

Στατιστική Ι

Γιώργος Τσιρογιάννης

Τμήμα Διοίκησης Επιχειρήσεων Αγροτικών
Προϊόντων και Τροφίμων,
Πανεπιστήμιο Πατρών



Διάλεξη 8η

Βασικές διακριτές κατανομές



6.1- 6.5

Κατανομή Bernoulli

- Πείραμα τύχης με μόνο δύο, αμοιβαίως αποκλειόμενα, δυνατά αποτελέσματα.
- Συνήθως το ένα αποτέλεσμα ονομάζεται επιτυχία (success) και το άλλο αποτυχία (failure).
- Σε μια δοκιμή *Bernoulli* ο αριθμός των επιτυχιών προφανώς είναι ή μια ή καμία.
- Αν η πιθανότητα επιτυχίας είναι p και η πιθανότητα αποτυχίας είναι q , προφανώς $p + q = 1$ ή $q = 1 - p$

Παραδείγματα

- Η επιλογή ενός ζώου και η εξέτασή του για να διαπιστωθεί αν έχει προσβληθεί από μια συγκεκριμένη ασθένεια είναι δοκιμή Bernoulli γιατί τα δυνατά αποτελέσματα είναι μόνο δύο: το ζώο είτε έχει προσβληθεί (επιτυχία) είτε δεν έχει προσβληθεί (αποτυχία).

Παραδείγματα

- Η λήψη δείγματος ελαιολάδου και ο έλεγχος για το αν είναι νοθευμένο είναι δοκιμή Bernoulli γιατί το ελαιόλαδο είτε είναι νοθευμένο (επιτυχία) είτε δεν είναι νοθευμένο (αποτυχία).
- Η επιλογή και ο έλεγχος ενός φυτού για το αν η ξηρή μάζα του ξεπερνάει τα 150gr είναι δοκιμή Bernoulli γιατί το φυτό έχει ξηρή μάζα είτε μεγαλύτερη από 150gr (επιτυχία) είτε το πολύ 150gr (αποτυχία).

Παραδείγματα

- Φυτεύουμε ένα σπόρο πιπεριάς και ελέγχουμε αν βλάστησε. Πρόκειται για *δοκιμή Bernoulli* γιατί ο σπόρος είτε θα βλαστήσει (*επιτυχία*) είτε δε θα βλαστήσει (*αποτυχία*).
- Η επιλογή και ο έλεγχος ενός φυτού για το αν έχει λιγότερα από 6 φύλλα είναι *δοκιμή Bernoulli* γιατί το φυτό έχει είτε λιγότερα από 6 (*επιτυχία*) είτε τουλάχιστον 6 φύλλα (*αποτυχία*).

Παραδείγματα

- Η επιλογή ενός μαθητή Α΄ Δημοτικού και ο έλεγχος της συγκέντρωσης της προσοχής του κατά τα πρώτα 20 λεπτά της πρώτης ώρας του προγράμματος μπορεί (με συγκεκριμένα κριτήρια) να χαρακτηριστεί είτε ικανοποιητική είτε όχι ικανοποιητική. Είναι επομένως μια δοκιμή Bernoulli.

Κατανομή Bernoulli

- Έστω X η διακριτή τυχαία μεταβλητή που εκφράζει τον αριθμό των επιτυχιών σε μια δοκιμή Bernoulli με πιθανότητα επιτυχίας p . Η κατανομή της X ονομάζεται κατανομή Bernoulli (Bernoulli distribution) με παράμετρο p και συμβολίζεται με $X \sim b(p)$.
- Συνάρτηση πιθανότητας

$$f(x) = P(X = x) = \begin{cases} p, & x = 1 \\ 1 - p, & x = 0 \end{cases} \quad \text{ή} \quad f(x) = P(X = x) = p^x (1 - p)^{1-x}, \quad x = 0, 1.$$

Μέση τιμή και διακύμανση

- Η μέση τιμή μ της X είναι

$$\mu = E(X) = 0 \cdot f(0) + 1 \cdot f(1) = p$$

- Η διακύμανση σ^2

$$\sigma^2 = E(X^2) - [E(X)]^2 = 0^2 \cdot f(0) + 1^2 \cdot f(1) - p^2 = p - p^2 = p(1 - p) = pq$$

Παράδειγμα



- Ο γεωπόνος ενός φυτώριου ισχυρίζεται ότι οι βολβοί τουλίπας που παράγονται στο φυτώριο βλαστάνουν σε ποσοστό 90%.
- Ένας αγρότης που είχε προμηθευθεί από το συγκεκριμένο φυτώριο μεγάλο αριθμό βολβών τουλίπας θέλησε να ελέγξει τον ισχυρισμό του γεωπόνου.
- Επέλεξε τυχαία 52 από τους βολβούς που φύτεψε και διαπίστωσε ότι βλάστησαν μόνο οι 38.
- Είναι δικαιολογημένες οι αμφιβολίες του αγρότη;

Παράδειγμα



- Η τυχαία επιλογή και ο έλεγχος ενός βολβού για το αν βλάστησε ή όχι, είναι μια δοκιμή Bernoulli.
- Ο έλεγχος 52 βολβών είναι μια σειρά-ακολουθία 52 δοκιμών Bernoulli.
- Επιτυχία ονομάζουμε το αποτέλεσμα «ο βολβός βλάστησε» και αποτυχία το αποτέλεσμα «ο βολβός δεν βλάστησε»

Παράδειγμα



- **Σκεπτικό:** Αν αποδεχθούμε τον ισχυρισμό του γεωπόνου, πόσο λογικό-πιθανό είναι το αποτέλεσμα του ελέγχου που έκανε ο αγρότης;
- Με δεδομένο ότι το ποσοστό των βολβών που βλαστάνουν είναι 90%, πόσο πιθανό είναι από τους 52 τυχαία επιλεγμένους βολβούς να βρεθούν να έχουν βλαστήσει οι 38;

Παράδειγμα



- Ποια είναι η πιθανότητα, σε 52 επαναλήψεις μιας δοκιμής Bernoulli να συμβούν 38 επιτυχίες, με δεδομένο ότι σε κάθε επανάληψη η πιθανότητα επιτυχίας p είναι ίση με 0.9;
- Ή γενικότερα:
- Αν X ο αριθμός των επιτυχιών σε μια ακολουθία n ανεξάρτητων δοκιμών Bernoulli με σταθερή πιθανότητα επιτυχίας p σε όλες τις δοκιμές, πώς μπορούμε να υπολογίσουμε τις πιθανότητες

$$P(X = x), \quad x = 0, 1, 2, \dots, n$$

Παράδειγμα



$$P(X = x) = \binom{v}{x} p^x (1-p)^{v-x}, \quad x = 0, 1, 2, \dots, v$$

$$P(X = 38) = \binom{52}{38} 0.9^{38} (1-0.9)^{14} \cong 0.0003$$


Η πιθανότητα να βλαστήσουν μόνο 38 από τους 52 βολβούς είναι πολύ μικρή (περίπου μηδενική)

Συνδυασμοί (γενικά)

- Έστω X σύνολο με v στοιχεία και ακέραιος $k \leq v$. Συνδυασμός των v στοιχείων ανά k ονομάζεται κάθε μη διατεταγμένο υποσύνολο του X με k στοιχεία.
- Το πλήθος των διαφορετικών συνδυασμών είναι

$$\binom{v}{k} = \frac{v!}{k!(v-k)!} = \frac{v!}{k!(v-k)!}$$

Διωνυμική κατανομή



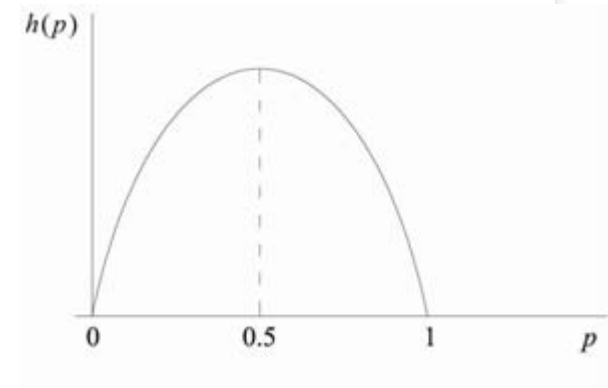
Διωνυμική κατανομή

- Αν X ο αριθμός των επιτυχιών σε μια ακολουθία n ανεξάρτητων δοκιμών Bernoulli με σταθερή πιθανότητα επιτυχίας p σε όλες τις δοκιμές, τότε η κατανομή της τυχαίας μεταβλητής X ονομάζεται διωνυμική κατανομή (binomial distribution) με παραμέτρους n και p και συμβολίζεται με $B(n, p)$ ή $X \sim B(n, p)$.
- Η συνάρτηση πιθανότητας είναι:

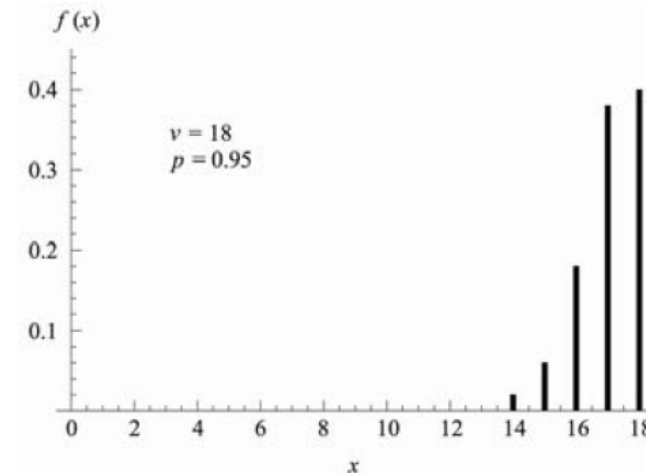
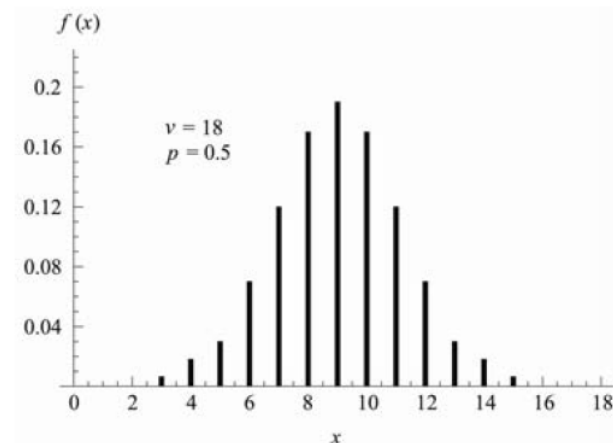
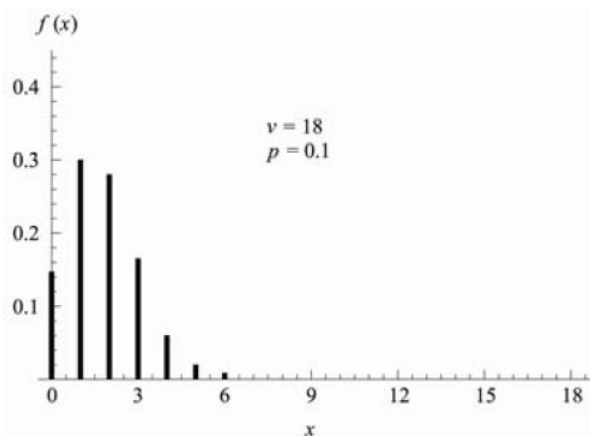
$$f(x) = P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}, \quad x = 0, 1, 2, \dots, n$$

$$(\alpha + \beta)^n = \sum_{x=0}^n \binom{n}{x} \alpha^x \beta^{n-x}, \quad n=1, 2, \dots$$

Μέση τιμή και η διακύμανση



- Η μέση τιμή και η διακύμανση είναι: $\mu = np$ και $\sigma^2 = np(1-p) = npq$
- Η διακύμανση ελαττώνεται όσο το p πλησιάζει το 0 ή το 1 και μεγιστοποιείται για $p = 1/2$



$$f(x) = P(X = x) = \binom{v}{x} p^x (1-p)^{v-x}, \quad x = 0, 1, 2, \dots, v$$

Παράδειγμα



- Μια βιομηχανία κατασκευάζει μεταλλικά ελατήρια για να αντέχουν σε συγκεκριμένη καταπόνηση.
- Σύμφωνα με τις προδιαγραφές παραγωγής, κάθε τέτοιο έλασμα αντέχει στη συγκεκριμένη καταπόνηση με πιθανότητα 0.8.
- Επιλέγουμε τυχαία 9 τέτοια ελάσματα και τα υποβάλλουμε στη συγκεκριμένη καταπόνηση.
- Ποια είναι η πιθανότητα να αντέξουν α) το πολύ 2 ελάσματα β) περισσότερα από 7 ελάσματα γ) τουλάχιστον 2 ελάσματα και δ) λιγότερα από 6 και τουλάχιστον 4 ελάσματα.

Παράδειγμα



- Έστω X ο αριθμός των ελασμάτων που θα αντέξουν την καταπόνηση
- $X \sim B(9, 0.8)$
- α) το πολύ 2 ελάσματα:

$$\begin{aligned} P(X \leq 2) &= P(X = 0) + P(X = 1) + P(X = 2) = \\ &= \binom{9}{0} \cdot 0.8^0 0.2^9 + \binom{9}{1} \cdot 0.8^1 0.2^8 + \binom{9}{2} \cdot 0.8^2 0.2^7 = 0.0003 \end{aligned}$$

Διωνυμική κατανομή

- Αν X ο αριθμός των επιτυχιών σε μια ακολουθία n ανεξάρτητων δοκιμών Bernoulli με σταθερή πιθανότητα επιτυχίας p σε όλες τις δοκιμές, τότε η κατανομή της τυχαίας μεταβλητής X ονομάζεται διωνυμική κατανομή (binomial distribution) με παραμέτρους n και p και συμβολίζεται με $B(n, p)$ ή $X \sim B(n, p)$.
- Η συνάρτηση πιθανότητας είναι:

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

Παράδειγμα



- β) περισσότερα από 7 ελάσματα:

$$P(X > 7) = P(X = 8) + P(X = 9) = \binom{9}{8} \cdot 0.8^8 0.2^1 + \binom{9}{9} \cdot 0.8^9 0.2^0 = 0.436207$$

- γ) τουλάχιστον 2 ελάσματα:

$$P(X \geq 2) = 1 - P(X < 2) = 1 - P(X = 0) - P(X = 1) = \dots = 1 - 0.000019 = 0.999981$$

- δ) λιγότερα από 6 και τουλάχιστον 4 ελάσματα:

$$P(4 \leq X < 6) = P(X = 4) + P(X = 5) = \binom{9}{4} \cdot 0.8^4 0.2^5 + \binom{9}{5} \cdot 0.8^5 0.2^4 = 0.0825754$$

Πιο πιθανή τιμή μιας διωνυμικής τ.μ.

- Αν $X \sim B(n, p)$, ποια είναι η πιο πιθανή από τις τιμές $0, 1, 2, \dots, n$ που μπορεί να πάρει η X ;

Όταν το γινόμενο $(n + 1)p$ δεν είναι ακέραιος αριθμός, η πιο πιθανή τιμή x_0 της $X \sim B(n, p)$ είναι το ακέραιο μέρος του, ενώ όταν το γινόμενο $(n + 1)p$ είναι ακέραιος αριθμός, οι πιο πιθανές τιμές είναι δύο, η $x_0 = (n + 1)p$ και η $x_0 = (n + 1)p - 1$.

Παράδειγμα (συν)



- Ο γεωπόνος ενός φυτώριου ισχυρίζεται ότι οι βολβοί τουλίπας που παράγονται στο φυτώριο βλαστάνουν σε ποσοστό 90%.
- Ένας αγρότης που είχε προμηθευθεί από το συγκεκριμένο φυτώριο μεγάλο αριθμό βολβών τουλίπας θέλησε να ελέγξει τον ισχυρισμό του γεωπόνου.
- Επέλεξε τυχαία 52 από τους βολβούς που φύτεψε και διαπίστωσε ότι βλάστησαν μόνο οι 38.
- Είναι δικαιολογημένες οι αμφιβολίες του αγρότη;

Παράδειγμα (συν)




- Το γινόμενο $(n + 1) p = 53 \cdot 0.9 = 47.7$ δεν είναι ακέραιος
- Οπότε η πιο πιθανή τιμή είναι το 47

Πιο πιθανή τιμή μιας διωνυμικής τ.μ.

- Αν $X \sim B(n, p)$, ποια είναι η πιο πιθανή από τις τιμές $0, 1, 2, \dots, n$ που μπορεί να πάρει η X ;

Όταν το γινόμενο $(n + 1)p$ δεν είναι ακέραιος αριθμός, η πιο πιθανή τιμή x_0 της $X \sim B(n, p)$ είναι το ακέραιο μέρος του, ενώ όταν το γινόμενο $(n + 1)p$ είναι ακέραιος αριθμός, οι πιο πιθανές τιμές είναι δύο, η $x_0 = (n + 1)p$ και η $x_0 = (n + 1)p - 1$.

Κατανομή Poisson



Κίνητρο - παράδειγμα



- Κτηνοτροφική περιοχή με 300000 αιγοπρόβατα
- Όλα τα αιγοπρόβατα εμβολιάζονται κάθε χρόνο
- Μικρή πιθανότητα $2 \cdot 10e^{-5}$ το εμβόλιο να προκαλέσει μια πολύ σοβαρή παρενέργεια
- Ποια η πιθανότητα να εμφανίσουν x ζώα την παρενέργεια

Κίνητρο - παράδειγμα



- Ο αριθμός X των ζώων που εμφανίζουν την παρενέργεια στη συγκεκριμένη περιοχή σε ένα έτος είναι διωνυμική τυχαία μεταβλητή $X \sim B(300000, 2 \cdot 10^{-5})$
- Μας ενδιαφέρουν οι πιθανότητες $P(X = x)$, όπου $x = 0, 1, \dots, 300000$

Διωνυμική κατανομή

- Αν X ο αριθμός των επιτυχιών σε μια ακολουθία n ανεξάρτητων δοκιμών Bernoulli με σταθερή πιθανότητα επιτυχίας p σε όλες τις δοκιμές, τότε η κατανομή της τυχαίας μεταβλητής X ονομάζεται διωνυμική κατανομή (binomial distribution) με παραμέτρους n και p και συμβολίζεται με $B(n, p)$ ή $X \sim B(n, p)$.

- Η συνάρτηση πιθανότητας είναι:

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

$$P(X = 0) = \binom{300000}{0} \cdot (2 \cdot 10^{-5})^0 (1 - 2 \cdot 10^{-5})^{300000} = 1 \cdot 1 \cdot (0.99998)^{300000} = 0.0024786$$

Κίνητρο - παράδειγμα



$$P(X = 2) = \binom{300000}{2} \cdot (2 \cdot 10^{-5})^2 (1 - 2 \cdot 10^{-5})^{299998} = \frac{300000!}{2! \cdot 299998!} \cdot (0.00002)^2 \cdot (0.99998)^{299998} =$$
$$= \frac{299999 \cdot 300000}{2} \cdot (0.00002)^2 \cdot (0.99998)^{299998} = 0.0446165.$$

$$P(X = 15) = \binom{300000}{15} \cdot (2 \cdot 10^{-5})^{15} (1 - 2 \cdot 10^{-5})^{299985} =$$
$$= \frac{300000!}{15! \cdot 299985!} \cdot (0.00002)^{15} \cdot (0.99998)^{299985} =$$
$$= \frac{299986 \cdot 299987 \cdot \dots \cdot 300000}{1 \cdot 2 \cdot 3 \cdot \dots \cdot 15} \cdot (0.00002)^{15} \cdot (0.99998)^{299985} = \dots = 0.0008912.$$

Εξαιρετικά δύσκολος ο υπολογισμός

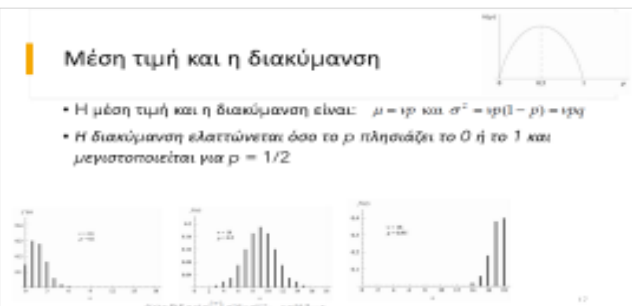
Οριακό θεώρημα Poisson

- Έστω ότι η τυχαία μεταβλητή X ακολουθεί τη διωνυμική κατανομή $B(n, p)$ με συνάρτηση πιθανότητας

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

Αν για $n \rightarrow +\infty$ το $p \rightarrow 0$ έτσι ώστε η μέση τιμή της X να συγκλίνει προς μια θετική σταθερά λ , δηλαδή, έτσι ώστε $np \rightarrow \lambda$, τότε

$$\lim_{n \rightarrow +\infty} \binom{n}{x} p^x (1-p)^{n-x} = e^{-\lambda} \frac{\lambda^x}{x!}, \quad x = 0, 1, 2, \dots$$



Η κατανομή Poisson

- Έστω X μια διακριτή τυχαία μεταβλητή με συνάρτηση πιθανότητας

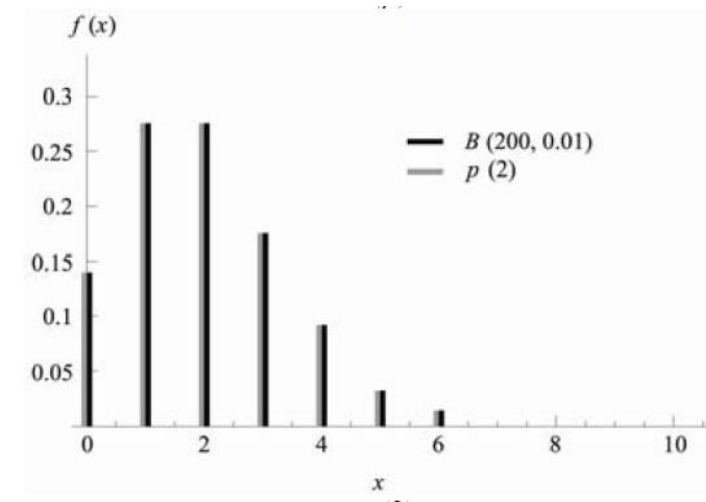
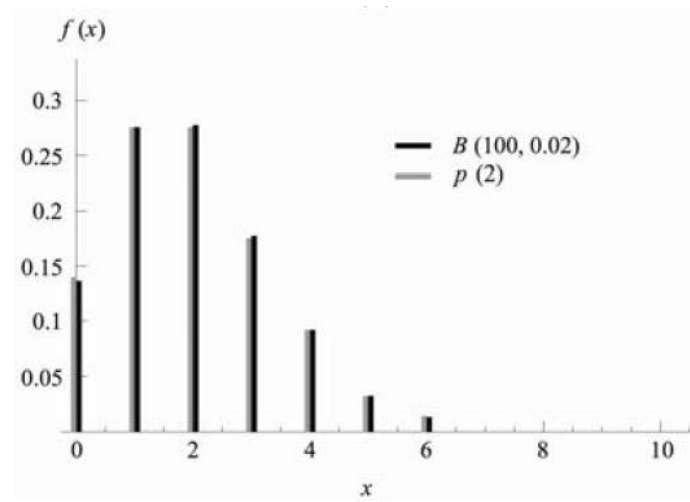
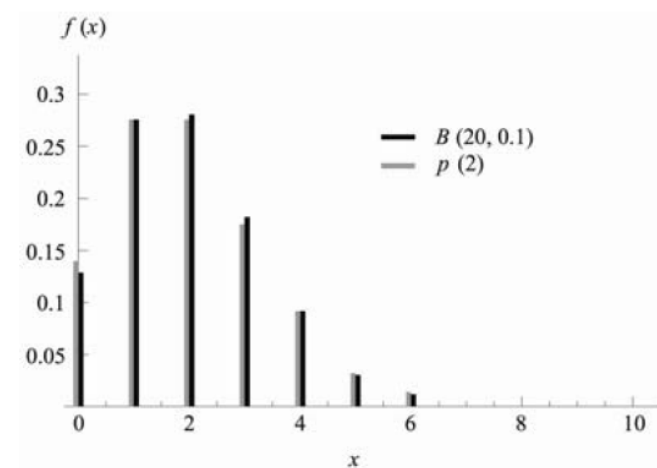
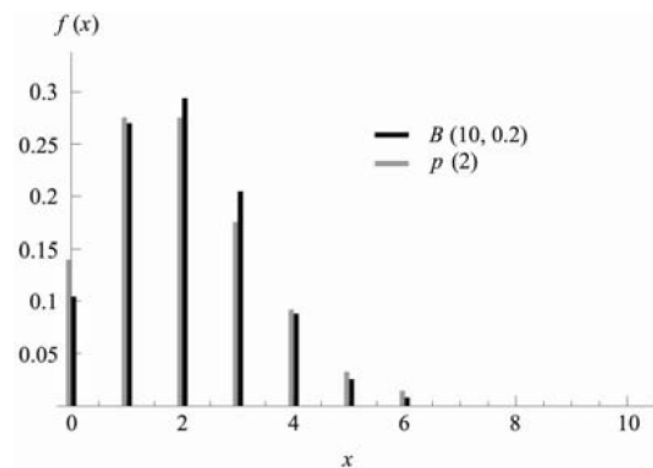
$$f(x) = P(X = x) = e^{-\lambda} \frac{\lambda^x}{x!}, \quad x = 0, 1, 2, \dots$$

όπου $\lambda > 0$. Η κατανομή της τυχαίας μεταβλητής X ονομάζεται κατανομή Poisson με παράμετρο λ και συμβολίζεται με $P(\lambda)$.

- Η **διωνυμική κατανομή προσεγγίζεται από την κατανομή Poisson** αν για n μεγάλο (θεωρητικά $n \rightarrow \infty$), η πιθανότητα επιτυχίας p συγκλίνει στο 0 (δηλ. $p \rightarrow 0$) έτσι ώστε η μέση τιμή της να συγκλίνει σε μια θετική σταθερά λ ($np \rightarrow \lambda$).

Σχόλια

- Πόσο μεγάλο πρέπει να είναι το n και πόσο μικρό το p ;
- $n \geq 20$ και $p \leq 10/n$
- $\lambda = np = 2$ όταν $n = 10, 20, 100, 200$ και $p = 0.2, 0.1, 0.02, 0.01$ αντίστοιχα



Μέση τιμή και διακύμανση

- Η μέση τιμή και η διακύμανση της κατανομής Poisson με παράμετρο λ , αντίστοιχα είναι

$$\mu = E(X) = \lambda$$

και

$$\sigma^2 = Var(X) = \lambda$$



Χρήσεις

- Χρησιμοποιείται για τη μοντελοποίηση «διωνυμικών καταστάσεων» όπου ενδιαφέρει ο αριθμός εμφανίσεων σπάνιων ενδεχομένων σε μεγάλους πληθυσμούς
- Γνωστή ως κατανομή των σπάνιων ενδεχομένων

Χρήσεις

- Ο αριθμός X των τροχαίων ατυχημάτων σε ένα τμήμα (με μεγάλη κυκλοφορία) του οδικού δικτύου μιας χώρας στη διάρκεια ενός Σαββατοκύριακου (ή μιας ημέρας, ή μιας εβδομάδας, ή ενός μήνα κτλ.)
- Ο αριθμός X των τυπογραφικών λαθών σε μια δακτυλογραφημένη σελίδα (ή σε ένα σύνολο σελίδων).

Χρήσεις

- Ο αριθμός X των κλήσεων στο help desk ενός μεγάλου Internet provider σε μια ημέρα (ή σε μια ώρα, ή σε μια εβδομάδα κτλ.).
- Ο αριθμός X των βλαβών μιας μηχανής σε μια ημέρα (ή σε μια εβδομάδα κτλ.).
- Ο αριθμός X των εργατικών ατυχημάτων που συμβαίνουν σε μια ημέρα σε συγκεκριμένη βιομηχανική ζώνη.
- Ο αριθμός X των ατόμων ενός πληθυσμού που ζουν περισσότερα από 100 χρόνια.

Χρήσεις

- Ο αριθμός X των παιδιών ενός πληθυσμού που θα γίνουν ψηλότερα από 1.95 μέτρα.
- Ο αριθμός X των πελατών ενός super market σε μια ημέρα, που θα αγοράσουν σοκολατάκια για σκύλους.
- Ο αριθμός X των φυσαλίδων σε υαλοπίνακα συγκεκριμένης επιφάνειας.
- Ο αριθμός X των επιβατών μιας αεροπορικής πτήσης που ενώ έχουν κάνει κράτηση θέσης δεν εμφανίζονται την ώρα αναχώρησης.

Πολυωνυμική κατανομή



Πολυωνυμική κατανομή

- Αναφέρεται σε πείραμα με $k \geq 2$ αμοιβαίως αποκλειόμενα δυνατά αποτελέσματα.
- Γενίκευση της διωνυμικής κατανομής.

Πολυωνυμική κατανομή

- Έστω ένα πείραμα τύχης που αποτελείται από n ανεξάρτητες πολυωνυμικές δοκιμές, όπου σε κάθε δοκιμή τα δυνατά αποτελέσματα είναι k , έστω $r_1, r_2, r_3, \dots, r_k$ και η πιθανότητα να εμφανισθεί το αποτέλεσμα r_i , $i = 1, 2, \dots, k$ σε μια (οποιαδήποτε) δοκιμή είναι p_i .
- Αν X_i ($i = 1, 2, \dots, k$) τυχαία μεταβλητή που εκφράζει πόσες φορές εμφανίσθηκε το αποτέλεσμα r_i στις n ανεξάρτητες επαναλήψεις της πολυωνυμικής δοκιμής, τότε η τυχαία μεταβλητή $(X_1, X_2, X_3, \dots, X_k)$ λέμε ότι ακολουθεί την πολυωνυμική κατανομή (polynomial distribution), με παραμέτρους n και $p_1, p_2, p_3, \dots, p_k$.
- Συμβολίζεται με $(X_1, X_2, X_3, \dots, X_k) \sim M(n, p_1, p_2, p_3, \dots, p_k)$

Πολυωνυμική κατανομή

Συνάρτηση πιθανότητας

$$P(X_1 = v_1, X_2 = v_2, \dots, X_k = v_k) = \frac{v!}{v_1! v_2! \dots v_k!} p_1^{v_1} p_2^{v_2} \dots p_k^{v_k}$$

$$\text{με } \sum_{i=1}^k v_i = v \text{ και } \sum_{i=1}^k p_i = 1$$

Μέση τιμή και διακύμανση

$$E(X_i) = vp_i \text{ και } \text{Var}(X_i) = vp_i(1 - p_i).$$

Παράδειγμα



- Σύμφωνα με ένα μοντέλο κληρονομικότητας, οι τρεις τύποι απογόνων A, B και Γ που προκύπτουν από μια ορισμένη διασταύρωση πειραματόζων, βρίσκονται σε αναλογία 9:3:4, αντίστοιχα.
- Αν στο πλαίσιο ενός πειράματος προέκυψαν από μια τέτοια διασταύρωση 64 απόγονοι, πόσοι αναμένεται να είναι τύπου A, πόσοι τύπου B και πόσοι τύπου Γ;

Παράδειγμα



- Η ταξινόμηση ενός απογόνου σε (ακριβώς) μια από τρεις κατηγορίες (τύπος Α, τύπος Β, τύπος Γ) είναι μια πολυωνυμική δοκιμή με $k = 3$ δυνατά αποτελέσματα.
- Ένα πείραμα που αποτελείται από $n = 64$ πολυωνυμικές δοκιμές η κάθε μια από τις οποίες έχει $k = 3$ δυνατά αποτελέσματα:
 - τύπος Α
 - τύπος Β
 - τύπος Γ

Παράδειγμα



- Η πιθανότητα πραγματοποίησης καθενός από τα τρία δυνατά αποτελέσματα είναι αντίστοιχα:
 - τύπος Α έχει πιθανότητα $p_1 = \frac{9}{16}$
 - τύπος Β έχει πιθανότητα $p_2 = \frac{3}{16}$
 - τύπος Γ έχει πιθανότητα $p_3 = \frac{4}{16}$
- Έστω $X_i, i = 1, 2, 3$ τυχαίες μεταβλητές που αντίστοιχα εκφράζουν πόσες φορές στις 64 επαναλήψεις εμφανίσθηκε απόγονος τύπου Α, τύπου Β και τύπου Γ.

Πολυωνυμική κατανομή

- Έστω ένα πείραμα τύχης που αποτελείται από n ανεξάρτητες πολυωνυμικές δοσμές, όπου σε κάθε δοσμή τα δυνατά αποτελέσματα είναι k , έστω r_1, r_2, \dots, r_k και η πιθανότητα να εμφανισθεί το αποτέλεσμα $r_i, i = 1, 2, \dots, k$ σε μια (στοιχειώδη) δοσμή είναι p_i .
- Αν $X_i (i = 1, 2, \dots, k)$ τυχαία μεταβλητή που εκφράζει πόσες φορές εμφανισθεί το αποτέλεσμα r_i στις n ανεξάρτητες επαναλήψεις της πολυωνυμικής δοσμής, τότε η τυχαία μεταβλητή (X_1, X_2, \dots, X_k) λέμε ότι ακολουθεί την πολυωνυμική κατανομή (polynomial distribution) με παραμέτρους n και p_1, p_2, \dots, p_k .
- Συμβολίζεται με $(X_1, X_2, \dots, X_k) \sim \text{PM}(n, p_1, p_2, \dots, p_k)$

Παράδειγμα



- Πόσοι αναμένεται να είναι τύπου Α, πόσοι τύπου Β και πόσοι τύπου Γ;
- Ζητάμε την μέση τιμή
$$E(X_1) = 64(9/16) = 36$$
$$E(X_2) = 64(3/16) = 12$$
$$E(X_3) = 64(4/16) = 16$$
- Από τους 64 απογόνους οι 36 αναμένεται να είναι τύπου Α, οι 12 τύπου Β και οι 16 τύπου Γ.

Πολυωνομική κατανομή

Συνάρτηση πιθανότητας

$$P(X_1 = v_1, X_2 = v_2, \dots, X_k = v_k) = \frac{n!}{v_1! v_2! \dots v_k!} p_1^{v_1} p_2^{v_2} \dots p_k^{v_k}$$

με $\sum_{i=1}^k v_i = n$ και $\sum_{i=1}^k p_i = 1$

Μέση τιμή και διακύμανση

$$E(X_i) = np_i \text{ και } \text{Var}(X_i) = np_i(1 - p_i)$$

Backup

Συνδυασμοί (γενικά)

- Έστω X σύνολο με n στοιχεία και ακέραιος $k \leq n$. Συνδυασμός των n στοιχείων ανά k ονομάζεται κάθε μη διατεταγμένο υποσύνολο του X με k στοιχεία.
- Το πλήθος των διαφορετικών συνδυασμών είναι

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