Διωνυμική κατανομή

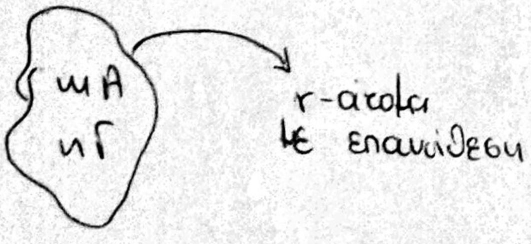
Εφαρμόζεται σε προκαθορισμένο αριθμό (n) ανεξάρτητων δοκιμών Bernulli με αμετάβλητη $P(E)=p$

Διωνυμική Κατανομή X = περιγραφή πλήθους του E στις n-επαναλήψεις

$X \sim B(n,p)$

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

(Λύση τελευταίου παραδείγματος προηγούμενου μαθήματος)



$$P\left(\begin{matrix}\text{να υπάρχουν} \\ k\text{-ακριβώς } A \\ 0 \le k \le r\end{matrix}\right) = \frac{m^k n^{r-k} \binom{r}{k \quad r-k}}{(m+n)^r} = \frac{\binom{r}{k} m^k n^{r-k}}{(m+n)^r}$$

Έστω επιτυχία $E = \{$να επιλεγεί A σε οποιαδήποτε επανάληψη$\}$

Έστω X = πλήθος των A στις r-επιλογές

$$X \sim B\left(r, p = P(E) = \frac{m}{m+n}\right)$$

Αμετάβλητη λόγω της επανάθεσης

$$P_X(x) = \binom{r}{x} p^x q^{r-x}, \quad x=0,1,\dots,r$$

$$P_X(x) = \binom{r}{x}\left(\frac{w}{w+n}\right)^{x}\left(1-\frac{w}{w+n}\right)^{r-x}$$

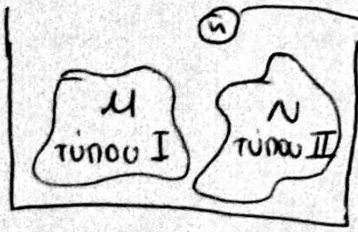
$$P(\text{να } \exists \ k\text{-A}) = P(X=k) = P_X(k) = \binom{r}{k}\left(\frac{w}{w+n}\right)^{k}\left(\frac{n}{w+n}\right)^{r-k}$$

Συμπέρασμα

Η διωνυμική κατανομή εφαρμόζεται σε δειγματοληψίες με επανάθεση.

Υπεργεωμετρική Κατανομή

Μοντέλο που εφαρμόζεται σε δειγματοληψίες χωρίς επανάθεση



χωρίς επανάθεση.

Ενδιαφερόμαστε για το πλήθος του τύπου I αντικειμένων στα n που επέλεξα.

Έστω X πλήθος του τύπου I στα n

Το X είναι μια τ.μ. Τιμές X : $0, 1, \dots, n$

Πρόκειται για διακριτή τ.μ. και ζητώ την σ.π της X

Έστω x μια τιμή της τ.μ X

Ζητώ $P_X(x) \overset{\text{op}}{=} P(X=x) = P\left(\begin{array}{l}\text{να υπάρχουν } x \text{ τύπου I} \\ \text{στα } n \text{ που επέλεξα}\end{array}\right) =$

$$= \frac{\binom{\mu}{x}\binom{N}{n-x}}{\binom{\mu+N}{n}}$$

Αρκεί η P_X να είναι σ.π δηλ.

1) $P_X(x) \geq 0$

2) $\sum_x P_X(x) = 1$

Αποδεικνύεται ότι ισχύουν.

Για να έχουν έννοια οι συνδυασμοί πρέπει $0 \le x \le M$ και $0 \le n-x \le N$

Άρα θα πρέπει $\max\{0, n-N\} \le x \le \min\{M, n\}$

ΟΡΙΣΜΟΣ

Η τ.μ. X λέγεται υπεργεωμετρική με παραμέτρους $M, N, n \in \mathbb{N}$ με $n \le M+N$ αν οι δυνατές τιμές της τ.μ. X είναι φυσικοί x με $\max\{0, n-N\} \le x \le \min\{M, n\}$ και η σ.π. της X δίνεται από τη σχέση

$$P_X(x) = \frac{\binom{M}{x}\binom{N}{n-x}}{\binom{M+N}{n}}$$

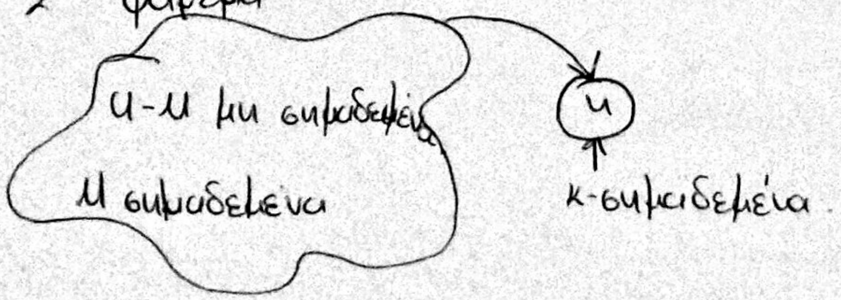
Συμβολισμός $X \sim Hg(M, N, n)$

Παράδειγμα (Εκτίμηση Αριθμού Ψαριών Λίμνης / Μοντέλα Πιάνω - Ξαναπιάνω)

Σε μια λίμνη υπάρχει άγνωστος αριθμός από n ψάρια. Ψαρεύονται από την λίμνη M ψάρια, σημαδεύονται με μια ανεξίτηλη μπογιά και τοποθετούνται ξανά στη λίμνη. Μετά από κάποιο χρονικό διάστημα ψαρεύονται n-ψάρια και k από αυτά είναι σημαδεμένα με μπογιά. Να εκτιμηθεί ο άγνωστος αριθμός n-ψαριών (Υποθέτουμε ότι δεν γεννιούνται ούτε πεθαίνουν ψάρια στο μεσοδιάστημα)

Λύση

2ο ψάρεμα



Έστω τ.μ. X που περιγράφει το πλήθος των σημαδεμένων (τύπου I) στα n που ψαρεύτηκαν την $2^η$ φορά.

$X \sim Hg(M, N-M, n)$

Άρα $p_X(x) = \dfrac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}$

Γνωρίζω ότι k στα n είναι σημαδεμένα

Άρα $P(X=k) = p_X(k)$ ← είναι όσο μεγαλύτερη μπορεί να γίνει

Άρα αρκεί να βρω το N που μεγιστοποιεί την σ.π. στο k.

$p_X(k) = \dfrac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}$ .... κάνοντας μεγιστοποίηση $N = \left[\dfrac{nM}{k}\right]$ ← ακέραιο μέρος

Μοντέλα πιάνω - ξαναπιάνω ⟶ capture - recapture

<u>Γεωμετρική Κατανομή</u>

Θεωρώ τ.π. μια ακολουθία (ανεξάρτητων) δοκιμών Bernulli, δηλαδή τ.π. τέτοια ώστε

Γ1) Αποτελούνται από μια ακολουθία επαναλήψεων της ίδιας διαδικασίας

Γ2) Σε κάθε επανάληψη υπάρχουν 2 αποτελέσματα E (Επιτυχία) και A (Αποτυχία)

Σε ένα τέτοιο τ.π. ενδιαφέρει συνήθως το πλήθος των επαναλήψεων μέχρι την πρώτη E

Έστω Χ το πλήθος των επαναλήψεων μέχρι για πρώτη φορά να έρθει Ε.

Το Χ είναι τ.μ. με τιμές: $x = 1, 2, 3, \dots, n,$

Αναζητώ την σ.π p_X

$$p_X(x) \overset{\text{ορ}}{=} P(X=x) = P(\overbrace{AA \dots A}^{x-1} E)$$

Υποθέτω επιπλέον **Γ3)** Οι επαναλήψεις είναι ανεξάρτητες

Άρα $p_X(x) = \overbrace{P(A) \cdot P(A) \cdots P(A)}^{x-1} \cdot P(E)$

Γ4) Υποθέτω ότι $P(E) = p$ αμετάβλητη. οπότε $P(A) = 1-p = q$ αμετάβλητη από δοκιμή σε δοκιμή

Άρα $p_X(x) = q^{x-1} \cdot p \qquad x = 1, 2, \dots, \quad q = 1-p \qquad 0<p<1$

Αν ισχύουν οι Γ1) - Γ4) τότε $\boxed{p_X(x) = q^{x-1} \cdot p}$

Είναι η p_X σ.π? Ναι γιατί i) $p_X(x) \geq 0$ προφανές

ii) $\sum_{x=1}^{\infty} p_X(x) = 1$

Απόδειξη ii)

$$\sum_{x=1}^{\infty} p_X(x) = \sum_{x=1}^{\infty} q^{x-1} p = p \sum_{x=1}^{\infty} q^{x-1} = p \sum_{v=0}^{\infty} q^v$$

Άθροισμα απείρων όρων φθίνουσας γεωμ. προόδου $\sum_{v=1}^{\infty} a^v = \frac{1}{1-q} \quad |a|<1$

Άρα $\sum_{x=1}^{\infty} p_X(x) = 1$

ΟΡΙΣΜΟΣ

Η τ.μ. X λέγεται γεωμετρική με παράμετρο p, $0<p<1$ αν οι δυνατές τιμές της X είναι $x=1,2,\ldots$ και η σ.π. είναι

$$p_X(x) = q^{x-1}p, \quad 0<p<1, \quad q=1-p, \quad x=1,2,\ldots$$

Συμβολισμός $X \sim Geo(p)$

Παράδειγμα.

Ζάρι ρίχνεται επανειλημμένα. Ποια η πιθανότητα το πρώτο 

α) Να εμφανιστεί στην $6^η$ ρίψη.

β) " μετά την $6^η$ ρίψη.

Λύση

$E=\{$ να έρθει

$\}$

Έστω X τ.μ. παριστά πλήθος ρίψεων (επαναλήψεων) μέχρι να έρθει για πρώτη φορά E.

$X \sim Geo\left(p=P(E)=\frac{1}{6}\right)$

α) $P(X=6) = p_X(6) = \left(\frac{5}{6}\right)^{6-1}\cdot\frac{1}{6} = 0,067$

β) $P(X>6) = 1-P(X\leq 6) = 1-\sum_{x=1}^{6} p_X(x) = 1-\sum_{x=1}^{6}\left(\frac{5}{6}\right)^{x-1}\cdot\frac{1}{6} = \left(\frac{5}{6}\right)^6$

Άλλος τρόπος

$$P(X>6) = \sum_{x=7}^{\infty} p_X(x) = \sum_{x=7}^{\infty}\left(\frac{5}{6}\right)^{x-1}\cdot\frac{1}{6} = \frac{1}{6}\sum_{x=7}^{\infty}\left(\frac{5}{6}\right)^{x-1} = \frac{1}{6}\sum_{x-7=0}^{\infty}\left(\frac{5}{6}\right)^{x-7}\left(\frac{5}{6}\right)^6 =$$

$$\underset{x-7=v}{=} \frac{1}{6}\left(\frac{5}{6}\right)^6\sum_{v=0}^{\infty}\left(\frac{5}{6}\right)^v = \cancel{\frac{1}{6}}\cdot\left(\frac{5}{6}\right)^6\cdot\frac{1}{\cancel{1-5/6}} = \left(\frac{5}{6}\right)^6$$

Ιδιότητα Αμνησίας Γεωμετρικής Κατανομής

ΠΡΟΤΑΣΗ

Αν η τ.μ. $X \sim Geo(p)$ τότε $P(X > m+n / X > m) = P(X > n)$

$m, n = 1, 2, 3, \dots$

Παράδειγμα.

Ο Παναγιώτης σημαδεύει συνεχώς έναν ~~στόχο~~ στόχο με $P(\text{να πετύχει}) = 0,1$
Ποια η πιθανότητα να πετύχει τον στόχο μετά από 5 προσπάθειες αν έχει ήδη προσπαθήσει ανεπιτυχώς περισσότερες από 3 φορές?

Λύση

Έστω $E = \{$να πετύχει τον στόχο σε μια οποιαδήποτε προσπάθεια$\}$
Έστω τ.μ. X παριστά το πλήθος των προσπαθειών μέχρι να πετύχει τον στόχο δηλ. την 1η E.

$$X \sim Geo(p = P(E) = 0,1)$$

$$P(X>5/X>3) = P(X>2+3/X>3) = P(X>2) = 1 - P(X \leq 2) = 1 - P(X=1) - P(X=2) =$$

$$= 1 - 0,9^0 \cdot 0,1 - 0,9 \cdot 0,1 = 0,81$$