Διακριτές Κατανομές

Bernoulli

Συμβ.: X~Bernoulli(p)

*Με τον όρο δοκιμή **Bernoulli**, αναφερόμαστε σε ένα πείραμα με δύο δυνατά αποτελέσματα: Επιτυχία (Συμβ. ε) και αποτυχία (Συμβ. α).Ορίζουμε ένα ενδεχόμενο ως επιτυχία με μια πιθανότητα p και αντίστοιχα το μη-ευνοϊκό ενδεχόμενο με πιθανότητα 1-p.*

Υποθέσεις

-Τα πειράματα είναι ανεξάρτητα μεταξύ τους έτσι ώστε το αποτέλεσμα οποιουδήποτε πειράματος να μην επηρρεάζει τα αποτελέσματα των υπολοίπων.

-Σε κάθε επανάληψη του πειράματος μπορούν να εμφανισθούν δύο μόνο δυνατά αποτελέσματα τα οποία θα χαρακτηρίζουμε σαν επιτυχία (ε) ή αποτυχία (α)

-Η πιθανότητα επιτυχίας (και αποτυχίας) δεν μεταβάλλεται από πείραμα σε πείραμα.

[Θυμίζουμε οτι: 1-p=q, 0≤p,q≤1 και p+q=1]!

Ω={α,ε}

$R_X = \{0,1\}$

Ορισμός 1.1.1

*Έστω X ο αριθμός των επιτυχιών σε μια δοκιμή Bernoulli με πιθανότητα επιτυχίας p και πιθανότητα αποτυχίας q. Η κατανομή της δίτιμης [0,1] τυχαίας μεταβλητής X καλείται **κατανομή Bernoulli**.*

$$f(x) = P(X = x) = \begin{cases} p, & x = 1 \\ q = 1 - p, & x = 0 \end{cases}$$

ή ισοδύναμα

$$f(x) = p^x q^{1-x} \quad , x=0,1$$

Μεση Τιμή

$$\mu = \sum_{x=0}^{\infty} \left(xf(x)\right)$$

$$\mu = \sum_{x=0}^{1} \left(xf(x)\right) = 0f(0) + 1f(1) = f(1) = p$$

Δεύτερη Ροπή

$$\mu_2 = E(X^2) = \sum_{x=0}^{1} \left(x^2 f(x)\right) = 0^2 f(0) + 1^2 f(1) = f(1) = p$$

Διακύμανση

$$\sigma^2 = V(X) = E(X^2) - [E(X)]^2 = p - p^2 = p(1-p) = pq$$

Διωνυμική Κατανομή

Συμβ. X ~ b(n , p)

Όταν ένα πείραμα με δύο δυνατά αποτελέσματα (δοκιμή Bernoulli) εξαγόμενα επαναλαμβάνεται συγκεκριμένο αριθμό φορών εκείνο το οποίο συνήθως μας ενδιαφέρει είναι ο αριθμός των επιτυχιών ή αποτυχιών που εμφανίστηκαν. Η τυχαία μεταβλητή X που δίνει τον αριθμό επιτυχιών σε ν δοκιμές Bernoulli με κοινή πιθανότητα επιτυχίας p λέγεται Διωνυμική τυχαία μεταβλητή.

Ορισμός 1.1.2

Έστω X ο αριθμός των επιτυχιών σε μια ακολουθία n (ανεξάρτητων) δοκιμών Bernoulli με σταθερή πιθανότητα επιτυχίας p και πιθανότητα αποτυχίας q=1-p. Η κατανομή της τυχαίας μεταβλητής X καλείται ***διωνυμική κατανομή*** *με παραμέτρους n και p και συμβολίζεται με b(n,p).*

$R_X = \{0,1,\ldots,n\}$

Πρόταση

Η συνάρτηση πιθανότητας της διωνυμικής κατανομής με παραμέτρους n και p δίνεται απο τον τύπο:

$$f(x) = P(X = x) = \binom{n}{k} p^k q^{n-k} \qquad x = 0,1,2,\ldots,n$$

Παραδείγματα Αναγνώρισης Κατανομής

Σε μια κάλπη υπάρχουν α άσπρες και β μαύρες σφαίρες. Εξάγεται μια σφαίρα, σημειώνεται το χρώμα της και στη συνέχεια η σφαίρα ***επιστρέφεται*** *στη κάλπη. Η διαδικασία αυτή* ***επαναλαμβάνεται ν φορές*** *και συμβολίζουμε με Χ τον αριθμό των μαύρων σφαιρών που εξήχθησαν. Τότε η τυχαία μεταβλητή Χ ακολουθεί τη διωνυμική κατανομή.*

Σε μια κάλπη υπάρχουν α άσπρες και β μαύρες σφαίρες. Εξάγεται μια σφαίρα, σημειώνεται το χρώμα της και στη συνέχεια η σφαίρα ***επιστρέφεται*** *στη κάλπη. Η διαδικασία αυτή* ***επαναλαμβάνεται ν φορές*** *και συμβολίζουμε με Χ τον αριθμό των άσπρων σφαιρών που εξήχθησαν. Τότε η τυχαία μεταβλητή Χ ακολουθεί τη διωνυμική κατανομή.*

Παράδειγμα 1.1.1

Μια αεροπορική εταιρεία γνωρίζει οτι το 5% των ατόμων που κάνουν κράτηση για να ταξιδέψουν, δεν εμφανίζονται. Αν η εταιρεία κάνει κράτηση για 52 άτομα σε μια πτήση που γίνεται σε ένα αεροσκάφος χωρητικότητας 50 ατόμων, ποια η πιθανότητα να υπάρχει διαθέσιμο κάθισμα για καθένα επιβάτη που εμφανίζεται για να ταξιδέψει;

Λυση

Έστω Χ ο αριθμός των ατόμων που θα εμφανιστούν απο τα 52

$$f(x) = P(X = x) = \sum_{x=0}^{52} \binom{52}{x} (0.95)^x (0.05)^{52-x} \qquad x = 0,1,2,\ldots,52$$

Η πιθανότητα να υπάρχει διαθέσιμο κάθισμα για κάθε ένα απιβάτη που εμφανίζεται για να ταξιδέψει, είναι ίση με την πιθανότητα να εμφανιστούν ***το πολυ*** *50 ατομα, δηλαδη,*

$$P(X \leq 50) = \sum_{x=0}^{52} \binom{52}{x} (0.95)^x (0.05)^{52-x}$$

$P(X>50)= P(X=51)+ P(X=52)=\binom{52}{51}(0.95)^{51}(0.05)^{52-51}+\binom{52}{52}(0.95)^{52}(0.05)^{52-52} \cong 0.26$

Και η ζητούμενη πιθανότητα θα είναι ίση με $P(X\leq 50) \cong 1-0.26=0.74$

Μέση Τιμή και Διακύμανση

μ=E(X)=np $\sigma^2 = V(X) = npq$

Γεωμετρική Κατανομή

Συμβ. X ~ G(p)

Ορισμός 1.1.3

Θεωρούμε μια ακολουθία (ανεξάρτητων) δοκιμών Bernoulli με πιθανότητα επιτυχίας p και πιθανότητα αποτυχίας q=1-p ***σταθερή για όλες τις δοκιμές****. Εστω X ο αριμός των δοκιμών μέχρι την εμφάνιση της πρώτης επιτυχίας. Η κατανομή της τυχαίας μεταβλητής X καλείται* ***γεωμετρική κατανομή*** *με παραμετρο p.*

Πρόταση

Η συνάρτηση πιθανότητας της γεωμετρικής κατανομής G(p) δίνεται απο την τύπο:

$$f(x) = P(X = x) = q^{x-1}p \qquad x = 0,1,2, \dots$$

¨Σχηματική Απόδειξη¨

$\boxed{\alpha\alpha\alpha\alpha\alpha\alpha\alpha\alpha\alpha\alpha}\ \varepsilon \rightarrow \boxed{qqqqqqqqqq}\ p = q^{x-1}p$

x-1 φορές

Παραδείγματα Αναγνώρισης Κατανομής

Θεωρούμε μια ακολουθία δοκιμών Bernoulli με σταθερή πιθανότητα επιτυχίας p. Αν X είναι ο αριθμός των δοκιμών ***μέχρι να εμφανιστεί η πρώτη επιτυχία*** *τότε η τυχαία μεταβλητή X ακολουθεί τη γεωμετρική κατανομή.*

Θεωρούμε μια ακολουθία δοκιμών Bernoulli με σταθερή πιθανότητα επιτυχίας p. Αν X είναι ο αριθμός των δοκιμών μέχρι να εμφανιστεί ***η πρώτη αποτυχία*** *τότε η τυχαία μεταβλητή X ακολουθεί τη γεωμετρική κατανομή.*

Μέση Τιμή και Διακύμανση

μ=E(X)= $1/p$, $\sigma^2 = V(X) = \frac{q}{p^2}$

Μέση Τιμή Γεψμετρικής Κατανομής

Αν ρίχνουμε ένα ζάρι συνεχώς, ο ***αναμενόμενος*** *αριθμός δοκιμών μέχρι να φέρουμε για πρώτη φορά π.χ άσσο είναι* $1/p = {}^{1}\!\big/_{\frac{1}{6}}$ *= 6.*

Αρνητική Διωνυμική Κατανομή

Συμβ. Χ ~ ΝΒ(r,p)

Ορισμός 1.1.4

Θεωρούμε μια ακολουθία (ανεξάρτητων) δοκιμών Bernoulli με πιθανότητα επιτυχίας p και πιθανότητα αποτυχίας q=1-p ***σταθερή για όλες τις δοκιμές****. Εστω Χ ο αριμός των δοκιμών* ***μέχρι την εμφάνιση της r - επιτυχίας****. Η κατανομή της τυχαίας μεταβλητής Χ καλείται* ***Αρνητική Διωνυμική κατανομή*** *με παραμέτρους r και p.*

$R_X = \{r, r+1, r+2, \dots\}$

$$f(x) = P(X = x) = \binom{x-1}{r-1} p^r q^{x-r}, \qquad x = r, r+1, \dots$$

Παραδείγματα Αναγνώρισης Κατανομής

Ένα αμερόληπτο ζάρι ρίχνεται συνεχώς και συμβολίζουμε με Χ τον αριθμό των δοκιμών ***μέχρι να εμφανιστεί για τρίτη φορά*** *η ένδειξη 3 ή 4.*

Υπεργεωμετρική Κατανομή

Συμβ. Χ ~ h(n,α,β)

Ορισμός 1.1.4

Εστω μια κάλπη που περιέχει α στον αριθμό κόκκινες σφαίρες και β στον αριθμό μαύρες σφαίρες. Εστω Χ ο αριθμός τψν άσπρω σφαιρών που περιέχονται στο δείγμα. Η κατανομή της τυχαίας μεταβλητής Χ καλείται ***Υπεργεωμετρική Κατανομή*** *με παραμέτρους α,β και n.*

Πρόταση

Η συνάρτηση πιθανότητας της υπεργεωμετρικής κατανομής h(n,α,β) δίνεται απο την τύπο:

$$f(x) = P(X = x) = \frac{\binom{\alpha}{x}\binom{\beta}{n-x}}{\binom{\alpha+\beta}{n}}, \qquad x = max(0, n-\beta), \dots, \min(n, \alpha).$$

Παραδείγματα Αναγνώρισης Κατανομής

*Σε μια κάλπη υπάρχουν α άσπρες και β μαύρες σφαίρες. Εξάγεται μια σφαίρα, σημειώνεται το χρώμα της. Στη συνέχεια η διαδικασία αυτή **επαναλαμβάνεται ν φορές χωρίς η σφαίρα να επιστέφεται κάθε φορά στην κάλπη** και συμβολίζουμε με X τον αριθμό των μαύρων σφαιρών που εξήχθησαν. Τότε η τυχαία μεταβλητή X ακολουθεί την υπεργεωμετρική κατανομή.*

Μέση Τιμή και Διακύμανση

$$\mu = E(X) = n\left(\frac{\alpha}{\alpha+\beta}\right), \quad \sigma^2 = V(X) = n\frac{a\beta}{(a+\beta)^2}\left(1 - \frac{n-1}{a+b-1}\right).$$

Κατανομή Poisson

Συμβ. X ~ P(λ)

Εστω X μια διακριτλη τυχαία μεταβλητή με σύνολο τιμών $R_X = \{0,1,2,\dots\}$ και συνάρτητη πιθανότητας

$$f(x) = P(X=x) = e^{-\lambda}\frac{\lambda^x}{x!}, \quad x = 0,1,2,\dots$$

όπου λ>0. Η κατανομή της τυχαίας μεταβλητής X λέγεται κατανομή Poisson με παράμετρο λ.

Εστω οτι η τυχαία μεταβλητή X ακολουθεί την Διωνυμική κατανομη με συναρτηση πιθανότητας

$$f(x) = P(X = x) = \binom{n}{k} p^k q^{n-k} \qquad x = 0,1,2,\dots,n$$

Αν για n→∞, ηπιθανότητα επιτυχίας P συγκλίνει στο 0 (p→ 0), έτσι ώστε το γινόμενο μ=E(X)=np να συγκλίνει προς μια θετική σταθερά λ (np → λ > 0) τότε

$$\lim_{n\to\infty} f(x) = e^{-\lambda}\frac{\lambda^x}{x!}, \quad x = 0,1,\dots$$

$$b(n,p) \approx P(np)$$

Μέση Τιμή και Διακύμανση

$\mu = E(X) = \lambda, \quad \sigma^2 = V(X) = \lambda$

Παραδείγματα Αναγνώρισης Κατανομής

Αριθμός τηλεφωνικών κλήσεων που φτάνουν σε ένα τηλεφωνικό κέντρο σε μια δεδομένη χρονική περίοδο.

Αριθμός πελατών που επισκέπτονται ένα κατάστημα σε μια χρονική περίοδο ή αριθμός πελατών ενός καταστήματος οι οποίοι αγοράζουν ενα συγκεκριμένο προϊόν κατα την διάρκεια των εκπτώσεων.

Αριθμός παιδιών ενός πληθυσμού τα οποία ζουν περισσότερο απο 90 χρόνια

Αριθμός ελαττωματικών προϊόντων που παράγονται απο μια συγκεκριμένη γραμμή παραγωγής σε ένα χρονικό διαστημα.