4.

Primjeri diskretnih razdioba

Sadržaj poglavlja

U ovom ćemo se poglavlju upoznati s najvažnijim diskretnim slučajnim varijablama: geometrijskom, binomnom i Poissonovom. Opisat ćemo pripadne razdiobe, objasniti u kojim se modelima one pojavljuju, izračunati njihove karakteristične funkcije i pripadne numeričke karakteristike.

4.1. Geometrijska razdioba

Neka je pri izvođenju nekog pokusa p vjerojatnost realizacije događaja A. Ponavljamo taj pokus u nepromijenjenim uvjetima *do prve realizacije* tog događaja. Neka slučajna varijabla X mjeri broj pokusa *u kojem se realizirao događaj* A. Onda kažemo da X ima **geometrijsku razdiobu** s parametrom p i pišemo $X \sim \mathscr{G}(p)$.

Odredimo zakon razdiobe ove slučajne varijable. Najprije, X poprima vrijednosti u skupu $\{1, 2, 3, \dots\}$. Odredimo vjerojatnost

$$p_k = \boldsymbol{P}(X = k).$$

Ako se realizirao događaj $\{X = k\}$, to znači da se u prvih $k-1$ pokusa A nije pojavio, a pojavio se u k-tom pokusu, pa je

$$p_k = \boldsymbol{P}(X = k) = p(1-p)^{k-1}.$$

Stavimo $q := 1 - p$. Primjetimo da vrijedi

$$\boldsymbol{P}(X > k) = (1-p)^k = q^k$$

h jer se tada događaj A nije ostvario u prvih k pokusa.

4.1.1. Karakteristična funkcija, očekivanje i disperzija

Odredimo najprije karakterističnu funkciju geometrijske razdiobe:

$$\vartheta(t) = \sum_{k=1}^{\infty} e^{itk} \cdot pq^{k-1} = pe^{it} \sum_{k=0}^{\infty} (qe^{it})^k = \frac{pe^{it}}{1 - qe^{it}}.$$

Iskoristimo tu funkciju da izračunamo očekivanje ove razdiobe. Vrijedi

$$\vartheta'(t) = \frac{ipe^{it}}{(1 - qe^{it})^2}$$

pa je

$$\boldsymbol{E}(X) = -i\vartheta'(0) = \frac{p}{(1-q)^2} = \frac{1}{p}.$$

Dakle, očekivanje je jednako recipročnoj vrijednosti parametra — vjerojatnosti p. Taj je rezultat u skladu s iskustvom. Pri bacanju jedne kocke, vjerojatnost pojavljivanja šestice jednaka je $p = \frac{1}{6}$. Broj bacanja kocke do pojave šestice mjeri slučajna varijabla koja ima geometrijsku razdiobu. Očekivani broj ponavljanja jednak je $\boldsymbol{E}(X) = 1/p = 6$.

Kolika je vjerojatnost da će se šestica zaista pojaviti u tih prvih šest bacanja? Odgovor na to pitanje daje zakon razdiobe slučajne varijable:

$$\boldsymbol{P}(X \leqslant 6) = 1 - \boldsymbol{P}(X > 6) = 1 - \left(\frac{5}{6}\right)^6 = 0.665\,.$$

Teorem 4.1. Odsustvo pamćenja geometrijske razdiobe

Slučajna varijabla X koja poprima vrijednosti u skupu $\{1, 2, 3, \dots\}$ ima geometrijsku razdiobu onda i samo onda ako vrijedi za sve $k, m \geqslant 1$

$$\boldsymbol{P}(X = k + m \mid X > k) = \boldsymbol{P}(X = m). \tag{1}$$

Dokaz. Jedan je smjer jednostavan: ako X ima geometrijsku razdiobu, onda je

$$\begin{aligned}\boldsymbol{P}(X = k + m \mid X > k) &= \frac{\boldsymbol{P}(X = k + m, X > k)}{\boldsymbol{P}(X > k)} = \frac{\boldsymbol{P}(X = k + m)}{\boldsymbol{P}(X > k)} \\ &= \frac{p(1-p)^{k+m-1}}{(1-p)^k} = p(1-p)^{m-1} = \boldsymbol{P}(X = m).\end{aligned}$$

Pokažimo obrat. Iz (1) slijedi, zbrajanjem svih nejednakosti po prirodnim brojevima većim od m:

$$\boldsymbol{P}(X > k + m \mid X > k) = \boldsymbol{P}(X > m).$$

Uvjetna vjerojatnost na lijevoj strani može se napisati u obliku:

$$\boldsymbol{P}(X > k + m \mid X > k) = \frac{\boldsymbol{P}(X > k + m, X > k)}{\boldsymbol{P}(X > k)} = \frac{\boldsymbol{P}(X > k + m)}{\boldsymbol{P}(X > k)}.$$

Tako smo dobili identitet

$$\boldsymbol{P}(X > k + m) = \boldsymbol{P}(X > k)\boldsymbol{P}(X > m).$$

To znači da funkcija $Q(k) := \boldsymbol{P}(X > k)$ zadovoljava funkcionalnu jednadžbu

$$Q(k + m) = Q(k)Q(m), \qquad \forall k, m \geqslant 0. \tag{2}$$

Nadalje, vrijedi $Q(0) = \boldsymbol{P}(X > 0) = 1$, $Q(1) = \boldsymbol{P}(X > 1) = 1 - p =: q$. Stavimo u (2) $k = 1$:

$$Q(m + 1) = Q(1)Q(m) \implies Q(m + 1) = q \cdot Q(m).$$

Iterirajući ovu relaciju dobivamo

$$Q(k) = q^k Q(0) = q^k.$$

Dakle, $P(X > k) = q^k$. Zato je

$$\boldsymbol{P}(X = k) = \boldsymbol{P}(X > k - 1) - \boldsymbol{P}(X > k) = q^{k-1} - q^k = (1 - q)q^{k-1} = pq^{k-1}.$$

Time smo dokazali tvrdnju. ◄

— ◆ —

Na što ovaj teorem ukazuje? Rekli smo da X mjeri broj realizacija pokusa *do pojavljivanja* nekog događaja. Ako promatramo bacanje kocke i šestica se nije pojavila u prvih pet bacanja, Kolika je vjerojatnost da će se ona pojaviti u sljedeća dva bacanja? Koliki je očekivani broj bacanja do pojave šestice u tom trenutku?

Odgovori na ova pitanja su: ista (isti) kao i na početku bacanja. Niti se mijenjaju vjerojatnosti, niti se mijenja očekivani broj bacanja. Ako se šestica nije pojavila u prvih pet bacanja, očekivani broj (novih) bacanja do njezine pojave je opet jednak 6. Kažemo da *geometrijska razdioba nema pamćenje*.

4.1.2. Pokusi s dva nezavisna obilježja

Pretpostavimo da pri realizaciji pokusa promatramo dva nezavisna događaja A_1 i A_2. Pokus ponavljamo do realizacije *bilo kojeg od njih*.

Opišimo ovaj pokus koristeći aparat slučajnih varijabli.

Neka slučajna varijabla X_1 prati pojavljivanje događaja A_1. Ona ima geometrijsku razdiobu s parametrom $p_1 = \boldsymbol{P}(A_1)$.

Na isti način, slučajna varijabla X_2 koja bilježi realizaciju događaja A_2 ima geometrijsku razdiobu s parametrom $p_2 = \boldsymbol{P}(A_2)$.

Te dvije slučajne varijable su nezavisne, jer su A_1 i A_2 nezavisni.

Slučajna varijabla X koja registrira prvo pojavljivanje bilo kojeg od ovih događaja može se zapisati formulom $X = \min(X_1, X_2)$. Dokazat ćemo da ona ima također geometrijsku razdiobu i odrediti njezin parametar.

Vrijedi, zbog nezavisnosti

$$\begin{aligned}\boldsymbol{P}(X > k) &= \boldsymbol{P}(X_1 > k,\ X_2 > k) \\ &= \boldsymbol{P}(X_1 > k)\boldsymbol{P}(X_2 > k) = q_1^k q_2^k =: q^k\end{aligned}$$

gdje je $q = q_1 q_2$. Zato

$$\begin{aligned}\boldsymbol{P}(X = k) &= \boldsymbol{P}(X > k - 1) - \boldsymbol{P}(X > k) \\ &= q^{k-1} - q^k = pq^{k-1}\end{aligned}$$

te X ima geometrijsku razdiobu s parametrom $p = 1 - q = 1 - (1 - p_1)(1 - p_2)$.

Korisno je izračunati vjerojatnost događaja $A = A_1 \cup A_2$ i usporediti s ovim rezultatom.

Primjer 4.1.

Izlazna centrala poduzeća ima N linija. U jednom trenutku, sve su linije zauzete. Duljina razgovora, mjerena u jedinicama vremena, ima geometrijsku razdiobu s očekivanjem μ. Izračunaj očekivano vrijeme do prve slobodne linije.

► Neka su $X_1, \dots, X_n$ duljine razgovora preko pojedinih linija. Tad je $X = \min\{X_1, \dots, X_n\}$. Ako je p parametar geometrijske razdiobe, onda vrijedi $\mu = \dfrac{1}{p}$ te je $p = \dfrac{1}{\mu}$. Slučajna varijabla X ima i sama geometrijsku razdiobu s parametrom $1 - q^N$. Njezino je očekivanje

$$\boldsymbol{E}(X) = \frac{1}{1 - q^N} = \frac{1}{1 - \left(1 - \dfrac{1}{\mu}\right)^N}.$$

Za veliki μ, približna vrijednost ovog očekivanja je $\dfrac{\mu}{N}$.

U interpretaciji ovog rezultata treba biti vrlo oprezan. Nije svejedno u kojim se jedinicama mjeri vrijeme. Uzmemo li da je na primjer $\mu = 120$ sekunda, dobit ćemo potpuno različite rezultate od pretpostavke $\mu = 2$ minute. Evo tablica vrijednosti očekivanja $\boldsymbol{E}(X)$ za razne vrijednosti broja N:

N	1	2	3	4	5	6	7	8	9	10
$\mu = 120$ sec	120	60.3	40.3	30.4	24.4	20.4	17.6	15.4	13.8	12.5
$\mu = 2$ min	2.00	1.33	1.14	1.07	1.03	1.02	1.01	1.00	1.00	1.00

Zbog čega dolazi do ovako različitih rezultata? ◄

4.2. Binomna razdioba

Vjerojatno najvažnija diskretna razdioba jest binomna.

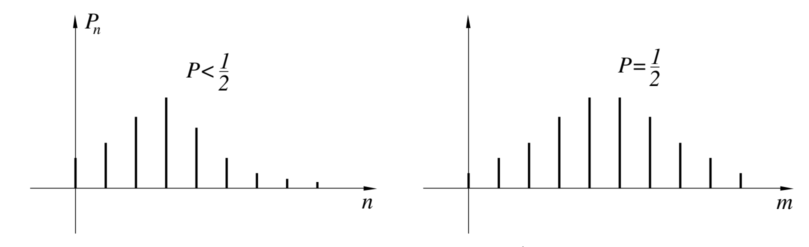
Neka je p vjerojatnost realizacije događaja A pri izvođenju nekog pokusa. Pretpostavimo da istovjetan pokus ponavljamo n puta. Neka slučajna varijabla X mjeri broj pojavljivanja događaja A. Onda kažemo da X ima **binomnu razdiobu** s parametrima n i p i pišemo $X \sim \mathscr{B}(n,p)$.

X poprima vrijednosti u skupu $\{0, 1, 2, \dots n\}$. Odredimo $p_k = \boldsymbol{P}(X = k)$.

Ako se realizirao događaj $\{X = k\}$, to znači da se u n pokusa A ostvario točno k puta, a $n-k$ puta se nije ostvario. Broj različitih mogućnosti za odabir pokusa u kojima se A ostvario je $\binom{n}{k}$. Zato je

$$p_k = \boldsymbol{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

Uobičajeno je označiti $q := 1 - p$.



Sl. 4.1. Razdioba binomne slučajne varijable, za $p < \frac{1}{2}$ (lijevo) i u simetričnom slučaju $p = \frac{1}{2}$ (desno).

Primjer 4.2.

Što je vjerojatnije u igri s ravnopravnim protivnikom: dobiti 3 partije od 4 ili 5 partija od 8? (Igra nema neriješenog ishoda.)

► Broj dobivenih partija ravna se po binomnoj razdiobi. U prvom slučaju to je zakon $B(4, \frac{1}{2})$, a u drugom slučaju $B(8, \frac{1}{2})$. Zato imamo

$$\boldsymbol{P}\{3 \text{ partije od } 4\} = \binom{4}{3} \cdot \frac{1}{2^3} \cdot \frac{1}{2} = \frac{8}{32},$$
$$\boldsymbol{P}\{5 \text{ partija od } 8\} = \binom{8}{5} \cdot \frac{1}{2^5} \cdot \frac{1}{2^3} = \frac{7}{32}. \blacktriangleleft$$

Primjer 4.3.

Pokus se sastoji u bacanju triju kocki. Izračunaj vjerojatnost da se u 10 nezavisnih pokusa 4 puta pojave točno 2 jedinice.

► Označimo s A događaj { pri bacanju triju kocaka pojavile su se točno 2 jedinice }.

Broj pojavljivanja jedinica je binomna slučajna varijabla $B(3, \frac{1}{6})$, pošto se bacaju 3 kocke a vjerojatnost pojavljivanja jedinice iznosi $\frac{1}{6}$. Zato je vjerojatnost događaja A jednaka

$$p = \boldsymbol{P}(A) = \binom{3}{2}\left(\frac{1}{6}\right)^2 \frac{5}{6} = \frac{5}{72}.$$

Broj realizacija dogadaja A pri ponavljanju 10 pokusa je binomna slučajna varijabla $B(10, p)$. Vjerojatnost da se on pojavi točno 4 puta iznosi

$$\binom{10}{4} p^4 q^6 = \binom{10}{4} \left(\frac{5}{72}\right)^4 \left(\frac{67}{72}\right)^6 = 3.17 \cdot 10^{-3}. \blacktriangleleft$$

4.2.1. Karakteristična funkcija, očekivanje i disperzija

Odredimo najprije karakterističnu funkciju binomne razdiobe. Neka je $X \sim \mathscr{B}(n,p)$. Tada imamo

$$\begin{aligned}\vartheta(t) &= \sum_{k=0}^{n} e^{itk} p_k = \sum_{k=0}^{n} e^{itk} \binom{n}{k} p^k q^{n-k} \\ &= \sum_{k=0}^{n} \binom{n}{k} (pe^{it})^k q^{n-k} = (pe^{it} + q)^n.\end{aligned}$$

Sada imamo

$$\begin{aligned}\vartheta'(t) &= n(pe^{it} + q)^{n-1} pe^{it} i, \\ \vartheta'(0) &= n(p+q)^{n-1} pi = npi \implies \boldsymbol{E}(X) = np.\end{aligned}$$

Na isti način dobivamo

$$\vartheta''(0) = n(n-1)p^2 i^2 + npi^2 = -n(n-1)p^2 - np$$

te je

$$\boldsymbol{D}(X) = -\vartheta''(0) + \vartheta'(0)^2 = n(n-1)p^2 + np - n^2p^2 = np - np^2 = npq.$$

Binomna razdioba, definicija i numeričke karakterisike

Kažemo da slučajna varijabla X ima **binomnu razdiobu** s parametrima n i p i pišemo $X \sim \mathscr{B}(n,p)$, ako X poprima vrijedosti unutar skupa $\{0, 1, 2, \dots, n\}$ s vjerojatnostima

$$p_k = \boldsymbol{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

Očekivanje i disperzija binomne razdiobe su

$$m_X = np, \qquad \sigma_X^2 = npq.$$

4.2.2. Stabilnost binomne razdiobe

Ako su $X_1 \sim \mathscr{B}(n_1, p)$ i $X_2 \sim B(n_2, p)$ nezavisne binomne slučajne varijable, onda je $X_1 + X_2$ binomna slučajna varijabla. Odredimo njezine parametre.

Vrijedi $\vartheta_{X_1}(t) = (q + pe^{it})^{n_1}$, $\vartheta_{X_2}(t) = (q + pe^{it})^{n_2}$ te je zbog nezavisnosti od X_1 i X_2

$$\vartheta_{X_1+X_2}(t) = \vartheta_{X_1}(t)\vartheta_{X_2}(t) = (q + pe^{it})^{n_1+n_2}$$

a to je karakteristična funkcija binomne razdiobe $\mathscr{B}(n_1 + n_2, p)$.

Ovaj je rezultat prirodan. Riječ je o tome da smo promatrali isti pokus, s time da smo ga podijelili na dvije cjeline: u prvoj smo pokus ponovili n_1 puta, a u drugoj n_2 puta. Broj realizacija događaja A u cijelom pokusu jednak je zbroju tih realizacija u pojedinim cjelinama.

4.2.3. Bernoullijeve slučajne varijable

Poseban, najjednostavniji primjer binomne slučajne varijable je **Bernoullijeva** ili **indikatorska** slučajna varijabla: ona poprima samo dvije vrijednosti: 1 s vjerojatnošću p i 0 s vjerojatnošću $q = 1 - p$. Ona bilježi realizaciju događaja A u jednom pokusu.

Ako su X_i Bernoullijeve nezavisne varijable s istim parametrom p, tada je njihov zbroj $X_1 + X_2 + \ldots + X_n$ binomna slučajna varijabla $\mathscr{B}(n,p)$. Ova tvrdnja slijedi zbog svojstva stabilnosti binomnih slučajnih varijabli.

Na temelju toga možemo lakše izračunati očekivanje i disperziju binomne slučajne varijable. naime, za indikatorsku slučajnu varijablu vrijedi

$$\boldsymbol{E}(X_i) = 0 \cdot q + 1 \cdot p = p, \qquad \boldsymbol{D}(X_i) = 0^2 \cdot q + 1^2 \cdot p - p^2 = pq, \qquad \forall i,$$

pa je, zbog nezavisnosti

$$\begin{aligned}\boldsymbol{E}(X) &= \boldsymbol{E}(X_1) + \ldots + \boldsymbol{E}(X_n) = np, \\ \boldsymbol{D}(X) &= \boldsymbol{D}(X_1) + \ldots + \boldsymbol{D}(X_n) = npq.\end{aligned}$$

— ◆ —

Primjer 4.4. ■ Najvjerojatnija realizacija

Slučajna varijabla X ima binomnu razdiobu $\mathscr{B}(n,p)$. Koja je najvjerojatnija realizacija slučajne varijable X?

► Tražimo takav k za koji je

$$p_k = \boldsymbol{P}(X = k) = \binom{n}{k} p^k q^{n-k}$$

najveće. Tada će biti $p_0 \leqslant p_1 \leqslant \ldots \leqslant p_k$ i $p_k \geqslant p_{k+1} \geqslant \ldots \geqslant p_n$. Dovoljno je stoga promatrati dvije nejednakosti $p_{k-1} \leqslant p_k$ i $p_k \geqslant p_{k+1}$ i pronaći k za koji su one ispunjene. Vrijedi

$$\frac{p_k}{p_{k+1}} = \frac{\binom{n}{k} p^k q^{n-k}}{\binom{n}{k+1} p^{k+1} q^{n-k-1}} = \frac{q}{p} \cdot \frac{k+1}{n-k}$$

i ovaj kvocijent je veći od 1 ako je $q(n-k) \geqslant p(k+1)$. Odavde dobivamo $k \geqslant (n+1)p - 1$. Slično

$$\frac{p_k}{p_{k-1}} = \frac{\binom{n}{k} p^k q^{n-k}}{\binom{n}{k-1} p^{k-1} q^{n-k+1}} = \frac{p}{q} \cdot \frac{n-k+1}{k} \geqslant 1$$

daje $k \leqslant (n+1)p$.

Prema tome, najvjerojatnija realizacija slučajne varijable $\mathscr{B}(n,p)$ je onaj cijeli broj uklješten između $(n+1)p - 1$ i $(n+1)p$. Ako je $(n+1)p$ cijeli broj, postoje tada dvije takve vrijednosti. ◄

Primjer 4.5.

Koliko puta moramo baciti kocku da bi najvjerojatniji broj pojavljivanja šestice bio 10?

► Neka je n broj bacanja kocke. Broj pojavljivanja šestice u n bacanja je binomna varijabla $\mathscr{B}(n, \frac{1}{6})$. Po prošlom zadatku mora biti

$$(n+1)\frac{1}{6} - 1 \leqslant 10 \leqslant (n+1)\frac{1}{6}.$$

Dakle, $n + 1 \leqslant 66$ i $n + 1 \geqslant 60$, odnosno, $59 \leqslant n \leqslant 65$. ◄

Primjer 4.6.

Vjerojatnost kvara brodskog motora u toku jednog dana iznosi p. Ukoliko je motor bio u kvaru m puta, vjerojatnost da mora otići u remont jednaka je $P(m) = 1 - \left(1 - \frac{1}{\omega}\right)^m$, gdje je ω neki parametar (srednji broj kvarova do odlaska broda u remont). Dokaži da je vjerojatnost da će brod morati otići na remont nakon n dana plovidbe jednaka $P_n = 1 - \left(1 - \frac{p}{\omega}\right)^n$.

► Neka je $P_{n,m}$ vjerojatnost da će u toku n dana motor biti u kvaru m puta. Po formuli potpune vjerojatnosti, dobivamo

$$P_n = \sum_{m=0}^{n} P_{n,m} P(m).$$

Vjerojatnost kvara u toku svakog dana jednaka p. Broj kvarova unutar n dana ravna se po binomnoj razdiobi. Zato je vjerojatnost $P_{n,m}$ dana sa

$$P_{n,m} = \binom{n}{m} p^m (1-p)^{n-m}.$$

Uvrštavanjem u gornju formulu dobivamo

$$\begin{aligned}P_n &= \sum_{m=0}^{n} \binom{n}{m} p^m (1-p)^{n-m} \left[1 - \left(1 - \frac{1}{\omega}\right)^m\right] \\ &= \sum_{m=0}^{n} \binom{n}{m} p^m (1-p)^{n-m} - \sum_{m=0}^{n} \binom{n}{m} \left(p - \frac{p}{\omega}\right)^m (1-p)^{n-m} \\ &= 1 - \left(p - \frac{p}{\omega} + 1 - p\right)^n = 1 - \left(1 - \frac{p}{\omega}\right)^n. \blacktriangleleft\end{aligned}$$

4.3. Poissonova razdioba

Poissonovu razdiobu možemo dobiti kao granični slučaj binomne, kad broj pokusa neograničeno raste. Ulogu vjerojatnosti p pojavljivanja događaja zamjenjuje *intenzitet* λ pojavljivanja događaja.

Promotrimo sljedeći primjer.

Unutar smjese od koje će se ispeći $m = 25$ kolačića stavljeno je $n = 100$ zrna grožđica. Neka je X slučajna varijabla: broj zrna unutar jednog kolačica. Kakva je razdioba slučajne varijable X?

Pretpostavljamo da se svako zrno može neovisno jedno o drugom naći s jednakom vjerojatnošću unutar bilo kojeg kolačića. Vjerojatnost da se zrno nađe unutar jednog odabranog kolačića je $p = \frac{1}{m}$. Broj zrna unutar tog kolačića je binomna slučajna varijabla s parametrima n i $p = \frac{1}{m}$.

Primjetimo da je očekivani broj zrna unutar jednog kolačića jednak $\boldsymbol{E}(X) = np = \frac{n}{m}$. Označimo tu veličinu s λ. Ona označava *intenzitet* pojavljivanja zrna unutar nekog kolačića.

Model sličan ovom pojavljuje se pri promatranju broja poziva koji će stići na neku telefonsku centralu u nekoj jedinici vremena. Ako za $m = 25$ minuta na centralu stigne *u prosjeku* $n = 100$ poziva, tada je broj poziva unutar jedne minute — baš kao u prijašnjem primjeru s grožđicama — binomna razdioba s parametrima $n = 100$, $p = \frac{1}{25}$. Primjetimo da je očekivani broj poziva $\lambda = 4$.

Razlika između ova dva primjera jest u tome što je broj grožđica bio unaprijed poznat, i ograničen odozgo. Ukupan broj poziva u drugom primjeru nije poznat, već je dan kao *statistička veličina*. Sasvim je razumno pretpostaviti, barem u teoriji, da taj broj nije ograničen odozgo.

4.3.1. Od binomne razdiobe prema Poissonovoj

Teorem 4.2. Aproksimacija binomne razdiobe

Neka je n velik a p malen. Označimo $\lambda = np$. tad vrijedi aproksimacija

$$\binom{n}{k} p^k q^{n-k} \approx \frac{\lambda^k}{k!} e^{-\lambda}. \tag{3}$$

Dokaz. Označimo $m = \frac{1}{p}$ i transformirajmo izraz slijeva:

$$\binom{n}{k} p^k (1-p)^{n-k} = \binom{n}{k} \left(\frac{1}{m}\right)^k \left(1 - \frac{1}{m}\right)^{n-k}$$

Sada je $\lambda = np = \frac{n}{m}$. Pustimo da broj n neograničeno raste:

$$\begin{aligned}\binom{n}{k}\left(\frac{\lambda}{n}\right)^k\left(1-\frac{\lambda}{n}\right)^{n-k} &= \frac{1}{k!}\frac{n(n-1)\cdots(n-k+1)}{n^k}\lambda^k\left(1-\frac{\lambda}{n}\right)^{n-k}\\ &= \frac{1}{k!}\cdot 1\cdot\left(1-\frac{1}{n}\right)\cdots\left(1-\frac{k-1}{n}\right)\lambda^k\left(1-\frac{\lambda}{n}\right)^{n-k}\\ &\to \frac{\lambda^k}{k!}e^{-\lambda}.\end{aligned}$$

Greška koja se čini ovakvom aproksimacijom približno je jednaka izrazu

$$r_n(k) = \frac{k - (k-np)^2}{2n} + \frac{kp^2}{2},$$

što nećemo ovdje dokazivati.

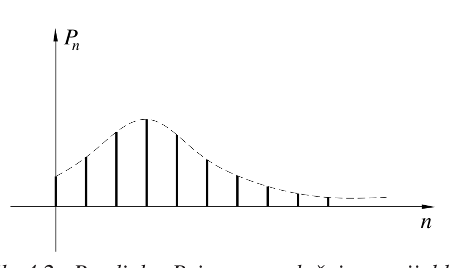
Poissonova razdioba, definicija i numeričke karakteristike

Kažemo da slučajna varijabla X ima **Poissonovu razdiobu** s parametrom $\lambda > 0$ i pišemo $X \sim \mathscr{P}(\lambda)$ ako ona poprima vrijednosti unutar skupa $\{0, 1, 2, \dots\}$ s vjerojatnostima

$$p_k = \boldsymbol{P}(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}.$$

Za očekivanje i disperziju ove razdiobe vrijedi

$$m_X = \lambda, \qquad \sigma_X^2 = \lambda.$$



Sl. 4.2. Razdioba Poissonove slučajne varijable

4.3.2. Karakteristična funkcija, očekivanje i disperzija

Odredimo najprije karakterističnu funkciju Poissonove razdiobe $\mathscr{P}(\lambda)$.

$$\vartheta(t) = \sum_{k=0}^{\infty} e^{itk} \frac{\lambda^k}{k!} e^{-\lambda} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{(\lambda e^{it})^k}{k!} = e^{-\lambda} e^{\lambda e^{it}} = e^{\lambda(e^{it}-1)}.$$

Odavde, na temelju veze karakteristične funkcije i momenata slučajne varijable, dobivamo $\boldsymbol{E}(X) = \lambda$, $\boldsymbol{D}(X) = \lambda$.

Primjer 4.7.

U telefonskoj centrali tijekom jednog sata bilo je 240 poziva. Odredi vjerojatnost da tijekom jedne minute **a**) nije bilo nijednog poziva, **b**) bilo je barem dva poziva.

► Neka je X slučajna varijabla: broj poziva u jednoj (bilo kojoj) minuti. To je Poissonova varijabla s intenzitetom λ koji je jednak očekivanoj vrijednosti, $\lambda = \frac{240}{60} = 4$.

$$\boldsymbol{P}(X = 0) = \frac{\lambda^0}{0!} e^{-\lambda} = e^{-4} = 0.018,$$

$$\boldsymbol{P}(X \geqslant 2) = 1 - \boldsymbol{P}(X = 0) - \boldsymbol{P}(X = 1) = 1 - e^{-4} - 4e^{-4} = 0.908. \blacktriangleleft$$

4.3.3. Stabilnost Poissonove razdiobe

Ako su $X_1 \sim \mathscr{P}(\lambda_1)$ i $X_2 \sim \mathscr{P}(\lambda_2)$ nezavisne slučajne varijable, onda je $X_1 + X_2$ također Poissonova slučajna varijabla. Dokažimo tu tvrdnju i odredimo parametar ove razdiobe.

Karakteristična funkcija Poissonove razdiobe je

$$\vartheta_{X_k}(t) = e^{\lambda_k(e^{it}-1)}, \qquad k = 1, 2$$

te slijedi

$$\vartheta_{X_1+X_2}(t) = e^{(\lambda_1+\lambda_2)(e^{it}-1)}$$

a to je karakteristična funkcija Poissonove razdiobe $P(\lambda_1 + \lambda_2)$.

— ◆ —

Zamislimo li varijablu X_1 kao broj poziva na prvi telefon centrale nekog poduzeća, a X_2 kao broj poziva na drugi telefon, tada, po gornjem, ukupan broj poziva ima također Poissonovu razdiobu.

Ako je poznata vrijednost tog zbroja, pogledajmo što se može reći o vrijednostima pojedinih varijabli.

Primjer 4.8.

Neka su X_1 i X_2 nezavisne slučajne varijable, s Poissonovim zakonom $P(\lambda_1)$, odnosno $P(\lambda_2)$. Poznato je da je njihov zbroj $X_1 + X_2$ poprimio vrijednost n. Dokaži da je tada vrijednost od X_1 raspoređena po binomnom zakonu s parametrima n i $p = \frac{\lambda_1}{\lambda_1 + \lambda_2}$, tj.

$$\boldsymbol{P}(X_1 = k \mid X_1 + X_2 = n) = \binom{n}{k} p^k q^{n-k}.$$

► Jednostavni račun daje

$$\begin{aligned}\boldsymbol{P}(X_1 = k \mid X_1 + X_2 = n) &= \frac{\boldsymbol{P}(X_1 = k, X_2 = n-k)}{\boldsymbol{P}(X_1 + X_2 = n)}\\ &= \frac{\dfrac{\lambda_1^k}{k!}e^{-\lambda_1} \cdot \dfrac{\lambda_2^{n-k}}{(n-k)!}e^{-\lambda_2}}{\dfrac{(\lambda_1+\lambda_2)^n}{n!}e^{-\lambda_1-\lambda_2}} = \binom{n}{k}\left(\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^k\left(\frac{\lambda_2}{\lambda_1+\lambda_2}\right)^{n-k}. \blacktriangleleft\end{aligned}$$

Primjer 4.9.

Centrala poduzeća ima dva pozivna broja. Na prvi stiže oko 20% poziva više nego na drugi broj. Ako je u protekloj minuti stiglo 5 poziva, kolika je vjerojatnost da je češće pozivan prvi broj?

► Pozivi na pojedine brojeve su nezavisne Poissonove varijable s parametrima λ_1 i λ_2 pri čemu je $\lambda_1 = 1.2\lambda_2$. Iskoristit ćemo rezultat prethodnog zadatka: razdioba varijable X_1 uz uvjet $X_1 + X_2 = 5$ je binomna, s parametrima n i p, pri čemu je

$$p = \frac{\lambda_1}{\lambda_1 + \lambda_2} = \frac{6}{11}.$$

Zato je

$$\begin{aligned}\boldsymbol{P}(X_1 \geqslant 3 \mid X_1 + X_2 = 5) &= \binom{5}{3}\left(\frac{6}{11}\right)^3\left(\frac{5}{11}\right)^2 + \binom{5}{4}\left(\frac{6}{11}\right)^4 \cdot \frac{5}{11} + \left(\frac{6}{11}\right)^5\\ &= 0.585. \blacktriangleleft\end{aligned}$$

4.3.4. Aproksimacija binomne razdiobe Poissonovom

Za veliko n i maleno p, binomna razdioba $\mathscr{B}(n, p)$ može se aproksimirati Poissonovom razdiobom $\mathscr{P}(np)$.

Primjer 4.10.

Proizvodi neke velike serije, koja sadrži 0.7% škarta, pakiraju se u kutije po 100 komada. Koliki će postotak kutija biti bez ijednog škarta, a koliki sa dva ili više škartova?

► Broj škartnih proizvoda u jednoj kutiji je slučajna varijabla X distribuirana po binomnom zakonu $\mathscr{B}(100, 0.007)$. Zato

$$\begin{aligned}\boldsymbol{P}(X = 0) &= \binom{100}{0} 0.007^0 0.993^{100} = 0.4954,\\ \boldsymbol{P}(X \geqslant 2) &= 1 - \boldsymbol{P}(X = 0) - \boldsymbol{P}(X = 1)\\ &= 1 - 0.993^{100} - \binom{100}{1} 0.007 \cdot 0.993^{99} = 0.1554.\end{aligned}$$

Možemo aproksimirati $X \approx \boldsymbol{P}(0.7)$. Jednostavniji račun daje

$$\begin{aligned}\boldsymbol{P}(X = 0) &= e^{-0.7} = 0.4966,\\ \boldsymbol{P}(X \geqslant 2) &= 1 - e^{-0.7} - 0.7e^{-0.7} = 0.1558. \blacktriangleleft\end{aligned}$$

Primjer 4.11.

Na lovački zrakoplov ispaljeno je 5000 metaka. Vjerojatnost pogotka za svaki metak je 0.001. Ako je zrakoplov pogođen, vjerojatnost da će metak prouzročiti pad zrakoplova je 0.05. Izračunaj vjerojatnost da će zrakoplov biti srušen.

► Neka je X slučajna varijabla: broj pogodaka u zrakoplov. Tada je $X \sim \mathscr{B}(5000, 0.001)$. Označimo

$$\begin{aligned}H_k &= \{\text{zrakoplov je pogođen sa } k \text{ metaka}\},\\ A &= \{\text{zrakoplov je srušen}\}.\end{aligned}$$

Vrijedi

$$\boldsymbol{P}(A) = \sum_{k=1}^{5000} \boldsymbol{P}(A|H_k)\boldsymbol{P}(H_k).$$

Pri tom je

$$\boldsymbol{P}(A|H_k) = 1 - \boldsymbol{P}(\overline{A}|H_k) = 1 - 0.95^k$$

(vjerojatnost da zrakoplov neće biti srušen ukoliko je pogođen s k metaka jednaka je 0.95^k).

$$\boldsymbol{P}(H_k) = \boldsymbol{P}(X = k) = \binom{5000}{k} 0.001^k 0.999^{5000-k}$$

što je nepraktično za daljnji račun. Zato aproksimiramo $\mathscr{B}(5000, 0.001) \approx P(5)$ i dobivamo

$$\boldsymbol{P}(H_k) = \frac{5^k}{k!} e^{-5}.$$

Sada je

$$\begin{aligned}\boldsymbol{P}(A) &= \sum_{k=1}^{5000} (1 - 0.95^k) \frac{5^k}{k!} e^{-5}\\ &= e^{-5}\left(\sum_{k=1}^{5000} \frac{5^k}{k!} - \sum_{k=1}^{5000} \frac{4.75^k}{k!}\right)\\ &= e^{-5}(e^5 - e^{4.75}) = 0.22. \blacktriangleleft\end{aligned}$$