Κατανομές συνεχών τυχαίων μεταβλητών

Συναρτήσεις πιθανότητας συνεχών τυχαίων μεταβλητών

- Έστω η συνεχής τυχαία μεταβλητή X. Τότε η συνάρτηση πυκνότητας πιθανότητας συμβολίζεται με f(x) και ισχύει:
- $0 \leq \mathrm{f(x)} \leq 1$
- $\int_{-\infty}^{+\infty} \mathrm{f(x)dx} = 1$

- Η συνάρτηση παίρνει τιμές στο [0,1] και το ορισμένο ολοκλήρωμά της στο πεδίο ορισμού της ισούται με μονάδα. Η συνάρτηση πυκνότητας πιθανότητας αποκτά έτσι ιδιότητες που προσιδιάζουν στον ορισμό της πιθανότητας.

Συναρτήσεις πιθανότητας συνεχών τυχαίων μεταβλητών

- Για δύο τιμές της X, τις x_1 και x_2, η πιθανότητα της X στο διάστημα $[x_1,x_2]$ ισούται με

$$P(x_1 \leq X \leq x_2) = \int_{x_1}^{x_2} f(x)dx$$

- Αθροιστική Συνάρτηση πιθανότητας ή Συνάρτηση Κατανομής μιας συνεχούς τυχαίας μεταβλητής

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)dt$$

Μέση τιμή, Διασπορά και Τυπική Απόκλιση συνεχών τυχαίων μεταβλητών

$$E(X) = \mu = \int_{-\infty}^{+\infty} xf(x)dx$$

$$Var(X) = \sigma^2 = E[X - E(X)]^2 = \int_{-\infty}^{+\infty} [x - E(X)]^2 f(x)dx$$

$$\sigma = \sqrt{\sigma^2}$$

Παράδειγμα

- Δίδεται η συνάρτηση

$$f(x) = \begin{cases} 0 & \text{για } X < 0 \text{ ή } X > 1 \\ 4x^3 & \text{για } 0 \leq X \leq 1 \end{cases}$$

i. Να αιτιολογηθεί ότι αυτή είναι συνάρτηση πυκνότητας πιθανότητας της τυχαίας μεταβλητής Χ.

ii. Να υπολογιστούν οι μέση τιμή της, η διασπορά και η τυπική της απόκλιση.

iii. Να υπολογιστεί η πιθανότητα $P(0 \leq X \leq 0,1)$.

Παράδειγμα

i. Για την f(x) ισχύει ότι $0 \leq f(x) \leq 1$. Θα πρέπει να αποδείξουμε ότι:

$$\int_{-\infty}^{+\infty} f(x)dx = 1$$

Πράγματι, υπολογίζουμε το ολοκλήρωμα από 0 έως 1 επειδή οπουδήποτε αλλού η τιμή της f(x) είναι μηδέν:

$$\int_{0}^{1} f(x)dx = \int_{0}^{1} 4x^3dx = 4\int_{0}^{1} x^3dx = 4\left[\frac{x^4}{4}\right]_0^1 = [x^4]_0^1 = 1^4 - 0^4$$
$$= 1$$

Παράδειγμα

$$\mu = \int_{-\infty}^{+\infty} \mathrm{x}f(\mathrm{x})d\mathrm{x}$$

$$= \int_{0}^{1} \mathrm{x}4\mathrm{x}^{3}d\mathrm{x} = 4\int_{0}^{1} \mathrm{x}^{4}d\mathrm{x} = 4\left[\frac{\mathrm{x}^{5}}{5}\right]_{0}^{1} = \frac{4\cdot 1^{5}}{5} - \frac{4\cdot 0^{5}}{5} = \frac{4}{5}$$

Παράδειγμα

- $\sigma^2 = E[X - E(X)]^2 = \int_{-\infty}^{+\infty}[x - E(X)]^2 f(x)dx = \int_0^1 \left(x - \frac{4}{5}\right)^2 4x^3 dx = \int_0^1 \left(x^2 - \frac{8}{5}x + \frac{16}{25}\right) 4x^3 dx = \left[\frac{4x^6}{6}\right]_0^1 - \left[\frac{32x^5}{25}\right]_0^1 + \left[\frac{16x^4}{25}\right]_0^1 = \frac{4\cdot 1^6}{6} - \frac{32\cdot 1^5}{25} + \frac{16\cdot 1^4}{25} = \frac{3}{5} - \frac{18}{16} + \frac{27}{48} = 0{,}027$

- $\sigma = \sqrt{\sigma^2} = \sqrt{0{,}027} = 0{,}16$

Παράδειγμα

iii. $P(0 \leq X \leq 0{,}1) = \int_0^{0{,}1} f(x)dx = [x^4]_0^{0{,}1} =$
$= 0{,}1^4 - 0^4 = 0{,}0001$

Συνεχείς τυχαίες μεταβλητές

Οι κατανομές που χρησιμοποιούνται περισσότερο είναι:

- Η Κανονική κατανομή
- Η κατανομή t του Student
- Η χ^2 κατανομή
- Η F κατανομή

Κανονική κατανομή

- Η Συνάρτηση Πυκνότητας Πιθανότητας της κανονικής κατανομής δίνεται από τον τύπο:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad , \qquad \text{για} -\infty < x < +\infty,$$

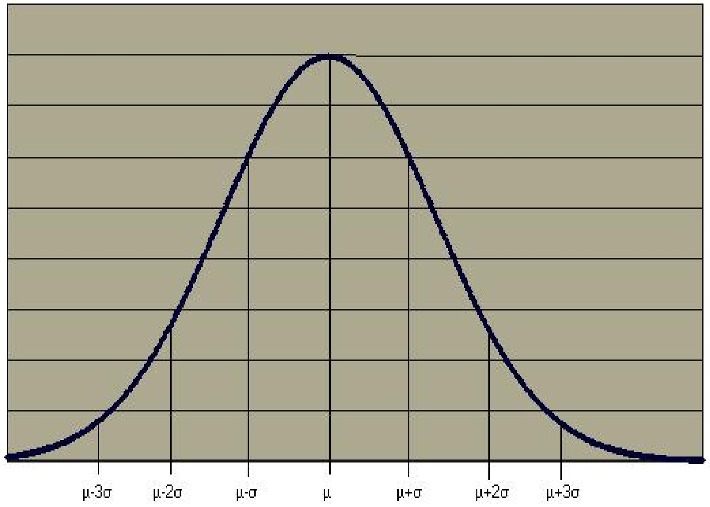
- Η Συνάρτηση Κατανομής δίνεται από τον τύπο:

$$F_X(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt = P(X < x)$$

Κανονική κατανομή

- Αν γνωρίζουμε την μέση τιμή και την διασπορά της Κανονικής κατανομής, γνωρίζουμε πλήρως και την ίδια την κατανομή. Αν γνωρίζουμε δηλαδή τις δύο παραμέτρους της Κανονικής κατανομής γνωρίζουμε απολύτως την κατανομή.
- Συμβολίζουμε με $N(\mu, \sigma^2)$ την Κανονική κατανομή και γράφουμε $X \sim N(\mu, \sigma^2)$ για να δηλώσουμε ότι η τυχαία μεταβλητή Χ ακολουθεί την *Κανονική κατανομή* ή κατανέμεται σύμφωνα με την *Κανονική κατανομή*.

Ιδιότητες της κανονικής κατανομής

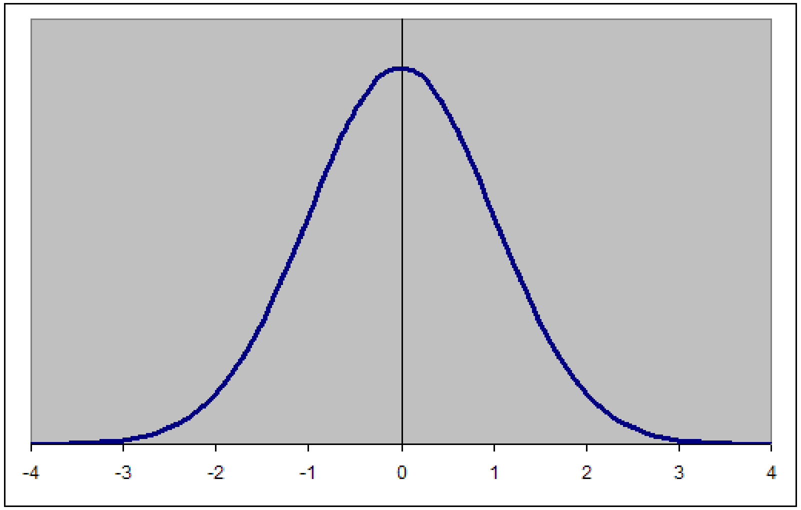


- Η συνάρτηση πυκνότητας πιθανότητας έχει συμμετρική γραφική παράσταση περί την μέση τιμή.
- Η μέση τιμή, η διάμεσος και η επικρατούσα τιμή ταυτίζονται.
- Το σχήμα της κατανομής είναι σαν καμπάνα (κωδωνοειδές).
- Η συμμετρία και η ταύτιση μέσης τιμής, διαμέσου και επικρατούσας τιμής δεν είναι ιδιότητα μόνον της Κανονικής κατανομής. Ωστόσο οι εξής τρεις ιδιότητες είναι:
- στο διάστημα (μ-σ, μ+σ) ανήκει περίπου το 68% των τιμών της X
- στο διάστημα (μ-2σ, μ+2σ) ανήκει περίπου το 95% των τιμών της X
- στο διάστημα (μ-3σ, μ+3σ) ανήκει περίπου το 99% των τιμών της X

Τυπική Κανονική Κατανομή

- Η κανονική κατανομή με παραμέτρους **$\mu=0$** και **$\sigma^2=1$** ονομάζεται *Τυπική Κανονική Κατανομή* και είναι ιδιαίτερα σημαντική, και συμβολίζεται με $N(0,1)$. Η τυχαία μεταβλητή Z που ακολουθεί την Τυπική κανονική κατανομή ονομάζεται Τυπική τυχαία μεταβλητή.
- Ένας λόγος, για τον οποίο είναι σημαντική η τυπική κανονική κατανομή, είναι το ότι κάθε κανονική κατανομή ανάγεται στην τυπική κανονική κατανομή μέσω του *Μετασχηματισμού Τυποποίησης* ή απλά της *Τυποποίησης*.
- Με το μετασχηματισμό αυτόν αφαιρούμε από κάθε τιμή της αρχικής τυχαίας μεταβλητής X (που κατανέμεται κανονικά με παραμέτρους μ και σ^2) τη μέση τιμή μ και διαιρούμε με την τυπική απόκλιση σ.

Τυπική κανονική κατανομή



- Το αποτέλεσμα της τυποποίησης είναι:

$$Z = \frac{X - \mu}{\sigma}$$

- Για την τυπική κανονική κατανομή έχουμε ότι:
- στο διάστημα (-1, 1) ανήκει περίπου το 68% των τιμών της Z
- στο διάστημα (-2, 2) ανήκει περίπου το 95% των τιμών της Z
- στο διάστημα (-3, 3) ανήκει περίπου το 99% των τιμών της Z

Παράδειγμα
Πώς γίνεται η τυποποίηση

- Έστω X~N(25,9). Δηλαδή $\mu=25$ και $\sigma^2=9$, άρα $\sigma=3$. Επομένως έχουμε:

$$\mathrm{Z} = \frac{\mathrm{X} - \mu}{\sigma} = \frac{\mathrm{X} - 25}{3}$$

Συναρτήσεις πιθανότητας της τυπικής κανονικής κατανομής

- Συνάρτηση πυκνότητας πιθανότητας της Τυπικής μεταβλητής
- $f(z) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{z^2}{2\sigma^2}} \quad \gamma\iota\alpha - \infty < z < +\infty$
- Συνάρτηση κατανομής της τυπικής μεταβλητής
- $F_Z(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} \exp\left(-\frac{1}{2}t^2\right) dt = P(Z < z)$

Ιδιότητες

- Όταν X~N(μ, σ²) τότε $F_X(x) = F_Z\left(\frac{x-\mu}{\sigma}\right)$
- δηλαδή η αθροιστική πιθανότητα της X ισούται με την αθροιστική πιθανότητα της τυπικής μεταβλητής της X στη οποία μετασχηματίστηκε μέσω του μετασχηματισμού τυποποίησης. Και επειδή για οποιαδήποτε τυχαία μεταβλητή X ισχύει:
- $P(x_1 \leq X \leq x_2) = F_X(x_2) - F_X(x_1)$
- $P(x_1 \leq X \leq x_2) = F_Z\left(\frac{x_2-\mu}{\sigma}\right) - F_Z\left(\frac{x_1-\mu}{\sigma}\right)$

Ιδιότητες

- Από αυτό το σημείο και μετά ανατρέχουμε στους πίνακες της αθροιστικής τυπικής κανονικής κατανομής.
- Αν πρέπει να υπολογίσουμε την αθροιστική πιθανότητα για αρνητικές τιμές της Z έχουμε:

- $\mathrm{F_Z}(-\mathrm{z}^*) = \mathrm{P}(\mathrm{Z} \leq -\mathrm{z}^*) = \mathrm{P}(\mathrm{Z} \geq \mathrm{z}^*) = 1 - \mathrm{F_Z}(\mathrm{z}^*)$

- Το παραπάνω ισχύει λόγω της συμμετρίας της συνάρτησης πυκνότητας πιθανότητας της κανονικής κατανομής. Πάντα βοηθάει να κάνουμε μια γραφική παράσταση ώστε να έχουμε εικόνα της συνάρτησης πυκνότητας πιθανότητας του σημείου z ή x και της πιθανότητας που αναζητούμε.

Εύρεση της πιθανότητας από τον πίνακα πιθανοτήτων της τυπικής κανονικής κατανομής

θέλουμε να υπολογίσουμε την πιθανότητα P(Z≤1,65)

z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,5	0,504	0,508	0,512	0,516	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	0,591	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	0,648	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,67	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,695	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,719	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	0,758	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,791	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	0,834	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	0,877	0,879	0,881	0,883
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,898	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	0,937	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	0,975	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	0,983	0,9834	0,9838	0,9842	0,9846	0,985	0,9854	0,9857
2,2	0,9861	0,9864	0,9868	0,9871	0,9875	0,9878	0,9881	0,9884	0,9887	0,989
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	0,992	0,9922	0,9925	0,9927	0,9929	0,9931	0,9932	0,9934	0,9936
2,5	0,9938	0,994	0,9941	0,9943	0,9945	0,9946	0,9948	0,9949	0,9951	0,9952
2,6	0,9953	0,9955	0,9956	0,9957	0,9959	0,996	0,9961	0,9962	0,9963	0,9964
2,7	0,9965	0,9966	0,9967	0,9968	0,9969	0,997	0,9971	0,9972	0,9973	0,9974
2,8	0,9974	0,9975	0,9976	0,9977	0,9977	0,9978	0,9979	0,9979	0,998	0,9981
2,9	0,9981	0,9982	0,9982	0,9983	0,9984	0,9984	0,9985	0,9985	0,9986	0,9986
3,0	0,9987	0,9987	0,9987	0,9988	0,9988	0,9989	0,9989	0,9989	0,999	0,999
3,1	0,999	0,9991	0,9991	0,9991	0,9992	0,9992	0,9992	0,9992	0,9993	0,9993
3,2	0,9993	0,9993	0,9994	0,9994	0,9994	0,9994	0,9994	0,9995	0,9995	0,9995
3,3	0,9995	0,9995	0,9995	0,9996	0,9996	0,9996	0,9996	0,9996	0,9996	0,9997
3,4	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9998

Μερικές χρήσιμες συμβουλές

- Υπολογίζουμε πιθανότητες μόνο για διαστήματα και όχι για συγκεκριμένες τιμές της X. Η πιθανότητα $P(X=k)$ ισούται με μηδέν
- Στις συνεχείς τυχαίες μεταβλητές, δεν επηρεάζει τα αποτελέσματα αν χρησιμοποιούμε ανισότητες ή ανισοϊσότητες όταν καθορίζουμε τα διαστήματα για τα οποία θέλουμε να υπολογίσουμε την πιθανότητα
- Για να υπολογίσουμε μια πιθανότητα σχετικά με την τυχαία μεταβλητή που ακολουθεί κανονική κατανομή $N(\mu, \sigma^2)$ εργαζόμαστε ως εξής:
  - ✓ πρώτα κάνουμε τον μετασχηματισμό τυποποίησης-υπολογίζουμε δηλαδή τις z-τιμές,
  - ✓ μετά φροντίζουμε η φορά της ανισότητας να είναι προς τα αριστερά (δηλ $Z<z$), και
  - ✓ τέλος αν το z είναι αρνητικό κάνουμε το μετασχηματισμό $P(Z<z)=1-P(Z<-z)$ ώστε να έχουμε τη θετική τιμή $-z$
- Επίσης για την περίπτωση που θέλουμε να υπολογίσουμε την πιθανότητα $P(z_1<Z<z_2)$ χρησιμοποιούμε την ιδιότητα $P(z_1<Z<z_2)=P(Z<z_2)-P(Z<z_1)$.

Παράδειγμα

Αν η τυχαία μεταβλητή X ακολουθεί την κανονική κατανομή N(10,4) υπολογίστε τις πιθανότητες:

α) $P(X<10)$,

β) $P(X<12)$,

γ) $P(X<8)$,

δ) $P(7<X<12)$,

ε) $P(X>8)$,

στ) $P(X=12)$,

ζ) $P(X>13)$.

Παράδειγμα

- α) P(X<10)=0,5 αφού το 10 είναι η μέση τιμή και διάμεσος της X.

- β) $\mathrm{P}(\mathrm{X} < 12) = \mathrm{P}\left(\frac{\mathrm{X}-\mu}{\sigma} < \frac{12-\mu}{\sigma}\right) = \mathrm{P}\left(\mathrm{Z} < \frac{12-10}{2}\right) = \mathrm{P}(\mathrm{Z} < 1) =$

  $\mathrm{F_Z}(1) = 0{,}8413$

- γ) $\mathrm{P}(\mathrm{X} < 8) = \mathrm{P}\left(\frac{\mathrm{X}-\mu}{\sigma} < \frac{8-\mu}{\sigma}\right) = \mathrm{P}\left(\mathrm{Z} < \frac{8-10}{2}\right) = \mathrm{P}(\mathrm{Z} < -1) = 1 -$

  $\mathrm{P}(\mathrm{Z} < 1) = 1 - \mathrm{F_Z}(1) = 1 - 0{,}8413 = 0{,}1587$

Παράδειγμα

- δ) $P(7 < X < 12) = P\left(\frac{7-\mu}{\sigma} < \frac{X-\mu}{\sigma} < \frac{12-\mu}{\sigma}\right) = P\left(\frac{7-10}{2} < \frac{X-\mu}{\sigma} < \frac{12-10}{2}\right) = P(-1{,}5 < Z < 1) = P(Z < 1) - P(Z < -1{,}5) = P(Z < 1) - [1 - P(Z < 1{,}5)] = P(Z < 1) - 1 + P(Z < 1{,}5) = F_Z(1) + F_Z(1{,}5) - 1 = 0{,}8413 + 0{,}9332 - 1 = 0{,}7745$

Παράδειγμα

- ε) $P(X > 8) = P\left(\frac{X-\mu}{\sigma} > \frac{8-\mu}{\sigma}\right) = P\left(Z > \frac{8-10}{2}\right) = P(Z > -1) =$ $P(Z < 1) = 0{,}8413$

- στ) $P(X = 12) = 0$

- ζ) $P(X > 13) = P\left(\frac{X-\mu}{\sigma} > \frac{13-\mu}{\sigma}\right) = P\left(Z > \frac{13-10}{2}\right) = P(Z > 1{,}5) =$ $1 - P(Z \leq 1{,}5) = 1 - F_Z(1{,}5) = 1 - 0{,}9332 = 0{,}0668$

Παράδειγμα

Το ύψος των κατοίκων μιας πόλης με πληθυσμό 1000, ακολουθεί την κανονική κατανομή με μέση τιμή μ=170cm και τυπική απόκλιση σ=7cm. Υπολογίστε τον αριθμό των κατοίκων με ύψος

α. κάτω των 183cm

β. μεταξύ 165cm και 177cm

γ. κάτω των 160cm ή άνω των 180cm

<u>Λύση</u>

Έστω X τυχαία μεταβλητή η οποία εκφράζει το ύψος των κατοίκων της πόλης. Τότε $X \sim N(170,7^2)$. Θα υπολογίσουμε αρχικά τις πιθανότητες:

α. $P(X<183)$

β. $P(165<X<177)$

γ. $P(X<160$ ή $X>180)$

Παράδειγμα

- α) $P(X < 183) = P\left(\frac{X-\mu}{\sigma} < \frac{183-\mu}{\sigma}\right) = P\left(Z < \frac{183-170}{7}\right) = P(Z < 1{,}86) = F_Z(1{,}86) = 0{,}9686$

- β) $P(165 < X < 177) = P(X < 177) - P(X \leq 165) = P\left(Z < \frac{177-170}{7}\right) - P\left(Z < \frac{165-170}{7}\right) = P(Z < 1) - P(Z < -0{,}71) = P(Z < 1) - \left(1 - P(Z < 0{,}71)\right) = P(Z < 1) + P(Z < 0{,}71) - 1 = F_Z(1) + F_Z(0{,}71) - 1 = 0{,}8413 + 0{,}7611 - 1 = 0{,}6024$

Παράδειγμα

$$\gamma)\ \ \mathrm{P}(\mathrm{X} < 160 \text{ ή } \mathrm{X} > 180) = \mathrm{P}(\mathrm{X} < 160) + \mathrm{P}(\mathrm{X} > 180)$$
$$= \mathrm{P}(\mathrm{X} < 160) + 1 - \mathrm{P}(\mathrm{X} \leq 180)$$
$$= 1 + \ \mathrm{P}\left(\mathrm{Z} < \frac{160 - 170}{7}\right) - \mathrm{P}\left(\mathrm{Z} < \frac{180 - 170}{7}\right)$$
$$= 1 + \mathrm{P}(\mathrm{Z} < -1{,}43) - \mathrm{P}(\mathrm{Z} < 1{,}43)$$
$$= 1 + \big(1 - \mathrm{P}(\mathrm{Z} < 1{,}43)\big) - \mathrm{P}(\mathrm{Z} < 1{,}43)$$
$$= 2 - 2\mathrm{P}(\mathrm{Z} < 1{,}43) = 2 - 2 \cdot 0{,}9236 = 0{,}1528$$

Παράδειγμα

Επομένως:

α) κάτοικοι με ύψος κάτω των 183cm είναι:

$1.000 \times 0{,}9686 = 968{,}6 \approx 969$

β) κάτοικοι με ύψος μεταξύ 165cm και 177cm

είναι: $1.000 \times 0{,}6024 = 602{,}4 \approx 602$

γ) κάτω των 160cm ή άνω των 180cm είναι:

$1.000 \times 0{,}1528 = 152{,}8 \approx 153$