ΑΡΙΣΤΟΤΕΛΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΟΝΙΚΗΣ

ΓΕΩΠΟΝΙΚΗ ΣΧΟΛΗ Α.Π.Θ.

Εφαρμογές Στατιστικής: Πιθανότητες-Κατανομές

Επιστημονική Επιμέλεια:

Δρ. Γεώργιος Μενεξές

Τομέας Φυτών Μεγάλης Καλλιέργειας
και Οικολογίας
Εργαστήριο Γεωργίας

Η παρουσίαση βασίζεται σε υλικό από το έργο του:

- Μάνος, Β. (1990). Γεωργική Οικονομική Στατιστική. Εκδόσεις ΖΗΤΗ, Θεσσαλονίκη.

Άσκηση 1

Παίρνουμε τυχαία ένα δείγμα 10 γεωργών και τους ρωτάμε αν έχουν δικό τους ελκυστήρα. Είναι γνωστό από προηγούμενες μελέτες ότι στην περιοχή που γίνεται η έρευνα το 65% των γεωργών έχουν δικό τους ελκυστήρα.

α) Ποια είναι η πιθανότητα σε τυχαίο δείγμα 10 γεωργών οι 8 να έχουν δικό τους ελκυστήρα;

β) Ποια είναι η πιθανότητα να ρωτήσουμε 3 γεωργούς για να προκύψει η πρώτη θετική απάντηση;

γ) Ποια είναι η πιθανότητα να ρωτήσουμε 5 γεωργούς για να πάρουμε 3 θετικές απαντήσεις;

Λύση

α) Η ζητούμενη πιθανότητα υπολογίζεται από τον τύπο της Διωνυμικής Κατανομής.

$$P(r) = \binom{n}{r} p^r (1-p)^{n-r}, \quad r = 0, 1, 2, \ldots, n$$

Συνεπώς:

Έστω X η τυχαία μεταβλητή που μετρά τον αριθμό των γεωργών που έχουν δικό τους θεριστήρα. Τότε η X ακολουθεί τη Διωνυμική Κατανομή με παραμέτρους $p = 0,65$ και $n = 10$.

$P($8 θετικές απαντήσεις στους 10 γεωργούς$) = P(X=8) =$

$$= \binom{10}{8} 0,65^{8} \cdot 0,35^{10-8} = \frac{10!}{2!\,8!} \cdot 0,65^{8} \cdot 0,35^{2} = 0,176.$$

β) Η ζητούμενη πιθανότητα υπολογίζεται από τον τύπο της Γεωμετρικής Κατανομής:

$$P(r) = (1-p)^r p \qquad r = 0, 1, 2, \dots$$

Έστω X η τυχαία μεταβλητή που μετρά τον αριθμό των γεωργών ώσπου να προκύψει η πρώτη θετική απάντηση. ↳ που πρέπει να ρωτήσουμε

P(3 γεωργοί για την πρώτη θετική απάντηση) =

$$= P(X=3) = 0{,}35^3 \cdot 0{,}65 = 0{,}028 \text{ (ή } 2{,}8\%\text{)}$$

γ) Η ζητούμενη πιθανότητα υπολογίζεται από τον τύπο της Αρνητικής Διωνυμικής Κατανομής.

$$P(r) = \binom{r-1}{k-1} p^{k} (1-p)^{r-k}, \quad r = k, k+1, \dots$$

Έστω X η τυχαία μεταβλητή που μετρά τον αριθμό των γεωργών που πρέπει να ρωτήσουμε ώστε να έχουμε k θετικές απαντήσεις.

P(5 γεωργοί για 3 θετικές απαντήσεις) =

$$= P(X=5) = \binom{5-1}{3-1} 0,65^{3} \cdot 0,35^{5-3} =$$

$$= \frac{4!}{2!\,2!} 0,65^{3} \cdot 0,35^{2} = 0,20. \quad (\text{ή } 20\%.)$$

Άσκηση (2).

Μια μεγάλη ποσότητα μήλων περιέχει 8% χαλασμένα μήλα. Παίρνουμε τυχαία 4 μήλα. Να βρεθούν:

α) Η πιθανότητα ακριβώς ένα μήλο να είναι χαλασμένο.

β) Η πιθανότητα κανένα μήλο να είναι χαλασμένο.

γ) Η πιθανότητα τουλάχιστον ένα μήλο να είναι χαλασμένο.

δ) Η μέση τιμή και η διακύμανση (παραλλακτικότητα) του αριθμού των χαλασμένων μήλων.

Λύση

α) Έστω X η τυχαία μεταβλητή που μετρά τον αριθμό των χαλασμένων μήλων.

$$P(X=1) = \binom{4}{1} 0{,}02^1 \cdot 0{,}098^3 = \frac{4!}{3!\,1!}\, 0{,}02 \cdot 0{,}98^3 =$$

$$= 0{,}0753 \text{ (ή } 7{,}5\% \text{ περίπου).}$$

β) $$P(X=0) = \binom{4}{0} 0{,}02^0 \cdot 0{,}098^4 = \frac{4!}{3!\,1!}\, 0{,}098^4 =$$

$$= 0{,}9224 \text{ (ή } 92{,}2\% \text{ περίπου).}$$

γ) $$P(X \geq 1) = P(X=1) + P(X=2) + P(X=3) + P(X=4).$$

2ος Τρόπος

Υπολογίζουμε τις 4 πιθανότητες.

$P(X=1)=0{,}0753$, $P(X=2)=0{,}0083$, $P(X=3)=0{,}0026$

$P(X=4)=0{,}0000002$. Επομένως η ζητούμενη

πιθανότητα είναι:

$P(X\geq 1)=0{,}0753+0{,}0083+0{,}0026+0{,}0000002$

$=0{,}0792$. (ή $7{,}9\%$ περίπου).

2ος Τρόπος.

$$P(X \geq 1) = 1 - P(X < 1) = 1 - P(X = 0)$$

$$= 1 - 0,9224 = 0,0776 \quad (\text{ή } 7,8\% \text{ περίπου}).$$

Η απόκλιση στον υπολογισμό οφείλεται στη στρογγυλοποίηση των αριθμών.

δ) $\mu = \sum r P(r) = 0 \cdot P(X=0) + 1 \cdot P(X=1) + 2P(X=2)$
$+ 3P(X=3) + 4P(X=4) =$

$= 0{,}0753 + 0{,}0046 + 0{,}0048 + 0{,}0000008 =$
$= 0{,}0847$

$\sigma^2 = \sum r^2 P(r) - \mu^2 = 0^2 \cdot P(X=0) + 1^2 P(X=1) + 2^2 P(X=2)$
$+ 3^2 P(X=3) + 4^2 \cdot P(X=4) - 0{,}0847^2$

$= 0{,}0917.$

Επομένως $\sigma = \sqrt{\sigma^2} \Rightarrow \sigma = \sqrt{0{,}0917} = 0{,}3028.$

Άσκηση (3)

Οι πιθανότητες β και γ της προηγούμενης άσκησης να υπολογισθούν με τον τύπο της προσέγγισης της Διωνυμικής Κατανομής από την Κανονική.

Λύση:

(*) Θα πρέπει $n \geq 30$ και $np \geq 5$.

Ζητάμε τις πιθανότητες $P(X \leq 0)$ και $P(X \geq 1)$.

$p = 0{,}02$ $1-p = 0{,}98$, $n = 4$ $np = 0{,}08$ επομένως η εκτίμηση δεν θα είναι καλή, διότι δεν ισχύουν οι προϋποθέσεις (*).

Υπολογίζουμε

$$np(1-p) = 4 \cdot 0,02 \cdot 0,98 = 0,0784.$$

$$\sqrt{np(1-p)} = 0,28$$

Γνωρίζουμε ότι:

$$P(X \leq r) = \Phi\left(\frac{r-np}{\sqrt{np(1-p)}}\right)$$ και

$P(X \geq r) = 1 - P(X$ [redacted] $r)$

από τους πίνακες της κανονικής κατανομής

Συνεπώς: $$P(X \leq 0) = \Phi\left(\frac{0-0,08}{0,28}\right) = \Phi(-0,29) =$$

$$= 1 - \Phi(0,29) = 1 - 0,6141 = 0,3859 \text{ (ή } 38,6\%)$$

$$P(X \geq 1) = 1 - \Phi\left(\frac{1-0,08}{0,28}\right) = 1 - \Phi(3,29)$$

$$= 1 - 0,99952 = 0,00048.$$

Παράδειγμα 1: Παίρνουμε τυχαία ένα δείγμα 10 γεωργών και τους ρωτάμε αν έχουν δικό τους ελκυστήρα. Είναι γνωστό από προηγούμενες μελέτες, ότι στην περιοχή που γίνεται η έρευνα το 65% των γεωργών έχουν δικό τους ελκυστήρα και το 35% δεν έχουν.

Στις δοκιμές αυτές του Bernoulli αν με p παραστήσουμε την πιθανότητα θετικής απάντησης τότε

$$p = 0,65 \qquad \text{και} \qquad 1-p = 0,35$$

Η πιθανότητα να έχουν 8 από τους 10 γεωργούς δικό τους ελκυστήρα βρίσκεται από τον τύπο της διωνυμικής κατανομής

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

Έτσι

P(8 θετικές απαντήσεις στους 10 γεωργούς που ρωτήσαμε)=

$$= \binom{10}{8} 0,65^{8} \cdot 0,35^{10-8} =$$

$$= \frac{10!}{2!8!} 0,65^{8} \cdot 0,35^{2} = 45.(0,03186).(0,1225) =$$

$$= 0,176$$

Η πιθανότητα να ρωτήσουμε 3 γεωργούς για να προκύψει η πρώτη θετική απάντηση βρίσκεται από τον τύπο της Γεωμετρικής κατανομής

$$p(r) = (1-p)^{r} p, \qquad r = 0,1,2,\dots$$

Έτσι

P(3 ερωτήσεις για την πρώτη θετική απάντηση)=

$$= 0,35^{3} \cdot 0,65 = 0,028$$

Η πιθανότητα να ρωτήσουμε 5 γεωργούς για να πάρουμε 3 θετικές απαντήσεις βρίσκεται από τον τύπο της Αρνητικής διωνυμικής κατανομής

$$p(r) = \binom{r-1}{\kappa-1} p^{\kappa}(1-p)^{r-\kappa}, \qquad r=\kappa,\kappa+1,\dots$$

Έτσι

P(5 ερωτήσεις για 3 θετικές απαντήσεις)=

$$= \binom{5-1}{3-1} 0,65^3 \cdot 0,35^{5-3} = \frac{4!}{2!2!} 0,65^3 \cdot 0,35^2 = 0,20$$

<u>Παράδειγμα 2</u>: Μια μεγάλη ποσότητα μήλων περιέχει 2% χαλασμένα μήλα. Παίρνουμε 4 μήλα τυχαία. Να βρεθούν:

α) Η πιθανότητα ακριβώς ένα μήλο να είναι χαλασμένο

β) Η πιθανότητα κανένα μήλο δεν είναι χαλασμένο

γ) Η πιθανότητα τουλάχιστο ένα μήλο είναι χαλασμένο

δ) Η μέση τιμή και η διακύμανση του αριθμού των χαλασμένων μήλων

Είναι φανερό πως πρόκειται για δοκιμασίες Bernoulli. Αν με p συμβολίσουμε την πιθανότητα το μήλο που παίρνουμε είναι χαλασμένο τότε

$p = 0,02$ και $1-p = 0,98$

Η πιθανότητα ακριβώς ένα μήλο είναι χαλασμένο βρίσκεται από τον τύπο της Διωνυμικής κατανομής

P(1 μήλο στα 4 που πήραμε είναι χαλασμένο)=

$$p(1) = \binom{4}{1} 0,02^1 \cdot 0,98^{4-1} = \frac{4!}{3!1!} 0,02 \cdot 0,98^3 = 0,0753$$

Η πιθανότητα κανένα μήλο δεν είναι χαλασμένο βρίσκεται από τον ίδιο τύπο

P(0 μήλα στα 4 που πήραμε είναι χαλασμένα)=

$$p(0) = \binom{4}{0} 0,02^0 \cdot 0,98^{4-0} = \frac{4!}{4!1!} 0,98^4 = 0,9224$$

Η πιθανότητα τουλάχιστο ένα μήλο είναι χαλασμένο είναι το άθροισμα των πιθανοτήτων $p(1)+p(2)+p(3)+p(4)$

όμως

$$p(1) = 0,0753$$

$$p(2) = \binom{4}{2} 0,02^2 \cdot 0,98^{4-2} = \frac{4!}{2!2!} 0,02^2 \cdot 0,98^2 = 0,0023$$

$$p(3) = \binom{4}{3} 0,02^3 \cdot 0,98^{4-3} = \frac{4!}{3!1!} 0,02^3 \cdot 0,98 = 0,0016$$

$$p(4) = \binom{4}{4} 0,02^4 \cdot 0,98^{4-4} = \frac{4!}{4!0!} 0,02^4 = 0,0000002$$

οπότε $p(1)+p(2)+p(3)+p(4) = 0,0792$

Επίσης έχουμε

$$\mu = \Sigma r p(r) = 0.p(0)+1.p(1)+2.p(2)+3.p(3)+4.p(4) =$$
$$= 0,0753+0,0046+0,0048+0,0000008 = 0,0847$$

$$\sigma^2 = \Sigma r^2 p(r) - \mu^2 = 0^2.p(0)+1^2.p(1)+2^2.p(2)+3^2.p(3)+4^2.p(4) - \mu^2 =$$
$$= 0,0753+0,0092+0,0144+0,0000032-0,0847^2 = 0,0917$$

$$\sigma = 0,3029$$

<u>Παράδειγμα 3</u>: Οι πιθανότητες β και γ του προηγούμενου παραδείγματος να υπολογισθούν με τον τύπο της προσέγγισης της Διωνυμικής κατανομής.

Αν X είναι η τυχαία μεταβλητή με Διωνυμική κατανομή που δίνει τον αριθμό των χαλασμένων μήλων, τότε οι παραπάνω πιθανότητες είναι οι

$$P(X \leq 0) \qquad \text{και} \qquad P(X \geq 1)$$

Στην περίπτωσή μας είναι

$$p = 0{,}02 \qquad 1-p = 0{,}98 \qquad n = 4 \qquad np = 0{,}08$$

$$np(1-p) = 0{,}0784 \qquad \text{και} \quad \sqrt{np(1-p)} = 0{,}28$$

Χρησιμοποιούμε τους τύπους

$$P(X \leq r) = \Phi\left(\frac{r-np}{\sqrt{np(1-p)}}\right)$$

και

$$P(X \geq r) = 1 - P(X \leq r)$$

οπότε έχουμε

$$P(X \leq 0) = \Phi\left(\frac{0-0{,}08}{0{,}28}\right) = \Phi(-0{,}29) =$$

$$= 1-\Phi(0{,}29) = 1-0{,}6141 = 0{,}3859$$

$$P(X \geq 1) = 1-\Phi\left(\frac{1-0{,}08}{0{,}28}\right) = 1-\Phi(3{,}29) = 1-0{,}99952 = 0{,}00048$$