Φροντιστήριο 2ο, 03,08/03/2019

Κατανομές από Μεταφορά

Οι τυχαίες μεταβλητές μεταφέρουν τις κατανομές πιθανότητας στους πραγματικούς. Αν θέλουμε να προσάψουμε πιθανότητα σε ένα μετρήσιμο υποσύνολο των πραγματικών, βρίσκουμε την αντίστροφη εικόνα αυτού μέσω της X και αποδίδουμε σε αυτό πιθανότητα μέσω της κατανομής που υπάρχει ήδη στον Ω.

Ορισμός

Έστω ο χώρος πιθανότητας $(\Omega, \Sigma_\Omega, P)$, ο μετρήσιμος χώρος $(\mathbb{R}, \Sigma_\mathbb{R})$ και τυχαία μεταβλητή X. Το ζεύγος P, X προσδιορίζει μονοσήμαντα κατανομή στο $\mathbb{R}$, έστω P^* που ορίζεται ως εξής:

$$\text{Αν } A \in \Sigma_\mathbb{R}, \text{ τότε } P^*(A) := P(X^{-1}(A))$$

Παρατηρήσεις

i. Η P^* είναι καλώς ορισμένη κατανομή στο $\mathbb{R}$ εξαιτίας του ότι η P είναι καλώς ορισμένη κατανομή στο Ω. Ονομάζεται κατανομή από μεταφορά της P στο $\mathbb{R}$, μέσω της X.

ii. Συνήθως, η P^* ονομάζεται (αγνοώντας την υφιστάμενη P) ως η κατανομή που ακολουθεί η X $(X \sim P^*)$.

iii. Εφόσον μπορούμε να μεταφέρουμε κατανομές στους πραγματικούς μέσω τυχαίων μεταβλητών, μπορούμε καταρχάς να μελετούμε κατανομές πιθανότητας στο $\mathbb{R}$, όπου υπάρχει πλούσια μαθηματική δομή

Παράδειγμα

Να βρεθεί το P^* όταν $\underline{\Omega}=\{K,\Gamma\}$, $P(\{K\})=\frac{1}{3}$, $X(K)=3$, $X(\Gamma)=4$

Βήμα 1 : Βρίσκουμε $\Sigma_{\underline{\Omega}}$, $P(\{\Gamma\})$

$$\Sigma_{\underline{\Omega}} = \{\emptyset, \underline{\Omega}, \{K\}, \{\Gamma\}\}$$

$$P(\underline{\Omega}) = P(\{K,\Gamma\}) = P(\{K\}\cup\{\Gamma\}) = P(\{K\}) + P(\{\Gamma\}) \Rightarrow$$

$$P(\underline{\Omega}) = P(\{K\}) + P(\{\Gamma\}) \Leftrightarrow 1 = \frac{1}{3} + P(\{\Gamma\}) \Rightarrow P(\{\Gamma\}) = \frac{2}{3}$$

Βήμα 2 : Υπολογίζουμε τις αντίστροφες εικόνες

$$X^{-1}(A) = \begin{cases} \emptyset \in \Sigma_{\underline{\Omega}}, & \text{αν } 3,4 \notin A \\ \underline{\Omega} \in \Sigma_{\underline{\Omega}}, & \text{αν } 3,4 \in A \\ \{K\} \in \Sigma_{\underline{\Omega}}, & \text{αν } 3\in A,\ 4\notin A \\ \{\Gamma\} \in \Sigma_{\underline{\Omega}}, & \text{αν } 3\notin A,\ 4\in A \end{cases}$$

Βήμα 3 : Υπολογίζουμε το $P^*(A)$

$$P^*(A) := P(X^{-1}(A)), \quad \forall A \in \Sigma_{\mathbb{R}}$$

$$P^*(A) = \begin{cases} P(\emptyset) = 0, & \text{αν } 3,4 \notin A \\ P(\underline{\Omega}) = 1, & \text{αν } 3,4 \in A \\ P(\{K\}) = \frac{1}{3}, & \text{αν } 3\in A,\ 4\in A \\ P(\{\Gamma\}) = \frac{2}{3}, & \text{αν } 3\notin A,\ 4\in A \end{cases}$$

Η έννοια του στηρίγματος

Για τα παρακάτω θεωρούμε χωρίς μεγάλη ακρίβεια ότι κάποιο υποσύνολο του $\mathbb{R}$ θεωρείται κλειστό αν και μόνο αν το συμπλήρωμά του θεωρείται ανοικτό. Παραδείγματα είναι το $\mathbb{R}$, το $\emptyset$, τα μονοσύνολα, τα κλειστά διαστήματα, τα διακριτά υποσύνολα του $\mathbb{R}$, πεπερασμένες ενώσεις αυθαιρέτου πλήθους, τομές αυτών κλπ. Αναλόγως, όταν $A, B \subseteq \mathbb{R}$ τότε το A θα ονομάζεται μικρότερο του B αν και μόνο αν $A \subset B$ (το A είναι γνήσιο υποσύνολο του B).

Ορισμός

Έστω κατανομή πιθανότητας P στο $\mathbb{R}$. Το στήριγμα (supp) αυτής είναι το μικρότερο κλειστό υποσύνολο του $\mathbb{R}$ στο οποίο η P αποδίδει μοναδιαία πιθανότητα.

Χρησιμότητα του στηρίγματος

Το στήριγμα διευκολύνει τη διαδικασία περιγραφής μιας κατανομής πιθανότητας στον $\mathbb{R}$, αφού $P(A) = P(A \cap supp)$

Πόρισμα

Αν $A \in \Sigma_{\mathbb{R}}$ τότε $P(A) = P(A \cap supp)$

Απόδειξη

$$A = (A \cap supp) \cup (A \cap supp') \Rightarrow$$

$$P(A) = P[(A \cap supp) \cup (A \cap supp')] \Rightarrow$$

$$P(A) = P(A \cap supp) + P(A \cap supp')$$

Γενικά ισχύει ότι $supp' \supseteq A \cap supp' \underset{\text{ΜΟΝΟΤΟΝΙΑ}}{\Rightarrow} P(supp') \geq P(A \cap supp')$

Εφόσον $P(supp) = 1, \Rightarrow P(supp') = 0$. Άρα $P(A \cap supp') \leq 0$ όμως δεν υφίσταται αρνητική πιθανότητα, συνεπώς $P(A \cap supp') = 0$. Άρα $P(A) = P(A \cap supp)$ (QED)

Για $supp = \{x_1, x_2, \dots, x_n\}$, $P(A) = P(A \cap supp) = P(A \cap \{x_1, x_2, \dots, x_n\})$

$$= P((A \cap \{x_1\}) \cup (A \cap \{x_2\}) \cup \dots)$$

ξένα ενδεχόμενα $\quad = P(A \cap \{x_1\}) + \dots + P(A \cap \{x_n\}) + \dots$

Ανάλογα με τη μορφή που μπορεί να πάρει, το στήριγμα μας παρέχει την παρακάτω κατηγοριοποίηση των κατανομών στο $\mathbb{R}$

α. Η P ονομάζεται διακριτή αν το στήριγμά της είναι διακριτό υποσύνολο του $\mathbb{R}$ (π.χ. πεπερασμένο, $\mathbb{N}$, $\mathbb{Z}$, κ.α.)

β. Η P θα ονομάζεται συνεχής αν το στήριγμά της είναι διάστημα.

γ. Η P θα ονομάζεται μικτή σε κάθε άλλη περίπτωση

Διακριτές Κατανομές

Διακριτή ονομάζεται όποια κατανομή πιθανότητας στο $\mathbb{R}$ έχει διακριτό στήριγμα (π.χ. Bernoulli, Binomial, Poisson)

Ισχυριζόμαστε ότι αυτές τις κατανομές μπορούμε να τις περιγράψουμε "εύκολα" χρησιμοποιώντας μόνο τον ορισμό, χωρίς να χρειάζονται καινούριες έννοιες. Αυτό συμβαίνει διότι εφόσον είναι διακριτή θα έχει και διακριτό στήριγμα, δηλαδή της μορφής $supp = \{x_1, x_2, \dots, x_n, \dots\}$

Ερώτηση

Είναι δυνατόν μια ΔΙΑΚΡΙΤΗ κατανομή να αποδίδει μηδενική πιθανότητα σε κάποιο σημείο του στηρίγματός της;

Απάντηση

όχι, διότι αυτό το $x_i \in supp$, $P(x_i) = 0$ θα μπορούσαμε να το αφαιρέσουμε από το supp. Συνεπώς, το supp δεν θα ήταν το μικρότερο, κλειστό υποσύνολο του $\mathbb{R}$ με $P(supp) = 1$.

Πόρισμα

Μια διακριτή κατανομή αποδίδει αυστηρά θετική πιθανότητα σε κάθε στοιχείο του στηρίγματός της, $supp = \{x_1, \dots, x_n\}$

Απόδειξη

Έστω $P(\{x_1\}) = 0$. Συνεπώς $P(\{x_2, x_3, \dots, x_n\}) = 1$. Το $\{x_2, \dots, x_n\}$ είναι διακριτό και άρα κλειστό, οπότε το $supp = \{x_1, \dots, x_n\}$ δεν είναι το μικρότερο, κλειστό υποσύνολο του $\mathbb{R}$ στο οποίο αποδίδεται μοναδιαία πιθανότητα. ΑΤΟΠΟ (Q.E.D.)

Προηγουμένως, δείξαμε ότι

$$P(A) = P(A \cap supp) \Leftrightarrow P(A) = P(A \cap \{x_1, \dots, x_n\})$$

$$\Leftrightarrow P(A) = P[(A \cap \{x_1\}) \cup (A \cap \{x_2\}) \cup \dots \cup (A \cap \{x_n\})]$$

και επειδή τα ενδεχόμενα είναι μεταξύ τους ξένα

$$P(A) = P(A \cap \{x_1\}) + P(A \cap \{x_2\}) + \dots + P(A \cap \{x_n\})$$

Υπάρχουν δύο δυνατές περιπτώσεις:

- $x_i \in A \Rightarrow A \cap \{x_i\} = \{x_i\}$
- $x_i \notin A \Rightarrow A \cap \{x_i\} = \emptyset$

$$, \quad A \cap \{x_i\} = \begin{cases} \emptyset, & \text{αν } x_i \notin A \\ \{x_i\}, & \text{αν } x_i \in A \end{cases}$$

όπου x_i είναι οποιοδήποτε στοιχείο του supp

Συνεπώς, $P(A \cap \{x_i\}) = \begin{cases} P(\emptyset) = 0, & \text{αν } x_i \notin A \\ P(\{x_i\}) > 0, & \text{αν } x_i \in A \end{cases}$

Συνοψίζοντας, για να περιγράψουμε μια διακριτή κατανομή στους πραγματικούς θα πρέπει

1. Να ορίσουμε το στήριγμά της
2. Να ορίσουμε τι πιθανότητα αποδίδει η κάθε κατανομή σε κάθε στοιχείο του στηρίγματος.

Αν γνωρίζουμε τα παραπάνω μπορούμε να υπολογίσουμε την πιθανότητα που αποδίδει η εκάστοτε κατανομή σε κάθε ενδεχόμενο A της συλλογής μετρήσιμων υποσυνόλων Σ του Ω.

Αντίστροφα, αν έχουμε τα 1. και 2. για να επιβεβαιώσουμε ότι η κατανομή είναι διακριτή στο $\mathbb{R}$ θα πρέπει:

1. Το στήριγμά της να είναι διακριτό
2. $\forall x_i \in supp$ θα πρέπει να ισχύει $P(\{x_i\}) > 0$
3. $P(supp) = 1$

Αν ισχύουν τα 1., 2., 3., τότε η κατανομή που μας έχει δοθεί είναι μια καλώς ορισμένη κατανομή στο $\mathbb{R}$.

Παραδείγματα

1. Εκφυλισμένη κατανομή στο $\mathbb{R}$

Δίνονται : $supp = \{1\}$

$P(\{1\}) = 1$

Ορίζουν τα δεδομένα μια διακριτή κατανομή στο $\mathbb{R}$;

Απάντηση

1. Το στήριγμα είναι διακριτό ως μονοσύνολο
2. Η πιθανότητα σε κάθε στοιχείο του supp είναι αυστηρά θετική

$P(\{1\}) > 0$

3. $P(supp) = 1$

Συνεπώς, η δοθείσα κατανομή είναι καλώς ορισμένη στο $\mathbb{R}$

Μπορούμε να υπολογίσουμε την πιθανότητα $P((0, 1/2))$;

Ναι, εφόσον $(0, 1/2) \in \Sigma_{\mathbb{R}}$

$$P((0, 1/2)) = P\big((0, 1/2) \cap supp\big) = P\big((0, 1/2) \cap \{1\}\big) = P(\emptyset) = 0$$

2. Bernoulli

Εστω $q \in (0,1)$. Το στήριγμα της Bernoulli είναι $supp = \{0,1\}$

$$Ber(q) : \begin{cases} P(\{0\}) = 1-q \\ P(\{1\}) = q \end{cases}$$

Είναι μια καλώς ορισμένη διακριτή κατανομή στο $\mathbb{R}$?

Απάντηση

1. $supp = \{0, 1\}$ διακριτο
2. $P(\{0\}) > 0$, $P(\{1\}) > 0$ διότι $q \in (0,1)$
3. $P(supp) = P(\{0,1\}) = P(\{0\}) + P(\{1\}) = 1$

Επίσης, $P((0, 1/2)) = P\big((0, 1/2) \cap supp\big) = P\big((0, 1/2) \cap \{0,1\}\big) = P(\emptyset) = 0$

και

$$P([0, 1/2)) = P\big([0, 1/2) \cap supp\big) = P\big([0, 1/2) \cap supp\big) = P(\{0\}) = 1-q$$

3. Διωνυμική Κατανομή (Binomial)

Στο supp $= \{0, \dots, n\}$ ως προς $q \in (0,1)$

$$P(\{i\}) = \binom{n}{i} q^i (1-q)^{n-i}, \text{ όπου } i \in \{0,1,2,\dots,n\}$$

και $\binom{n}{i} = \dfrac{n!}{i!(n-i)!} \in \mathbb{Q}$: συνδυασμοί n ανά i

n: μας δίνει το μέγεθος του στηρίγματος

q: —"— έναν τρόπο να υπολογίζουμε πιθανότητες

Είναι η παραπάνω μια καλώς ορισμένη κατανομή στο $\mathbb{R}$;

Απάντηση

1. Προφανώς το στήριγμα είναι διακριτό (supp $= \{0, \dots, n\}$)
2. Δεν ορίζεται παραγοντικό σε αρνητικούς, δηλαδή $n! > 0$.
   Επίσης, $0! = 1$. Συνεπώς, $\binom{n}{i} > 0$.
   Επιπλέον, $q \in (0,1)$, δηλαδή $q \neq 0$ κ' $q \neq 1$. Άρα, $q^i(1-q)^{n-i} > 0$
   Τελικά, $P(\{i\}) > 0$, $\forall i \in$ supp
3. $P(\text{supp}) = P(\{0,1,\dots,n\}) = P(\{0\} \cup \{1\} \cup \dots \cup \{n\}) = P(\{0\}) + \dots + P(\{n\})$
   $$= \sum_{i=0}^{n} P(\{i\}) = \sum_{i=0}^{n} \binom{n}{i} q^i (1-q)^{n-i}$$
   Αρκεί να αποδείξουμε ότι το άθροισμα ισούται με ένα.

Διωνυμικό Ανάπτυγμα

$$(a+b)^n = \sum_{i=0}^{n} \binom{n}{i} a^i b^{n-i} \quad a, b \in \mathbb{R},\ n \in \mathbb{N}$$

π.χ. $n=2$: $\displaystyle\sum_{i=0}^{2} \binom{2}{i} a^i \cdot b^{2-i} = \binom{2}{0} a^0 \cdot b^2 + \binom{2}{1} a^1 \cdot b^1 + \binom{2}{2} a^2 \cdot b^0$

$$= \frac{2!}{0!(2-0)!} a^0 \cdot b^2 + \frac{2!}{1!(2-1)!} a \cdot b + \frac{2!}{2!(2-2)!} a^2 \cdot b^0 = a^2 + 2ab + b^2$$
$$= (a+b)^2$$

Συνεπώς, από διωνυμικό ανάπτυγμα έχουμε

$$P(\text{supp}) = \sum_{i=0}^{n} P(\{i\}) = \sum_{i=0}^{n} \binom{n}{i} q^i (1-q)^{n-i} = (q+1-q)^n = 1$$

Συνεπώς, η κατανομή είναι καλώς ορισμένη στο $\mathbb{R}$

Σημείωση: Η διωνυμική γενικεύει την Bernoulli.
Να ελεγχθεί ο ισχυρισμός.

4. Κατανομή Poisson ως προς παράμετρο $\lambda>0$

$supp = \mathbb{N} = \{0,1,\dots,\dots\}$, $P(\{i\}) = \frac{\lambda^i}{i!} e^{-\lambda}$, $i\in\mathbb{N}$, $\lambda>0$

Είναι η παραπάνω κατανομή μια καλώς ορισμένη κατανομή στο $\mathbb{R}$;

Απάντηση

1. Το στήριγμα είναι πράγματι διακριτό
2. Εφόσον ο παρονομαστής είναι διάφορος του μηδέν ($i!\neq 0$, δεν απειρίζεται) τότε $P(\{i\})>0$
3. $P(supp) = P(\{0,1,2,\dots\}) = P(\{0\}\cup\{1\}\cup\dots)$

$$= \underbrace{P(\{0\}) + P(\{1\}) + \dots}$$

άπειρο άθροισμα που το διαχειριζόμαστε σαν πεπερασμένο, π.χ. βγάζουμε κοινό παράγοντα ονομάζεται σειρά και προκύπτει από τα κατάλληλα όρια των πεπερασμένων αθροισμάτων

$$= \sum_{i=0}^{\infty} P(\{i\}) = \sum_{i=0}^{\infty} \frac{\lambda^i}{i!} e^{-\lambda} = e^{-\lambda} \sum_{i=0}^{\infty} \frac{\lambda^i}{i!}$$

Αρκεί να δείξουμε ότι $\sum_{i=0}^{\infty} \frac{\lambda^i}{i!} = e^{\lambda}$

Ανάπτυγμα Maclaurin

$$e^x = \sum_{i=0}^{\infty} \frac{x^i}{i!} \quad \forall x\in\mathbb{R}$$

Ισχύει για κάθε $x\in\mathbb{R}$, συνεπώς και για $x=\lambda$

$$\Rightarrow \sum_{i=0}^{\infty} \frac{\lambda^i}{i!} = e^{\lambda}$$

Συνεπώς, από ανάπτυγμα Maclaurin έχουμε

$$P(supp) = \sum_{i=0}^{\infty} P(\{i\}) = \sum_{i=0}^{\infty} \frac{\lambda^i}{i!} e^{-\lambda} = e^{-\lambda} \cdot \sum_{i=0}^{\infty} \frac{\lambda^i}{i!} = e^{-\lambda} \cdot e^{\lambda} = 1$$

Άρα η κατανομή είναι καλώς ορισμένη στο $\mathbb{R}$