

Εισαγωγικά στοιχεία

Ω (δειγματικός χώρος): Το σύνολο των αποτελεσμάτων ενός πειράματος

Ενδεχόμενο: Κάθε υποσύνολο του δειγματικού χώρου, το οποίο έχει άμεσα ή έμμεσα, προσδιορίσιμη πιθανότητα.

$\emptyset$: Το κενό σύνολο (αδύνατο ενδεχόμενο)

$$P(\Omega) = 1 \mid P(\emptyset) = 0$$

- Για το τυχαίο ενδεχόμενο A ισχύει: $0 \leq P(A) \leq 1$

- Συμπληρωματικά (αντίθετα) ενδεχόμενα:
2 ενδεχόμενα A, B, για τα οποία $\rightarrow A \cup B = \Omega$
 $A \cap B = \emptyset$

Προσοχή στον συμβολισμό!! $A = B^c \mid B = A^c$

Πράξεις ενδεχομένων

$(A \cup B)$ $\rightarrow$ ένωση ενδεχομένων (πραγματοποιείται το A ή το B ή και τα 2 ταυτόχρονα)

$(A \cap B) = (AB)$ $\rightarrow$ τομή ενδεχομένων (πραγματοποιούνται ταυτόχρονα τα A και B)

Οι πράξεις ενδεχομένων ισχύουν και για περισσότερα από 2 ενδεχόμενα (σύνολα).

- Ασυμβίβαστα ενδεχόμενα (ξένα): 2 ενδεχόμενα για τα οποία ισχύει $A \cap B = \emptyset \Rightarrow P(A \cap B) = P(\emptyset) = 0$
- Ανεξάρτητα ενδεχόμενα: 2 ενδεχόμενα A, B ονομάζονται ανεξάρτητα όταν: $P(AB) = P(A)P(B)$
Γενικεύεται η έννοια και για 3 ενδεχόμενα:
 $P(AB\Gamma) = P(A)P(B)P(\Gamma)$ (και ανά 2 ανεξάρτητα)

* Αν ισχύει: $P(AB) \neq P(A)P(B) \Rightarrow$ Εξαρτημένα ενδεχόμενα

Δεσμευμένη πιθανότητα

Έστω 2 ενδεχόμενα A, B όπου είναι δεδομένο ότι πραγματοποιείται το B. Ονομάζεται δεσμευμένη πιθανότητα του ενδεχομένου A δεδομένου του B, ή αλλιώς “πιθανότητα να πραγματοποιηθεί το ενδεχόμενο A, υπό την προϋπόθεση ότι πραγματοποιείται το B”, η πιθανότητα που δίνεται από τη σχέση:

$$P(A|B) = \frac{P(AB)}{P(B)}$$

- Αν 2 ενδεχόμενα A, B είναι ανεξάρτητα:

$$P(A|B) = \frac{P(A)P(B)}{P(B)} \Rightarrow \\ \Rightarrow P(A|B) = P(A) \text{ (ή) } P(B|A) = \frac{P(A)P(B)}{P(A)} = P(B)$$

- Όταν τα ενδεχόμενα A, B είναι ασυμβίβαστα $\Rightarrow P(AB) = 0 \Rightarrow P(A|B) = 0 = P(B|A)$

Νόμοι De Morgan

- $(A^c \cap B^c) = (A \cup B)^c$

Εκφράζει την
πραγματοποίηση κανενός
από τα 2 ενδεχόμενα

- $(A^c \cup B^c) = (A \cap B)^c$

Θεώρημα Ολικής Πιθανότητας

- $P(A) = \sum_{i=1}^n P(A \cap B_i) = \sum_{i=1}^n P(A|B_i) P(B_i)$

B_i : n ασυμβίβαστα ενδεχόμενα όπου η ένωσή τους είναι ο δειγματικός χώρος Ω
+ Τύπος Bayes (σελ. 5)

Συνεχείς κατανομές

- Συνεχής μεταβλητή: μπορεί να πάρει οποιαδήποτε τιμή, θεωρητικά άπειρες, δε δίνεται η περιγραφή του πειράματος αλλά εκφράζεται μέσω της συνάρτησης πυκνότητας πιθανότητας:

$$f(x) \geq 0, \forall x \in \mathbb{R} \quad (1)$$

→ Συνθήκη κανονικοποίησης: $\int_{-\infty}^{+\infty} f(x) dx = 1 \quad (2)$

Αν μια δοσμένη συνάρτηση ικανοποιεί τις συνθήκες (1) και (2) αποτελεί συνάρτηση πυκνότητας πιθανότητας. Η συνθήκη (2) ισχύει για κάθε σ.π.π. και επομένως όταν ο τύπος της $f(x)$ περιέχει κάποια πραγματική παράμετρο, αξιοποιείται για τον υπολογισμό της παραμέτρου.

$F(x)$: συνάρτηση κατανομής (εκφράζει πιθανότητα) $\Rightarrow$

$$\Rightarrow 0 \leq F(x) \leq 1$$

$$F'(x) = f(x) \Rightarrow F(x) = \int_{-\infty}^x f(t) dt \quad (3)$$

Όταν δίνεται μια συνάρτηση και μας ζητείται να ελέγξουμε αν αποτελεί συνάρτηση κατανομής, τότε εξετάζουμε αν ισχύουν οι παρακάτω προϋποθέσεις:

1) Η $F(x)$ είναι δεξιά συνεχής, δηλαδή: $F(x) = \lim_{t \rightarrow x^+} F(t)$

2) $0 \leq F(x) \leq 1 \rightarrow$ άρα βρίσκουμε το σύνολο τιμών

3) Η $F(x)$ είναι μη φθίνουσα.

Υπολογισμός Πιθανότητας

- $P(a < X < b) = \int_a^b f(x) dx = F(b) - F(a)$
- $P(X < a) = P(-\infty < X < a) = F(a) - \lim_{x \rightarrow -\infty} F(x)$
- $P(X > b) = P(b < X < +\infty) = \lim_{x \rightarrow +\infty} F(x) - F(b)$

Παράμετροι κατανομής

A) Μέση τιμή: $\mu = E[X] = \int_{-\infty}^{+\infty} x f(x) dx$

→ Γενίκευση: $E[g(X)] = \int_{-\infty}^{+\infty} g(x) f(x) dx$

B) Διασπορά (κεντρική ροπή 2^{ης} τάξης):

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

Γ) Κεντρική ροπή ν-τάξης:

$$\mu_\nu = E[(X - \mu)^\nu] = \int_{-\infty}^{+\infty} (x - \mu)^\nu f(x) dx$$

Δ) Ροπή ν-τάξης περί την αρχή:

$$\mu_\nu = E[(X)^\nu] = \int_{-\infty}^{+\infty} (x)^\nu f(x) dx$$

Βασικές κατηγορίες συνεχών κατανομών

- Ομοιόμορφη κατανομή: $f(x) = \begin{cases} c, & x \in [a, b] \\ 0, & x < a \text{ και } x > b \end{cases}$

$$\int_{-\infty}^{+\infty} f(x) dx = 1 \Rightarrow c = \frac{1}{b-a}$$

- Εκθετική κατανομή: $f(x) = \begin{cases} 0, & x < 0 \\ c e^{-cx}, & x \geq 0 \end{cases}$
με $c \in \mathbb{R}$ και $c > 0$.

Αν εφαρμόσουμε τη συνθήκη κανονικοποίησης:

$$\int_{-\infty}^{+\infty} f(x) dx = 1, \text{ προκύπτει } \underline{\text{ταυτοτική σχέση.}}$$

Για να υπολογίσουμε την τιμή της σταθεράς $c \in \mathbb{R}$ θα πρέπει να μας δίνεται η τιμή κάποιας παραμέτρου της κατανομής.

- Κανονική κατανομή (Super SOS!!!): Αποτελεί τη βασικότερη κατηγορία συνεχούς κατανομής, αλλά δε δουλεύει με τυπολόγιο ολοκληρωμάτων αλλά με τον πίνακα τυποποιημένης κανονικής κατανομής, που αποτελεί ξεχωριστή περίπτωση μεθοδολογίας η οποία θα αναλυθεί εκτενώς παρακάτω.

Δισδιάστατη συνεχής κατανομή

- $f(x,y)$: από κοινού συνάρτηση πυκνότητας πιθανότητας

→ Συνθήκη κανονικοποίησης: $\iint f(x,y) dx dy = 1$

- Για $a \leq x \leq b$ και για $\gamma \leq y \leq \delta$ ($a, b, \gamma, \delta \in \mathbb{R}$) ολοκληρώνουμε πρώτα ως προς όποια μεταβλητή θέλουμε.

- Για $a \leq x \leq b$ και για $f(x) \leq y \leq g(x)$, το χωρίο είναι απλό ως προς x. Ολοκληρώνουμε πρώτα ως προς y.
- Για $\gamma \leq y \leq \delta$ και για $f(y) \leq x \leq g(y)$, το χωρίο είναι απλό ως προς y. Ολοκληρώνουμε πρώτα ως προς x.

→ Υπολογισμός πιθανότητας:

$$P(a \leq X \leq b, \gamma \leq y \leq \delta) = \int_{\gamma}^{\delta} \left[\int_a^b f(x, y) dx \right] dy$$

Τη διαδικασία ολοκλήρωσης την ακολουθούμε με κριτήριο τα άκρα όπως αναφέρθηκε παραπάνω. Για τον υπολογισμό των πιθανοτήτων, η χρήση συμβόλου ανισότητας ($<$ ή $>$), ή ανισότητας ($\leq$ ή $\geq$), δεν επηρεάζει καθόλου το τυπολόγιο υπολογισμού της πιθανότητας.

Περιθώριες κατανομές 2-διάστατης μεταβλητής

- $f_X(x) = \int_{-\infty}^{+\infty} f(x, y) dy$ → περιθώρια συνάρτηση πυκνότητας πιθανότητας ως προς x
- $f_Y(y) = \int_{-\infty}^{+\infty} f(x, y) dx$ → περιθώρια συνάρτηση πυκνότητας πιθανότητας ως προς y

→ Ανεξάρτητες μεταβλητές X, Y:

→
όταν ισχύει

$$f(x, y) = f_X(x) \cdot f_Y(y)$$

Παράμετροι 2- διάστατης κατανομής

$$E[g(x, y)] = \iint g(x, y) f(x, y) dx dy$$

- Συνδιακύμανση (ή μικτή κεντρική ροπή 2^{ης} τάξης):

$$\text{Cov}(X, Y) = E(XY) - E(X) E(Y) \rightarrow \text{SOS!!}$$

- Συντελεστής συσχέτισης των μεταβλητών X, Y

$$\rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y}$$

Όταν $\text{Cov}(X, Y) = 0 \Rightarrow \rho(X, Y) = 0 \Rightarrow$ οι μεταβλητές X, Y λέγονται ασυσχέτιστες ή ορθογώνιες.

$$\sigma_X^2 = E[X^2] - E^2[X] \quad | \quad \sigma_Y^2 = E[Y^2] - E^2[Y]$$

Δεσμευμένη κατανομή

- Ορίζουμε τη δεσμευμένη σ.π.π. της συνεχούς τ.μ. X όταν $Y=y$ ως:
 $f_{X|Y}(x|y) = \frac{f(x,y)}{f_Y(y)}$, όπου f_Y είναι η περιθώρια σ.π.π. της μεταβλητής Y.
- Ορίζουμε τη δεσμευμένη σ.π.π. της συνεχούς τ.μ. Y όταν $X=x$ ως:
 $f_{Y|X}(y|x) = \frac{f(x,y)}{f_X(x)}$, όπου f_X είναι η περιθώρια σ.π.π. της μεταβλητής X.

→ Βασική παρατήρηση: → SOS!!

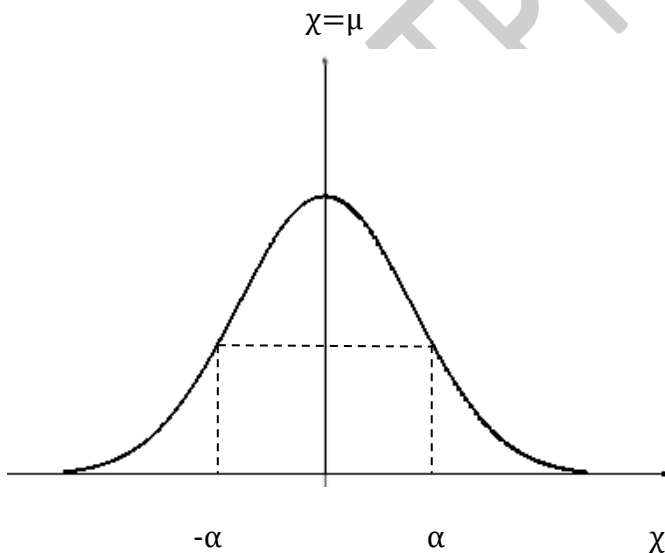
Όταν ζητείται ο υπολογισμός πιθανότητας και δεν είναι σαφώς καθορισμένα τα όρια τιμών των μεταβλητών X, Y (άρα ούτε και τα άκρα ολοκλήρωσης του διπλού ολοκληρώματος) και δίνεται απλά μια ανισοτική σχέση μεταξύ των X, Y , (π. χ. $(X + Y) > 1$), τότε θα πρέπει να κατασκευάσουμε το χωρίο που αντιστοιχεί στο γεωμετρικό τόπο της δοσμένης ανισότητας και στη συνέχεια με βάση τα σύνορα του χωρίου να καθορίσουμε τα άκρα ολοκλήρωσης των X, Y .

ΚΑΝΟΝΙΚΗ ΚΑΤΑΝΟΜΗ

(Κατανομή Gauss)

→ Βασική διαφορά με τις υπόλοιπες κατανομές:

Η σ.π.π. είναι ιδιαίτερη σύνθετη στη χρήση της με ολοκληρώματα επομένως δε “δουλεύει” με ολοκληρώματα, αλλά με χρήση του Πίνακα της Τυποποιημένης Κανονικής Κατανομής.



$\int_{-\infty}^{+\infty} f(x) dx = 1 \Rightarrow$ Το εμβαδόν μεταξύ του άξονα των x και της γραφικής παράστασης της f ισούται με 1.

→ Προσοχή!!! Ο πίνακας που χρησιμοποιούμε αφορά την τυποποιημένη κανονική κατανομή, όπου η μέση τιμή έχει μηδενική τιμή και η τυπική απόκλιση σ ισούται με 1.

Λόγω της συμμετρίας του σχήματος, θα ισχύει:

$$P(X > a) = P(X < -a)$$

η κάθε πιθανότητα εκφράζει το εμβαδόν του αντίστοιχου χωρίου που περικλείεται από την κατακόρυφη ευθεία και την καμπύλη

- $$P(a \leq X \leq b) = P(X \leq b) - P(X \geq a)$$

ΧΡΗΣΗ ΠΙΝΑΚΑ

- Ισχύει:
$$\Phi(Z) = P(X < Z) = \int_{-\infty}^Z f(u)du \quad (A)$$

Για δεδομένη τιμή του z , η πιθανότητα υπολογίζεται με χρήση του πίνακα της τυποποιημένης κανονικής κατανομής. Η σχέση (A) ισχύει για $X < z$.

Παρατηρήσεις

1) Ο πίνακας αποτελείται από γραμμές και στήλες, όπου για το z απαιτούνται τιμές και από τις δύο (η στήλη μας δίνει το δεύτερο δεκαδικό της τιμής του z και η γραμμή το ακέραιο μέρος και το πρώτο δεκαδικό της τιμής του z).

2) Οι τιμές στο εσωτερικό του πίνακα αποτελούν τις τιμές της $\Phi(z)$.

3) Στις ασκήσεις θα μας δίνεται μια τυχαία μεταβλητή X ακολουθεί κανονική (όχι τυποποιημένη) κατανομή, επομένως για να προχωρήσουμε σε χρήση του πίνακα, θα πρέπει πρώτα να ορίσουμε τη μεταβλητή:

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

όπου μ η μέση τιμή της X και σ η τυπική απόκλιση. Η Z ακολουθεί τυποποιημένη κανονική κατανομή, οπότε πλέον μπορούμε να χρησιμοποιήσουμε τον πίνακα.

Συχνές περιπτώσεις ζητούμενων πιθανοτήτων

A) Αν ζητείται η πιθανότητα: $P(X \geq Z) = 1 - P(X < Z) \Rightarrow$
 $\Rightarrow P(X \geq Z) = 1 - \Phi(Z)$

B) Αν ζητείται η πιθανότητα $P(X < -Z)$, με $(-z) < 0$, επειδή ο πίνακας περιέχει μόνο θετικές τιμές της z , θα πρέπει να χρησιμοποιήσουμε τη συμμετρία της γραφικής παράστασης και την ισότητα εμβαδών:

$$P(X < -z) = P(X > z) \Rightarrow P(X < -z) = 1 - P(X < z) = 1 - \Phi(z)$$

Γ) Οι πιθανότητες στην περίπτωση κανονικής κατανομής εκφράζουν το εμβαδόν του χωρίου που περικλείεται μεταξύ του γεωμετρικού τόπου του χωρίου που εκφράζει η ανισότητα του x και της γραφικής παράστασης της καμπύλης.

$$\Delta) P(a \leq Z \leq b) = \underbrace{\Phi(b) - \Phi(a)}_{\text{χρήση πίνακα}}$$