3.2 Τυπική κανονική κατανομή

Έστω η μεταβλητή $X \sim N(\mu,\sigma)$. Η μεταβλητή $Z=\frac{X-\mu}{\sigma} \sim N(0,1)$. Η συνάρτηση πυκνότητας πιθανότητας της μεταβλητής Z δίνεται από τον τύπο:

$$f(z)=\frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}$$

και η συνάρτηση αθροιστικής κατανομής από

$$F(z)=\int_{-\infty}^{z}\frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}dz$$

Η συνάρτηση αθροιστικής κατανομής της τυπικής κανονικής κατανομής δεν εξαρτάται από καμία παράμετρο. Συνεπώς γνωρίζοντας μόνο το z μπορούμε να βρούμε το F(z). Για το λόγο αυτό υπάρχουν έτοιμοι πίνακες της τυπικής αθροιστικής κατανομής.

Σημειώνουμε ότι αν γνωρίζουμε την κατανομή που ακολουθεί μία μεταβλητή, τότε μπορούμε να βρούμε ποια πιθανότητα έχουν οι τιμές της να ανήκουν σε ένα διάστημα. Επίσης έχουμε τη δυνατότητα να αντιστρέψουμε τη διαδικασία, δηλ. μπορούμε να υπολογίσουμε το διάστημα που περιέχει ένα συγκεκριμένο ποσοστό παρατηρήσεων της μεταβλητής. Για παράδειγμα, αν μία μεταβλητή ακολουθεί την τυπική κανονική κατανομή, τότε γνωρίζουμε ότι στο διάστημα (-1, 1) αντιστοιχεί στη πιθανότητα 1-2Χ0,1587=0,6826, αλλά και ότι η πιθανότητα 0,39 αντιστοιχεί στο διάστημα (-0,51, 0,51).

Παράδειγμα 3.1

Έστω ότι ο δείκτης IQ ενός συνόλου ανθρώπων ακολουθεί κανονική κατανομή $N(100,100)$. Πόσο IQ θα πρέπει να έχει κανείς για ανήκει στο ανώτερο 2% αυτού του πληθυσμού; Πόσο για να ανήκει στο κατώτερο 5%; Σε ποιο διάστημα τιμών (γύρω από τη μέση τιμή) θα πρέπει να περιλαμβάνεται το IQ κάποιου έτσι ώστε να μπορούμε να πούμε ότι ανήκει στο (μεσαίο) 39% των περιπτώσεων;

Αφού η μεταβλητή $X \sim N(100,100)$, η μεταβλητή $Z=\frac{X-100}{10} \sim N(0,1)$.

Προφανώς από τη σχέση $Z=\frac{X-100}{10}$ προκύπτει ότι $X=100+10Z$.

Από τον πίνακα τυπικής κανονικής κατανομής παίρνουμε:

$Z_{0,98}$=2,05 Άρα x=100+10*2,05=120,5

$Z_{0,05}$=-1,64 Άρα x=100-10*1,64=83,6

$Z_{-0,305}$=-0,51 Άρα x=100-10*0,51=94,9

$Z_{0,305}$=0,51 Άρα x=100+10*0,51=105,1

Γενικότερα αν $X \sim N(\mu,\sigma^2)$ και θέλουμε να υπολογίσουμε τη τιμή της X:

- για την οποία το a% των περιπτώσεων είναι μικρότερο, τότε $x=\mu+z_a\sigma$
- για την οποία το a% των περιπτώσεων είναι μεγαλύτερο, τότε $x=\mu+z_{1-a}\sigma$

Επίσης αν θέλουμε να υπολογίσουμε το διάστημα στο οποίο ανήκει το (1-a)% των περιπτώσεων τότε $(\mu+z_{a/2}\sigma,\ \mu+z_{1-a/2}\sigma)$