ΚΕΦΑΛΑΙΟ 4 - ΕΙΔΙΚΕΣ ΔΙΑΚΡΙΤΕΣ ΚΑΤΑΝΟΜΕΣ

Πειράματα Bernoulli: πείραμα με δύο πιθανά αποτελέσματα: επιτυχία (με πιθανότητα p) και αποτυχία (με πιθανότητα $1-p=q$)

π.χ. Ρίψη νομίσματος, εξέταση μαθήματος, επίδοση σε αθλητικούς αγώνες.

Διαδοχικά πειράματα Bernoulli είναι ανεξάρτητα

1) Διωνυμική κατανομή

Επαναλαμβάνουμε ενα πείραμα Bernoulli με πιθανότητα επιτυχίας p, n φορές.

Δεδομένα: n, p

Ορίζουμε την τ.μ $X = \#$ επιτυχιών

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

* $E(X) = n \cdot p$

* $V(X) = n \cdot p(1-p) = n \cdot p \cdot q$

$$\binom{n}{x} = \frac{n!}{x!(n-x)!}$$

Συμβολίζουμε με $X \sim Bin(n, p)$

π.χ

15 ερωτήσεις πολλαπλής επιλογής
5 πιθανές απαντήσεις στην καθεμια
Μόνο 1 σωστη
Ένας φοιτητής απαντάει την καθεμια στην τύχη
Να βρεθούν οι πιθανότητες:
α) Να απαντήσει ακριβώς 8 ερωτήσεις σωστά
β) » » περισσότερες από 3 και λιγότερες από 8 ερωτήσεις σωστά

Πείραμα Bernoulli: επιλογή απάντησης σε μια ερώτηση
15 πειράματα Bernoulli με πιθαν. επιτυχίας $1/5$
X = # σωστών απαντήσεων του φοιτητή
$X \sim Bin(15, 1/5)$

α) $P(X=8) = f(8) = \binom{15}{8}\left(\frac{1}{5}\right)^{8} \cdot \left(\frac{4}{5}\right)^{15-8} = \dots = \boxed{0{,}00348}$.

β) $P(3 < X < 8)$
$= P(X=4) + P(X=5) + P(X=6) + P(X=7) = f(4) + f(5) + f(6) + f(7)$
$= \binom{15}{4}\left(\frac{1}{5}\right)^{4} \cdot \left(\frac{4}{5}\right)^{15-4} \dots = \boxed{0{,}3476}$

Αναμενόμενος βαθμός του φοιτητή: $E(X) = n \cdot p = 15 \cdot \frac{1}{5} = 3$

2) Γεωμετρική κατανομή

Επαναλαμβάνουμε ένα πείραμα Bernoulli με πιθ. επιτυχίας p μέχρι να σημειώσουμε την πρώτη επιτυχία
Δεδομένα: p

Έστω η τ.μ. X = # δοκιμών μέχρι την 1η επιτυχία

* $f(x) = P(X=x) = (1-p)^{x-1} \cdot p = q^{x-1} \cdot p$, $x = 1, 2 \dots$
* $F(x) = 1 - q^{x}$, $x = 1, 2, 3 \dots$
* $E(X) = \frac{1}{p}$
* $V(X) = \frac{q}{p^2}$

Συμβολίζουμε με $X \sim Geo(p)$

Παράδειγμα

Έστω ότι ένας υπάλληλος fast food έχει πιθανότητα 0,9 να ψήσει ένα μπιφτέκι σωστά.

Να υπολογιστούν:

α) Η πιθανότητα να ψήσει λάθος το πολύ 5 μπιφτέκια μέχρι να ψήσει ένα σωστά

β) Ο μέσος όρος μπιφτεκιών που θα ψήσει λάθος μέχρι να ψήσει ένα σωστά

Πείραμα Bernoulli : ψήσιμο μπιφτεκιών

πιθ. επιτυχίας $= 0,9$

$X =$ # μπιφτεκιών μέχρι να ψηθεί ένα σωστά

$X \sim Geo(0,9)$

α) $P(X \leq 5) = F(5) = 1 - 0,1^5 = 0,999\ldots$

β) $E(X) = \frac{1}{p} = \frac{1}{0,9} = 1,111\ldots$

3) Κατανομή Pascal ή Αρνητική Διωνυμική

Επαναλαμβάνουμε ένα πείραμα Bernoulli με πιθ. επιτυχίας p μέχρι να σημειωθούν r επιτυχίες

Δεδομένα: p, r

$X =$ # δοκιμών μέχρι την r-οστή επιτυχία

* $f(x) = \binom{x-1}{r-1} p^r q^{x-r}, \quad x = r, r+1 \ldots$

$E(X) = \frac{r}{p}$ $V(X) = \frac{rq}{p^2}$

Συμβολίζουμε με $X \sim NB(r,p)$

Τι κάναμε;

Πειράματα Bernoulli

1) Διωνυμική κατανομή (# επιτυχιών σε n δοκιμές) Bin(n,p)
2) Γεωμετρική κατανομή (# δοκιμών μέχρι την 1η επιτυχία) Geo(p)
3) Αρν. διωνυμική κατανομή (# δοκιμων μέχρι την r-οστή επιτυχία) NB(r,p)

Παράδειγμα: Αθλητής άλματος εις ύψος, υπερπηδά στην προπόνηση το ύψος των 225 cm με πιθανότητα $p = 0,4$

α) Ποια η πιθανότητα να περάσει το ύψος στην 4η προσπάθεια για 1η φορά;
β) Ποια η αναμενόμενη τιμή των προσπαθειών μέχρι να περάσει το ύψος για πρώτη φορά;
γ) Ποια η πιθανότητα σε 10 προσπάθειες να έχει 6 επιτυχημένες;
δ) Ποια η πιθανότητα σε 10 προσπάθειες να υπερπηδήσει το ύψος για 1η φορά στην 6η προσπάθεια;
ε) Ποια η πιθανότητα να χρειαστεί 7 προσπάθειες για να υπερπηδήσει το ύψος 3 φορές;

Λύση

α) $X =$ # άλματα μέχρι να περάσει το εμπόδιο

$X \sim Geo(0,4)$

$P(X=4) = f_X(4) = (q^{x-1}p) = 0,6^3 \cdot 0,4 = \boxed{0,0864}$

β) Αναμενόμενη τιμή προσπαθειών

$= E(X) = \frac{1}{p} = \frac{1}{0,4} = \boxed{2,5}$

γ) Y = # επιτυχημένων αλμάτων σε 10 προσπάθειες

$Y \sim Bin(10, 0{,}4)$

$$P(Y=6) = f_Y = \binom{n}{x} p^x q^{n-x} = \binom{10}{6} 0{,}4^6 \cdot 0{,}6^4 = \boxed{0{,}1115}$$

δ) $P(X=6) = f_X(6) = 0{,}6^5 \cdot 0{,}4 = \boxed{0{,}0311}$

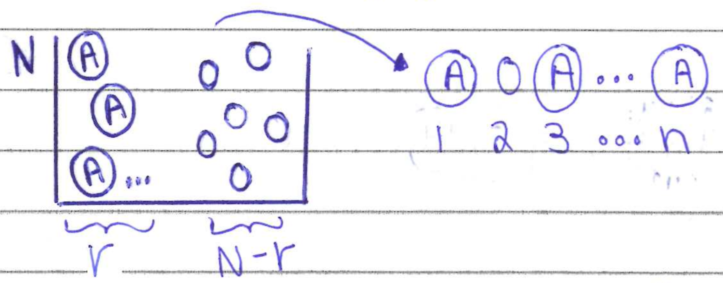
ε) Z = # αλμάτων μέχρι να υπερπηδήσει το ύψος 3 φορές

$Z \sim NB(3, 0{,}4)$

$$P(Z=7) = f_Z(7) = \binom{x-1}{r-1} p^r q^{x-r} = \binom{6}{2} 0{,}4^3 \cdot 0{,}6^3 = \boxed{0{,}1244}$$

4) Υπεργεωμετρική κατανομή

Βασίζεται σε δειγματοληψία χωρίς επανάθεση



Έχουμε N αντικείμενα από τα οποία:
- r έχουν ένα κοινό χαρακτηριστικό A
- τα υπόλοιπα δεν το έχουν

Επιλέγουμε ένα δείγμα μεγέθους n

Δεδομένα: N, r, n

Τ.μ X = # αντικειμένων με χαρακτηριστικό A στο δείγμα μεγέθους n

$$* \quad P(X=x) = f(x) = \frac{\binom{r}{x}\binom{N-r}{n-x}}{\binom{N}{n}}, \quad x = 0, 1, \dots r$$

$E(X) = \frac{nr}{N}$ Συμβολ. $X \sim H(N, r, n)$
Hyp.

$V(X) = \frac{nr}{N} \cdot \frac{N-r}{N} \cdot \frac{N-n}{N-1}$

παράδειγμα: Σε μια κλήρωση Lotto τοποθετούνται στην κληρωτίδα σφαιρίδια αριθμημένα από το 1 ως το 49 και επιλέγονται στην τυχη 6 αριθμοί.

α) Ποια η πιθανότητα να περιέχονται 4 τυχεροί αριθμοί ανάμεσα σε 20 αριθμούς που έχουμε σημειώσει;

β) Ποιο είναι το αναμενόμενο πλήθος τυχερών αριθμών στους 20 που έχουμε σημει

Λύση:

X = # τυχερών αριθμών στους 20 που εχουμε σημειώσει
$N = 49$, $r = 6$, $n = 20$
$X \sim H(49, 6, 20)$

α) $P(X=4) = f(4) = \frac{\binom{6}{4}\binom{43}{16}}{\binom{49}{20}} = \boxed{0,14}$

β) $E(X) = \frac{20 \cdot 6}{49} = \boxed{2,45}$

Παράδειγμα

Ενας σκοπευτής ρίχνει 10 βολές προς έναν στόχο. Η πιθαν. να τον πετύχει 4 φορες είναι επταπλάσια από την πιθαν. να τον πετύχει 3 φορες.

α) Να υπολογιστεί η ευστοχία του σκοπευτή (πιθανοτ. να πετύχει το στόχο σε τυχαία βολή.

β) Πιθαν. να πετύχει το στόχο 2 φορές σε 5 βολές;
γ) Πιθ. να χρειαστεί 5 βολές ακριβώς για να πετύχει το στόχο 2 φορές;
δ) αναμενόμενος αριθμός άστοχων βολών μέχρι την πρώτη εύστοχη;
ε) Πόσες φορές πρέπει να ρίξει για να είναι βέβαιος ότι κατά 95% θα πετύχει τον στόχο;
στ) πιθ. σε 5 βολές να πετύχει τον στόχο το πολύ 4 φορές αν είναι γνωστό ότι τον πέτυχε τουλάχιστον 2 φορές

<u>Λύση</u>

α) $X = \#$ εύστοχων βολών στις 10
$X \sim Bin(10, p)$

$P(X=4) = 7P(X=3) \Rightarrow f_X(4) = 7 \cdot f_X(3)$

$\Rightarrow \binom{10}{4} p^4 (1-p)^6 = 7 \binom{10}{3} p^3 (1-p)^7$

$= \frac{\cancel{10!}}{4!6!} p^{4} \cancel{(1-p)^6} = 7 \cdot \frac{\cancel{10!}}{3!7!} \cancel{p^3} (1-p)^{\cancel{7}}$

$= \frac{1}{4!6!} p = 7 \cdot \frac{1}{3!7!} (1-p) \Rightarrow \frac{3!7!}{4!6!} p = 7(1-p) \Rightarrow \frac{\cancel{7}}{4} p = \cancel{7}(1-p)$

$\Rightarrow p = 4-4p \Rightarrow 5p = 4 \Rightarrow p = \frac{4}{5} = \boxed{0,8}$

β) $Y = \#$ εύστοχων βολών στις 5
$Y \sim Bin(5, 0,8)$
$P(Y=2) = \binom{5}{2} 0,8^2 \cdot 0,2^3 = \boxed{0,0512}$

γ) $Z = \#$ βολών μέχρι τη 2η εύστοχη βολή

$Z \sim NB(2, 0{,}8)$

$P(Z=5) = f_Z(5) = \binom{4}{1} 0{,}8^2 \cdot 0{,}2^2 = \ldots$

δ) $T = \#$ βολών μέχρι την 1η εύστοχη

$T \sim Geo(0{,}8)$

$E(T) = \frac{1}{p} = \frac{1}{0{,}8} = \boxed{1{,}25}$

ε) αριθμό βολών μέχρι την 1η εύστοχη $= x$

$P(T \leq x) \geq 0{,}95$

$\Rightarrow F_T(x) \geq 0{,}95 \Rightarrow 1 - q^x \geq 0{,}95$

$1 - 0{,}2^x \geq 0{,}95$

$0{,}2^x \leq 0{,}05$

$x \cdot \log 0{,}2 \leq \log 0{,}05$

$\log 0{,}2 < 0 \quad \Rightarrow x \geq \frac{\log 0{,}05}{\log 0{,}2} \Rightarrow \boxed{x \geq 1{,}86}$

στ)

$P(Y \leq 4 / Y \geq 2) = \frac{P(Y \leq 4 \text{ και } Y \geq 2)}{P(Y \geq 2)} = \frac{P(2 \leq Y \leq 4)}{P(Y \geq 2)}$

• $P(2 \leq Y \leq 4)$

$= P(Y=2) + P(Y=3) + P(Y=4)$.

$= f_Y(2) + f_Y(3) + f_Y(4) \ldots$

• $P(Y \geq 2) = 1 - P(Y < 2)$

$= 1 - (P(Y=0) + P(Y=1))$

$= 1 - (f_Y(0) + f_Y(1)) = \ldots$

Συνέχεια μαθήματος

5) Κατανομή Poisson

Ενδιαφερόμαστε για γεγονότα που επαναλαμβάνονται ανά ταυτά χρονικά διαστήματα ή σε διάφορα σημεία στον χώρο

$X = \#$ εμφανίσεων του γεγονότος

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

- $E(X) = V(X) = \lambda$ • Συμβολίζουμε με $X \sim P(\lambda)$

Θεώρημα

Αν $X \sim Bin(n,p)$ και $n \to \infty$, $p \to 0$,
$np = \lambda =$ σταθ. $\Rightarrow$ τότε προσεγγιστικά: $X \sim P(\lambda)$

Πρακτικά: Για $n \geq 30$, $p \leq 0{,}1$ και $np \leq 20$,
εφαρμόζουμε το θεώρημα

παράδειγμα: Ο αριθμός κλήσεων ανά ώρα σε ένα τηλεφ. κέντρο ακολουθεί κατανομή Poisson με μέση τιμή 2

Να βρεθούν οι πιθανότητες:

α) Να γίνουν 2 κλήσεις μέσα σε 1 ώρα
β) Να γίνουν το πολύ 2 κλήσεις " "
γ) Να γίνει τουλάχιστον μια κλήση μέσα σε μια ώρα
δ) Μέσος όρος κλήσεων ανά ώρα

α) $X = \#$ κλήσεων ανά ώρα (μεση τιμή $= 2 = \lambda$)

$X \sim P(2) = P(X=2)$

$= f(2) = \frac{e^{-2} \cdot 2^2}{2!} = \boxed{2e^{-2}}$

β) $P(X \leq 2) = P(X=0) + P(X=1) + P(X=2)$

$= \frac{e^{-2} \cdot 2^0}{0!} + \frac{e^{-2} \cdot 2^1}{1!} + \frac{e^{-2} \cdot 2^2}{2!} = e^{-2}(1+2+2) = \boxed{5e^{-2}}$

γ) $P(X \geq 1) = 1 - P(X<1) = 1 - P(X=0) = 1 - \frac{e^{-2} \cdot 2^0}{0!} = \boxed{1 - e^{-2}}$

δ) $E(X) = 2 =$ μέσος όρος κλήσεων ανά ώρα

ε) Η πιθανότητα να δεχθεί 2 κλήσεις σε ένα τρίωρο

$Y = \#$ κλήσεων ανά τρίωρο $\Rightarrow Y \sim P(6)$ $2 \cdot 3$

$P(Y=2) = \frac{e^{-6} \cdot 6^2}{2!} = \boxed{18e^{-6}}$

Παράδειγμα

Έχει παρατηρηθεί ότι το 1% του πληθυσμού έχει ληξιπρόθεσμα δάνεια σε 100 άτομα, ποια η πιθανότητα να βρούμε 2 άτομα με ληξιπ. δάνειο;

πιθαν. ένα τυχαίο άτομο να έχει ληξ. δάνειο $= 1\% = 0{,}01$ (πείραμα Bernoulli)

$X = \#$ ατόμων (από τα 100) που έχουν ληξ. δάνειο

$X \sim Bin(100, 0{,}01)$

$P(X=2) =$

Για να δω αν εφαρμόζεται: ① $n = 100 \geq 30$ ② $p = 0{,}01 \leq 0{,}1$

③ $np = 100 \cdot 0{,}01 = 1 \leq 20$

Άρα προσεγγιστικά: $X \sim P(1)$

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