Zlatniczki Ádám
adam.zlatniczki@cs.bme.hu

Nevezetes diszkrét eloszlások

Szűk elméleti összefoglaló

Binomiális eloszlás:

Jelölés: $X \sim B(n,p)$ vagy $X \in B(n,p)$

Tipikus használata: Egy kétféle kimenetelű (valami beteljesül vagy sem) kísérletet elvégzünk *n*-szer, ahol minden egyes kísérletben az esemény beteljesülésének a valószínűsége fixen *p*. Mi a valószínűsége, hogy az *n* kísérletből *k* darab beteljesül?

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

$$F(k;n,p) = P(X \leq k) = \sum_{i=0}^{\lfloor k \rfloor} \binom{n}{i} p^i (1-p)^{n-i}$$

Poisson eloszlás:

Jelölés: $X \sim Po(\lambda)$ vagy $X \in Po(\lambda)$

Tipikus használata: Tudjuk, hogy egy adott időtartam alatt átlagosan λ-szor következik be egy esemény. Mi a valószínűsége, hogy ez az esemény pontosan *k*-szor következik be ugyanebben az időtartamban?

$$f(k;\lambda) = P(X=k) = \frac{\lambda^k}{k!} e^{-\lambda}$$

$$F(k;\lambda) = P(X \leq k) = e^{-\lambda} \sum_{i=0}^{\lfloor k \rfloor} \frac{\lambda^i}{i!}$$

Ha $t_1$időegység alatt *j*-szer következett be egy esemény, akkor $\lambda_{t_1} = \frac{j}{t_1}$. Az időegységet megváltoztathatjuk t_2-re, ekkor $\lambda_{t_2} = \lambda \times \frac{t_2}{t_1}$. Ehhez a transzformációhoz persze ugyanabban a mértékegységben kell lennie az időtartamoknak!

Pl.: óránként 3 autó hajt át a piroson, de a kérdés arra szól, hogy mi a valószínűsége, hogy egy nap alatt 100 autó hajt át. Ekkor $t_1 = 1[óra]$, $\lambda_{t_1} = \frac{3}{1}$, $t_2 = 1[nap]$. Ezekből $\lambda_{t_2} = \lambda_{t_1} \frac{t_2}{t_1} = 3 \frac{1[nap]}{1[óra]} = 3 \frac{24[óra]}{1[óra]} = 72$.

Zlatniczki Ádám
adam.zlatniczki@cs.bme.hu

Geometriai eloszlás:

Jelölés: $X \sim G(p)$ vagy $X \in G(p)$

Tipikus használata: Egy kísérletet addig ismétlünk, amíg egy p valószínűségű esemény be nem következik. Mi a valószínűsége annak, hogy az esemény pontosan a k-adik alkalommal következik be?

$$f(k;p) = P(X = k) = (1-p)^{k-1}p$$

$$F(k;p) = P(X \leq k) = 1 - (1-p)^{k}$$

ZH és vizsga feladatok

1. Példa

Egy teherautó 50 láda tojást szállít, mindegyik ládában pontosan 1000 tojással. Szállításkor minden tojás 0,002 valószínűséggel összetörhet (a többitől függetlenül). A megrendelő akkor vesz át egy ládát, ha abban az összetört tojások száma nem haladja meg a 10-et. Mennyi a valószínűsége, hogy legfeljebb két ládát nem vesz át?

Megoldás:

Jelölje X az át nem vett ládák számát. A kérdés $P(X \leq 2)$. A ládák egymástól függetlenül, azonos valószínűséggel eshetnek ebbe a kategóriába, tehát $X \sim B(50, p_1)$, $k_1 = 2$. Ehhez azonban szükségünk van p_1-re.

Tulajdonképpen $p_1 = P(Y \leq 10)$, ahol Y az egy ládában összetört tojások száma. A ládában a tojások egymástól függetlenül, ugyanakkora $p_2 = 0{,}002$ valószínűséggel törhetnek össze, így $Y \sim B(1000, p_2)$ és $k_2 = 10$. Ebből megkapjuk, hogy

$$p_1 = P(Y > 10) = 1 - P(Y \leq 10) = 1 - \sum_{i=0}^{10} \binom{1000}{i} p_2^i (1-p_2)^{1000-i}$$

Ezáltal megkapjuk, hogy

$$P(X \leq 2) = \sum_{i=0}^{2} \binom{50}{i} p_1^i (1-p_1)^{50-i}$$

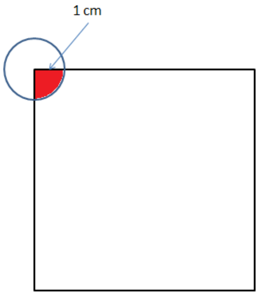
2. Példa

Egy 20 cm oldalhosszúságú négyzetrácsos hálózatra leejtünk 10 db. 2 cm átmérőjű pénzérmét. Mennyi a valószínűsége, hogy legalább 5 darab valamelyik négyzet csúcsát fogja lefedni?

Megoldás:

Az egyes pénzérmék elejtése egymástól független, így azt kell meghatározni, mi annak a valószínűsége, hogy egy pénzérme egy négyzet valamelyik csúcsát lefedi. Ahhoz, hogy az érme egy csúcsot lefedhessen a középpontja a csúcstól legföljebb 1 cm távolságra kerülhet. Az alábbi ábrán pirossal van feltüntetve ez a terület.

Zlatniczki Ádám
adam.zlatniczki@cs.bme.hu



Ez a terület $\frac{1^2 \times \pi}{4}$. Mivel azonban négy sarka van a négyzetnek, így ezt 4-gyel megszorozva π lesz a megfelelő terület. A valószínűsége annak, hogy egy érme egy négyzet valamelyik csúcsát fedje egy négyzetnek $p = \frac{\pi}{20^2}$.

Jelölje X_i, hogy hány érme fed csúcsot. Látható, hogy $X_i \in B(10, p)$. Annak a valószínűsége, hogy legalább 5 darab érme fed csúcsot tehát

$$P(X \geq 5) = \sum_{i=5}^{10} P(X_i)$$

$$P(X_i) = \binom{10}{i} p^i (1-p)^{10-i}$$

3. Példa

Addig dobunk kockával, amíg 6-ost nem kapunk. Mi annak a valószínűsége, hogy a szükséges dobások száma legalább 5 és kevesebb mint 8?

Megoldás:

Jelölje X a 6-oshoz szükséges dobások számát. Vegyük észre, hogy ez geometriai eloszlást követ, $X \sim G\left(\frac{1}{6}\right)$.

A kérdés $P(5 \leq X < 8) = P(X = 5) + P(X = 6) + P(X = 7) = \left(1 - \frac{1}{6}\right)^{5-1} \left(\frac{1}{6}\right) + \left(1 - \frac{1}{6}\right)^{6-1} \left(\frac{1}{6}\right) + \left(1 - \frac{1}{6}\right)^{7-1} \left(\frac{1}{6}\right) = 0{,}0804 + 0{,}0670 + 0{,}0558 = 0{,}2032$

Vagy az eloszlásfüggvénnyel:

$$P(5 \leq X < 8) = P(X < 8) - P(X < 5) = P(X \leq 7) - P(X \leq 4) = F_X(7) - F_X(4)$$

Ha valaki nem biztos benne, hogy diszkrét esetben az eloszlásfüggvény hol kisebb, vagy hol kisebb-vagy-egyenlő, akkor írja inkább föl az egyes P(X=k) valószínűségek összegeként, mert a pont akkor is ugyanúgy jár a megoldásért, ráadásul nehezebb elrontani.

4. Példa

Novemberben a hónap 30 napjából átlagosan 12-őn esik. Mi a valószínűsége, hogy egy héten három nap is esni fog?

Zlatniczki Ádám
adam.zlatniczki@cs.bme.hu

Megoldás:

Adott időtartamon belül az átlagos esetek számát megadták, így Poisson eloszlást tudunk alkalmazni. Azonban a kérdés 1 hétre vonatkozik, az adatunk pedig 1 hónapra vonatkozik, így át kell skálázni a gyakoriságot.

$$t_1 = 1\ [hónap]$$

$$\lambda_{t_1} = \frac{12}{1} = 12$$

$$\lambda_{t_2} = 12 \times \frac{7}{30} = 2,8$$

Ebből

$$P(X = 3) = \frac{2,8^3}{3!} e^{-2,8} = 0,222$$

5. Példa

Egy kommunikációs csatorna 0 vagy 1 számjegyeket tud továbbítani. Sajnos hálózati zavarok miatt minden számjegy továbbításába 0,2 valószínűséggel hiba csúszik (egymástól függetlenül). Tegyük fel, hogy egy fontos egybites információt szeretnénk küldeni. Hogy a hiba esélyét csökkentsük, a 0 helyett 00000, az 1 helyett 11111 sort küldjük, a vevő oldalon pedig többségi értelmezéssel dekódolják. Mennyi a valószínűsége, hogy ilyen kódolás mellett is hibásan dekódolják az üzenetünket?

Megoldás:

Jelölje X a hibásan megérkező bitek számát. Ha X > 2, akkor a túloldal hibásan fogja dekódolni amit küldtünk. Mivel függetlenül, azonos valószínűséggel hibázik az 5 kísérletből a csatorna, így $X \sim B(5, 0,2)$.

$$P(X > 2) = 1 - P(X \le 2) = 1 - \sum_{i=0}^{2} \binom{5}{i} 0,2^i (1 - 0,2)^{5-i} = 0,058$$

6. Példa

A [-1,1]x[-1,1] négyzeten egymás után sorsolunk ki véletlen pontokat. Akkor állunk meg, ha az első kisorsolt pont beleesik az $(x - 0,5)^2 + (y - 0,5)^2 = 0,25$ körbe. Adja meg a pontok eloszlását!

Megoldás:

Jelölje Y a szükséges dobások számát. Adja magát, hogy $Y \sim G(p)$. Már csak *p*-re van szükségünk az eloszlás felírásához. Ehhez a geometriai módszert alkalmazhatjuk, ahogyan tettük azt **Hiba! A hivatkozási forrás nem található.**. Tehát $p = \frac{\pi}{16}$, így

$$F_Y(k) = 1 - \left(1 - \frac{\pi}{16}\right)^k, k = 1,2, \dots$$

Zlatniczki Ádám
adam.zlatniczki@cs.bme.hu

7. Példa

Egy teherautó 200 láda tojást szállít, mindegyik ládában pontosan 100 tojással. Szállításkor minden tojás 0,01 valószínűséggel összetörhet, egymástól függetlenül. A megrendelő akkor vesz át egy ládát, ha abban az összetört tojások száma nem haladja meg a 2-t. Mennyi annak a valószínűsége, hogy legfeljebb 2 ládát nem fog átvenni? Kerülje el a binomiális eloszlás használatát!

Megoldás:

A feladat alapvetően binomiális eloszlással jól leírható, de itt most nem alkalmazhatjuk. Használjuk hát ki, hogy a binomiális eloszlás közelíthető Poisson eloszlással!

Egy ládában átlagosan 0,01*100=1 tojás a törött. Jelölje X egy ládában a törött tojások számát. Ekkor $X \sim Po(1)$, valamint $p = P(X > 2) = 1 - P(X \leq 2) = 1 - \sum_{k=0}^{2} \frac{1^k}{k!} e^{-1}$

200 ládából így átlagosan 200*p kerül visszautasításra. Jelölje Y, hogy hány ládát nem fogadnak el. $Y \sim Po(200p)$.

$$P(Y \leq 2) = \sum_{i=0}^{2} \frac{200p^i}{i!} e^{200p}$$

8. Példa

A boltban az izzók 2%-a hibás. Ha veszünk 50 darabot, akkor hány darab lesz benne rossz a legnagyobb valószínűséggel, és mekkora ez a valószínűség?

Megoldás:

A hibás izzók száma leírható mind binomiális, mind Poisson eloszlással. Nézzük meg mindkettőt a gyakorlás kedvéért.

<u>Binomiális:</u>

$$X \sim B(50, 0.02)$$

A binomiális eloszlás a várható értékét veszi föl a legmagasabb valószínűséggel, azaz X=50*0,02=1-nél, $P(X = 1) = \binom{50}{1} 0{,}02^1 (1 - 0{,}02)^{49}$

Mivel még nem tudjuk mi az a várható érték, így kihasználhatjuk azt, hogy a binomiális eloszlásnak egyetlen maximumhelye van, tehát rendre meg kellene néznünk P(X=0), P(X=1), P(X=2)-t, ahol már látnánk, hogy P(X=1) a maximum.

<u>Poisson:</u>

A globális maximum megkeresésének az elvét kellene itt is alkalmazni, csak $X \sim Po(1)$ mellett.

9. Példa

Az egyetemen sok szerver van, melyek egymástól függetlenül romlanak el azonos valószínűséggel. Az év 360 napján átlagosan 18 olyan nap van, hogy egyetlen készülék sem romlik el. Várhatóan hány olyan nap lesz az évben, amikor 3-nál több szerver romlik el?

Zlatniczki Ádám
adam.zlatniczki@cs.bme.hu

Megoldás:

Adott időtartam alatt egy esemény átlagos bekövetkezésének a számát kaptuk meg ➔Poisson eloszlást tudunk fölírni. Jelölje X az egy nap alatti meghibásodások számát.

Tudjuk, hogy $X \sim Po(\lambda)$, viszont nem ismerjük λ-t, viszont tudjuk becsülni, mivel azt a valószínűséget, hogy X=0 megkaptuk.

$$P(X=0) \approx \frac{18}{360}$$

$$P(X=0) = \frac{\lambda^0}{0!}e^{-\lambda} = e^{-\lambda} \approx 18/360$$

$\frac{1}{e^\lambda} \approx \frac{1}{20}$➔$e^\lambda \approx 20$➔$\lambda \approx \ln 20$

Annak a valószínűsége, hogy egy nap több mint 3 szerver romoljon el

$$p = P(X \geq 4) = 1 - P(X \leq 3) = 1 - \sum_{i=0}^{3} \frac{\lambda^i}{i!}e^{-\lambda}$$

Várhatóan $360 \times p$ ilyen nap lesz.

10. Példa

Két kockával addig dobunk egyszerre, amíg mindkét érték azonos nem lesz. Mennyi a valószínűsége, hogy legföljebb 10-szer kell dobni?

Megoldás:

Egy kísérletet ismétlünk, amíg be nem következik egy elvárt esemény ➔geometriai eloszlás kell.

Annak a valószínűsége, hogy a két kockával ugyanakkorát dobunk

$$p = \frac{6}{36}$$

Tehát az X: szükséges dobások száma $X \sim G(p)$, így

$$P(X \leq 10) = \sum_{i=1}^{10} P(X=i) = \sum_{i=1}^{10} (1-p)^{i-1}p = 1 - (1-p)^i$$

11. Példa

Az A könyvben az egy oldalra eső sajtóhibák száma $X \in Po(\lambda)$, míg a B könyvben ugyanez $Y \in Po(\mu)$. Igaz lehet-e az alábbi két állítás egyszerre:

- az A könyvben háromszor annyi a hiba
- B-ben ötször akkora egy hibamentes oldalnak a valószínűsége

Megoldás:

Első állítás:

Zlatniczki Ádám
adam.zlatniczki@cs.bme.hu

$$P(X=3Y)=(X,Y\ független ek)=\sum_{k=0}^{\infty}P(X=3k)P(Y=k)=\sum_{k=0}^{\infty}\frac{\mu^k\lambda^k}{k!\,(3k)!}e^{-(\mu+\lambda)}$$

Ez az érték mindig nagyobb mint 0, tehát ez nem zárja ki a két állítás bekövetkezhetőségét.

Második állítás:

$$P(Y=0)=5P(X=0)$$

$$P(Y=0)=e^{-\mu}$$

$$P(X=0)=e^{-\lambda}$$

$\frac{e^{-\mu}}{e^{-\lambda}}=5 \rightarrow e^{-\mu-\lambda}=5$

Ez nyilván teljesülhet megintcsak, amennyiben $\lambda-\mu=\ln 5$. Ezzel egy időben teljesülhet az első állítás is.

12. Példa

Hányszor dobjunk egy kockával, ha azt akarjuk, hogy 0,5-nél ne legyen kisebb az esélye, hogy a 6-os dobások száma legalább 2?

Megoldás:

$$X\in B\left(\frac{1}{6},n\right)$$

$$P(X\geq 2|n)\geq 0{,}5$$

$$P(X\geq 2|n)=1-P(X<2|n)=1-\sum_{i=0}^{1}P(X=i|n)$$

$$P(X=i|n)=\binom{n}{i}\left(\frac{1}{6}\right)^i\left(1-\frac{1}{6}\right)^{n-i}$$

$1-\sum_{i=0}^{1}P(X=i|n)\geq 0{,}5$ egyenletből következik, hogy

$\sum_{i=0}^{1}P(X=i|n)<0{,}5$, azaz

$$\binom{n}{0}\left(\frac{1}{6}\right)^0\left(1-\frac{1}{6}\right)^{n-0}+\binom{n}{1}\left(\frac{1}{6}\right)^1\left(1-\frac{1}{6}\right)^{n-1}<0{,}5$$

$$\left(\frac{5}{6}\right)^n+n\times\frac{1}{6}\times\left(\frac{5}{6}\right)^{n-1}\leq 0{,}5$$

$$\left(\frac{5}{6}\right)^{n-1}\left(\frac{5}{6}+\frac{n}{6}\right)\leq 0{,}5$$

$$\left(\frac{5}{6}\right)^{n-1}\leq\frac{3}{5+n}$$

Zlatniczki Ádám
adam.zlatniczki@cs.bme.hu

Ez leghamarabb az n=10-nél fog teljesülni, tehát legalább 10-szer kell dobni.