

Uporabna statistika

Gregor Dolinar

Fakulteta za elektrotehniko
Univerza v Ljubljani

16. oktober 2013

Geometrijska porazdelitev

Primer

Mama hoče imeti sina, oče hoče imeti čim manj otrok. Kolikšna je verjetnost, da bosta imela 4 otroke?

Slučajna spremenljivka X je enaka številu otrok.

$$P[X = 4] = P[DDDF] = 0.485^3 \cdot 0.515 = 0.059.$$

Definicija

Naj bo poskus sestavljen iz Bernoullijevih poskusov, ki so med seboj neodvisni in vsak Bernoullijevi poskus ima samo 2 možna izida, verjetnost uspeha v vsakem izmed Bernoullijevih poskusov je enaka p .

Slučajno spremenljivko, ki je enaka številu Bernoullijevih poskusov do prvega uspešnega poskusa, imenujemo geometrijska slučajna spremenljivka.

Definicija

Gostota verjetnosti geometrijske slučajne spremenljivke X je enaka

$$f(x) = (1 - p)^{x-1} \cdot p, \quad x = 1, 2, \dots$$

Torej je

$$X : \begin{pmatrix} 1 & \dots & k & \dots \\ p & \dots & (1 - p)^{k-1} \cdot p & \dots \end{pmatrix}$$

Primer. Digitalni prenos signalov. Verjetnost, da je signal slab je 0.1. Naj bo vrednost slučajne spremenljivke X enaka številu prenosov do prvega slabega signala. Kolikšne so verjetnosti:

$$P[X = 1], P[X = 10], P[X = 20]?$$

$$P[X = 1] = 0.1$$

$$P[X = 10] = 0.9^9 \cdot 0.1 = 0.03874$$

$$P[X = 20] = 0.9^{19} \cdot 0.1 = 0.01351$$

Trditev

Če je X geometrijska slučajna spremenljivka s parametrom p , potem je

$$\mu = E(X) = \frac{1}{p}.$$

Varianca je

$$V(X) = \frac{1-p}{p^2}.$$

Verjetnost za napako ostaja enaka. Sistem se ne obrablja. Sistem nima spomina.

Verjetnost, da bo v danem trenutku število prenosov do prvega slabega signala enako pet, je enaka, če je bilo do danega trenutka uspešno prenešenih sto signalov ali pa nobeden.

Negativna binomska porazdelitev

Primer

Mama hoče imeti dva sina, oče hoče imeti čim manj otrok.

Kolikšna je verjetnost, da bosta imela 4 otroke?

Slučajna spremenljivka X je enaka številu otrok.

$$\begin{aligned}P[X = 4] &= P[FDDF] + P[DFDF] + P[DDFF] \\ &= 3 \cdot 0.485^2 \cdot 0.515^2 = 0.187.\end{aligned}$$

Definicija

Naj bo poskus sestavljen iz Bernoullijevih poskusov, ki so med seboj neodvisni in vsak Bernoullijevi poskus ima samo 2 možna izida, verjetnost uspeha v vsakem izmed Bernoullijevih poskusov je enaka p .

Slučajno spremenljivko, ki je enaka številu Bernoullijevih poskusov do r -tega uspešnega poskusa, imenujemo negativna binomska slučajna spremenljivka.

Definicija

Gostota verjetnosti negativne binomske slučajne spremenljivke X je enaka

$$f(x) = \binom{x-1}{r-1} (1-p)^{x-r} \cdot p^r, \quad x = r, r+1, \dots$$

Trditev

Če je X negativna binomska slučajna spremenljivka s parametroma p in r , potem je

$$\mu = E(X) = \frac{r}{p}.$$

Varianca je

$$V(X) = \frac{r(1-p)}{p^2}.$$

Elektronska tehtnica na avtomatski polnilni liniji ustavi proizvodnjo, če so tri embalaže napolnjeni manj od zahtevane vrednosti. Verjetnost, da je polnjenje prelahko, je 0.001. Vsa polnjenja so neodvisna.

(a) Kolikšna je verjetnost, da se proizvodnja ustavi po 3, po 100, po 1000, po 3000 napolnjenih embalažah?

(b) Koliko embalaž je v povprečju napolnjenih, preden se proizvodnja ustavi?

Hipergeometrijska porazdelitev

Definicija

Naj bo v množici z N elementi K elementov slabih. Iz množice naključno izberemo n elementov. Slučajno spremenljivko, ki je enaka številu slabih elementov v vzorcu, imenujemo hipergeometrijska slučajna spremenljivka.

Definicija

Gostota verjetnosti hipergeometrijske slučajne spremenljivke X je enaka

$$f(x) = \frac{\binom{K}{x} \binom{N-K}{n-x}}{\binom{N}{n}},$$

pri čemer je $x = \max\{0, n + K - N\}, \dots, \min\{K, n\}$ in $p = \frac{K}{N}$.

Trditev

$$E(X) = np, \quad V(X) = np(1 - p) \frac{N - n}{N - 1}$$

Primer

V skladišču je 100 aparatov, ki so bili proizvedeni v matični tovarni in 20 aparatov, ki so bili proizvedeni v obratu v tujini.

V razstavnici so pripeljali 3 aparate iz skladišča. Kolikšna je verjetnost, da so bili vsi proizvedeni v matični tovarni?

$$N = 120, \quad K = 100, \quad n = 3, \quad x = 3.$$

Poissonova porazdelitev

Naj bo X binomska slučajna spremenljivka za n Bernoullijevih poskusov z verjetnostjo p . Označimo $E(X) = np = \lambda$. Potem je

$$\begin{aligned} P[X = x] &= \binom{n}{x} p^x (1-p)^{n-x} \\ &= \binom{n}{x} \left(\frac{\lambda}{n}\right)^x \left(1 - \frac{\lambda}{n}\right)^{n-x} \end{aligned}$$

Trditev

$$\lim_{n \rightarrow \infty} P[X = x] = \frac{e^{-\lambda} \lambda^x}{x!}, \quad x = 0, 1, \dots$$

Poissonov proces. Na danem intervalu se naključno pojavljajo števila. V povprečju se jih pojavi λ . Denimo, da lahko interval razdelimo na podintervale, tako da velja:

- ▶ verjetnost, da se na podintervalu pojavi več kot 1 število, je nič
- ▶ verjetnost, da se na podintervalu pojavi število, je enaka za vse podintervale in sorazmerna z dolžino podintervala
- ▶ če se število pojavi na nekem podintervalu, je to neodvisno od ostalih podintervalov

Definicija

Slučajna spremenljivka X , ki je enaka številu pojavljenih števil na intervalu pri Poissonovem procesu, se imenuje Poissonova slučajna spremenljivka s parametrom λ .

Definicija

Gostota verjetnosti Poissonove slučajne spremenljivke X je enaka

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

Trditev

$$E(X) = \lambda, \quad V(X) = \lambda$$

Primer. Pri proizvodnji optičnega diska se na 1 kvadratnem centimetru v povprečju pojavi 0.1 delec nečistoče. Proizvajamo disk s površino 100 kvadratnih centimetrov.

Kolikšna je verjetnost, da se na disku pojavi 12 delcev nečistoče? Optični disk zavržemo, če ima več kot 3 delce nečistoče. Kolikšna je verjetnost, da izdelani disk zavržemo?

Nekaj primerov

Pri proizvodnji izdelka je 1 % izdelkov neprimernih.

- ▶ Kako velik je vzorec, v katerem je v povprečju 5 izdelkov neprimernih?
- ▶ Koliko je standardni odklon za število neprimernih izdelkov v vzorcu v zgornjem primeru, če je bilo proizvedenih 1000 izdelkov?

(hipergeometrijska porazdelitev)

V povprečju se na magnetnem disku, na katerega shranjujemo podatke, pojavi 1 napaka na vsakih 10^5 bitov. Disk je sestavljen iz sektorjev, ki so veliki 4096 bytov.

- ▶ Kolikšna je verjetnost, da je v posameznem sektorju več kot 1 napaka?
- ▶ Koliko sektorjev moramo v povprečju pregledati, da odkrijemo sektor z napako?

(Poissonova porazdelitev, 1 byte je 8 bitov, $\lambda = \frac{4096 \cdot 8}{10^5} = 0.32768$ za en sektor)

Proizvajalec, ki ima naslove od 500 kupcev, želi opraviti anketo. Pričakuje, da bo anketo izpolnilo 25 % kupcev.

- ▶ Kolikšna je verjetnost, da bo dobil vsaj 2 izpolnjeni ankete, če pošlje anketo 10 kupcem?
- ▶ Kolikšna je verjetnost, da bo dobil vsaj 5 izpolnjenih anket, če pošlje anketo 25 kupcem?

(binomska porazdelitev)