ΚΕΦΑΛΑΙΟ 4

ΒΑΣΙΚΕΣ ΣΥΝΕΧΕΙΣ ΚΑΤΑΝΟΜΕΣ

1. ΟΜΟΙΟΜΟΡΦΗ ΚΑΤΑΝΟΜΗ

Η απλούστερη συνεχής κατανομή πιθανότητας είναι η ομοιόμορφη η οποία εκχωρεί ίσες (ομοιόμορφες) πιθανότητες στα στοιχειώδη δυνατά αποτελέσματα ενός τυχαίου (στοχαστικού) πειράματος με συνεχή (μη απαριθμητό) δειγματικό χώρο *Ω*. Συγκεκριμένα, ας θεωρήσουμε μια συνεχή τυχαία μεταβλητή *X* ορισμένη στον *Ω* με πεδίο τιμών το διάστημα $[\alpha, \beta]$, όπου $\alpha < \beta$ πραγματικοί αριθμοί. Η ομοιόμορφη εκχώρηση πιθανότητας εκφράζεται από τη σχέση

$$P(x_1 < X \leq x_2) = c(x_2 - x_1)\,,\ \alpha \leq x_1 < x_2 \leq \beta\,, \tag{1.1}$$

όπου *c* προσδιοριστέα σταθερή. Θέτοντας $x_1 = \alpha$, $x_2 = \beta$ και χρησιμοποιώντας τη σχέση $P(\alpha < X \leq \beta) = P(\alpha \leq X \leq \beta) = 1$ συμπεραίνουμε ότι

$$c = \frac{1}{\beta - \alpha}. \tag{1.2}$$

Σημειώνουμε ότι στην περίπτωση αυτή, στην οποια η τυχαία μεταβλητή *X* είναι συνεχής, οπότε $P(X = x) = 0$ για κάθε $x \in R$, η εκχώρηση πιθανότητας δεν γίνεται σε σημεία αλλά σε διαστήματα και είναι ανάλογη του μήκους των. Τούτο είναι ισοδύναμο με το ότι διαστήματα του ιδίου μήκους είναι ισοπίθανα.

Η συνάρτηση κατανομής της τυχαίας μεταβλητής *X*, όπως προκύπτει από τις (1.1) και (1.2), δίδεται από την

$$F(x) = \begin{cases} 0, & -\infty < x < \alpha \\ \dfrac{x - \alpha}{\beta - \alpha}, & \alpha \leq x < \beta \\ 1, & \beta \leq x < \infty. \end{cases} \tag{1.3}$$

Η συνάρτηση αυτή είναι συνεχής και έτσι παραγωγίζοντάς την συνάγουμε τη συνάρτηση πυκνότητας της τυχαίας μεταβλητής *X*:

$$f(x) = \begin{cases} \dfrac{1}{\beta - \alpha}, & \alpha \leq x \leq \beta \\ 0, & x < \alpha \ \text{ ή } \ x > \beta. \end{cases} \tag{1.4}$$

Ορισμός 1.1. *Έστω X μια συνεχής τυχαία μεταβλητή με συνάρτηση πυκνότητας την (1.4). Η κατανομή της τ.μ. X συμβολίζεται με* $U(\alpha, \beta)$ *και καλείται ομοιόμορφη ή ορθογώνια στο διάστημα* $[\alpha, \beta]$. *Τα σημεία α και β είναι παράμετροι της κατανομής.*

Σχετικά με τις ροπές της ομοιόμορφης κατανομής αποδεικνύουμε το επόμενο θεώρημα.

Θεώρημα 1.1. *Έστω ότι η τυχαία μεταβλητή X έχει την ομοιόμορφη κατανομή $U(\alpha,\beta)$. Τότε η μέση τιμή και η διασπορά αυτής δίδονται από τις*

$$\mu = E(X) = \frac{\alpha+\beta}{2},\ \sigma^2 = V(X) = \frac{(\beta-\alpha)^2}{12}. \tag{1.5}$$

Απόδειξη. Η μέση τιμή της τ.μ. X, σύμφωνα με τον ορισμό, είναι

$$\mu = E(X) = \frac{1}{\beta-\alpha}\int_\alpha^\beta x dx = \left[\frac{x^2}{2(\beta-\alpha)}\right]_\alpha^\beta = \frac{\beta^2-\alpha^2}{2(\beta-\alpha)}$$

και επειδή $(\beta^2-\alpha^2) = (\beta-\alpha)(\beta+\alpha)$,

$$\mu = E(X) = \frac{\alpha+\beta}{2}.$$

Επίσης είναι

$$E(X^2) = \frac{1}{\beta-\alpha}\int_\alpha^\beta x^2 dx = \left[\frac{x^3}{3(\beta-\alpha)}\right]_\alpha^\beta = \frac{\beta^3-\alpha^3}{3(\beta-\alpha)}$$

και επειδή $\beta^3-\alpha^3 = (\beta-\alpha)(\beta^2+\alpha\beta+\alpha^2)$,

$$E(X^2) = \frac{\alpha^2+\alpha\beta+\beta^2}{3}.$$

Η διασπορά της τ.μ. X είναι τότε

$$\sigma^2 = V(X) = E(X^2) - [E(X)]^2 = \frac{\alpha^2+\alpha\beta+\beta^2}{3} - \frac{\alpha^2+2\alpha\beta+\beta^2}{4} = \frac{(\beta-\alpha)^2}{12}.$$

Παράδειγμα 1.1. Ας θεωρήσουμε ένα όργανο μέτρησης με ακρίβεια τριών δεκαδικών ψηφίων. Το παρεχόμενο από το όργανο αυτό τέταρτο δεκαδικό ψηφίο αποτελεί στρογγύλευση προς τον πλησιέστερο ακέραιο. Τα σφάλματα που προκύπτουν από τη στρογγύλευση της μέτρησης δύνανται να θεωρηθούν ότι έχουν την ομοιόμορφη κατανομή $U(\alpha,\beta)$ με $\alpha = -10^{-4}/2$, $\beta = 10^{-4}/2$. Να υπολογισθούν (α) η πιθανότητα όπως το σφάλμα μέτρησης μιας ποσότητας είναι κατ' απόλυτη τιμή μεγαλύτερο του $10^{-4}/3$ και (β) η μέση τιμή και η διασπορά του σφάλματος μέτρησης.

(α) Χρησιμοποιώντας την (1.3) με $\alpha = -10^{-4}/2$, $\beta = 10^{-4}/2$ παίρνουμε

$$P(|X| > 10^{-4}/3) = 1 - P(|X| \le 10^{-4}/3) = 1 - [F(10^{-4}/3) - F(-10^{-4}/3)]$$
$$= 1 - \frac{2}{3} = \frac{1}{3}.$$

(β) Σύμφωνα με τις (1.5) έχουμε

$$\mu = E(X) = 0,\ \sigma^2 = V(X) = 10^{-8}/12.$$

Παράδειγμα 1.2. Έστω ότι ο σειρμός φθάνει σε συγκεκριμένο σταθμό του υπογείου σιδηροδρόμου κάθε 10 λεπτά, αρχίζοντας τα δρομολόγιά του στις 5 π.μ. Αν ένας επιβάτης φθάνει στο σταθμό σε χρόνο ο οποίος κατανέμεται ομοιόμορφα στο διάστημα 7:20 ως 7:40 να υπολογισθούν οι πιθανότητες να περιμένει το σειρμό (α) το πολύ 4 λεπτά και (β) τουλάχιστο 7 λεπτά.

Έστω X ο χρόνος άφιξης του επιβάτη στο σταθμό, μετρούμενος σε λεπτά με αρχή τη χρονική στιγμή 7:20. Τότε η τ.μ. X έχει την ομοιόμορφη κατανομή στο διάστημα $[0,20]$ και έτσι

$$F(x)=\begin{cases}0, & x<0\\ \dfrac{x}{20}, & 0\le x<20\\ 1, & x\ge 20.\end{cases}$$

(α) Το ενδεχόμενο A ο επιβάτης να περιμένει το πολύ 4 λεπτά είναι ισοδύναμο με το ενδεχόμενο να φθάσει στο σταθμό στο διάστημα 7:26 ως 7:30 ή στο διάστημα 7:36 ως 7:40. Επομένως

$$P(A)=P(6<X\le 10)+P(16<X\le 20)=\{F(10)-F(6)\}+\{F(20)-F(16)\}=\frac{2}{5}.$$

(β) Το ενδεχόμενο B ο επιβάτης να περιμένει τουλάχιστο 7 λεπτά είναι ισοδύναμο με το ενδεχόμενο να φθάσει στο σταθμό στο διάστημα 7:20 ως 7:23 ή 7:30 ως 7:33. Επομένως

$$P(B)=P(0<X\le 3)+P(10<X\le 13)=\{F(3)-F(0)\}+\{F(13)-F(10)\}=\frac{3}{10}.$$

2. ΕΚΘΕΤΙΚΗ ΚΑΤΑΝΟΜΗ ΚΑΙ ΚΑΤΑΝΟΜΗ ERLANG

2.1. Εκθετική κατανομή

Ορισμός 2.1. *Έστω X μια συνεχής τυχαία μεταβλητή με συνάρτηση πυκνότητας*

$$f(x)=\begin{cases}\theta e^{-\theta x}, & 0\le x<\infty\\ 0, & -\infty<x<0,\end{cases} \tag{2.1}$$

όπου $0<\theta<\infty$. Η κατανομή της τ.μ. X καλείται εκθετική με παράμετρο θ.

Σημειώνουμε ότι η συνάρτηση (2.1) είναι μη αρνητική και

$$\int_{-\infty}^{\infty} f(x)dx=\int_0^{\infty}\theta e^{-\theta x}dx=\left[-e^{-\theta x}\right]_0^{\infty}=1,$$

όπως απαιτείται από τον ορισμό της συνάρτησης πυκνότητας.

Η συνάρτηση κατανομής της τ.μ. X, σύμφωνα με την (2.10) του Κεφ. 2, είναι η

$$F(x)=\begin{cases}0, & -\infty<x<0\\ 1-e^{-\theta x}, & 0\le x<\infty.\end{cases} \tag{2.2}$$

Θεώρημα 2.1. *Έστω ότι η τυχαία μεταβλητή X έχει την εκθετική κατανομή με συνάρτηση πυκνότητας τη (2.1). Τότε η μέση τιμή και η διασπορά αυτής δίδονται από τις*

$$\mu=E(X)=\frac{1}{\theta},\ \sigma^2=V(X)=\frac{1}{\theta^2}. \tag{2.3}$$

Απόδειξη. Η μέση τιμή της τ.μ. X, σύμφωνα με τον ορισμό, δίδεται από την

$$\mu = E(X) = \int_{-\infty}^{\infty} x f(x) dx = \int_{0}^{\infty} \theta x e^{-\theta x} dx = \frac{1}{\theta}\int_{0}^{\infty} y e^{-y} dy,$$

όπου χρησιμοποιήθηκε ο μετασχηματισμός $y = \theta x$. Εφαρμόζοντας την ολοκλήρωση κατά παράγοντες το τελευταίο ολοκλήρωμα είναι

$$\int_{0}^{\infty} y e^{-y} dy = -\int_{0}^{\infty} y de^{-y} = -[ye^{-y}]_0^{\infty} + \int_{0}^{\infty} e^{-y} dy = -[ye^{-y} + e^{-y}]_0^{\infty} = 1$$

και έτσι

$$\mu = E(X) = \frac{1}{\theta}.$$

Ομοίως

$$E(X^2) = \int_{-\infty}^{\infty} x^2 f(x) dx = \int_{0}^{\infty} \theta x^2 e^{-\theta x} dx = \frac{1}{\theta^2}\int_{0}^{\infty} y^2 e^{-y} dy$$

και επειδή

$$\int_{0}^{\infty} y^2 e^{-y} dy = -\int_{0}^{\infty} y^2 de^{-y} = -[y^2 e^{-y}]_0^{\infty} + 2\int_{0}^{\infty} y e^{-y} dy = -[y^2 e^{-y} + 2ye^{-y} + 2e^{-y}]_0^{\infty} = 2$$

παίρνουμε

$$E(X^2) = \frac{2}{\theta^2}.$$

Επομένως

$$V(X) = E(X^2) - [E(X)]^2 = \frac{1}{\theta^2}.$$

Η ιδιότητα του *αμνήμονος* είναι χαρακτηριστική της εκθετικής κατανομής. Την ιδιότητα αυτή αποδεικύουμε στο επόμενο θεώρημα.

Θεώρημα 2.2. *Έστω ότι η τυχαία μεταβλητή X έχει την εκθετική κατανομή με συνάρτηση πυκνότητας τη (2.1). Τότε*

$$P(X > x + y \mid X > x) = P(X > y), \quad x \geq 0, \quad y \geq 0. \tag{2.4}$$

Απόδειξη. Η δεσμευμένη πιθανότητα του ενδεχομένου $\{X > x + y\}$ δεδομένου του ενδεχομένου $\{X > x\}$, λαμβάνοντας υπόψη ότι $\{X > x + y\} \subseteq \{X > x\}$ και χρησιμοποιώντας την (2.2), είναι ίση με

$$P(X > x + y \mid X > x) = \frac{P(X > x + y, X > x)}{P(X > x)} = \frac{P(X > x + y)}{P(X > x)}$$

$$= \frac{1 - F(x + y)}{1 - F(x)} = \frac{e^{-\theta(x+y)}}{e^{-\theta x}} = e^{-\theta y}$$

και επειδή

$$P(X > y) = 1 - F(y) = e^{-\theta y}$$

έπεται η (2.4).

Παρατήρηση 2.1. Ας θεωρήσουμε μια ανέλιξη Poisson X_t, $t \geq 0$ με μέση τιμή $E(X_t) = \theta t$ (βλ. Παρατήρηση 5.2 του Κεφ. 3) και ας παραστήσουμε με *T* το χρόνο αναμονής μέχρι την πραγματοποίηση της πρώτης επιτυχίας (εμφάνισης του ενδεχομένου *A*). Επειδή το ενδεχόμενο $\{T > t\}$, όπως η πρώτη επιτυχία πραγματοποιηθεί μετά τη χρονική στιγμή *t*, είναι ισοδύναμο με το ενδεχόμενο $\{X_t = 0\}$, όπως ο αριθμός των επιτυχιών μέχρι τη χρονική στιγμή *t* είναι μηδέν, χρησιμοποιώντας την (5.6) του Κεφ. 3, συνάγουμε τη σχέση

$$P(T > t) = P(X_t = 0) = e^{-\theta t}, \ t \geq 0$$

και από αυτή τη συνάρτηση κατανομής της τ.μ. *T*,

$$F(t) = \begin{cases} 0, & -\infty < t < 0 \\ 1 - e^{-\theta t}, & 0 \leq t < \infty. \end{cases} \tag{2.5}$$

Επομένως, σύμφωνα με τη (2.2) ο χρόνος αναμονής *T* μέχρι την πραγματοποίηση της πρώτης επιτυχίας σε μια ανέλιξη Poisson έχει εκθετική κατανομή. Γενικότερα δύναται να δειχθεί ότι οι ενδιάμεσοι χρόνοι μεταξύ διαδοχικών επιτυχιών σε μια ανέλιξη Poisson έχουν εκθετική κατανομή.

Παράδειγμα 2.1. Έστω ότι η διάρκεια σε λεπτά ενός τηλεφωνήματος, σ' ένα δημόσιο τηλεφωνικό θάλαμο, ακολουθεί την εκθετική κατανομή με μέση τιμή 10 λεπτά. Επίσης, έστω ότι τη στιγμή που κάποιος μπαίνει στον τηλεφωνικό αυτό θάλαμο για ένα τηλεφώνημα ένας άλλος φθάνει εκεί και δεν συναντά κανένα να περιμένει. Να υπολογισθούν οι πιθανότητες ο δεύτερος να περιμένει (α) περισσότερο από 10 λεπτά (β) μεταξύ 10 και 20 λεπτών.

Αν *X* είναι η διάρκεια του τηλεφωνήματος του πρώτου ατόμου, τότε

$$F(x) = \begin{cases} 0, & x < 0 \\ 1 - e^{-x/10}, & x \geq 0 \end{cases}$$

και οι ζητούμενες πιθανότητες είναι (α)

$$P(X > 10) = 1 - F(10) = e^{-1} = 0{,}3679,$$

και (β)

$$P(10 < X \leq 20) = F(20) - F(10) = e^{-1} - e^{-2} = 0{,}3679 - 0{,}1353 = 0{,}2326.$$

2.2. Κατανομή Erlang

Ορισμός 2.2. *Έστω X μια συνεχής τυχαία μεταβλητή με συνάρτηση πυκνότητας*

$$f(x) = \begin{cases} \dfrac{\theta^\nu}{(\nu-1)!} x^{\nu-1} e^{-\theta x}, & 0 \leq x < \infty \\ 0, & -\infty < x < 0, \end{cases} \tag{2.6}$$

όπου ν θετικός ακέραιος και $0 < \theta < \infty$. *Η κατανομή της τ.μ. X καλείται κατανομή Erlang με παραμέτρους ν και θ.*

Σημειώνουμε ότι η συνάρτηση (2.6) είναι μη αρνητική και επειδή

$$I_\nu = \int_0^\infty x^{\nu-1} e^{-x} dx = (\nu-1)!, \ \nu = 1, 2, \ldots, \tag{2.7}$$

συμπεραίνουμε ότι

$$\int_{-\infty}^{\infty} f(x)dx = \frac{\theta^{\nu}}{(\nu-1)!}\int_{0}^{\infty} x^{\nu-1}e^{-\theta x}dx = \frac{1}{(\nu-1)!}\int_{0}^{\infty} y^{\nu-1}e^{-y}dy = 1\,,$$

όπως απαιτείται από τον ορισμό της συνάρτησης πυκνότητας.

Το ολοκλήρωμα I_{ν}, $\nu = 1, 2, \ldots$ δύναται να υπολογισθεί, εφαρμόζοντας την ολοκλήρωση κατά παράγοντες, ως εξής:

$$I_{\nu+1} = \int_{0}^{\infty} x^{\nu}e^{-x}dx = -\int_{0}^{\infty} x^{\nu}de^{-x} = -[x^{\nu}e^{-x}]_{0}^{\infty} + \nu\int_{0}^{\infty} x^{\nu-1}e^{-x}dx$$

και έτσι

$$I_{\nu+1} = \nu I_{\nu}\,,\ \nu = 1, 2, \ldots. \tag{2.8}$$

Εφαρμόζοντας διαδοχικά την αναγωγική αυτή σχέση και επειδή

$$I_1 = \int_{0}^{\infty} e^{-x}dx = 1\,,$$

συνάγουμε τη (2.7).

Θεώρημα 2.3. *Έστω ότι η τυχαία μεταβλητή X έχει την κατανομή Erlang με συνάρτηση πυκνότητας τη (2.6). Τότε η μέση τιμή και η διασπορά αυτής δίδονται από τις*

$$\mu = E(X) = \frac{\nu}{\theta}\,,\ \sigma^2 = V(X) = \frac{\nu}{\theta^2}\,. \tag{2.9}$$

Απόδειξη. Η μέση τιμή της τυχαίας μεταβλητής X, δίδεται από την

$$\mu = E(X) = \int_{-\infty}^{\infty} xf(x)dx = \frac{\theta^{\nu}}{(\nu-1)!}\int_{0}^{\infty} x^{\nu}e^{-\theta x}dx = \frac{1}{\theta(\nu-1)!}\int_{0}^{\infty} y^{\nu}e^{-y}dy$$

και χρησιμοποιώντας την (2.7) συνάγουμε την

$$E(X) = \frac{\nu!}{\theta(\nu-1)!} = \frac{\nu}{\theta}\,.$$

Ομοίως

$$E(X^2) = \int_{-\infty}^{\infty} x^2 f(x)dx = \frac{\theta^{\nu}}{(\nu-1)!}\int_{0}^{\infty} x^{\nu+1}e^{-\theta x}dx = \frac{1}{\theta^2(\nu-1)!}\int_{0}^{\infty} y^{\nu+1}e^{-y}dy$$

και

$$E(X^2) = \frac{(\nu+1)!}{\theta^2(\nu-1)!} = \frac{(\nu+1)\nu}{\theta^2}\,.$$

Επομένως η διασπορά της τ.μ. X είναι

$$\sigma^2 = V(X) = E(X^2) - [E(X)]^2 = \frac{(\nu+1)\nu}{\theta^2} - \frac{\nu^2}{\theta^2} = \frac{\nu}{\theta^2}\,.$$

Παρατήρηση 2.2. Ας θεωρήσουμε μια ανέλιξη Poisson X_t, $t \geq 0$ με μέση τιμή $E(X_t) = \theta t$ (βλ. παρατήρηση 5.2 του Κεφ. 3) και ας παραστήσουμε με T_ν το χρόνο αναμονής μέχρι την πραγματοποίηση της ν-οστής επιτυχίας (εμφάνισης του ενδεχομένου A). Επειδή το ενδεχόμενο $\{T_\nu > t\}$, όπως η ν-οστή επιτυχία πραγματοποιηθεί μετά τη χρονική στιγμή t είναι ισοδύναμο με το ενδεχόμενο $\{X_t < \nu\}$, όπως ο αριθμός των επιτυχιών μέχρι τη χρονική στιγμή t είναι μικρότερος του ν, χρησιμοποιώντας την (5.6) του Κεφ. 3, συνάγουμε τη σχέση

$$P(T_\nu > t) = P(X_t < \nu) = \sum_{\kappa=0}^{\nu-1} P(X_t = \kappa) = \sum_{\kappa=0}^{\nu-1} e^{-\theta t} \frac{(\theta t)^\kappa}{\kappa!}, \ t \geq 0 .$$

Η συνάρτηση κατανομής της τ.μ. T_ν δίδεται τότε από την

$$F(t) = 1 - e^{-\theta t} \sum_{\kappa=0}^{\nu-1} \frac{(\theta t)^\kappa}{\kappa!}, \ t \geq 0, \tag{2.10}$$

με $F(t) = 0, \ t < 0$. Παραγωγίζοντας αυτήν ως προς t παίρνουμε

$$\frac{d}{dt} F(t) = \theta e^{-\theta t} \sum_{\kappa=0}^{\nu-1} \frac{(\theta t)^\kappa}{\kappa!} - e^{-\theta t} \sum_{\kappa=1}^{\nu-1} \frac{\theta (\theta t)^{\kappa-1}}{(\kappa-1)!}$$

και επομένως η συνάρτηση πυκνότητας της τ.μ. T_ν είναι η

$$f(t) = \frac{\theta^\nu}{(\nu-1)!} t^{\nu-1} e^{-\theta t}, \ 0 \leq t < \infty .$$

Η κατανομή αυτή μελετήθηκε από το Δανό μαθηματικό A. K. Erlang (1878-1929). Σημειώνουμε ότι η σχέση (2.10), επειδή

$$F(t) = \int_0^t \frac{\theta^\nu}{(\nu-1)!} x^{\nu-1} e^{-\theta x} dx$$

συνεπάγεται τη χρήσιμη στις εφαρμογές σχέση

$$F(t) = \int_0^t \frac{\theta^\nu}{(\nu-1)!} x^{\nu-1} e^{-\theta x} dx = 1 - e^{-\theta t} \sum_{\kappa=0}^{\nu-1} \frac{(\theta t)^\kappa}{\kappa!} . \tag{2.11}$$

Παράδειγμα 2.2. Η ημερήσια κατανάλωση ηλεκτρικής ενέργειας σε εκατομύρια κιλοβατόρες σε μια πόλη είναι μια τυχαία μεταβλητή X η οποία ακολουθεί την κατανομή Erlang με μέση τιμή $E(X) = 10$ εκατομύρια κιλοβατόρες και διασπορά $V(X) = 5$ εκατομύρια κιλοβατόρες. Η μέγιστη ποσότητα ενέργειας που μπορεί να δοθεί στην πόλη σε μια μέρα είναι 15 εκατομύρια κιλοβατόρες. Να υπολογισθεί η πιθανότητα να μη ικανοποιηθούν οι ημερήσιες ανάγκες της πόλης σε ηλεκτρική ενέργεια.

Η τυχαία μεταβλητή X έχει συνάρτηση πυκνότητας

$$f(x) = \frac{\theta^\nu}{(\nu-1)!} x^{\nu-1} e^{-\theta x}, \ 0 \leq x < \infty$$

όπου ν θετικός ακέραιος και $0 < \theta < \infty$. Η μέση τιμή και διασπορά της X, σύμφωνα με τις (2.9) είναι

$$E(X) = \frac{\nu}{\theta}, \ V(X) = \frac{\nu}{\theta^2}$$

και επειδή $E(X)=10$, $V(X)=5$, παίρνουμε

$$\nu=4\,,\ \theta=2/5\,.$$

Επομένως

$$f(x)=\frac{(2/5)^4}{3!}x^3e^{-2x/5}\,,\ 0\le x<\infty$$

και η πιθανότητα να μη ικανοποιηθούν οι ημερήσιες ανάγκες της πόλης σε ηλεκτρική ενέργεια δίδεται από την

$$P(X>15)=\int_{15}^{\infty}\frac{(2/5)^4}{3!}x^3e^{-2x/5}dx\,.$$

Χρησιμοποιώντας τη (2.11) και τον Πίνακα 2 της συνάρτησης πιθανότητας της κατανομής παίρνουμε

$$P(X>15)=\sum_{\kappa=0}^{3}e^{-6}\frac{6^{\kappa}}{\kappa!}=0{,}1422\,.$$

Παράδειγμα 2.3. Έστω ότι ο αριθμός των τραυματιών σε αυτοκινητιστικά δυστυχήματα με σοβαρά κατάγματα που εισάγονται σε νοσοκομεία των Αθηνών ακολουθεί την κατανομή Poisson με μέση τιμή 8 άτομα ανά ημέρα. Να υπολογισθούν (α) η πιθανότητα όπως ο χρόνος αναμονής μέχρι την άφιξη του τρίτου τραυματία, μετρούμενος από την αρχή της ημέρας, είναι τουλάχιστο 12 ώρες και (β) ο μέσος χρόνος αναμονής μέχρι την άφιξη του τρίτου τραυματία.

(α) Ο αριθμός X_t των τραυματιών σε χρονικό διάστημα t ωρών ακολουθεί την κατανομή Poisson με μέση τιμή $E(X_t)=\theta t$, όπου $\theta=8/24=1/3$. Ο χρόνος αναμονής T_3 ακολουθεί την κατανομή Erlang με συνάρτηση κατανομής

$$F(t)=1-e^{-t/3}\sum_{\kappa=0}^{2}\frac{(t/3)^{\kappa}}{\kappa!}\,.$$

Επομένως

$$P(T_3>12)=1-F(12)=1-e^{-4}\sum_{\kappa=0}^{2}\frac{4^{\kappa}}{\kappa!}$$

και χρησιμοποιώντας τον Πίνακα 2 της συνάρτησης πιθανότητας της κατανομής Poisson παίρνουμε

$$P(T_3>12)=1-(0{,}0183+0{,}0733+0{,}1465)=0{,}7619\,.$$

(β) Η μέση τιμή της T_3, σύμφωνα με την πρώτη από τις (2.9), είναι

$$E(T_3)=\frac{3}{\theta}=9\,.$$

3. ΚΑΝΟΝΙΚΗ ΚΑΤΑΝΟΜΗ

Η σημαντικότερη κατανομή πιθανότητας τόσο από θεωρητική άποψη όσο και από άποψη εφαρμογών είναι η κανονική κατανομή. Η κατανομή αυτή χρησιμοποιήθηκε αρχικά από τους De Moivre και Laplace για την προσέγγιση της διωνυμικής κατανομής για μεγάλο ν. Ο Gauss, ο οποίος διατύπωσε τη θεωρία των σφαλμάτων, χρησιμοποίησε την κανονική κατανομή ως προσεγγιστική της κατανομής των τυχαίων σφαλμάτων. Η ονομασία κανονική είναι σχετικά πρόσφατη και οφείλεται στον Karl Pearson.

Ορισμός 3.1. *Έστω Χ μια συνεχής τυχαία μεταβλητή Χ με συνάρτηση πυκνότητας*

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \quad -\infty < x < \infty, \tag{3.1}$$

όπου $-\infty < \mu < \infty$ *και* $0 < \sigma < \infty$ *είναι παράμετροι. Η κατανομή της τυχαίας μεταβλητής Χ συμβολίζεται με* $N(\mu, \sigma^2)$ *και καλείται κανονική κατανομή με παραμέτρους* μ *και* σ^2.

Σημειώνουμε ότι η συνάρτηση (3.1) είναι μη αρνητική και χρησιμοποιώντας διαδοχικά τους μετασχηματισμούς $z = (x-\mu)/\sigma$ και $u = z/\sqrt{2}$ και το ολοκλήρωμα του Euler,

$$\int_{-\infty}^{\infty} e^{-u^2} du = \frac{\sqrt{\pi}}{2},$$

συμπεραίνουμε ότι

$$\int_{-\infty}^{\infty} f(x)dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-z^2/2} dz = \frac{2}{\sqrt{\pi}} \int_{0}^{\infty} e^{-u^2} du = 1,$$

όπως απαιτείται από τον ορισμό της συνάρτησης πυκνότητας.

Παρατήρηση 5.1. Η εξέχουσα θέση την οποία κατέχει η κανονική κατανομή στη Θεωρία Πιθανοτήτων και Στατιστική οφείλεται και στη μεγάλη ποικιλία των εφαρμογών της. Συγκεκριμένα, τα τυχαία σφάλματα που εμφανίζονται σε διάφορες μετρήσεις έχουν κανονική κατανομή. Για το λόγο αυτό, η κανονική κατανομή αναφέρεται πολλές φορές και ως *κατανομή σφαλμάτων*. Επίσης, πολλές κατανομές τόσο διακριτές όσο και συνεχείς μπορούν κάτω από ορισμένες συνθήκες να προσεγγισθούν από την κανονική κατανομή. Το άθροισμα και ο μέσος όρος μεγάλου αριθμού παρατηρήσεων ακολουθεί κατά προσέγγιση κανονική κατανομή ανεξάρτητα από το ποιά κατανομή ακολουθούν οι αρχικές παρατηρήσεις. Ακόμη, πολλά πληθυσμιακά χαρακτηριστικά (π.χ. ύψος, βάρος, βαθμολογία σε τεστ κλπ) ακολουθούν (περιγράφονται ικανοποιητικά από) την κανονική κατανομή

Η μέση τιμή και η διασπορά της κανονικής κατανομής συνάγονται στο ακόλουθο θεώρημα.

Θεώρημα 3.1. *Έστω ότι η τυχαία μεταβλητή Χ έχει την κανονική κατανομή με συνάρτηση πυκνότητας την (3.1). Τότε η μέση τιμή και η διασπορά αυτής δίδονται από τις*

$$E(X) = \mu, \; V(X) = \sigma^2. \tag{3.2}$$

Απόδειξη. Η μέση τιμή της τυχαίας μεταβλητής *X* δίδεται από την

$$E(X) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} x e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx.$$

Θέτοντας $z = (x-\mu)/\sigma$, οπότε $x = \mu + \sigma z$, παίρνουμε

$$E(X)=\mu\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{-z^2/2}dz+\sigma\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}ze^{-z^2/2}dz=\mu+\sigma\frac{1}{\sqrt{2\pi}}\left[e^{-z^2/2}\right]_{-\infty}^{\infty}=\mu .$$

Η διασπορά της τυχαίας μεταβλητής X δίδεται από την

$$V(X)=\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty}(x-\mu)^2 e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}dx .$$

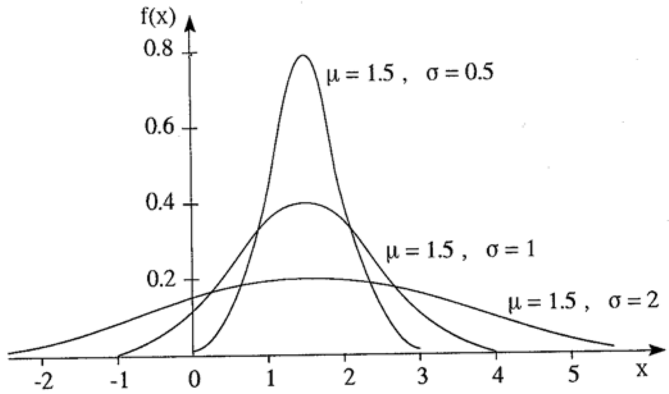
Θέτοντας $z=(x-\mu)/\sigma$ και ολοκληρώνοντας κατά παράγοντες παίρνουμε

$$V(X)=\sigma^2\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}z^2e^{-z^2/2}dz=-\sigma^2\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}zde^{-z^2/2}$$

$$=\sigma^2\frac{1}{\sqrt{2\pi}}\left[ze^{-z^2/2}\right]_{-\infty}^{\infty}+\sigma^2\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{-z^2/2}dz=\sigma^2 ,$$

και έτσι συμπληρώνεται η απόδειξη του θεωρήματος.

Το διάγραμμα της συνάρτησης πυκνότητας $f(x)$, $-\infty<x<\infty$ της κανονικής κατανομής είναι κωδωνοειδούς μορφής. Η παράμετρος μ προσδιορίζει τη θέση της καμπύλης ως προς τον άξονα των x και η παράμετρος σ το οξύ ή πεπλατισμένο του σχήματός της. Συγκεκριμένα, όσο πιο μικρό είναι το σ τόσο πιο οξεία είναι η καμπύλη και όσο πιο μεγάλο είναι το σ τόσο πιο πεπλατισμένη είναι αυτή (βλέπε Σχήμα 3.1).



Σχήμα 3.1. Η συνάρτηση πυκνότητας $f(x)$ της κανονικής κατανομής

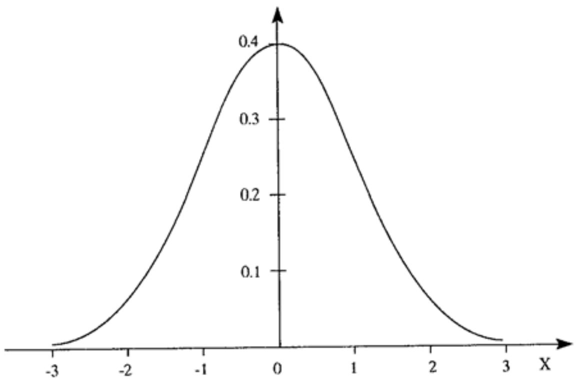
Ιδιαίτερο ενδιαφέρον παρουσιάζει η ειδική περίπτωση κανονικής κατανομής με μέση τιμή $\mu=0$ και διασπορά $\sigma^2=1$. Σχετικά σημειώνουμε ότι η τυποποιημένη τυχαία μεταβλητή $Z=(X-\mu)/\sigma$ έχει συνάρτηση πυκνότητας (βλέπε Παράδειγμα 3.1 του Κεφαλαίου 3)

$$\varphi(z)=\frac{1}{\sqrt{2\pi}}e^{-z^2/2},\ -\infty<z<\infty .$$

Η κατανομή της τυχαίας μεταβλητής Z, η οποία είναι $N(0,1)$, καλείται *τυποποιημένη κανονική κατανομή*. Η γραφική παράσταση της συνάρτησης πυκνότητας της τυποποιημένης κανονικής κατανομής δίδεται στο Σχήμα 3.2. Η χρησιμότητα της τυποποιημένης κανονικής κατανομής οφείλεται στην πινακοποίηση της κατανομής της

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-t^2/2} dt, \ -\infty < z < \infty .$$

Ο συνήθης πίνακας της τυποποιημένης κανονικής κατανομής δίδει της τιμές $\Phi(z)$ για z από 0 μέχρι 3 με βήμα 0,01. Για τον προσδιορισμό των τιμών της $\Phi(z)$ για z από -3 μέχρι 0 χρησιμοποιείται η ακόλουθη ιδιότητα της τυποποιημένης κακονικής κατανομής.



Σχήμα 3.1. Η συνάρτηση πυκνότητας της τυποποιημένης κανονικής κατανομής

Θεώρημα 3.2. *Η συνάρτηση κατανομής της τυποποιημένης κανονικής κατανομής* $\Phi(z)$, $-\infty < z < \infty$, *ικανοποιεί τη σχέση*

$$\Phi(z) + \Phi(-z) = 1, \ -\infty < z < \infty .$$

Απόδειξη. Η συνάρτηση $\Phi(-z)$

$$\Phi(-z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-z} e^{-t^2/2} dt ,$$

χρησιμοποιώντας το μετασχηματισμό $t = -u$, παίρνει τη μορφή

$$\Phi(-z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-z} e^{-t^2/2} dt = \frac{1}{\sqrt{2\pi}} \int_{z}^{\infty} e^{-u^2/2} du .$$

Επομένως

$$\Phi(z) + \Phi(-z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-t^2/2} dt + \frac{1}{\sqrt{2\pi}} \int_{z}^{\infty} e^{-u^2/2} du = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-t^2/2} dt = 1 .$$

και η απόδειξη ολοκληρώθηκε.

Η προσέγγιση της συνάρτησης πιθανότητας της διωνυμικής κατανομής για μεγάλο ν (θεωρητικά $\nu \to \infty$) από τη συνάρτηση πυκνότητας της κανονικής κατανομής, η οποία αποδείχθηκε αρχικά για $p = 1/2$ από τον De Moivre το 1733 και για οποιοδήποτε p από τον Laplace το 1812, δίδεται χωρίς απόδειξη στο επόμενο θεώρημα.

Θεώρημα 3.3. *Έστω ότι η τυχαία μεταβλητή X ακολουθεί τη διωνυμική κατανομή με συνάρτηση πιθανότητας*

$$f(x)=\binom{\nu}{x}p^x q^{\nu-x}, \ x=0,1,\ldots,\nu,$$

όπου $q=1-p$ *και* $0<p<1$. *Τότε για μεγάλο* ν (*θεωρητικά* $\nu\to\infty$) *ισχύει η προσέγγιση*

$$f(x)\cong\frac{1}{\sqrt{\nu pq}\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x-\nu p}{\sqrt{\nu pq}}\right)^2}, \ -\infty<x<\infty.$$

Χρήσιμο στις εφαρμογές είναι το ακόλουθο πόρισμα.

Πόρισμα 3.1. *Έστω ότι η τυχαία μεταβλητή Χ ακολουθεί τη διωνυμική κατανομή με παραμέτρους ν και p. Τότε για μεγάλο ν* (*θεωρητικά* $\nu\to\infty$) *ισχύει η προσέγγιση*

$$P(\alpha\le X\le\beta)\cong\Phi\left(\frac{\beta-\nu p}{\sqrt{\nu pq}}\right)-\Phi\left(\frac{\alpha-\nu p}{\sqrt{\nu pq}}\right), \ \alpha<\beta.$$

Παρατήρηση 3.1. *Διόρθωση συνέχειας*. Η προσέγγιση της διωνυμικής κατανομής από την κανονική κατανομή δεν είναι ικανοποιητική όταν το ν δεν είναι αρκετά μεγάλο. Γενικά, στις περιπτώσεις προσέγγισης μιας διακριτής κατανομής απο μια συνεχή έχει αποδειχθεί ότι οι προσεγγίσεις βελτιώνονται σημαντικά με τη χρησιμοποίηση της διόρθωσης συνέχειας. Σύμφωνα με αυτή η πιθανότητα $f(x)=P(X=x)$, $x=0,1,\ldots$ προσεγγίζεται από την πιθανότητα $P(x-1/2\le X\le x+1/2)$ και έτσι στην περίπτωση κανονικής προσέγγισης

$$P(X=x)\cong\Phi\left(\frac{x-\nu p+1/2}{\sqrt{\nu pq}}\right)-\Phi\left(\frac{x-\nu p-1/2}{\sqrt{\nu pq}}\right).$$

Γενικότερα

$$P(\alpha\le X\le\beta)\cong\Phi\left(\frac{\beta-\nu p+1/2}{\sqrt{\nu pq}}\right)-\Phi\left(\frac{\alpha-\nu p-1/2}{\sqrt{\nu pq}}\right), \ \alpha\le\beta.$$

Παράδειγμα 3.1. Ας υποθέσουμε ότι η διάρκεια κύησης X μιας γυναίκας ακολουθεί την κανονική κατανομή με μέση τιμή $\mu=270$ ημέρες και τυπική απόκλιση $\sigma=30$ ημέρες. Να υπολογισθεί η πιθανότητα όπως η διάρκεια κύησης παιδιού είναι μικρότερη από επτά μήνες.

Εισάγοντας την τυποποιημένη τυχαία μεταβλητή $Z=(X-\mu)/\sigma$, όπου $\mu=270$ και $\sigma=30$, συνάγουμε για τη ζητουμένη πιθανότητα την έκφραση

$$P(X<210)=P\left(\frac{X-270}{30}<\frac{210-270}{30}\right)=P(Z<-2)=\Phi(-2)=1-\Phi(2).$$

Από τον πίνακα της τυποποιημένης κανονικής κατανομής έχουμε $\Phi(2)=0{,}9772$ και έτσι

$$P(X<210)=1-0{,}9772=0{,}0228\cong 2\%.$$

Παράδειγμα 3.2. Έστω ότι η τυχαία μεταβλητή X ακολουθεί την κανονική κατανομή $N(\mu,\sigma^2)$. Να υπολογισθεί η πιθανότητα όπως η X απέχει από το μέσο μ το πολύ κ τυπικές αποκλίσεις σ, για $\kappa=1,2,3$.

Εισάγοντας την τυποποιημένη τυχαία μεταβλητή $Z = (X-\mu)/\sigma$ συνάγουμε για τη ζητουμένη πιθανότητα την έκφραση

$$P(|X-\mu| \le \kappa\sigma) = P(|Z| \le \kappa) = P(-\kappa \le Z \le \kappa) = \Phi(\kappa) - \Phi(-\kappa) = 2\Phi(\kappa) - 1 .$$

Από τον πίνακα της τυποποιημένης κανονικής κατανομής έχουμε $\Phi(1) = 0{,}8413$, $\Phi(2) = 0{,}9772$, $\Phi(3) = 0{,}9987$ και έτσι

$$P(\mu - \sigma \le X \le \mu + \sigma) = P(|X-\mu| \le \sigma) = 0{,}6826 \cong 68\% ,$$

$$P(\mu - 2\sigma \le X \le \mu + 2\sigma) = P(|X-\mu| \le 2\sigma) = 0{,}9544 \cong 95\% ,$$

$$P(\mu - 3\sigma \le X \le \mu + 3\sigma) = P(|X-\mu| \le 3\sigma) = 0{,}9974 \cong 99{,}7\% .$$

Συμπερασματικά, *το 68% περίπου των τιμών ενός κανονικού πληθυσμού βρίσκονται σε απόσταση το πολύ μιας τυπικής απόκλισης, το 95% περίπου σε απόσταση δύο τυπικών αποκλίσεων και το 99.7% περίπου σε απόσταση τριών αποκλίσεων από το μέσο του πληθυσμού.*

Παράδειγμα 3.3. Ας θεωρήσουμε ότι ο χρόνος εμφάνισης X ενός φωτογραφικού φιλμ ακολουθεί κανονική κατανομή με μέση τιμή $\mu = 30$ λεπτά και τυπική απόκλιση $\sigma = 1{,}2$ λεπτά. Να υπολογισθούν (α) η πιθανότητα όπως ο χρόνος εμφάνισης ενός φιλμ μη υπερβεί τα 28 λεπτά και (β) η πιθανότητα όπως σε τουλάχιστον 2 από 10 φιλμ ο χρόνος εμφάνισης μη υπερβεί τα 28 λεπτά.

(α) Εισάγοντας την τυποποιημένη τυχαία μεταβλητή $Z = (X-\mu)/\sigma$, όπου $\mu = 30$ και $\sigma = 1{,}2$, η πιθανότητα όπως ο χρόνος εμφάνισης ενός φιλμ μη υπερβεί τα 28 λεπτά ισούται με

$$P(X \le 28) = P\left(\frac{X-30}{1{,}2} \le \frac{28-30}{1{,}2}\right) = P(Z \le -1{,}67) = 1 - \Phi(1{,}67) = 0{,}0475 \cong 5\%$$

(β) Ο αριθμός Y των φιλμ με χρόνο εμφάνισης το πολύ 28 λεπτών ακολουθεί τη διωνυμική κατανομή με αριθμό δοκιμών $\nu = 10$ και πιθανότητα επιτυχίας $p = P(X \le 28) = 0{,}05$, οπότε

$$P(Y = y) = \binom{10}{y}(0{,}05)^y (0{,}95)^{10-y}, \; y = 0,1,...,10 .$$

Επομένως η πιθανότητα όπως σε τουλάχιστον 2 από 10 φιλμ ο χρόνος εμφάνισης μη υπερβεί τα 28 λεπτά είναι

$$P(Y \ge 2) = 1 - P(Y < 2) = 1 - P(Y = 0) - P(Y = 1) =$$

$$= 1 - \binom{10}{0}(0{,}05)^0 (0{,}95)^{10} - \binom{10}{1}(0{,}05)^1 (0{,}95)^9 = 0{,}086 .$$

Παράδειγμα 3.4. Έστω ότι ένα δείγμα ν ατόμων εκλέγεται τυχαία από ένα πληθυσμό για την εκτίμηση του ποσοστού p των ατόμων του πληθυσμού που πάσχουν από μια συγκεκριμένη ασθένεια. (α) Να υπολογισθεί το μέγεθος ν του δείγματος έτσι ώστε το ποσοστό των ατόμων στο δείγμα που πάσχουν από την ασθένεια να διαφέρει από το πραγματικό ποσοστό p κατ' απόλυτη τιμή λιγότερο από 1% με πιθανότητα τουλάχιστον 95%. (β) Αν είναι γνωστό ότι $p \le 0{,}03$ (δηλαδή ότι πρόκειται περί σπάνιας ασθένειας) ποιό πρέπει να είναι το μέγεθος ν του δείγματος;

(α) Αν X είναι ο αριθμός των ατόμων του δείγματος που πάσχουν από την ασθένεια, τότε η τυχαία μεταβλητή X ακολουθεί τη διωνυμική κατανομή με παραμέτρους ν και p. Το ποσοστό των ατόμων του δείγματος που πάσχουν από την ασθένεια είναι ίσο με X/ν, οπότε το ζητούμενο μπορεί να διατυπωθεί ως εξής

$$P\left(\left|\frac{X}{\nu}-p\right|\le 0{,}01\right)\ge 0{,}95\,.$$

Χρησιμοποιώντας την προσέγγιση της διωνυμικής κατανομής από την κανονική κατανομή το αριστερό μέλος της συνθήκης μετασχηματίζεται στο

$$P\left(\left|\frac{X}{\nu}-p\right|\le 0{,}01\right)=P\left(-0{,}01\le\frac{X}{\nu}-p\le 0{,}01\right)=P\left(\frac{-0{,}01\nu}{\sqrt{\nu\,pq}}\le\frac{X-\nu\,p}{\sqrt{\nu\,pq}}\le\frac{0{,}01\nu}{\sqrt{\nu\,pq}}\right)$$

$$\cong \Phi\left(\frac{0{,}01\sqrt{\nu}}{\sqrt{pq}}\right)-\Phi\left(-\frac{0{,}01\sqrt{\nu}}{\sqrt{\nu pq}}\right)=2\Phi\left(\frac{0{,}01\sqrt{\nu}}{\sqrt{pq}}\right)-1$$

και η συνθήκη παίρνει τη μορφή

$$2\Phi\left(\frac{0{,}01\sqrt{\nu}}{\sqrt{pq}}\right)-1\ge 0{,}95$$

ή ισοδύναμα

$$\Phi\left(\frac{0{,}01\sqrt{\nu}}{\sqrt{pq}}\right)\ge 0{,}975\,.$$

Από τον Πίνακα 3 της τυποποιημένης κανονικής κατανομής έχουμε $\Phi(1{,}96)=0{,}975$ και οπότε

$$\frac{0.01\sqrt{\nu}}{\sqrt{pq}}\ge 1{,}96$$

και επομένως

$$\nu\ge 38416p(1-q)\,.$$

Παρατηρούμε ότι η συνάρτηση $g(p)=p(1-p)=p-p^2$ μεγιστοποιείται για $p=0{,}5$, οπότε

$$p(1-p)\le 0{,}5(1-0{,}5)=0{,}25$$

και έτσι το απαιτούμενο μέγεθος του δείγματος είναι

$$\nu\ge 38416\cdot 0{,}25\cong 9604\,.$$

(β) Αν είναι γνωστό ότι $p\le 0{,}03$, τότε

$$p(1-p)\le 0{,}03(1-0{,}03)=0{,}0021$$

και έτσι το απαιτούμενο μέγεθος του δείγματος είναι

$$\nu\ge 38416\cdot 0{,}0021\cong 81\,.$$

6. ΑΣΚΗΣΕΙΣ

1. Έστω ότι η τυχαία μεταβλητή *X* έχει την ομοιόμορφη κατανομή στο διάστημα $[\alpha, \beta]$. Αν $E(X)=1$ και $V(X)=3$, (α) να υπολογισθούν οι σταθερές α και β, (β) να προσδιορισθεί η συνάρτηση πυκνότητας της τυχαίας μεταβλητής $Y=|X|$ και (γ) να βρεθούν οι $E(Y)$ και $V(Y)$.

2. Έστω ότι η τυχαία μεταβλητή *X* έχει την ομοιόμορφη κατανομή στο διάστημα $[0,1]$. Δείξετε ότι η συνάρτηση κατανομής της τυχαίας μεταβλητής

$$Y = \frac{1}{\theta}\log\frac{1}{1-X}$$

δίδεται από την

$$F_Y(y) = \begin{cases} 0, & -\infty < y < 0 \\ 1-e^{-\theta y}, & 0 \le y < \infty. \end{cases}$$

και συμπεράνετε ότι ακολουθεί την εκθετική κατανομή.

3. (*Συνέχεια*). Έστω $Z=h(X)$, όπου $h(x)=z$ για $1-q^z < x \le 1-q^{z+1}$, $z=0,1,\ldots$, με $0<q<1$. Δείξετε ότι η τυχαία μεταβλητή *Z* ακολουθεί τη γεωμετρική κατανομή με συνάρτηση πιθανότητας

$$f_Z(z) = pq^z,\ z=0,1,\ldots,$$

όπου $p=1-q$.

4. Έστω X_t ο αριθμός των θανάτων σε νοσοκομείο των Αθηνών από μια σπάνια ασθένεια σε χρονικό διάστημα *t* ωρών. Αν σε συγκεκριμένο χρονικό διάστημα $[0,s]$ συνέβη ένας θάνατος, δείξετε ότι η χρονική στιγμή *T* του θανάτου ακολουθεί την ομοιόμορφη κατανομή στο διάστημα $[0,s]$.

5. Ο χρόνος ζωής *X* σε ώρες μιας ορισμένης ηλεκτρονικής λυχνίας ακολουθεί την εκθετική κατανομή με μέση τιμή $E(X)=1000$ ώρες. Το εργοστάσιο που κατασκευάζει τις λυχνίες δίδει εγγύηση α ωρών στους πελάτες του. Να υπολογισθεί το α έτσι ώστε με πιθανότητα τουλάχιστο 0,95 οι λυχνίες να επιζούν του χρόνου εγγύησης.

6. Ο χρόνος ζωής *X* του ιού της γρίπης μέσα στον οργανισμό ενός ατόμου ακολουθεί την εκθετική κατανομή με μέση τιμή 3 μέρες. Να υπολογισθούν (α) η πιθανότητα όπως ένα άτομο που προσβλήθηκε από τον ιό γίνει καλά στο χρονικό διάστημα από 2 μέχρι 4 μέρες και (β) η δεσμευμένη πιθανότητα όπως ένα άτομο που προσβλήθηκε από τον ιό γίνει καλά σε λιγότερο από 5 συνολικά μέρες δεδομένου ότι έχει 2 μέρες άρρωστος. (γ) Επίσης να υπολογισθεί η πιθανότητα όπως 3 τουλάχιστο από 10 άτομα που προσβλήθηκαν από τον ιό γίνουν καλά στο χρονικό διάστημα από 2 μέχρι 4 μέρες.

7. Έστω ότι η ποσότητα *X* σε χιλιάδες λίτρα που πωλεί ένα πρατήριο βενζίνης σε μια μέρα πέραν των χιλίων λίτρων ακολουθεί την κατανομή Erlang με μέση τιμή 5 χιλιάδες λίτρα και τυπική απόκλιση 2,5 χιλιάδες λίτρα. Αν οι δεξαμενές του πρατηρίου μια συγκεκριμένη μέρα έχουν 8 χιλιάδες λίτρα να υπολογισθούν η πιθανότητα το πρατήριο να μη μπορέσει να ανταποκριθεί στη ζήτηση.

8. Το βάρος των νεογέννητων παιδιών σε μια συγκεκριμένη χώρα ακολουθεί την κανονική κατανομή με μέση τιμή 3,400 κιλά για τα αγόρια και 3,350 κιλά για τα κορίτσια και τυπική απόκλιση 0,400 κιλά για τα αγόρια και 0,350 κιλά για τα κορίτσια. Να υπολογισθούν οι πιθανότητες (α) ένα νεογέννητο αγόρι να έχει βάρος μεγαλύτερο των 4,000 κιλών και (β) ένα νεογέννητο κορίτσι να έχει βάρος μεγαλύτερο των 3,000 κιλών και μικρότερο των 3,700 κιλών.

9. Το ύψος X των ανδρών ενός πληθυσμού ακολουθεί την κανονική κατανομή με μέση τιμή 175 εκατοστόμετρα και τυπική απόκλιση 5 εκατοστόμετρα. Να υπολογισθούν τα ποσοστά του πληθυσμού των ανδρών με ύψος (α) μεγαλύτερο των 175, (β) μεγαλύτερο των 180 και (γ) μεταξύ των 170 και των 180. Να υπολογισθούν οι πιθανότητες όπως σε τυχαίο δείγμα 6 ανδρών (δ) όλοι είναι ύψους άνω των 180 και (ε) οι δύο είναι υψηλότεροι του μέσου και τέσσερεις χαμηλότεροι του μέσου.

10. Μία αυτόματη μηχανή κατασκευάζει βίδες των οποίων το μήκος σε χιλιοστόμετρα ακολουθεί την κανονική κατανομή με μέση τιμή 40 και τυπική απόκλιση 1. Αν το μήκος μιας βίδας είναι εκτός του διαστήματος $[39, 40]$, η βίδα θεωρείται ελαττωματική. Να υπολογισθούν (α) το ποσοστό των ελαττωματικών βίδων που παράγει η μηχανή, (β) η πιθανότητα όπως σε μια τυχαία επιλογή 5 βίδων μια το πολύ είναι ελαττωματική και (γ) η πιθανότητα όπως σε μια τυχαία επιλογή 100 βίδων 20 το πολύ είναι ελαττωματικές.

11. Έστω ότι η αντοχή X ενός υφάσματος (χιλιόγραμμα δύναμης) ακολουθεί την κανονική κατανομή με μέση τιμή 50 και διασπορά 4. Ένα τόπι 25 μέτρων του υφάσματος με $X > 47$ αποφέρει κέρδος 25 ευρώ. Αν $X \leq 47$ το ύφασμα πωλείται ως δεύτερης διαλογής και αποφέρει κέρδος 15 ευρώ. Να υπολογισθεί το αναμενόμενο κέρδος ανά τόπι.

12. (α) Αν X είναι μια κανονική τυχαία μεταβλητή με μέση τιμή μ και διασπορά σ^2 και c ένας πραγματικός αριθμός τέτοιος ώστε

$$P(X > c) = 2P(X \leq c),$$

να δειχθεί ότι

$$c + 0{,}43\sigma = \mu$$

(β) Αν οι τιμές του σιδήρου στο αίμα των ανδρών ενός πληθυσμού ακολουθούν την κανονική κατανομή με μέση τιμή 110 mg/dl και διασπορά 25 mg/dl να βρεθεί η τιμή c του σιδήρου για την οποία το ποσοστό ανδρών που την υπερβαίνει είναι διπλάσιο του ποσοστού που δεν την υπερβαίνει.

13. (α) Να υπολογισθεί η πιθανότητα όπως σε 40 ρίψεις ενός συνήθους νομίσματος εμφανιστούν 20 κεφαλές χρησιμοποιώντας κανονική προσέγγιση με διόρθωση συνέχειας. (β) Ποιά είναι η ακριβής τιμή της πιθανότητας αυτής;

14. Σε μια δίκη που αφορούσε την πατρότητα ενός παιδιού ο κατηγορούμενος μπόρεσε να αποδείξει ότι βρισκόταν εκτός της χώρας για το χρονικό διάστημα που άρχιζε 295 μέρες πριν τη γέννηση του παιδιού και τελείωνε 240 ημέρες πριν τη γέννηση. Αν υποθέσουμε ότι η διάρκεια κύησης ακολουθεί κανονική κατανομή με μέση τιμή 9 μήνες και τυπική απόκλιση 10 ημέρες, ποιά είναι η πιθανότητα ο κατηγορούμενος να είναι πράγματι ο πατέρας του παιδιού;

15. Η τιμή X της χοληστερόλης στο αίμα των ατόμων ενός συγκεκριμένου πληθυσμού ακολουθεί κατά προσέγγιση την κανονική κατανομή με μέση τιμή 250 και τυπική απόκλιση 50. (α) Να υπολογισθεί το ποσοστό των ατόμων του πληθυσμού των οποίων η τιμή χοληστερίνης είναι μεταξύ 200 και 260. (β) Να βρεθεί η τιμή της χοληστερόλης c τέτοια ώστε στο 10% των ατόμων του πληθυσμού η χοληστερόλη υπερβαίνει το c.

16. Έστω ότι ένα δείγμα ν ατόμων εκλέγεται τυχαία από ένα πληθυσμό για την εκτίμηση του ποσοστού p των καπνιστών. (α) Να υπολογισθεί το ν ώστε το ποσοστό των καπνιστών στο δείγμα να διαφέρει από το πραγματικό ποσοστό p κατ' απόλυτη τιμή λιγότερο του 0,05 με πιθανότητα τουλάχιστον 0,99. (β) Αν είναι γνωστό ότι το πραγματικό ποσοστό των καπνιστών είναι $p \le 0{,}3$, ποιό πρέπει να είναι το μέγεθος ν του δείγματος;

17. Έστω ότι η απόκλιση X της βολής ενός σκοπευτή από το κέντρο στόχου είναι μια συνεχής τυχαία μεταβλητή με συνάρτηση πυκνότητας

$$f(x) = 2(1-x),\ 0 \le x \le 1 .$$

Ο σκοπευτής κερδίζει το στοίχημα αν η απόκλιση της βολής από το κέντρο του στόχου δεν είναι μεγαλύτερη από το $1/2$. Να υπολογισθεί ο αριθμός ν των βολών που απαιτούνται έτσι ώστε η πιθανότητα να κερδίσει ο σκοπευτής το στοίχημα να είναι τουλάχιστο 0,98.