2. Συνεχείς Κατανομές

2.1 Κανονική Κατανομή (Normal ή Gaussian Distribution)

Η κανονική κατανομή ειναι η πιο σπουδαία κατανομη της θεωρίας πιθανοτήτων και της Στατιστικής, κυρίως για τον λόγο της χρησιμότητας της σε ενα μεγάλο πλήθος εφαρμογών. Μερικές περιπτώσεις που ακολουθούν την κανονική κατανομή ειναι:

- Πληθυσμιακά χαρακτηριστικά (Υψος,Βάρος,βαθμολογία σε τεστ κλπ)
- Τυχαία σφάλματα που εμφανίζονται σε διάφορες μετρήσεις. Για αυτόν τον λόγο η κανονική κατανομή αναφερεται και ως η κατανομή των σφαλμάτων.
- Το άθροισμα και ο μέσος όρος μεγαλου αριθμού παρατηρήσεων τυχαίας μεταβλητής ακολουθεί κατα προσεγγιση κανονική κατανομή ανεξάρτητα απο το ποια κατανομή ακολουθούν οι αρχικές παρατηρήσεις.
- Πολλές κατανομές τόσο διακριτές, όσο και συνεχείς μπορούν να προσεγγιθούν κάτω απο ορισμένες συνθήκες απο την κανονική κατανομη.
- Η κανονική κατανομή αποτελεί σύμφωνα με το κεντρικό οριακό θεώρημα (το άθροισμα ενός ικανοποιητικά μεγάλου αριθμού ανεξάρτητων και ισόνομων τυχαίων μεταβλητών προσεγγίζεται από την κανονική κατανομή) τη βάση της στατιστικής συμπερασματολογίας ή επαγωγικής στατιστικής.
- Οι ιδιότητες της κανονικής κατανομής αξιοποιούνται στη Στατιστική Συμπερασμασματολογία. Ουσιαστικά, η κανονική κατανομή, αποτελεί το θεμέλιο της Στατιστικής Συμπερασμασματολογίας.

Συμβολίζουμε οταν μια τυχαία μεταβλητή X ακολουθεί κανονική κατανομη ως

$$X \sim N(\mu, \sigma^2)$$

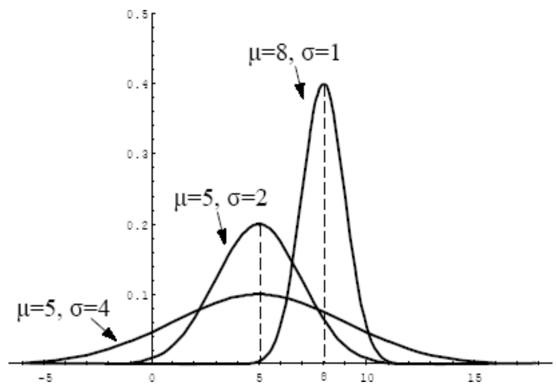
με παραμέτρους μ και σ^2, οπου μ = μέση τιμή και σ^2 η διασπορά.

Η Συνάρτηση Πυκνότητας Πιθανότητας (σ.π.π) της κανονικής κατανομής περιγράφεται απο την συναρτηση:

$$f(x) = \frac{1}{\sqrt{2\pi\,\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, x$$

$E(X) = \mu$ και $V(X) = \sigma^2 , sd(X) = \sqrt{V(X)} = \sigma.$

Η συνάρτηση είναι ορισμένη γύρω απο το μέσο της μ και ¨απλώνεται¨ με διασπορά σ^2. Ορισμένα απο τα χαρακτηριστικά της κατανομής είναι οτι ειναι συμμετρική.



Τυποποιημενη ή Τυπική Κανονική Κατανομή

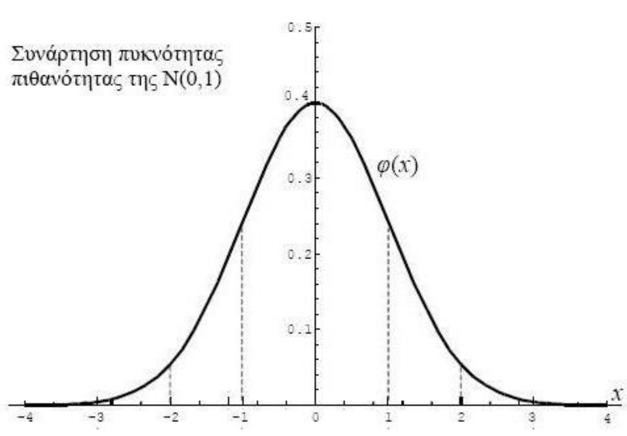
Προταση

Εστω $X \sim N(\mu, \sigma^2)$. Τοτε η $Z = \frac{X-\mu}{\sigma} \sim N(0,1)$.

Αυτό σημαίνει οτι για κάθε τιμή απο την τυχαία μεταβλητη Χ αφαιρούμε την μέση της τιμή και διαιρουμε με την τυπική απόκλιση της, οπου προκύπτει ένας νέος αριθμός. Η κατανομή αυτων των νέων αριθμών ακολουθεί **ξανά** την κανονική κατανομη με διαφορετικές παραμέτρους αυτη την φορά N~(μ=0,σ ^{2}=1). Εχουμε ακολουθήσει την διαδικασία τυποποίησης τηςτυχαίας μεταβλητης Χ για να την μετατρέψουμε σε μια νέα τυχαία μεταβλητη Ζ.

Η Συνάρτηση Πυκνότητας Πιθανότητας (σ.π.π) της τυπικής κανονικής κατανομής περιγράφεται απο την συναρτηση:

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



Η εύρεση πιθανοτήτων απο την τυποποιημένη κανονική κατανομή απαιτεί τον υπολογισμο της συνάρτησης κατανομής

$$\Phi(z) = P(Z \leq z) = \int_{-\infty}^{z} \varphi(y)\, dy = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy \; (\Sigma\chi.1)$$

για τις διάφορες τιμές του z ∈ ℝ. Στην πράξη χρησιμοποιούμε τους πίνακες της τυποποιημένης κανονικής κατανομής. Ενδεικτικά μερικές τιμές αποτελεσμάτων της συναρτησης κατανομής της τυπικής κανονικής κατανομής ειναι η παρακάτω:

z	-3	-2	-1	0	1	2	3
$\Phi(z)$	0.0013	0.0228	0.1587	0.5000	0.8413	0.9772	0.9987

Μερικές τιμές ακόμα για την συναρτηση κατανομής της τυπικής κανονικής κατανομής

$$\Phi(0.90) = 0{,}8195, \Phi(0.91) = 0{,}8186\,, \ldots, \Phi(0.99) = 0{,}8389$$

Στο Παραρτημα 1 δινονται όλες οι τιμες τις συναρτησης κατανομής της τυπικής κανονικής κατανομής.

Ιδιότητες

$$\Phi(z) + \Phi(-z) = 1 \; [\Sigma\chi.2]$$

Για z = 0 εφαρμόζοντας την Σχ.2

$$\Phi(0) + \Phi(0) = 1 \Leftrightarrow 2\,\Phi(0) = 1 \Leftrightarrow \Phi(0) = \frac{1}{2}$$

Επίσης

$$P(-1 \leq Z \leq 1) = P(Z \leq 1) - P(Z \leq -1) \overset{\Sigma\chi1}{\Leftrightarrow} \Phi(1) - \Phi(-1) \overset{\Sigma\chi2}{\Leftrightarrow} \Phi(1) - (1 - \Phi(1)) = 2\Phi(1) - 1. (1)$$

$$P(-2 \leq Z \leq 2) = P(Z \leq 2) - P(Z \leq -2) \overset{\Sigma\chi1}{\Leftrightarrow} \Phi(2) - \Phi(-2) \overset{\Sigma\chi2}{\Leftrightarrow} \Phi(2) - (1 - \Phi(2)) = 2\Phi(2) - 1. (2)$$

$$P(-3 \leq Z \leq 3) = P(Z \leq 3) - P(Z \leq -3) \overset{\Sigma\chi1}{\Leftrightarrow} \Phi(3) - \Phi(-3) \overset{\Sigma\chi2}{\Leftrightarrow} \Phi(3) - (1 - \Phi(3)) = 2\Phi(3) - 1. (3)$$

Και γενικα για $z = z_0$

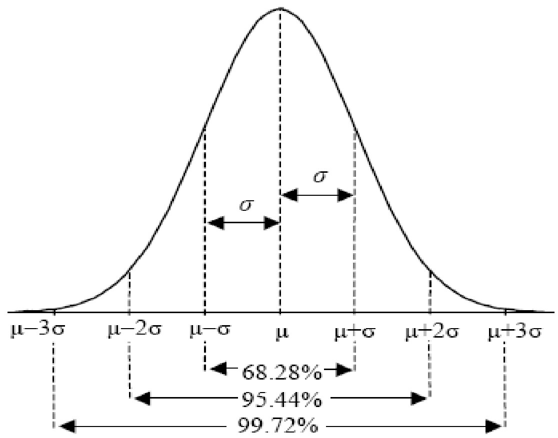
$$P(-z_0 \leq z \leq z_0) = 2\Phi(z_0) - 1.$$

Υπολογισμός των (1), (2), (3) απο τους πίνακες του παραρτήματος 1.

$$2\Phi(1) - 1 = 2(0.8483) - 1 = 0.6826 \cong 68\%$$

$$2\Phi(2) - 1 = 2(0.9772) - 1 = 0.9544 \cong 95\%$$

$$2\Phi(3) - 1 = 2(0.9987) - 1 = 0.9974 \cong 99.7\%$$



Παραδειγμα

Αν $Z \sim N(0,1)$ να βρεθεί ο αριθμός για τον οποίο ισχυεί

$$P(Z > z) = \alpha, \quad 0 < \alpha < 1.$$

Για α = 0.01, 0.05 και 0.10 .

Απάντηση

Εφόσον,

$$P(Z > z) = 1 - P(Z \leq z) = 1 - \Phi(z),$$

θα έχουμε

$$1 - \Phi(z) = \alpha \Leftrightarrow \Phi(z) = 1 - \alpha$$

Για α = 0.01 θα πρέπει να ισχυεί οτι

$$\Phi(z) = 1 - 0.01 = 0.99$$

απο τους πίνακες τιμών της τυπικής κανονικής κατανομης βρίσκουμε

$$z \cong 2.33.$$

Όμοια για α = 0.05

$$\Phi(z) = 1 - 0.05 = 0.95 \Leftrightarrow z \cong 1.645$$

τέλος για α = 0.10

$$\Phi(z) = 1 - 0.10 = 0.90 \Leftrightarrow z \cong 1.28$$

Ο αριθμός z για τον οποίο ισχύει $P(Z > z) = \alpha , 0 < \alpha < 1$ ονομάζεται ***άνω-α-(ποσοστιαίο) σημείο ή ποσοστημόριο.*** Συμφωνα με το προηγούμενο παράδειγμα θα συμβολίσουμε

z_0.01 ≅ 2.33 , z_0.05 ≅ 1.645, z_0.10 ≅ 1.28 **.**

2.2 Κατανομή Χι-Τετράγωνο (Chi-Square Distribution)

Ορισμός

Αν $Z_1, Z_2, \ldots ,Z_\nu$ ανεξάρτητες ν-τυποποιημένες κανονικές κατανομές με ${Z_i}^2 = \frac{X_i-\mu}{\sigma} \approx$ N(0,1) δηλαδή,

$${Z_1}^2 = \frac{X_1-\mu}{\sigma} \approx N(0,1)\,, \quad {Z_2}^2 = \frac{X_2-\mu}{\sigma} \approx N(0,1)\,, \ldots, \quad {Z_\nu}^2 = \frac{X_\nu-\mu}{\sigma} \approx N(0,1)\,.$$

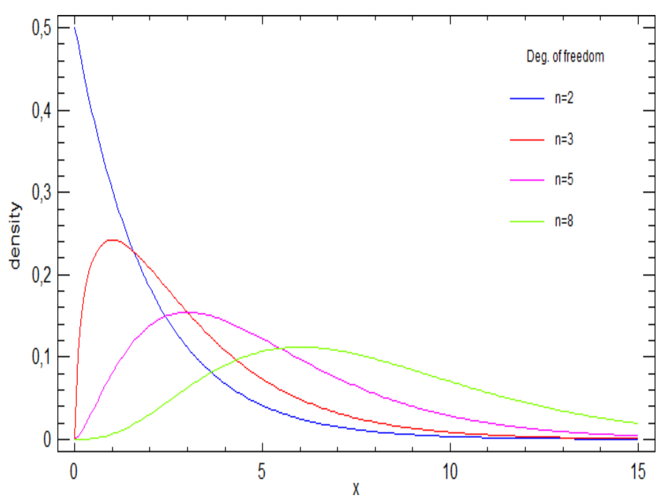
Τότε το άθροισμα των τετραγώνων των τυποποιημένων κανονικών κατανομών

$$\sum_{i=1}^{\nu} {Z_i}^2 = (Z_1 + Z_2 + , \ldots , + Z_\nu) \sim {X_\nu}^2$$

θα ακολουθεί την κατανομή Χι-Τετράγωνο με ν-βαθμούς ελευθερίας.

Συνάρτηση Πυκνότητας Πιθανότητας

$$f_{{X_\nu}^2}(x) = \frac{x^{(\frac{\nu}{2}-1)} e^{-\frac{x}{2}}}{2^{\nu/2}\, \Gamma(\frac{\nu}{2})}, \qquad x > 0.$$



Συνάρτηση Κατανομής

$$F_{X_\nu{}^2}(x) = \frac{1}{2^{\nu/2}\,\Gamma(\frac{\nu}{2})}\int_0^x t^{(\frac{\nu}{2}-1)}e^{(-\frac{t}{2})}, \qquad x > 0.$$

Ανω - α - ποσοστιαίο σημείο

$$P\big(\, X > X_\nu{}^2(a)\,\big) = a$$

Για ν>30 η κατανομή Χι-Τετράγωνο προσεγγίζεται ικανοποιητικά απο την κανονική καντανομε με μέση τιμή ν και διακύμανση 2ν, δηλαδή

$$X_\nu{}^2 \approx N(\nu, 2\nu)$$

Π.χ για $\nu = 50$ *:* $\qquad X_{50}{}^2 \approx N(50{,}100)$

2.3 t - Κατανομή ή Κατανομή Student (t-Distribution/t-Student Distribution)

Ορισμος:

Εστω οτι $Z = \frac{X-\mu}{\sigma} \sim N(0{,}1)$ τυχαία μεταβλητή και έστω S_ν κατανομή ανεξάρτητη απο την Z. Τότε η τυχαία μεταβλητή T_ν

$$T_\nu = \frac{Z}{\sqrt{\frac{S_\nu}{\nu}}} \sim t_\nu$$

Ακολουθεί την t-student κατανομή με ν-βαθμούς ελευθερίας.

Συνάρτηση Πυκνότητας Πιθανότητας της Y $= \sqrt{\frac{S_\nu}{\nu}}$

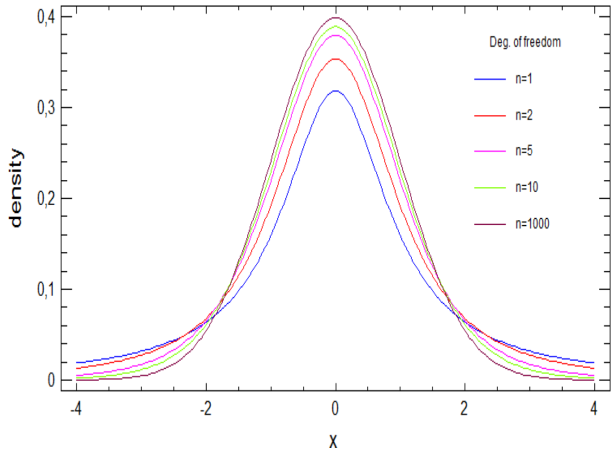
$$f_Y(y) = \frac{\nu^{\frac{\nu}{2}}\, y^{\nu-1}\, e^{-\frac{\nu y^2}{2}}}{2^{(\nu/2\,-\,1)}\,\Gamma(\frac{\nu}{2})}, \qquad y > 0.$$

Συνάρτηση Πυκνότητας Πιθανότητας της Z

$$f_Z(z) = \frac{e^{-\frac{z^2}{2}}}{\sqrt{2\pi}}, \qquad z \in \mathbb{R}.$$

Συνάρτηση Πυκνότητας Πιθανότητας της T

$$f_{T_\nu}(x) = \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\pi\nu}\,\Gamma\left(\frac{\nu}{2}\right)(1+\frac{x^2}{\nu})^{\frac{\nu+1}{2}}}, \qquad x \in \mathbb{R}.$$



Συνάρτηση Κατανομής

$$F_{T_\nu}(x) = \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\pi\nu}\,\Gamma\left(\frac{\nu}{2}\right)}\int_0^x \frac{1}{(1+\frac{t^2}{\nu})^{\frac{\nu+1}{2}}}\,dt, \qquad x > 0.$$

Ιδιότητες

$f_{T_\nu}(x) = f_{T_\nu}(-x), \quad \forall x \in \mathbb{R}$

$E(X) = 0$ και $V(X) = \frac{\nu}{\nu-2}$, $\nu > 2$

Ανω - α - ποσοστιαίο σημείο

$$P(\,X > t_\nu(a)\,) = a$$

Για $\nu > 30$ *η κατανομή t-student προσεγγίζεται ικανοποιητικά απο την* ***Κανονική Καντανομή*** *με μέση τιμή* $E(X) = 0$ *και διακύμανση* $V(X) = \frac{\nu}{\nu-2}$, *δηλαδή*

$$X_\nu{}^2 \approx N(0, \frac{\nu}{\nu-2})$$

Π.χ για $\nu = 50$ *:* $\qquad X_{50}{}^2 \approx N\left(0, \frac{50}{48}\right) = N(0\,, 1.04)$