσεγγιστικό τύπο:

$$P(Y \geq 150) \cong 1 - \Phi\left(\frac{150 - 0.5 - 142.35}{\sqrt{96.8}}\right) \cong$$

$\cong 1 - \Phi(0.77) = 0.2206$. Στη δεύτερη περίπτωση η προσεγγιστική τιμή είναι πολύ πιο κοντά στην ακριβή τιμή 0.220991.

#4 Σε μία πόλη **100.000** κατοίκων πρόκειται να κατασκευαστεί ένα νοσοκομείο. Αν το ποσοστό των ατόμων που χρειάζεται να εισαχθεί στο νοσοκομείο σε μία ημέρα είναι 2‰, πόσα κρεβάτια θα πρέπει να προβλεφθούν στο νοσοκομείο ώστε να μπορούν να εξυπηρετηθούν οι ημερήσιες ανάγκες με πιθανότητα (α) τουλάχιστον 95%; (β) τουλάχιστον 99%;

Λύση Έστω **X το** πλήθος των κατοίκων της πόλης που θα χρειαστεί να εισαχθούν στο νοσοκομείο σε μία ημέρα. Τότε $X \sim B(n = 100.000,\ p = 0.002)$. Σύμφωνα με το Κ.Ο.Θ

$$Z = \frac{X - np}{\sqrt{np(1-p)}} = \frac{X - 200}{\sqrt{199.6}} = \frac{X - 200}{14.13}$$

$\sim N(0,1)$
προσεγγιστικά

Αν υποθέσουμε ότι στο νοσοκομείο προβλέπονται **K** κρεβάτια, η πιθανότητα να εξυπηρετηθούν όλες οι ημερήσιες ανάγκες της πόλης θα είναι ίση με:

$$P(X\le K)=P\left(\frac{X-200}{14.13}\le\frac{K-200}{14.13}\right)\simeq\Phi\left(\frac{K-200}{14.13}\right)$$

(α) Ζητάμε να ισχύει $P(X\le K)\ge 0.95$ δηλαδή

$$\Phi\left(\frac{K-200}{14.13}\right)\ge 0.95$$

Αφού $\Phi(1.64)=0.9495$, $\Phi(1.65)=0.9505$ αρκεί

$$\frac{K-200}{14.13}\ge 1.65 \Rightarrow K\ge 200+(14.13)\cdot(1.65)$$
$$=223.3$$

δηλ. θα πρέπει να προβλεφθούν τουλάχιστον 224 κρεβάτια.

(β) Όμοια θα πρέπει να ισχύει: $\Phi\left(\frac{K-200}{14.13}\right)\ge 0.99$

από όπου βρίσκουμε όμοια ότι:

$$K\ge 200+(14.13)(2.33)=332.9.$$

#5 Δύο παίκτες Α, Β παίζουν το εξής παιχνίδι: Ρίχνουν διαδοχικά ένα ζάρι και αν η ένδειξη είναι 1, 3 ή 5, ο παίκτης Α δίνει στον Β ποσό 1, 2 ή 3 € αντίστοιχα. Αν η ένδειξη είναι 2, 4 ή 6, ο Β δίνει στον Α ποσό **3, 2 ή 1** € αντίστοιχα. Να υπολογιστεί η πιθανότητα σε 60 ρίψεις

(α) ο παίκτης Α να κερδίσει τουλάχιστον 17 €.

(β) ο παίκτης Α να κερδίσει το πολύ 19 €.

Λύση Έστω X_i το κέρδος του παίκτη Α κατά την i-οστή ρίψη του ζαριού ($i=1,\dots,60$). Τότε:

$$P(X_i=j)=\frac{1}{6},\ j=-3,-2,-1,1,2,3$$

ενώ το συνολικό κέρδος του παίκτη A σε $n=60$ ρίψεις θα δίνεται από την τ.μ. $S=\sum_{i=1}^{n} X_i$.

Από το Κ.Ο.Θ. η τυποποιημένη τ.μ. $Z=\frac{S-E(S)}{\sqrt{Var(S)}}$

θα ακολουθεί κατά προσέγγιση την τυποποιημένη κανονική κατανομή $N(0,1)$.

Όμως, $E[X_i]=\sum_{j} j\,P(X_i=j)=\frac{1}{6}[(-3)+(-2)+(-1)+1+2+3]$
$=0.$

$$Var(X_i)=E[X_i^2]-(E[X_i])^2=E[X_i^2]=\sum_{j} j^2 P(X_i=j)$$
$$=\frac{14}{3}.$$

Άρα

$$E(S)=\sum_{i=1}^{60} E(X_i)=0$$

$$Var(S)=\sum_{i=1}^{60} Var(X_i)=60\cdot\frac{14}{3}=280.$$

(α) $P(S\geq 17)=P\left(\frac{S-0}{\sqrt{280}}\geq\frac{17-0}{\sqrt{280}}\right)=P(Z\geq 1.02)$

$\simeq 1-\Phi(1.02)=0.1539$ (από πίνακες).

(β) Όμοια,

$P(S\leq 19)=P\left(\frac{S-0}{\sqrt{280}}\leq\frac{19-0}{\sqrt{280}}\right)=P(Z\leq 1.14)$

$\simeq \Phi(1.14)=0.8729$ (από πίνακες).

#6 Το ποσό της νικοτίνης που περιέχεται σε ένα τσιγάρο συγκεκριμένης μάρκας είναι τ.μ. με μέση τιμή 0.8 mgr και τυπική απόκλιση 0.1 mgr. Αν ένα άτομο καπνίζει 5 πακέτα τσιγάρα την εβδομάδα, ποια είναι η πιθανότητα το συνολικό ποσό νικοτίνης στο οποίο θα εκτεθεί να είναι τουλάχιστον 82 mgr; Το κάθε πακέτο περιέχει 20 τσιγάρα.

Λύση Το άτομο που εξετάζουμε θα καπνίσει $5 \cdot 20 = 100$ τσιγάρα την εβδομάδα.

Έστω X_i το ποσό της νικοτίνης που περιέχεται στο i-οστό τσιγάρο, $i = 1, \dots, 100$. Θα έχουμε $E[X_i] = 0.8$, $Var(X_i) = (0.1)^2 = 0.01$.

Το συνολικό ποσό νικοτίνης που περιέχεται στα 100 τσιγάρα είναι $S = X_1 + X_2 + \dots + X_{100}$

$$E(S) = 100 \cdot (0.8) = 80, \quad Var(S) = 100 \cdot (0.01) = 1.$$

Με χρήση του Κ.Ο.Θ. η τυποποιημένη τ.μ.

$$Z = \frac{S - E(S)}{\sqrt{Var(S)}} = \frac{S - 80}{1} \underset{\text{προσεγγιστικά}}{\sim} N(0,1).$$

Άρα

$$P(S > 82) = P\left(\frac{S-80}{1} > \frac{82-80}{1}\right) = P(Z > 2)$$

$$= 1 - P(Z \leq 2) \cong 1 - \Phi(2) = 0.0228 \text{ (από πίνακες)},$$

όπου Φ η σ.κ. της $N(0,1)$.