MEM-262 Παραμετρική Στατιστική

Τμήμα Μαθηματικών και Εφ. Μαθηματικών, Πανεπιστήμιο Κρήτη

Διδάσκων : Κώστας Σμαραγδάκης

Διάλεξη 1 : 02-10-2020

Πείραμα τύχης και δειγματικός χώρος Ω

Ορισμός : Πείραμα τύχης

Πείραμα τύχης είναι το πείραμα το αποτέλεσμα του οποίου δεν μπορεί να προβλεφτεί με βεβαιότητα ανεξάρτητα του αριθμού πραγματοποιήσεων του στο ίδιο περιβάλλον.

Ορισμός : Δειγματικός χώρος Ω

Το σύνολο των δυνατών αποτελεσμάτων του πειράματος τύχης καλείται **δειγματικός χώρος**.

Παραδείγματα

- Ρίψη ενός ζαριού. $\Omega = \{1,2,3,4,5,6\}$
- Ρίψη δύο νομισμάτων. $\Omega = \{KK, K\Gamma, \Gamma K, \Gamma\Gamma\}$
- Χρόνος μεταξύ διαδοχικών πραγματοποιήσεων ενός φαινομένου. $\Omega = [0, +\infty)$

Είναι σημαντικό να καθορίσουμε τι πληροφορία χρειαζόμαστε μετά από επαναλήψεις του πειράματος.

Περιλαμβάνει όλα τα ενδεχόμενα που μας χρειάζονται.
Στις εφαρμογές συνήθως χρησιμοποιούμε τη μικρότερη σ-άλγεβρα που περιέχει τη πληροφορία που μας ενδιαφέρει.

Ορισμός : σ-άλγεβρα στο Ω

Μια συλλογή $\mathcal{G}$ υποσυνόλων του Ω ονομάζεται σ-άλγεβρα εάν ισχύουν οι παρακάτω ιδιότητες:

1. $\Omega \in \mathcal{G}$
2. $A \in \mathcal{G} \;\Rightarrow\; A^c \in \mathcal{G}$
3. $A_i \in \mathcal{G}, i = 1, 2 \ldots \;\Rightarrow\; \cup_{j=1}^{\infty} A_j \in \mathcal{G}$

$(\bigcap_j A_j)^c = \bigcup_j A_j^c$

Θα δούμε στη συνέχεια γιατί οι παραπάνω ιδιότητες είναι απαραίτητες για να εξασφαλισθεί η συνέπεια της πληροφορίας που αντλούμε από το πείραμα τύχης.

Ορισμός : μεγαλύτερη σ-άλγεβρα στο Ω

Το **δυναμοσύνολο** του Ω (συλλογή όλων των υποσυνόλων) αποτελεί την μεγάλυτερη σ-άλγεβρα στο Ω. Το συμβολίζουμε ως 2^{Ω}.

Παράδειγμα

Θα γράψουμε το δυναμοσύνολο του $\Omega = \{1, 2, 3\}$ $\quad g = \{\emptyset, \Omega\}$

$2^{\Omega} = \{\{1\}, \{2\}, \{3\}, \{1,2\}, \{1,3\}, \{2,3\}, \{1,2,3\}, \emptyset\}$

Ορισμός : σ-άλγεβρα που παράγεται από τη συλλογή υποσυνόλων $\mathcal{A}$

Έστω $\mathcal{A}$ μια συλλογή υποσυνόλων του Ω. Συμβολίζουμε με $\sigma(\mathcal{A})$ την μικρότερη σ-άλγεβρα στο Ω η οποία περιέχει τη $\mathcal{A}$.

Παράδειγμα

Έστω $\Omega = \{1, 2, 3, 4\}$ και $\mathcal{A} = \{\{1, 2\}, \{3\}\}$. Θα γράψουμε την $\sigma(\mathcal{A})$

$\{1,2\}^c \quad \{3\}^c \quad \{1,2,3\}^c \quad \Omega$

$\sigma(\mathcal{A}) = \{\{1,2\}, \{3\}, \{1,2,3\}, \{3,4\}, \{1,2,4\}, \{4\}, \{1,2,3,4\}, \emptyset\}$

$\{1,2\} \cup \{3\}$

Έστω ένα πείραμα τύχης με δειγματικό χώρο το $\mathbb{R}$, δηλαδή $\Omega \doteq \mathbb{R}$.

Ορισμός : σ-άλγεβρα Borel

Η σ-άλγεβρα που παράγεται από τα ανοιχτά διαστήματα του $\mathbb{R}$ καλείται Borel.

$$\mathcal{B}(\mathbb{R}) \doteq \sigma(\{(a,b),\ -\infty \leq a \leq b \leq \infty\})$$

$(\alpha, b] = \bigcap_{n=1}^{\infty} (\alpha, b + \frac{1}{n})$

Μπορούμε να δείξουμε ότι ισχύουν ισοδύναμα:

$(\alpha, b) = \bigcup_{n=1}^{\infty} (\alpha, b - \frac{1}{n}]$

$$\mathcal{B}(\mathbb{R}) \doteq \sigma(\{(-\infty, x],\ x \in \mathbb{R}\})$$

$-\infty \leq a \leq b < \infty$ $\quad$ $-\infty < a \leq b \leq +\infty$ $\quad$ $-\infty < a \leq b < +\infty$

$$\mathcal{B}(\mathbb{R}) \doteq \sigma(\{(a,b]\}), \quad \mathcal{B}(\mathbb{R}) \doteq \sigma(\{[a,b)\}), \quad \mathcal{B}(\mathbb{R}) \doteq \sigma(\{[a,b]\})$$

Ερώτηση : *Πως μπορεί να ορισθεί η* $\mathcal{B}([0,1])$;

$\mathcal{B}([0,1]) = \sigma(\{(\alpha, b),\ 0 \leq \alpha \leq b \leq 1\})$

Παρουσιάζει ενδιαφέρον ότι υπάρχουν υποσύνολα του $\mathbb{R}$ *που δεν ανήκουν στην* $\mathcal{B}(\mathbb{R})$.
Δηλαδή $\mathcal{B}(\mathbb{R}) \neq 2^{\mathbb{R}}$.

Ορισμός : Μέτρο πιθανότητας $P: \mathcal{G} \to [0,1]$

Η απεικόνιση P από τη σ-άλγεβρα $\mathcal{G}$ στο διάστημα $[0,1]$ ονομάζεται μέτρο πιθανότητας εάν ισχύουν οι παρακάτω ιδιότητες:

1. $P(\Omega) = 1$
2. Αν $A_i \in \mathcal{G}, i = 1, 2 \ldots$ ξένα υποσύνολα στην $\mathcal{G}$ τότε $P(\cup_{j=1}^{\infty} A_j) = \sum_{j=1}^{\infty} P(A_j)$

Άσκηση

Αποδείξτε ότι για το μέτρο πιθανότητας ισχύουν οι παρακάτω ιδιότητες

1. Για κάθε $A \in \mathcal{G}$ ισχύει $P(A^c) = 1 - P(A)$. $P(\emptyset) = 0$
2. Για κάθε $A_1, A_2 \in \mathcal{G}$ με $A_1 \subseteq A_2$ ισχύει $P(A_1) \leq P(A_2)$.
3. Για κάθε $A_1, A_2 \in \mathcal{G}$ ισχύει $P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2)$.

Ορισμός : Χώρος πιθανότητας

Η τριπλέτα $(\Omega, \mathcal{G}, P)$ καλείται **χώρος πιθανότητας**.

Παράδειγμα - Ρίψη ενός ζαριού

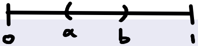
Έστω ότι θέλουμε να μελετήσουμε εάν το αποτέλεσμα είναι άρτιος ή περιττός

$\{1,3,5\}$ $\{2,4,6\}$

- Δειγματικός χώρος : $\Omega = \{1, 2, 3, 4, 5, 6\}$
- σ-άλγεβρα ενδεχομένων : $\mathcal{G} = \{\emptyset, \{1, 3, 5\}, \{2, 4, 6\}, \Omega\}$
- Μέτρο πιθανότητας : $P : \mathcal{G} \to [0, 1]$

$$P(\emptyset) = 0,\ P(\{1, 3, 5\}) = 0.5,\ P(\{2, 4, 6\}) = 0.5,\ P(\Omega) = 1$$

Ερώτηση : *Θα μπορούσαμε να είχαμε επιλέξει* $\mathcal{G} \doteq 2^{\Omega}$;



Παράδειγμα - Επιλογή αριθμού στο $[0, 1]$

Έστω ότι θέλουμε να μελετήσουμε πόσο πιθανό είναι ο επιλεγμένος αριθμός να ανήκει στα διάφορα Borel υποσύνολα του $[0, 1]$.

- Δειγματικός χώρος : $\Omega = [0,1]$
- σ-άλγεβρα ενδεχομένων : $\mathcal{G} = \mathcal{B}([0,1])$
- Μέτρο πιθανότητας : $P((\alpha,b)) = b-\alpha$

$$P([\alpha,b]) = b-\alpha$$

$$P([0,1]) = 1-0 = 1$$

$$P(1/2) = 0$$

Ορισμός : Ανεξάρτητα ενδεχόμενα

Δύο ενδεχόμενα $A_1, A_2 \in \mathcal{G}$ καλούνται **ανεξάρτητα** εάν για το μέτρο πιθανότητας ισχύει:

$$P(A_1 \cap A_2) = P(A_1)P(A_2)$$

.

Ορισμοί : Ανεξαρτησία σε ακολουθίες ενδεχομένων $\{A_1, A_2, \dots\}$ $A_i \in \mathcal{G}$

Έστω μια ακολουθία ενδεχομένων, δηλαδή $A_i \in \mathcal{G}, i = 1, \dots$. Λέμε ότι τα ενδεχόμενα (όροι της ακολουθίας) είναι:

1. Ανά δύο ανεξάρτητα εάν $P(A_i \cap A_j) = P(A_i)P(A_j)$ για κάθε $i, j = 1, 2, \dots$ με $i \neq j$.
2. Από κοινού ανεξάρτητα εάν $P(\cap_j A_j) = \prod_j P(A_j)$.

Παράδειγμα

Επιλέγουμε με ίση πιθανότητα ένα από τα παρακάτω σημεία του $\mathbb{R}^3$.

$$\overset{1/4}{(1,0,0)},\ \overset{1/4}{(0,1,0)},\ \overset{1/4}{(0,0,1)},\ \overset{1/4}{(1,1,1)}$$

Έστω τα ενδεχόμενα $A_i = \{$η i-οστή συνιστώσα του σημείου είναι $1\}$, $i = 1, 2, 3$.

1. Είναι τα ενδεχόμενα A_i, $i = 1, 2, 3$ ανά δύο ανεξάρτητα;
2. Είναι τα ενδεχόμενα A_i, $i = 1, 2, 3$ από κοινού ανεξάρτητα;

$A_1 = \{(1,0,0),(1,1,1)\}$ $\quad A_2 = \{(0,1,0),(1,1,1)\}$ $\quad A_3 = \{(0,0,1),(1,1,1)\}$

$A_1 \cap A_2 = \{(1,1,1)\} = A_1 \cap A_3 = A_2 \cap A_3$ $\quad P(A_1 \cap A_2) = 1/4 = P(A_1 \cap A_3) = P(A_2 \cap A_3)$

$A_1 \cap A_2 \cap A_3 = \{(1,1,1)\}$ $\quad P(A_1 \cap A_2 \cap A_3) = 1/4$ $\quad P(A_1) = P(A_2) = P(A_3) = 1/2$

$P(A_1 \cap A_2) = P(A_1) \cdot P(A_2)$

$\frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2}$

$P(A_1 \cap A_2 \cap A_3) \neq P(A_1) P(A_2) P(A_3)$

$1/4 \neq 1/2 \; 1/2 \; 1/2 = 1/8$

Θα ασχοληθούμε με τη πιθανότητα να πραγματοποιηθεί το $A \in \mathcal{G}$ εφόσον πραγματοποιείται το $B \in \mathcal{G}$. Συμβολίζουμε το παραπάνω ενδεχόμενο ως $A|B$.

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \ P(B) > 0$$

Άσκηση

Δείξτε ότι αν $A, B \in \mathcal{G}$ είναι ανεξάρτητα ενδεχόμενα τότε $P(A|B) = P(A)$.

$P(A|A) = ?$

Τυχαίες μεταβλητές

$\mathcal{G} = 2^{\Omega}$
$\Omega = \{K, \Gamma\}$
$K \to 1$
$\Gamma \to 0$

$X(\omega) = \begin{cases} 1, & \omega \in K \\ 0, & \omega \in \Gamma \end{cases}$ $1 \in (1/2, 3/2)$

$X^{-1}(1) = K \in \mathcal{G}$ $X^{-1}((1/2, 3/2)) = K \in \mathcal{G}$
$X^{-1}(0) = \Gamma \in \mathcal{G}$
$X^{-1}(\mathbb{R}) = \Omega \in \mathcal{G}$ $X^{-1}(2) = \emptyset \in \mathcal{G}$

Έστω $(\Omega, \mathcal{G}, P)$ χώρος πιθανότητας.

Ορισμός : Τυχαία μεταβλητή $X : \Omega \to \mathbb{R}$

Ονομάζεται κάθε απεικόνιση X από το δειγματικό χώρο στο $\mathbb{R}$ για την οποία ισχύει:

$$X^{-1}(A) = \{\omega : X(\omega) \in A\} \in \mathcal{G} \quad \text{για κάθε} \quad A \in \mathcal{B}(\mathbb{R}).$$

Επιπλέον:

X είναι τυχαία μεταβλητή ανν $\mathcal{B}(\mathbb{R}) = \sigma((-\infty, x])$

$$\{\omega : X(\omega) \le x\} \in \mathcal{G}, \quad \text{για κάθε} \quad x \in \mathbb{R}$$

Παράδειγμα

$$\Omega = \{1, 2, 3, 4, 5, 6\}, \ \mathcal{G} = \{\emptyset, \{1, 3, 5\}, \{2, 4, 6\}, \Omega\}$$

$$X : \Omega \to \mathbb{R}, \quad X(\omega) = \begin{cases} 1, & \omega \in \{1, 3, 5\} \\ 0, & \omega \in \{2, 4, 6\}. \end{cases}$$

Είναι η X τυχαία μεταβλητή;

$$X^{-1}(1) = \{1,3,5\} \in \mathcal{G}$$

$$X^{-1}(0) = \{2,4,6\} \in \mathcal{G}$$

Χώρος πιθανότητας παραγόμενος από τυχαία μεταβλητή

Ορισμός : Μέτρο πιθανότητας της τυχαίας μεταβλητής X

Ορίζουμε τη απεικόνιση P_X από το $\mathcal{B}(\mathbb{R})$ στο $[0,1]$ ως:

$$P_X(A) = P(X^{-1}(A)) = P(\{\omega \in \Omega : X(\omega) \in A\}) \quad \text{για κάθε} \quad A \in \mathcal{B}(\mathbb{R})$$

Θεώρημα

Η τριπλέτα $(\mathbb{R}, \mathcal{B}(\mathbb{R}), P_X)$ ορίζει χώρο πιθανότητας. $X : \Omega \to \mathbb{R}$

$P_X(A) = P(X^{-1}(A)) \in [0,1]$ ✓

1. $P_X(\mathbb{R}) = P(\{\omega : X(\omega) \in \mathbb{R}\}) = P(\Omega) = 1$ ✓

Έστω A_n ξένα υποσύνολα στη $\mathcal{B}(\mathbb{R})$

2. $P_X(\bigcup_n A_n) = P(\{\omega : X(\omega) \in \bigcup_n A_n\}) = P(\bigcup_n \{\omega : X(\omega) \in A_n\}) = \sum_n P(\{\omega : X(\omega) \in A_n\}) = \sum_n P_X(A_n)$ ✓

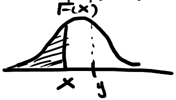
***Συμβολισμός** : Για ευκολία από εδώ και στο εξής θα γράφουμε και την P_X ως P.*

Συγκεντρωτική συνάρτηση κατανομής

Ορισμός : Συγκεντρωτική συνάρτηση κατανομής

Για την τυχαία μεταβλητή X ορίζουμε τη συγκεντρωτική συνάρτηση κατανομής ως:

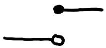
$$F(x) = P(X \leq x), \quad \text{για κάθε} \quad x \in \mathbb{R}.$$



Ιδιότητες

Για την συγκεντρωτική συνάρτηση κατανομής της τμ X ισχύουν:

1. $-\infty < x < y < \infty \quad \Rightarrow \quad F(x) \leq F(y)$.
2. $\lim_{x \to -\infty} F(x) = 0 \quad$ και $\quad \lim_{x \to \infty} F(x) = 1$.
3. Η F είναι δεξιά συνεχής.



Διακριτές τυχαίες μεταβλητές

$X(\omega) = \begin{cases} 1 \ (x_2), & \omega \in K \\ 0 \ (x_1), & \omega \in \Gamma \end{cases}$ $\quad S(X) = \{0, 1\}$

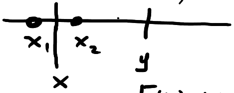
Συνάρτηση μάζας πιθανότητας

Μια τυχαία μεταβλητή X καλείται **διακριτή** αν $P(X = x) > 0$ ισχύει μόνο σε ένα αριθμήσιμο σύνολο $S(X) = \{x_1, x_2, \ldots\}$. Το σύνολο αυτό ονομάζεται **φορέας** της X.

Ορίζουμε:

$$p(x) = P(X = x), \quad \text{για κάθε} \quad x \in \mathbb{R}$$

$F(x) = p(x_1)$



$F(y) = p(x_1) + p(x_2)$

Η συνάρτηση p ονομάζεται **συνάρτηση μάζας πιθανότητας** της X.

Συγκεντρωτική συνάρτηση κατανομής

$$F(x) = P(X \leq x) = \sum_{x_j \leq x} p(x_j), \quad x \in \mathbb{R}$$

Συνάρτηση πυκνότητας πιθανότητας

Μια τυχαία μεταβλητή X καλείται **συνεχής** αν υπάρχει συνάρτηση $f : \mathbb{R} \to [0, \infty)$ τέτοια ώστε:

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(s)ds, \quad \text{για κάθε} \quad x \in \mathbb{R}.$$

$f(x) = F'(x)$

Η συνάρτηση f θα καλείτε **συνάρτηση πυκνότητας πιθανότητας** της τυχαίας μεταβλητής.

Πιθανότητα ενδεχομένου

Έστω $A \in \mathcal{B}(\mathbb{R})$ η πιθανότητα του ενδεχομένου $X \in A$ δίνεται ως:

$$P(X \in A) = \int_{A} f(s)ds$$

Ερώτηση : *Ποία είναι η πιθανότητα $X = x$ για οποιοδήποτε $x \in \mathbb{R}$;*

$P(x \leq X \leq x+\varepsilon) = P(X \leq x+\varepsilon) - P(X \leq x) \xrightarrow{\varepsilon \to 0} 0$