06.12.2019

ΠΙΘΑΝΟΤΗΤΕΣ

Κανονική Κατανομή ή Κατανομή Gauss

Ορισμός: Η τ.μ. X θα λέγεται κανονική με παραμέτρους μ και σ^2 $(-\infty < \mu < +\infty\ ,\ \sigma^2 > 0)$ αν το σύνολο τιμών της είναι $x \in (-\infty, +\infty)$ και η σ.π.π. της X είναι

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad , \quad \begin{array}{c} -\infty < x < +\infty \\ \sigma = +\sqrt{\sigma^2} \end{array}$$

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

Ιδιότητες της $N(\mu, \sigma^2)$

① Είναι η f_X σ.π.π.; ΝΑΙ i) $f_X(x) \geq 0\ ,\ \forall x \in \mathbb{R}$

ii) $\int_{-\infty}^{+\infty} f_X(x)\,dx = 1$

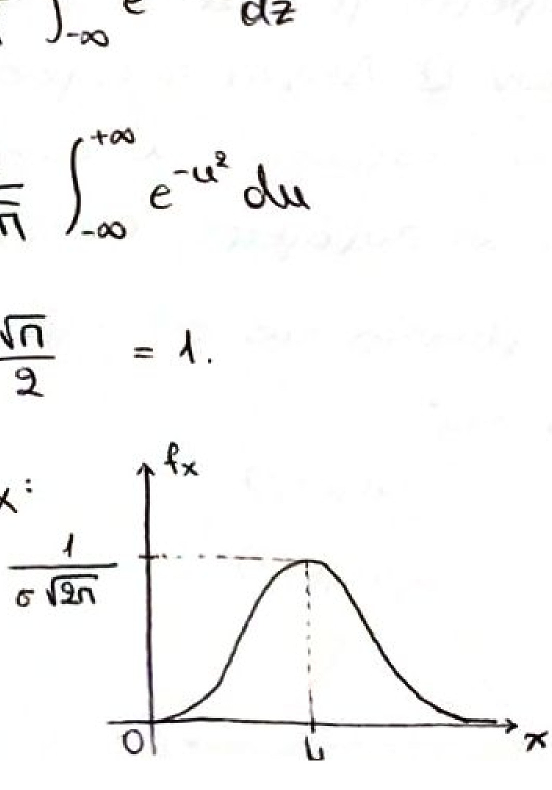
Απόδειξη: Ολοκλήρωμα Euler $\Gamma(1/2) = \sqrt{\pi}$ − $\int_0^{+\infty} e^{-y^2} dy = \frac{\sqrt{\pi}}{2}$.

- $$\int_{-\infty}^{+\infty} f_X(x)\,dx \overset{z = \frac{x-\mu}{\sigma}}{=\!=} \frac{1}{2\sqrt{\pi}} \int_{-\infty}^{+\infty} e^{-z^2/2} dz$$

$$\overset{u = z/\sqrt{2}}{=\!=} \frac{2}{\sqrt{\pi}} \int_{-\infty}^{+\infty} e^{-u^2} du$$

$$\overset{Euler}{=\!=} \frac{2}{\sqrt{\pi}} \cdot \frac{\sqrt{\pi}}{2} = 1.$$

② Γραφική παράσταση της f_X:



③ α.σ.κ. της $N(\mu,\sigma^2)$:

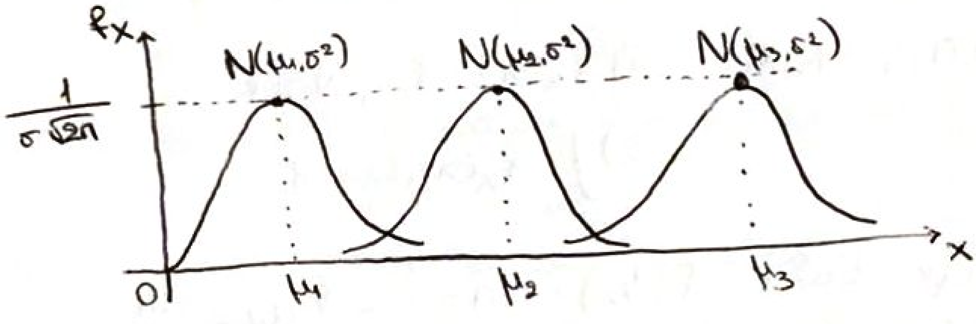
$\forall x\in\mathbb{R}$, $F_X(x) = P(X \le x) = \int_{-\infty}^{x} f_X(t)\,dt$

$$\Rightarrow F_X(x) = \int_{-\infty}^{x} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(t-\mu)^2}{2\sigma^2}}\,dt \quad , \forall x\in\mathbb{R}$$

Δεν υπολογίζεται αναλυτικά.

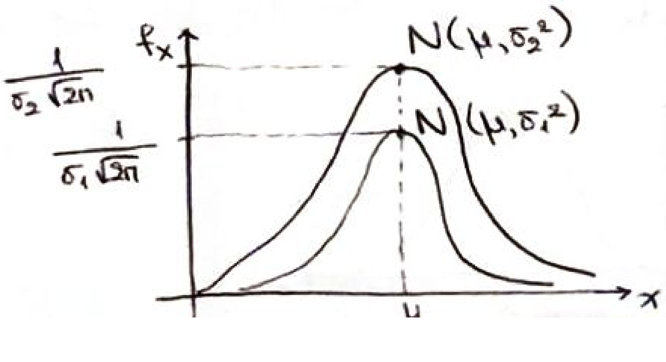
④ Φυσική ερμηνεία των παραμέτρων μ, σ^2.

- Φυσική ερμηνεία της μ: Θεωρούμε $N(\mu_1,\sigma^2)$, $N(\mu_2,\sigma^2)$ και $N(\mu_3,\sigma^2)$ με $\mu_1<\mu_2<\mu_3$.



Η παράμετρος μ είναι ενδεικτική της θέσης της κανονικής στον $\mathbb{R}$. Δείχνει το σημείο γύρω από το οποίο κατανέμεται η "καμπάνα" της κανονικής κατανομής (το μ είναι γνωστό ως παράμετρος θέσης)

- Φυσική ερμηνεία της σ^2: Θεωρούμε $N(\mu,\sigma_1^2)$, $N(\mu,\sigma_2^2)$ όπου $\sigma_1^2<\sigma_2^2$.



Η παράμετρος σ^2 εκφράζει τον βαθμό συγκέντρωσης της κατανομής γύρω από το μ (η σ^2 είναι γνωστή ως παράμετρος διασποράς)

⑤ Ιδιότητα αναλλοίωτου κανονικής κατανομής (Πρόταση)

Έστω τ.μ. $X \sim N(\mu, \sigma^2)$. Τότε η τ.μ. $y = ax + b$, $a > 0$, $b \in \mathbb{R}$ ακολουθεί κανονική κατανομή με παραμέτρους $a\mu + b$ και $a^2\sigma^2$.

Δηλαδή $y \sim N(a\mu + b, a^2\sigma^2)$, $a > 0$, $\mu \in \mathbb{R}$.

Απόδειξη: Στηρίζεται στους μετασχηματισμούς

Ειδική περίπτωση $\left.a = \frac{1}{\sigma} > 0, \quad b = -\frac{\mu}{\sigma}\right\}$

ο μετασχηματισμός $y = ax + b$ ονομάζεται τυπικός μετασχηματισμός

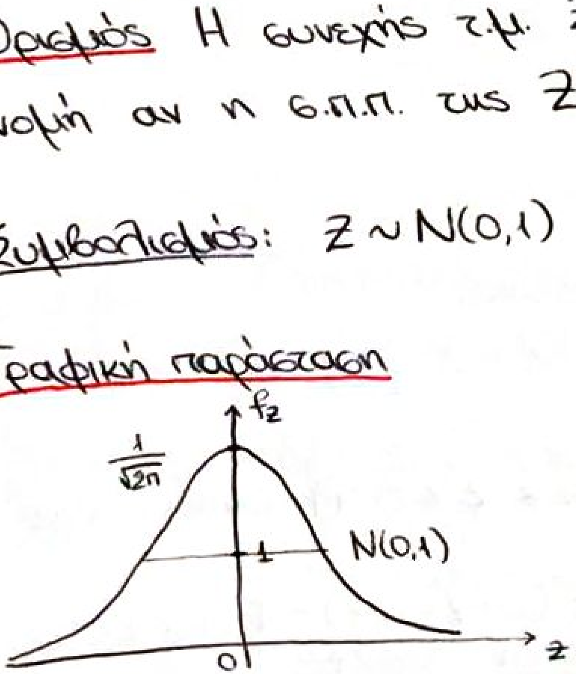
Πόρισμα Έστω $X \sim N(\mu, \sigma^2)$. Τότε η τ.μ. $z = \frac{1}{\sigma}x - \frac{\mu}{\sigma} = \frac{x-\mu}{\sigma} \sim N(0,1)$

Η $N(0,1)$ στην οποία μετασχηματίζονται όλες οι κανονικές κατανομές $N(\mu, \sigma^2)$ μέσω του τυπικού μετασχηματισμού ονομάζεται ΤΥΠΙΚΗ ΚΑΝΟΝΙΚΗ ΚΑΤΑΝΟΜΗ.

Ορισμός Η συνεχής τ.μ. Z ακολουθεί την τυπική κανονική κατανομή αν η σ.π.π. της Z είναι $f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}$, $z \in \mathbb{R}$.

Συμβολισμός: $Z \sim N(0,1)$

Γραφική παράσταση



__α.σ.κ.__ (Φ ή Φ_Z) : $\Phi(z) = P(Z \le z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$, $z \in \mathbb{R}$

Η Φ δεν υπολογίζεται αναλυτικά

__Ερώτημα:__ Αφού η Φ δεν είναι διαθέσιμη σε αναλυτική μορφή πως μπορώ να υπολογίσω πιθανότητες, αφού $P(a \le Z \le b) = \Phi(b) - \Phi(a)$; $Z \sim N(0,1)$

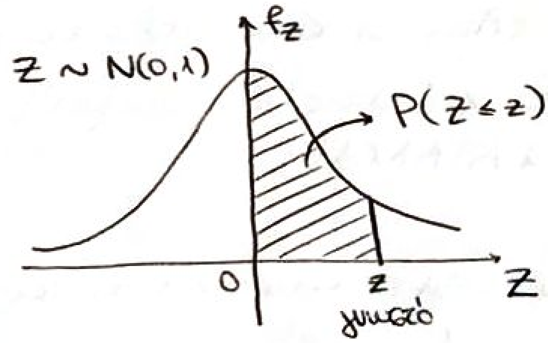
__Απάντηση:__ Αριθμητική ολοκλήρωση αφού $P(a \le Z \le b) = \Phi(b) - \Phi(a) = \int_a^b \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz$

ή

__ΠΙΝΑΚΕΣ ΤΥΠΙΚΗΣ ΚΑΝΟΝΙΚΗΣ__

δίνουν πιθανότητες της μορφής

$P(Z \le z)$ ή $P(\quad < Z < z)$ με z γνωστό και $Z \sim N(0,1)$



__Παράδειγμα Χρήσης Πινάκων N(0,1)__

Αν $Z \sim N(0,1)$ να υπολογιστούν οι πιθανότητες

$P(0 < Z \le 0{,}75) = 0{,}2734$

$P(0{,}4 \le Z < 1{,}46) = P(0 < Z \le 1{,}46) - P(0 < Z \le 0{,}4) = 0{,}4279 - 0{,}1554 = \dots$

$P(-0{,}7 < Z < -0{,}25) = P(0{,}25 < Z < 0{,}7) = P(0 < Z < 0{,}7) - P(0 < Z < 0{,}25)$
$= 0{,}258 - 0{,}0987 = \dots$

$$P(-1,24 \le Z \le 0,36) = P(0 < Z < 1,24) + P(0 < Z < 0,36) =$$
$$= 0,3925 + 0,1406 = \dots$$

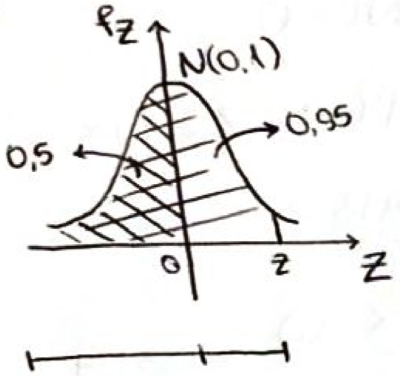
Αντίστροφη Χρήση Πινάκων

Π.χ. Αν $Z \sim N(0,1)$ να υπολογιστεί $z \in \mathbb{R}$ τ.ω. α) $P(Z \le z) = 0,95$

β) $P(Z \ge z) = 0,6772$

γ) $P(Z < z) = 0,87$

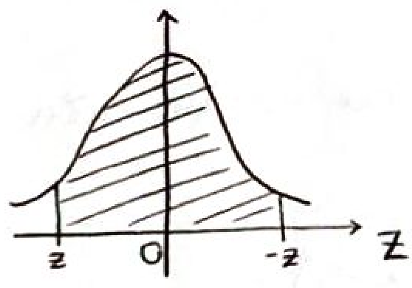
α)



$$0,95 = P(Z \le z) = P(-\infty < Z < 0) + P(0 < Z \le z)$$
$$= 0,5 + P(0 < Z < z)$$
$$\Rightarrow P(0 < Z < z) = 0,95 - 0,5 = 0,45$$

Άρα $P(0 < Z \le z) = 0,45 \Rightarrow z = 1,64$

β)



$$0,6772 = P(Z \ge z) = P(Z \le -z)$$
$$= P(-\infty < Z < 0) + P(0 < Z \le -z)$$
$$= 0,5 + P(0 < Z \le -z)$$
$$\Rightarrow P(0 < Z \le -z) = 0,1772$$
$$\Rightarrow -z = 0,46 \quad \Rightarrow z = -0,46.$$

Υπολογισμό Πιθανοτήτων της μορφής:

$P(a \le X \le b)$ με $X \sim N(\mu, \sigma^2)$ (α, β γνωστά)

$$P(a \le X \le b) = P\left(\frac{a-\mu}{\sigma} \le \frac{X-\mu}{\sigma} \le \frac{b-\mu}{\sigma}\right) = P\left(\frac{a-\mu}{\sigma} \le Z \le \frac{b-\mu}{\sigma}\right)$$

υπολογίζεται από τους πίνακες της $N(0,1)$

όπου $Z \sim N(0,1)$

π.χ. Επίδοση φοιτητών σε ένα μάθημα περιγράφεται από $N(5,1)$

(α) P(η επίδοση φοιτητή να είναι μεταξύ 4 και 6)

(β) Αν είναι ήδη γνωστό ότι η επίδοση φοιτητή είναι μεγαλύτερη του 5. Ποια η πιθανότητα να είναι μεγαλύτερη του 6;

(γ) Ποια τουλάχιστον επίδοση περιμένουμε να συμβεί με πιθανότητα 0,9;

Απάντηση: (α) Έστω X η επίδοση τότε $X \sim N(5,1)$.

$$P(\alpha) = P(4 \leq X \leq 6) = P\left(\frac{4-5}{1} \leq \frac{X-5}{1} \leq \frac{6-5}{1}\right) = P(-1 \leq Z \leq 1) =$$

$$= 2 \cdot P(0 \leq Z \leq 1) = 2 \cdot 0{,}3413 = \ldots$$

(β) $$P(X \geq 6 \mid X \geq 5) = \frac{P(X \geq 6 \text{ ή } X \geq 5)}{P(X \geq 5)} = \frac{P(X \geq 6)}{P(X \geq 5)} = 0{,}3174$$

(γ) Έστω x_0 η επίδοση που πάνω από αυτή θα έχουμε πιθανότητα 0,9. Ζητώ x_0 τέτοιο ώστε:

$$0{,}9 = P(X \geq x_0) = P\left(\frac{X-5}{1} \geq \frac{x_0-5}{1}\right) = P(Z \geq x_0 - 5) =$$

$$= P(-\infty < Z \leq -(x_0-5)) = 0{,}5 + P(0 < Z < -(x_0-5))$$

$$\Rightarrow P(0 < Z \leq -(x_0-5)) = 0{,}4 \quad \Rightarrow \quad -(x_0-5) \simeq 1{,}28 \quad \Rightarrow \quad x_0 \simeq 3{,}72.$$