

Ειδικές κατανομές πιθανότητας

Καρακωνσταντής Κωνσταντίνος (ΑΜ 9046)

Καναβάκης Γεώργιος (ΑΜ 9044)

Εφαρμογές 1

Στατιστική & ανάλυση δεδομένων

Περιγραφή δεδομένων

Χρησιμοποιούνται για να περιγράψουν και να αναλύσουν τα δεδομένα που συλλέγονται από πειράματα, έρευνες ή παρατηρήσεις

Αναγνώριση προτύπων

Εφαρμόζεται στην αναγνώριση προτύπων (pattern recognition) και στη μηχανική μάθηση, για να βοηθήσει στον προσδιορισμό αν τα δεδομένα ταιριάζουν σε μια συγκεκριμένη κατανομή

Εφαρμογές 2

Μηχανική & φυσική / Οικονομία & χρηματοοικονομικά

- Θερμοδυναμική και στατιστική φυσική
- Μηχανική σφαλμάτων
- Ελαστικότητα και ροή ρευστών

- Ανάλυση κινδύνου
- Θεωρία παιγνίων

Εφαρμογές 3

Βιοστατιστική & ιατρική / Ασφάλεια & βιομηχανία

- Εκτίμηση κινδύνου και υγείας
- Δοκιμές και έρευνες

- Αξιολόγηση κινδύνου
- Μοντελοποίηση αποτυχίας συστημάτων

Εφαρμογές 4

Μοντελοποίηση & προσομοιώσεις / Μηχανική μάθηση & τεχνητή νοημοσύνη

- Μοντελοποίηση τυχαίων φαινομένων
- Σενάρια τύχης και αβεβαιότητας
- Κατηγοριοποίηση και ανάλυση
- Μοντελοποίηση αστάθειας και αβεβαιότητας

Εφαρμογές 5

Αστρονομία & κοσμολογία / Κοινωνικές επιστήμες

- Κατανομή αστεριών και γαλαξιών
- Αναλύσεις παρατηρήσεων

- Μοντελοποίηση συμπεριφοράς
- Ανάλυση δημογραφικών δεδομένων

Εφαρμογές 6

Μάρκετινγκ & επιχειρήσεις

- Προβλέψεις πωλήσεων
- Διαχείριση αποθεμάτων

Ειδικές διακριτές κατανομές

- Διωνυμική κατανομή ή κατανομή Bernoulli
- Υπεργεωμετρική κατανομή
- Αρνητική διωνυμική κατανομή ή κατανομή Pascal
- Κατανομή Poisson

Διωνυμικό πείραμα

Ένα πείραμα αποτελούμενο από δοκιμές, η κάθε μία από τις οποίες έχει μόνο δυο δυνατά αποτελέσματα E (επιτυχία) ή A (αποτυχία)

Π.χ. Επιτυχία είναι, η υποστήριξη από ένα άτομο της θέσης του πολιτικού κόμματος και αποτυχία το αντίθετο

Π.χ. Κεφαλή (K) ή γράμματα (Γ)

Προϋποθέσεις

1. Η πιθανότητα επιτυχίας p ή αποτυχίας q παραμένει σταθερή σε όλες τις δοκιμές του πειράματος

$$0 \leq p \leq 1$$

2. Οι δοκιμές του πειράματος είναι στατιστικά ανεξάρτητες

Διωνυμική κατανομή

Έστω ένα διωνυμικό πείραμα.

X είναι η τυχαία μεταβλητή της οποίας οι τιμές x παριστάνουν τον αριθμό των επιτυχιών που θα σημειωθούν στην ακολουθία των n δοκιμών.

$$p_x(x) = \binom{n}{x} p^x p^{n-x}, \quad x = 0, 1, 2, \dots, n$$

Εφαρμογές

1. Στις διαδοχικές ρίψεις ενός νομίσματος
2. Στο βιομηχανικό έλεγχο ποιότητας, παράγονται συνεχώς είδη (π.χ. βίδες, λαμπτήρες, κ.λπ.) και κατά τον έλεγχο ποιότητας ταξινομούνται ως Ε, ή ως Α
3. Στις δειγματοληψίες

Υπεργεωμετρική κατανομή

Εμφανίζεται σε δειγματοληψία χωρίς επανατοποθέτηση από πεπερασμένο πληθυσμό N στοιχείων κάθε ένα από τα οποία είναι δυνατό να ανήκει σε μια από δυο κατηγορίες C και C' . Έστω Np ο αριθμός των στοιχείων που ανήκουν στην κατηγορία C και Nq ο αριθμός εκείνων που ανήκουν στη κατηγορία C' , με ποσοστά p , q .

$$p + q = 1$$

$$P\{X = x\} = \frac{\binom{Np}{x} \binom{Nq}{n-x}}{\binom{N}{n}}$$

Θέλω να εκτιμήσω τον αριθμό N των ψαριών μίας λίμνης.

Ψαρεύω Np ψάρια, τα σημαδεύω με μια κουκκίδα και τα αφήνω ελεύθερα.

Την άλλη ημέρα ψαρεύω n ψάρια και μετράω τον αριθμό των x ψαριών που έχουν την κουκκίδα.

Το υπεργεωμετρικό πρότυπο είναι τώρα εφαρμόσιμο και μπορώ να εκτιμήσω το N .

Γεωμετρική κατανομή

- Είναι η κατανομή του αριθμού των «δοκιμών αναμονής» ή «χρόνου αναμονής» μέχρι να εμφανιστεί η 1^η επιτυχία σε ένα διωνυμικό πείραμα.

- Η συνάρτηση πιθανότητας είναι

$$p_x(x) = pq^{x-1}, \quad x = 1, 2, 3, \dots$$

- Το **p** είναι η πιθανότητα επιτυχίας των δοκιμών Bernoulli και αποτελεί την παράμετρο της κατανομής.

- Στο παιχνίδι κορώνα–γράμματα, ενδιαφέρομαι για το χρόνο αναμονής μέχρι να εμφανιστεί η 1^η κορώνα.

- Γιατρός ενδιαφέρεται για την πιθανότητα ο 5^{ος} άρρωστος που θα τον επισκεφτεί μια ημέρα, να είναι ο 1^{ος} που πάσχει από κάποιο ιό γρίπης.

- Βιομήχανος ενδιαφέρεται για την πιθανότητα το 8^ο υπό κατασκευή αντικείμενο να είναι το 1^ο ελαττωματικό.

Αρνητική διωνυμική κατανομή ή κατανομή Pascal

- Αποτελεί γενίκευση της γεωμετρικής κατανομής
- Ενδιαφέρομαι για τον αριθμό X των δοκιμών (χρόνος αναμονής) μέχρι να εμφανιστεί η k επιτυχία
- Ισοδύναμος ορισμός: έστω Y ο αριθμός των αποτυχιών που χρειάζονται σε μια ακολουθία δοκιμών Bernoulli μέχρι να εμφανιστεί η k επιτυχία

- Έστω 20% η πιθανότητα ένα νήπιο εκτιθέμενο σε μια μεταδοτική νόσο να προσβληθεί από αυτή. Ποια η πιθανότητα το 12^ο νήπιο που εκτίθεται στη νόσο να είναι το 3^ο που θα προσβληθεί από αυτή;

$$\binom{x-1}{k-1} p^{k-1} q^{x-1-(k-1)} p$$

Κατανομή Poisson 1/2

- Έστω ραδιενεργή ουσία που με τη πάροδο του χρόνου εκπέμπει σωματίδια.
- Οι συνθήκες του πειράματος παραμένουν σταθερές με την πάροδο του χρόνου.
- Υποδιαιρώ το μοναδιαίο χρονικό διάστημα σε ένα μεγάλο αριθμό n μικρών ίσων υποδιαστημάτων.
- Έστω p η πιθανότητα εκπομπής, σταθερή για κάθε υποδιάστημα.

- Όσο μικρότερο είναι το υποδιάστημα, τόσο μικρότερο είναι το p .
- Η εκπομπή μοιάζει με μια σειρά δοκιμών Bernoulli με επιτυχία E την εκπομπή ενός σωματιδίου και σταθερό p .

$$\binom{n}{x} p^x q^{n-x}$$

Συχνά, η κατανομή Poisson χρησιμοποιείται για τον κατά προσέγγιση υπολογισμό των πιθανοτήτων της διωνυμικής κατανομής

Πιθανότητες					
X	Διωνυμικές	Poisson	X	Διωνυμικές	Poisson
0	0,1216	0,1353	7	0,0020	0,0034
1	0,2702	0,2707	8	0,0004	0,0009
2	0,2852	0,2707	9	0,0001	0,0002
3	0,1901	0,1804	10	0,0000	0,0000
4	0,0898	0,0902	11	0,0000	0,0000
5	0,0319	0,0361	12	0,0000	0,0000
6	0,0089	0,0120			

Κατανομή Poisson 2/2

- Είναι μια από τις βασικότερες στη θεωρία πιθανοτήτων και στη στατιστική, με τεράστιο εύρος εφαρμογών.
- Είναι απαραίτητο οι συνθήκες του πειράματος ή του φαινομένου να παραμένουν σταθερές με την πάροδο του χρόνου.
- Ακολουθούν ενδεικτικές εφαρμογές:

- Ραδιενεργές διασπάσεις
- Τροχαία ατυχήματα
- Πλήθος τηλεφωνημάτων που δέχεται ένα τηλεφωνικό κέντρο σε χρόνο t .
- Πλήθος αυτοκινήτων που διέρχονται από τα διόδια.
- Πλήθος πελατών που προσέρχονται σε super market.

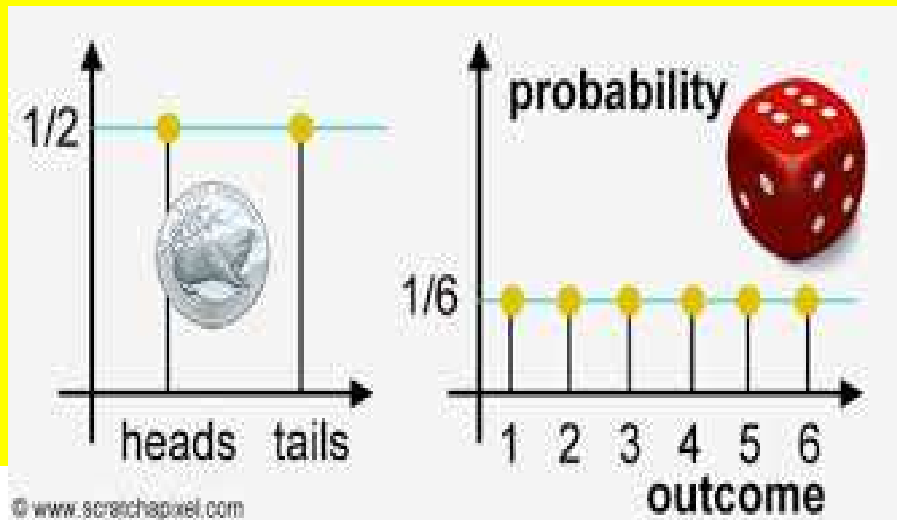
Ειδικές συνεχείς κατανομές

- Ομοιόμορφη ή ορθογώνια κατανομή
- Εκθετική κατανομή
- Κατανομή Γ
- Κατανομή B
- Κανονική κατανομή ή κατανομή Gauss
- Κατανομή Cauchy
- Κατανομή Laplace ή διπλή εκθετική κατανομή
- Λογαριθμοκανονική (long-normal) κατανομή
- Κατανομή Weibull
- Κατανομή Maxwell
- Κατανομή ακροτάτων τιμών ή του Gumbel
- Κατανομή Pareto

Ομοιόμορφη κατανομή /

Εκθετική κατανομή

Είναι η απλούστερη από τις συνεχείς κατανομές και εκφράζει το ισοπίθανο (ομοιόμορφο) των αποτελεσμάτων ενός πειράματος.



- Είναι η μόνη από τις συνεχείς κατανομές με την ιδιότητα της αμνησίας. Η μόνη από τις διακριτές κατανομές που έχει την ίδια ιδιότητα είναι η γεωμετρική κατανομή.
- Οι αφίξεις λεωφορείων σε μία στάση ακολουθούν τη διαδικασία Poisson με $\lambda=6$ αφίξεις ανά ώρα ή 0,1 αφίξεις ανά λεπτό. Ποια η πιθανότητα κάποιος να περιμένει στη στάση λιγότερο από επιπλέον 5 λεπτά, δοθέντος ότι έχει ήδη περιμένει 15 λεπτά;

Κατανομή Γ /

- Αποτελεί γενίκευση της εκθετικής και χρησιμοποιείται για την περιγραφή διαφόρων φυσικών φαινομένων

Κατανομή B

- Η ομοιόμορφη κατανομή στο διάστημα $(0, 1)$ είναι ειδική περίπτωση της κατανομής B .

Κανονική κατανομή ή κατανομή Gauss

- Αποτελεί τη θεμελιώδη κατανομή πάνω στην οποία στηρίζεται όλο το επιστημονικό οικοδόμημα της στατιστικής και των πιθανοτήτων.
- Ανακαλύφθηκε το 1733 από τον De Moivre, ως το όριο της διωνυμικής κατανομής.
- Το 1774 ο Laplace μελέτησε τις μαθηματικές της ιδιότητες.
- Αποδίδεται στον Gauss, ο οποίος την αναφέρει για πρώτη φορά σε μια εργασία του το 1809.
- Το μήκος, το βάρος, ο όγκος, το ύψος, η επίδοση, οι βαθμοί κλπ. ακολουθούν την κανονική κατανομή. Ο μέσος όρος μεγάλου αριθμού μετρήσεων, έχει κατανομή που μπορεί να προσεγγιστεί από την κανονική κατανομή.
- Η κανονική κατανομή είναι το όριο πολλών κατανομών πιθανότητας.

Κατανομή Cauchy / Κατανομή Laplace ή διπλή εκθετική κατανομή

- Συναντάται κυρίως σε φυσικές επιστήμες και μηχανική.
- μοιάζει με την κανονική κατανομή, αλλά με βαρύτερες ουρές.

- Χρησιμοποιείται σε μηχανολογικά προβλήματα
- Στη μηχανική
- Στην επεξεργασία σήματος & εικόνας
- Στην οικονομία
- Στη μοντελοποίηση δεδομένων

Λογαριθμοκανονική κατανομή /

Κατανομή Weibull

Χρησιμοποιείται σε προβλήματα

- φυσικής,
- βιομηχανίας,
- οικονομικών επιστημών,
- κοινωνιολογίας,
- βιολογίας,
- ανθρωπομετρίας

Περιγράφει ικανοποιητικά τη διάρκεια ζωής διαφόρων αντικειμένων, όπως εξαρτήματα, συσκευές κλπ.

Χρησιμοποιείται ευρέως στη στατιστική ανάλυση αξιοπιστίας, στη θεωρία διάρκειας ζωής και στις φυσικές επιστήμες και τη μηχανική.

Κατανομή Maxwell / Κατανομή ακροτάτων τιμών ή του Gumbel

Εφαρμόζεται σε προβλήματα της φυσικής (ταχύτητα μορίων)

Χρησιμοποιείται ευρέως σε στατιστική, μηχανική αξιοπιστίας και ανάλυση κινδύνου κυρίως για τη μοντελοποίηση ακραίων τιμών

