Φροντιστήριο 2 19/03/2018

Η έννοια του στηρίγματος

Για τα παρακάτω θεωρούμε χωρίς μεγάλη ακρίβεια ότι κάποιο υποσύνολο του $\mathbb{R}$ θεωρείται κλειστό 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 αυθαίρετου πλήθους, τομές αυτών κλπ.

Αναλόγως όταν $A, B \subseteq \mathbb{R}$ τότε το A θα ονομάζεται μικρότερο του B αν $A \subset B$ (το A γνήσιο υποσύνολο του B)

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

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

Απόδειξη: $A = (A \cap supp) \cup (A \cap supp')$

$$P(A) = P\big((A \cap supp) \cup (A \cap supp')\big)$$

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

Ισχύει $supp' \supseteq A \cap supp' \xRightarrow{\text{ΜΟΝΟΤΟΝΙΑ}} \underbrace{P(supp')}_{0} \geq P(A \cap supp')$

Δεν μπορεί να είναι αρνητική $\Rightarrow P(A \cap supp') = 0$

$$\Rightarrow P(A) = P(A \cap supp)$$

Για $supp = \{x_1, x_2, \dots, x_n, \dots\}$ [ΔΙΑΚΡΙΤΟ]

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

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

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

Κλειστά σύνολα

όλα τα διακριτά υποσύνολα του $\mathbb{R}$, τα κλειστά διαστήματα, ενώσεις κλειστών διαστημάτων κ.α.
π.χ. φυσικοί, ρητοί, ακέραιοι, ακέραιοι, $\{0,1\}$, $[0,1]$

Διακριτά Σύνολα

οποιοδήποτε υποσύνολο του $\mathbb{R}$ αποτελείται από απομονωμένους αριθμούς (δηλ. από αριθμούς που δεν αποτελούν διάστημα)
π.χ. φυσικοί $\mathbb{N}$, Ακέραιοι $\mathbb{Z}$, ρητοί $\mathbb{Q}$, $\{0,1,2\}$, κ.α.

Μικρότερο Σύνολο

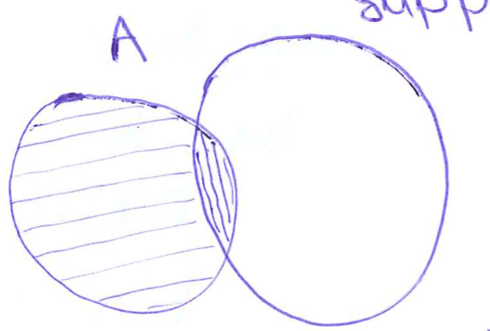
Αν αρχίσω και αφαιρώ στοιχεία του

Χρησιμότητα Στηρίγματος

Η έννοια του στηρίγματος είναι χρήσιμη επειδή διευκολύνει τη διαδικασία περιγραφής μιας κατανομής πιθανότητας στο $\mathbb{R}$, αφού

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

Απόδειξη από πίσω



supp' = συμπλήρωμα supp

Η έννοια του στηρίγματος ανάλογα με τη "μορφή" που αυτό μπορεί να πάρει 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

ορισμός : Διακριτή ονομάζεται όποια κατανομή πιθανότητας στο $\mathbb{R}$ της οποίας το supp είναι διακριτό.

π.χ. Bernoulli, Binomial, Poisson, κ.α.

Περιγραφή Διακριτών στους Πραγματικούς

Ισχυρισμός : Αυτές τις κατανομές μπορούμε να τις περιγράψουμε "εύκολα" χρησιμοποιώντας μόνο τον ορισμό, χωρίς να χρειάζονται καινούριες έννοιες.

Γιατί? : Αφού είναι διακριτή ξέρουμε ότι το στήριγμά της θα είναι διακριτό, δηλαδή θα έχει τη μορφή

$supp = \{x_1, x_2, \ldots, x_n, \ldots\}$

Ερώτηση : Είναι δυνατόν μια ΔΙΑΚΡΙΤΗ κατανομή να αποδίδει σε κάποιο σημείο του στηρίγματός της μηδενική πιθανότητα?

Απάντηση: όχι, γιατί αυτό το $x_i \in supp$, $P(x_i) = 0$ θα μπορούσα να το αφαιρέσω από το supp. Άρα το supp δε θα ήταν το μικρότερο, κλειστό υποσύνολο με $P = 1$

Πόρισμα 1

Μια διακριτή κατανομή αποδίδει αυστηρά θετική πιθανότητα σε κάθε στοιχείο του στηρίγματός της

Έστω $supp = \{x_1, \dots, x_n\}$

Απόδειξη:

Έστω $P(\{x_1\}) = 0 \iff P(\{x_2, x_3, \dots, x_n\}) = 1$

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

Προσοχή! Το παραπάνω ισχύει μόνο για τις διακριτές κατανομές

Όπως δείξαμε ισχύουν τα παρακάτω για μια διακριτή κατανομή:

i. $P(A) = P(A \cap supp)$

ii. $P(A) = P(A \cap \{x_1, x_2, \dots, x_n, \dots\})$

iii. $P(A) = P(\underbrace{(A \cap \{x_1\}) \cup (A \cap \{x_2\}) \cup \dots \cup (A \cap \{x_n\}) \cup \dots}_{\text{ΞΕΝΑ ΜΕΤΑΞΥ ΤΟΥΣ}})$

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

Υπάρχουν δύο δυνατές περιπτώσεις:

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

Έχουμε δηλαδή ότι,

$$A \cap \{x_i\} = \begin{cases} \emptyset, & \text{αν } x_i \notin A \\ \{x_i\}, & \text{αν } x_i \in A \end{cases}$$ όπου x_i είναι οποιοδήποτε στοιχείο του supp

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

Άρα, για να περιγράψω μια διακριτή στους πραγματικούς πρέπει:

1. Να ορίσω το στήριγμά της
2. Να ορίσω τι πιθανότητα παίρνει το κάθε στοιχείο του supp (τι πιθανότητα αποδίδει η κάθε κατανομή σε κάθε στοιχείο του supp)

Αν έχω τα παραπάνω μπορώ να βρω τι πιθανότητα αποδίδεται σε $\forall A \in \Sigma_{\mathbb{R}}$

<u>Αντίστροφα</u>

Αν έχω τα 1., 2., για να επιβεβαιώσω αν είναι είναι διακριτή μία κατανομή στο $\mathbb{R}$, ελέγχω τα παρακάτω:

1. Αν το supp είναι διακριτό
2. Αν για $\forall x_i \in$ supp, ισχύει $P(\{x_i\}) > 0$
3. Αν $P(\text{supp}) = 1$. Αν ισχύουν τα 1,2,3 τότε αυτό που μας έχει δοθεί είναι μια καλώς ορισμένη ~~συνάρτη~~ διακριτή κατανομή στο $\mathbb{R}$.

Παραδείγματα

Σε κάθε ένα από τα παρακάτω παραδείγματα θα δίνονται το supp και η πιθανότητα που αποδίδεται σε κάθε στοιχείο του supp και εμείς θα ελέγχουμε αν:

1. το supp είναι διακριτό
2. η πιθανότητα κάθε στοιχείου του supp > 0
3. $P(supp) = 1$

Αν ισχύουν, τότε αυτό που μας έχει δοθεί θα ορίζει μια καλώς ορισμένη διακριτή κατανομή στο $\mathbb{R}$.

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

Δίνονται: $supp = \{1\}$

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

ορίζει μια διακριτή κατανομή στο $\mathbb{R}$;

1. μονοσύνολο ✓
2. $P(\{1\}) > 0$ ✓
3. $P(\{1\}) = 1$ ✓

Άρα ναι

Αν έχοντας τα παραπάνω μου ζητηθεί να υπολογίσω την $P((0, 1/2))$ έχει νόημα να τη βρω αφού $(0, 1/2) \in \Sigma_{\mathbb{R}}$

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

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

2. Bernoulli

Έστω $q \in (0,1)$. Η κατανομή Bernoulli στο $\{0,1\}$ ως προς q ορίζεται ως:

1. $supp = \{0,1\}$ <= στήριγμα ✓

2. $P(\{0\}) = 1-q > 0$, $P(\{1\}) = q > 0$ πιθανότητα που αποδίδεται σε κάθε σημείο του στηρίγματος ✓

3. $P(supp) = P(\{0,1\}) = P(\{0\}) + P(\{1\}) = 1$ ✓

$P((0,\frac{1}{2})) =$

Εδώ $P((0,\frac{1}{2}) \cap supp) = P((0,\frac{1}{2}) \cap (\{0,1\})) = P(\varnothing) = 0$

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

3. Διωνυμική Κατανομή (Binomial)

Στο $\{0,1,\ldots,n\}$ ως προς $q \in (0,1)$

Έστω φυσικός > 0, και $q \in (0,1)$

"Συνδυασμοί n ανά i" $\binom{n}{i} = \frac{n!}{i!(n-i)!} \in \mathbb{Q}$

- ο n μας δίνει πόσο μεγάλο είναι το στήριγμα
- ο q μας —"— έναν τρόπο να υπολογίζουμε πιθανότητες (θα δίνεται)

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

Ερώτηση: Είναι καλώς ορισμένη η κατανομή?

Απάντηση: Πρέπει να ελέγξω 3 πράγματα

1. Προφανώς το supp είναι διακριτό
2. Δεν ορίζεται παραγοντικό σε αρνητικούς.
   Ακόμη $0! = 1$. Άρα $n > i \Rightarrow \binom{n}{i} > 0$. (Αυστηρά θετικός)
   Επίσης $q \in (0,1)$, δηλαδή $q \neq 0$ κ' $q \neq 1$.
   Συνεπώς $q^i (1-q)^{n-i} > 0$ Αυστηρά θετικό
   Άρα η πιθανότητα για $\forall i \in supp$ είναι αυστηρά θετική $\Rightarrow P(\{i\}) > 0$, για $\forall i \in supp$
3. Ελέγχω αν $P(supp) = 1$

$$P(supp) = P(\{0,1,\ldots,n\}) = P(\{0\} \cup \{1\} \cup \ldots \cup \{n\}) \overset{\text{προσθετικότητα}}{=\!=}$$

$$= P(\{0\}) + P(\{1\}) + \ldots + P(\{n\})$$

αθροίζω ως προς όλα τα στοιχεία του στηρίγματος

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

Αρκεί ν.δ.ο. είναι ίσο με 1.

Διωνυμικό Ανάπτυγμα (θα δίνεται)

το διωνυμικό ανάπτυγμα $a, b \in \mathbb{R}$, $n \in \mathbb{N}$

$$(a+b)^n = \sum_{i=0}^{n} \binom{n}{i} a^i b^{n-i}$$

π.χ.

για $n=2$ $(a+b)^2 = \sum_{i=0}^{2} \binom{2}{i} a^i \cdot b^{2-i}$

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

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

συνέχεια παραδείγματος διωνυμικής κατανομής

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

$$= 1^n = 1$$

$$\Rightarrow P(supp) = 1$$

Άρα η κατανομή είναι <u>καλώς</u> ορισμένη

Η διωνυμική γενικεύει την Bernoulli (το supp αντί για n+1 στοιχεία, έχει μόνο 2)

<u>Άσκηση</u>: Για n=1 ο τύπος της Binomial μας δίνει τον τύπο της Bernoulli. Ελέγξτε το

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

$supp = \mathbb{N} = \{0, 1, 2, \ldots\}$ (το στήριγμα αυτής της κατανομής είναι το σύνολο των φυσικών)

$P(\{i\}) = \frac{\lambda^i}{i!} e^{-\lambda}, \; i \in \mathbb{N}$ αυτή είναι η πιθανότητα που θα δίνεται σε κάθε i

$\lambda > 0$

Ελέγχω αν είναι καλώς ορισμένη

1. supp διακριτό
2. $i! \neq 0$, άρα καλώς ορισμένο (δεν απειρίζεται)
   $P(\{i\}) > 0$ (αυστηρά θετικό)
3. $P(supp) = 1$;

Η παράμετρος λ καθορίζει για ποια ακριβώς κατανομή Poisson μιλάμε.

Έχουμε $P(supp) = P(\{0,1,2,\dots,\}) = P(\{0\}\cup\{1\}\cup\dots)$

$$\overset{\text{προσθετικότητα}}{=} \underbrace{P(\{0\}) + P(\{1\}) + P(\{2\}) + \dots}$$

άπειρο άθροισμα. το διαχειρίζομαι όμως σαν πεπερασμένο, π.χ. βγάζω κοινό παράγοντα ονομάζεται σειρά και προκύπτει από τα κατάλληλα όρια των πεπερασμένων αθροισμάτων.

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

* Θα δίνεται ότι είναι καλώς ορισμένα κ' ότι μπορώ να τα διαχειρίζομαι ως πεπερασμένα

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

Αρκεί να δείξω ότι $\sum_{i=0}^{\infty} \frac{\lambda^i}{i!} = e^{\lambda}$ ώστε $P(supp)=1$

Δίνεται: Ανάπτυγμα MacLaurin της e^x

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

Το ανάπτυγμα ισχύει για κάθε x άρα για $x=\lambda$

$$\overset{\text{MacLaurin}}{\Longrightarrow} e^{\lambda} = \sum_{i=0}^{\infty} \frac{\lambda^i}{i!}$$

Αντικαθιστώ $P(supp) = e^{-\lambda} \cdot e^{\lambda} = 1 \Rightarrow P(supp) = 1$

- Οι τύποι των κατανομών έχουν προκύψει από άλλες γνωστές κατανομές, παίρνοντας κατάλληλα όρια
- όλα τα παραπάνω αποτελούν παραδείγματα καλώς ορισμ. διακριτών κατανομών στο $\mathbb{R}$
- μπορώ να υπολογίσω χρησιμοποιώντας τις παραπάνω κατανομές την πιθανότητα που αποδίδεται σε οποιοδήποτε διάστημα των πραγματικών $\subseteq \mathbb{R}$

Υπολογισμοί

π.χ. Έστω η Poisson. Ποια είναι η $P((-2,-1)\cup(1,3/2))$?

Έχει νόημα να αναρωτηθώ ποια είναι, εφόσον γνωρίζω ότι $(-2,-1)\cup(1,3/2)\in\Sigma_R$

* Αν δεν μπορώ να βρω την P, σημαίνει ότι η περιγραφή της κατανομής που μας έχει δωθεί μέσω των (1), (2) δεν είναι πλήρης.

Έχουμε: $P((-2,-1)\cup(1,3/2)) = P\big(((-2,-1)\cup(1,3/2))\cap supp\big)$

$supp = \{1,2,\dots,\}$

$$\Leftrightarrow P\big(\underbrace{((-2,-1)\cup(1,3/2))}_{\text{θα πρέπει να υπάρχουν φυσικοί}}\cap\{1,2,\dots\}\big) = P(\emptyset) = 0$$

Συνεπώς η κατανομή Poisson αποδίδει μηδενική πιθανότητα σε αυτό το διάστημα.

π.χ. Ποια είναι η $P((-2,-1)\cup[1,3/2))$?

$$P((-2,-1)\cup[1,3/2)) = P(\{1\}) = \frac{\lambda^1}{1!}e^{-\lambda} = \lambda\cdot e^{-\lambda}$$

και αν ξέραμε ποια ακριβώς Poisson έχουμε π.χ. την $\lambda=\frac{1}{2}$, βρίσκαμε ακριβώς την P.

$$P(\{1\}) = \frac{1}{2}\cdot e^{-\frac{1}{2}}$$

π.χ. Ποια είναι η $P((-2,-1)\cup[1,3])$;

$$P((-2,-1)\cup[1,3]) = P(\{1,2,3\}) \underset{\text{προσθετικότητα}}{\overset{= P(\{1\}\cup\{2\}\cup\{3\})}{}}$$

$$P(\{1\}) + P(\{2\}) + P(\{3\})$$

$$= \lambda \cdot e^{-\lambda} + \frac{\lambda^2}{2} \cdot e^{-\lambda} + \frac{\lambda^3}{3} \cdot e^{-\lambda} = \lambda \cdot e^{-\lambda}\left(1 + \frac{\lambda}{2} + \frac{\lambda^3}{6}\right)$$

<u>Ερώτηση</u>: Πόσες κατανομές Poisson υπάρχουν?

<u>Απάντηση</u>: Όσες είναι οι δυνατές τιμές που μπορεί να πάρει το $\lambda \in (0, \infty)$. Άρα αυτό που είδαμε είναι μια οικογένεια κατανομών με παράμετρο λ.

Στο στατιστικό πρόβλημα: αν γνωρίζω σε ποια οικογένεια ανήκει η άγνωστη κατανομή, π.χ. Poisson τότε πρέπει να βρω μόνο την παράμετρο (λ). Συνεπώς, το πρόβλημα εύρεσης της άγνωστης κατανομής ανάγεται σε πρόβλημα εύρεσης άγνωστης παραμέτρου.