A sztochasztika alapjai

5. feladatsor: nevezetes eloszlások, CHT

Megoldások

1. Egy könyvben az egyes oldalakon a sajtóhubák száma egymástól független, Poisson(2) eloszlást követ. Határozzuk meg annak a valószínűségét, hogy a 30. és 31. oldalon sincs hiba! Adjuk meg az ezeken az oldalakon található sajtóhubák várható értékét és szórását!

Megoldás. Először is felidézzük a Poisson-eloszlás definícióját (rosszabb esetben megnézzük a képletgyűjteményben). Eszerint X Poisson-eloszlású $\lambda = 2$ paraméterrel, ha

$$\mathbf{P}(X = k) = \frac{\lambda^k}{k!}e^{-\lambda} = \frac{2^k}{k!}e^{-2}, \quad k = 0, 1, 2, \ldots.$$

Jelölje X a 30. oldalon, Y a 31. oldalon levő hibák számát. A feladat szerint ezek független, Poisson-eloszlásúak 2 paraméterrel. Az, hogy egyik oldalon sincsen hiba, azt jelenti, hogy $X = 0$ és $Y = 0$. Ennek a valószínűsége, a függetlenség miatt

$$\mathbf{P}(X = 0, Y = 0) = \mathbf{P}(X = 0) \cdot \mathbf{P}(Y = 0) = e^{-2} \cdot e^{-2} = e^{-4}.$$

Tudjuk, hogy λ-paraméterű Poisson várható értéke és szórásnégyzete is λ, ezért

$$\mathbf{E}(X + Y) = \mathbf{E}(X) + \mathbf{E}(Y) = 4,$$

és mivel független változók összegének szórásnégyzete egyenlő a szórásnégyzetek összegével, ezért

$$\mathbf{D}^2(X + Y) = \mathbf{D}^2(X) + \mathbf{D}^2(Y) = 4,$$

azaz $\mathbf{D}(X + Y) = 2$.

2. Egy augusztusi éjszakán megfigyelhető csillaghullások száma Poisson–eloszlást követ. Annak a valószínűsége, hogy egy éjszaka egyetlen hullócsillagot sem látunk 0,1. Várhatóan hány hullócsillag figyelhető meg egy éjszaka?

Megoldás. Itt is Poisson-eloszlás van. Most ki kell találnunk az adatokból a paramétert, mert nincs expliciten megadva, mint az előbb. Legyen X a megfigyelt hullócsillagok száma egy este. Ekkor $X \sim$Poisson(λ). Az, hogy nem látunk hullócsillagot azt jelenti, hogy $X = 0$. Azaz, azt tudjuk, hogy

$$\mathbf{P}(X = 0) = 0,1.$$

Na de a formula szerint ez éppen $\lambda^0/0! \cdot e^{-\lambda} = e^{-\lambda}$. Tehát $e^{-\lambda} = 0,1$ ahonnan $\lambda = -\log 0,1 = \log 10$ (itt a $\log = \ln$ a természetes alapú logaritmus). A Poisson paramétere tetszőleges pozitív szám lehet, nem kell, hogy egész legyen! Poisson várható értéke a paramétere, azaz

$$\mathbf{E}(X) = \lambda = \log 10.$$

3. Egy biztosítótársaság felmérte, hogy egy év során egy családi ház 0,0002 valószínűséggel gyullad ki. Mennyi a valószínűsége, hogy 2008-ban egy faluban, ahol 15000 ház van, négynél kevesebb tűz üt ki? (Közelítsünk Poisson–eloszlással!)

Megoldás. Itt is Poisson-eloszlással kell számolni. A pontos eloszlás binomiális $n = 15000$ és $p = 0{,}0002$ paraméterekkel, de a binomiális jól közelíthető Poisson-eloszlással ha p kicsi és n nagy.

Először meg kell határoznunk a λ paramétert. Mivel a tűz valószínűsége egy háznál 0,0002, és a faluban 15000 ház van, ezért az összes tűz várható értéke

$$15000 \cdot 0{,}0002 = 3.$$

Tehát egy olyan Poisson-eloszlás kell nekünk, aminek a várható értéke 3. A várható érték éppen a paraméter, tehát $\lambda = 3$. Ezek szerint annak a valószínűsége, hogy négynél kevesebb tűz üt ki az

$$\begin{aligned}\mathbf{P}(X < 4) &= \mathbf{P}(X = 0) + \mathbf{P}(X = 1) + \mathbf{P}(X = 2) + \mathbf{P}(X = 3)\\ &= \sum_{k=0}^{3} \frac{3^k}{k!} e^{-3} = e^{-3}\left(1 + 3 + \frac{9}{2} + \frac{27}{6}\right)\\ &= e^{-3} 13.\end{aligned}$$

5. Egy villanykörte élettartama exponenciális eloszlású, átlagosan 2 évig működik. Mennyi a valószínűsége, hogy legalább egy évig fog működni egy új villanykörte? Mennyi a valószínűsége, hogy legalább még egy évig fog működni egy már fél éve működő? Mennyi időt él meg a villanykörték 90%-a?

Megoldás. Exponenciális eloszlásunk van, először meg kell határozni a paramétert.

Legyen egy villanykörte évben mért élettartama X. Ekkor a feladat szerint X exponenciális eloszlású és várható értéke 2. Tudjuk (vagy megnézzük a képletgyűjteményben), hogy az exponenciális várható értéke a paraméter reciproka. Azaz, ha $X \sim \mathrm{Exp}(\lambda)$, akkor $\mathbf{E}(X) = 1/\lambda$. Az, hogy egy villanykörte átlagosan 2 évig működik, éppen azt jelenti, hogy az élettartam várható értéke 2. Azaz $2 = 1/\lambda$, ahonnan kapjuk, hogy $\lambda = 1/2$.

Az, hogy egy új villanykörte legalább egy évig jó, azt jelenti, hogy az élettartam nagyobb, mint 1. Tehát a kérdés, hogy mennyi $\mathbf{P}(X > 1)$ (mivel folytonos eloszlásunk van, mindegy, hogy ≥ 1 vagy > 1). Az exponenciális eloszlásfüggvénye $\mathbf{P}(X \leq x) = 1 - e^{-\lambda x}$, ha $x \geq 0$, és 0 különben, ezért

$$\mathbf{P}(X > 1) = 1 - \mathbf{P}(X \leq 1) = e^{-\frac{1}{2}} \approx 0{,}6.$$

A második kérdésnél tudjuk, hogy már fél éve működik az izzónk. Ez egy feltételes valószínűség, hiszen van egy részinformációm, nevezetesen $\{X > 1/2\}$. A kérdés, hogy mennyi a valószínűsége, hogy még legalább 1 évig él, azaz $\{X > 3/2\}$. A feltételes valószínűség definíciója szerint

$$\mathbf{P}(X > 3/2 | X > 1/2) = \frac{\mathbf{P}(X > 3/2, X > 1/2)}{\mathbf{P}(X > 1/2)} = \frac{\mathbf{P}(X > 3/2)}{\mathbf{P}(X > 1/2)} = \frac{e^{-\frac{1}{2}\cdot\frac{3}{2}}}{e^{-\frac{1}{2}\cdot\frac{1}{2}}} = e^{-\frac{1}{2}}.$$

Most lényegében beláttuk, hogy az exponenciális eloszlás örökifjú, ami azt jelenti, hogy tetszőleges $s, t > 0$ esetén $\mathbf{P}(X > s + t | X > s) = \mathbf{P}(X > t)$, azaz ha a változóm már s időt megélt (feltétel!), akkor annak a valószínűsége, hogy még t-t megél az ugyanannyi, mint annak a valószínűsége, hogy t-t megél (nincs feltétel!).

A harmadik kérdésen kicsit el kell mélázni. Az, hogy a időt megél az izzók 90%-a azt jelenti, hogy ha egyszerre becsavarok 100 izzót, akkor a idő után még 90 db világít. Másképpen, 10 égett ki a-ig, azaz annak a valószínűsége, hogy egy izzó élettartam a-nál kisebb, az 0,1. Tehát a kérdés az az a szám, melyre

$$\mathbf{P}(X \leq a) = 0{,}1.$$

(Mivel folytonos eloszlásunk van, ezért mindegy, hogy hol van szigorú egyenlőtlenség, és hol nincs.) Beírva az eloszlásfüggvény képletét ($\lambda = 1/2$, ezt már kiszámoltuk!)

$$0{,}1 = 1 - e^{-\frac{a}{2}},$$

vagyis $a = 2 \log \frac{10}{9} \approx 0{,}21$. Azaz 0,21 évet él meg a villanykörték 90%-a.

8. A skót bakák mellkasának körmérete $N(88, 10)$ eloszlást követ. Mekkora hányaduk fér bele 84-es zubbonyba?

Megoldás. Végre egy normális eloszlás. Most mindkét paraméter meg van adva. A várható érték $\mu = 88$, a szórásnégyzet $\sigma^2 = 10$. Arra figyeljünk, hogy a szórásnégyzet a 10, nem a szórás! Azt fogjuk használni, hogy ha X normális eloszlású μ és σ^2 paraméterekkel, akkor $Z = (X - \mu)/\sigma$ standard normális eloszlású, ami meg táblázatba van szedve. Figyeljünk oda, hogy mi σ és mi σ^2!

Legyen X egy skót baka mellkasának körmérete. Erre úgy gondolunk, hogy a skót katonák összességéből véletlenszerűen kiválasztunk egyet. Ezzel a jelöléssel a feladat éppen a $\{X \leq 84\}$ esemény valószínűségét kérdezi. Ez pedig

$$\mathbf{P}(X \leq 84) = \mathbf{P}\left(\frac{X - 88}{\sqrt{10}} \leq \frac{84 - 88}{\sqrt{10}}\right) = \mathbf{P}(Z \leq -1{,}29) = \Phi(-1{,}26).$$

Az standard normális eloszlásfüggvény táblázatában negatív x-ek nincsenek. De nem baj, tudjuk, hogy a sűrűség páros függvény, amiből következik, hogy

$$\Phi(-x) = 1 - \Phi(x), \quad x > 0.$$

Tehát

$$\Phi(-1{,}26) = 1 - \Phi(1{,}26) = 1 - 0{,}8962 = 0{,}1038.$$

Azaz a skót bakák kb. 10%-a fér bele egy 84-es zubbonyba.

9. Egy munkadarabokat készítő gép 40 cm-re van beállítva. A hiba normális eloszlást követ 0 várható értékkel. Annak a valószínűsége, hogy egy munkadarab nagyobb, mint 40,5 cm, 0,05. Mennyi a szórás?

Megoldás. Jelölje X egy munkadarab hibáját (előjellel, tehát lehet negatív is), azaz a munkadarab hossza $40 + X$ cm. Tudjuk, hogy X normális eloszlású. Most nincs megadva mindkét paramétere a normális eloszlásnak. Azt tudjuk, hogy 0 a várható érték, azaz $\mu = 0$. Legyen σ^2 a szórásnégyzet. Azt tudjuk még, hogy

$$\mathbf{P}(X > 0{,}5) = 0{,}05,$$

hiszen a munkadarab pontosan akkor nagyobb, mint 40,5, ha a hiba nagyobb, mint 0,5. Megint a normális eloszlás skálázási tulajdonságát használjuk. Eszerint $X/\sigma = Z$ standard normális. Tehát

$$\mathbf{P}(X > 0{,}5) = \mathbf{P}(X/\sigma > 0{,}5/\sigma) = \mathbf{P}(Z > 0{,}5/\sigma) = 1 - \Phi(0{,}5/\sigma) = 0{,}05.$$

Innen adódik, hogy

$$\Phi(0{,}5/\sigma) = 0{,}95$$

Most azt kell megkeresni a normális eloszlás táblázatában, hogy hol veszi fel a 0,95 értéket. Ez valahol 1,64 és 1,65 között történik, legyen 1,65. Tehát azt kaptuk, hogy

$$\frac{0{,}5}{\sigma} = 1{,}65,$$

ahonnan

$$\sigma = 0{,}3.$$

Azaz a szórás 0,3.

13. Az FC Barcelona passzolási hatékonysága 2019. áprilisában $p = 0{,}85$ (azaz egy passz ekkora valószínűséggel sikeres). Adjuk meg annak a valószínűségét, hogy a Liverpool elleni 525 passzból legalább 460 sikeres.

Megoldás. Ez a legklasszikusabb példa a de Moivre–Laplace-tétel alkalmazására. Jelölje S a sikeres passzok számát. Mivel összesen 525 passz van, és mindegyik passz egymástól függetlenül $0,85$ valószínűséggel sikeres, ezért S eloszlása binomiális $n = 525$ és $p = 0{,}85$ paraméterekkel. A de Moivre–Laplace-tétel szerint

$$\frac{S - np}{\sqrt{np(1-p)}}$$

közelítőleg standard normális eloszlású, azaz

$$\mathbf{P}\left(a \leq \frac{S - np}{\sqrt{np(1-p)}} \leq b\right) \approx \Phi(b) - \Phi(a).$$

A pontos állítás az, hogy ha $n \to \infty$, akkor az $\approx$ helyett $=$ van. Ha csak felső korlát van, akkkor $a = -\infty$ és $\Phi(-\infty) = 0$, ha csak alsó, akkor $b = \infty$ és $\Phi(\infty) = 1$. A feladat kérdése a $\mathbf{P}(S \geq 460)$ valószínűség. Egyszerű átalakítással elérjük, hogy $(S - np)/\sqrt{np(1-p)}$ jelenjen meg. Az $np = 525 \cdot 0{,}85 \approx 446$ és $\sqrt{np(1-p)} \approx 8{,}2$ értékeket beírva

$$\mathbf{P}(S \geq 460) = \mathbf{P}\left(\frac{S - np}{\sqrt{np(1-p)}} \geq \frac{460 - np}{\sqrt{np(1-p)}}\right) = \mathbf{P}\left(\frac{S - np}{\sqrt{np(1-p)}} \geq 1{,}7\right)$$
$$\approx 1 - \Phi(1{,}7) = 1 - 0{,}955 = 0{,}045.$$

Tehát a keresett valószínűsége 0,045.

14. Chicago és Los Angeles között két vasútvonal van, melyek mindegyikén egy-egy vonat közlekedik. Mindkét vonat egyidőben indul, lényegében egyformán kényelmes és k személyes. Tegyük fel, hogy 1000 utas egymástól függetlenül $1/2 - 1/2$ valószínűséggel választ vonatot. Legalább mekkora legyen az ülőhelyek k száma, hogy 0,01-nél kisebb legyen annak a valószínűsége, hogy lesz olyan utas, akinek nem jut ülőhely?

Megoldás. Ez is de Moivre–Laplace. Jelölje S az A társasággal utazók számát. Ekkor a B-vel utazók száma $1000 - S$. Mivel $n = 1000$ ember egymástól függetlenül $1/2 - 1/2$ valószínűséggel dönt A ill. B mellett, ezért S binomiális eloszlású $n = 1000$ és $p = 1/2$ paraméterekkel. Legyen k az ülőhelyek száma (mindkét vonaton). Az, hogy lesz olyan, akinek nem jut hely, azt jelenti, hogy vagy az A társaságnál túl sokan vannak, vagy a B-nél. Azaz, vagy $S > k$ (A vonat betelik) vagy $1000 - S > k$ (B vonat betelik). A kérdés a legkisebb olyan k érték, melyre

$$\mathbf{P}(S > k \text{ vagy } S < 1000 - k) \leq 0{,}01.$$

Vegyük észre, hogy $k \geq 500$ kell legyen, és ekkor csak az egyik vonat telhet be, azaz a fenti valószínűségben szereplő két esemény egymást kizáró, ahonnan

$$\mathbf{P}(S > k \text{ vagy } S < 1000 - k) = \mathbf{P}(S > k) + \mathbf{P}(S < 1000 - k).$$

Használjuk a de Moivre–Laplace-tételt. Beírjuk az $np = 500$, $\sqrt{np(1-p)} = 5\sqrt{10} \approx 15{,}8$ értékeket. Ekkor

$$\mathbf{P}(S > k) = \mathbf{P}\left(\frac{S - np}{\sqrt{np(1-p)}} > \frac{k - np}{\sqrt{np(1-p}}\right) \approx 1 - \Phi((k - 500)/15{,}8),$$

és ugyanígy (csak felhasználjuk, hogy $\Phi(-x) = 1 - \Phi(x)$ ha $x > 0$, és hogy $500 - k \leq 0$)

$$\mathbf{P}(S < 1000 - k) = \mathbf{P}\left(\frac{S - np}{\sqrt{np(1-p)}} > \frac{1000 - k - np}{\sqrt{np(1-p}}\right) \approx 1 - \Phi((k - 500)/15{,}8).$$

Azaz, keressük azt a legkisebb k értéket melyre

$$2\,[1 - \Phi((k - 500)/15{,}8)] \leq 0{,}01.$$

Átrendezve,

$$\Phi((k - 500)/15{,}8) \geq 0{,}995.$$

Kikeressük a normális eloszlás táblázatából, hogy hol veszi fel Φ a 0,995 értéket. Ez 2,57, azaz

$$\frac{k - 500}{15{,}8} \geq 2{,}57,$$

amit átrendezve

$$k \geq 540{,}6.$$

Mivel k egész, ezért 541 a legkisebb olyan k, amire a feltétel teljesül.

16. Budapesten meg akarják állapítani a dohányosok p arányát. Ehhez kiválasztanak n egyént úgy, hogy minden választásnál mindenki ugyanakkora valószínűséggel kerül kiválasztásra, és csak ezek közt nézik meg a dohányosok k számát. Legalább mekkora legyen az n, hogy a kapott $p' = k/n$ arány legalább 0,95 valószínűséggel legfeljebb 0,005 hibával közelítse a valódi p arányt, akármi is $p \in (0, 1)$?

Megoldás. Ez már érdekesebb feladat, ugyanis semmi nincs megadva. Világos, hogy a feladat nagyon fontos, ugyanis a közvéleménykutatásokhoz pontosan ilyen típusú kérdést kell feltenni. Van valami ismeretlen valószínűség p, ami azt mutatja meg, hogy az emberek ilyen aránya szavazna az A párt jelöltjére. Nem tudjuk mi a p, de erről szeretnénk valamit mondani. Hány embert kell megkérdezni, hogy valami okosat mondhassunk?

Jelölje S a dohányzók számát a megkérdezettek között. Ekkor S binomiális eloszlású véletlen változó n (meghatározandó, de ismert) és p (ismeretlen) paraméterekkel. Világos, hogy az ismeretlen p értékre az S/n becslést adjuk. Elég összetett a kérdés, kicsit el kell rajta gondolkodni. A becslés hibája $|S/n - p|$. Azt akarjuk, hogy ez nagy valószínűséggel (0,95) kicsi legyen (0,005-nél kisebb), azaz olyan n értéket keresünk, amire

$$\mathbf{P}\left(\left|\frac{S}{n} - p\right| < 0{,}005\right) \geq 0{,}95.$$

(Annak a valószínűsége, hogy a hiba 0,005-nél kisebb, legalább 0,95.) A neheze megvan. Használjuk a de Moivre–Laplace-tételt. Eszerint

$$\mathbf{P}\left(a \leq \frac{S - np}{\sqrt{np(1-p)}} \leq b\right) \approx \Phi(b) - \Phi(a).$$

Tehát be kell erőltetni az $(S-np)/\sqrt{np(1-p)}$ kifejezést. Tegyük meg:

$$\begin{aligned}
\mathbf{P}\left(\left|\frac{S}{n}-p\right|<0{,}005\right) &= \mathbf{P}\left(\left|\frac{S-np}{n}\right|<0{,}005\right)\\
&= \mathbf{P}\left(\left|\frac{S-np}{\sqrt{np(1-p)}}\right|<0{,}005\frac{\sqrt{n}}{\sqrt{p(1-p)}}\right)\\
&\approx \Phi\left(0{,}005\frac{\sqrt{n}}{\sqrt{p(1-p)}}\right)-\Phi\left(-0{,}005\frac{\sqrt{n}}{\sqrt{p(1-p)}}\right)\\
&= 2\Phi\left(0{,}005\frac{\sqrt{n}}{\sqrt{p(1-p)}}\right)-1.
\end{aligned}$$

Annyit használtunk, hogy $|x|\leq a$ pontosan akkor, ha $-a<x<a$, és hogy $\Phi(-x)=1-\Phi(x)$. Ezek szerint az kell, hogy

$$2\Phi\left(0{,}005\frac{\sqrt{n}}{\sqrt{p(1-p)}}\right)-1\geq 0{,}95,$$

azaz

$$\Phi\left(0{,}005\frac{\sqrt{n}}{\sqrt{p(1-p)}}\right)\geq 0{,}975.$$

A táblázatból kikeresve azt kapjuk, hogy

$$0{,}005\frac{\sqrt{n}}{\sqrt{p(1-p)}}\geq 1{,}96.$$

Átrendezve

$$n\geq 392^2\cdot p(1-p). \tag{$*$}$$

Na de nem ismerjük p értékét. Úgy kell n-et választani, hogy a fenti egyenlőtlenség minden p-re igaz legyen. Tehát válasszuk p-t úgy, hogy a jobb oldal maximális legyen. Ez $p=1/2$-nél van, értéke $1/4$. Tehát, ha

$$n\geq 396^2\frac{1}{4}=38416,$$

akkor $(*)$ teljesül minden $p\in[0,1]$ esetén.

17. Legyen f folytonos függvény a $[0,1]$ intervallumon. A hozzátartozó n-edik *Bernstein-polinom* $B_n(f)(x)=\sum_{k=0}^{n}f(k/n)\binom{n}{k}x^k(1-x)^{n-k}$. Igazoljuk, hogy $B_n(f)$ egyenletesen konvergál az f függvényhez, azaz

$$\lim_{n\to\infty}\sup_{x\in[0,1]}|B_n(f)(x)-f(x)|=0.$$

Megoldás. Ez egy nehezebb elméleti feladat, ilyen nem lesz a dolgozatban. Azért próbálja mindenki megérteni a megoldást.

Ennek igazolásához vegyük észre, hogy $B_n(f)(x) = \mathbf{E}f(S_n/n)$, ahol $S_n = X_1 + \ldots + X_n$, és $X_1, \ldots, X_n$ független azonos eloszlású Bernoulli(x) véletlen változók (azaz $\mathbf{P}(X_1 = 1) = x = 1 - \mathbf{P}(X_1 = 0)$). Vagyis S_n pedig binomiális eloszlású n és $p = x$ paraméterekkel, azaz

$$\mathbf{P}(S_n = k) = \binom{n}{k} x^k (1-x)^{n-k}, \quad k = 0, 1, \ldots, n.$$

A várható érték tulajdonságai szerint

$$\mathbf{E}(f(S_n/n)) = \sum_{k=0}^{n} f(k/n)\mathbf{P}(S_n = k) = \sum_{k=0}^{n} f(k/n)\binom{n}{k} x^k (1-x)^{n-k} = B_n(f)(x).$$

A Csebisev-egyenlőtlenség szerint

$$\mathbf{P}\left(\left|\frac{S_n}{n} - x\right| > c\right) \le \frac{\mathbf{D}^2(S_n)}{n^2c^2} = \frac{x(1-x)}{nc^2}.$$

Legyen $\varepsilon > 0$ rögzített. Mivel folytonos függvény zárt intervallumon egyenletesen folytonos, ezért létezik olyan $\delta = \delta(\varepsilon) > 0$, hogy $|u - v| \le \delta$ esetén $|f(u) - f(v)| \le \varepsilon$. Így

$$\begin{aligned} |f(x) - B_n(f)(x)| &= |\mathbf{E}\left[f(x) - f(S_n/n)\right]| \le 2M\mathbf{P}(|S_n/n - x| > \delta) + \varepsilon \\ &\le 2M\frac{\mathbf{D}^2(S_n)}{n^2\delta^2} + \varepsilon \le \frac{2Mx(1-x)}{n\delta^2} + \varepsilon, \end{aligned}$$

ahol M az $|f|$ maximuma a $[0, 1]$ intervallumon. A kapott becslés x-ben egyenletes, ezért az állítást beláttuk.

Ez a Weierstrass-féle approximációs tétel egy konstruktív bizonyítása.

5. feladatsor 11. Valamely növényfajta magjaiból álló mintában a hibás magok száma λ paraméterű Poisson–eloszlású véletlen változó. Minden mintát 3 technikus vizsgál meg egymás után, hogy eltávolítsák a hibás magokat. Az i-edik technikus $p_i < 1$ valószínűséggel veszi észre a hibás magokat; döntései az egyes magokra nézve függetlenek, és az egyes technikusok is egymástól függetlenül döntenek. Határozzuk meg az el nem távolított hibás magok eloszlását!

Megoldás. Azt nézzük, hogy egy technikus után mi a helyzet. Legyen $p = p_1$. Jelölje X az eredeti mintában a hibás magok számát, és Y a technikus ellenőrzése után megmaradt hibás magok számát. Tudjuk, hogy X Poisson-eloszlású λ paraméterrel, azaz lehetséges értékei $0, 1, \ldots$. Világos, hogy Y-nak is ezek a lehetséges értékei.

Így nehéz a feladat. Az eredeti hibás magok száma is véletlen. Mi lenne, ha tudnánk a hibás magok számát. Tegyük fel, hogy az eredeti hibás magok száma k, azaz $X = k$.

Jön az első technikus, minden egyes hibás magot p valószínűséggel vesz észre, azaz $1-p$ valószínűséggel nem veszi észre. Ez minden hibás magra egymástól független, ezért a bennemaradt hibás magok száma binomiális eloszlású k és $1-p$ paraméterekkel. Tehát, az $X=k$ feltétel mellett (ez egy kutya közönséges feltételes valószínűség)

$$\mathbf{P}(Y=\ell|X=k)=\binom{k}{\ell}(1-p)^{\ell}p^{k-\ell}, \quad \ell=0,1,\dots,k.$$

Akkor alkalmazzuk a teljes valószínűség tételét az $\{X=0\}$, $\{X=1\}$, ..., teljes eseményrendszerrel. Eszerint

$$\mathbf{P}(Y=\ell)=\sum_{k=0}^{\infty}\mathbf{P}(Y=\ell|X=k)\mathbf{P}(X=k).$$

Írjuk be a Poisson-eloszlás formuláját, és a fent kapott formulát. Kicsit számolunk (persze ha $Y=\ell$ akkor X is legalább ℓ, azaz $\mathbf{P}(Y=\ell|X=k)=0$, ha $k<\ell$)

$$\begin{aligned}
\mathbf{P}(Y=\ell)&=\sum_{k=0}^{\infty}\mathbf{P}(Y=\ell|X=k)\mathbf{P}(X=k)\\
&=\sum_{k=\ell}^{\infty}\binom{k}{\ell}(1-p)^{\ell}p^{k-\ell}\frac{\lambda^k}{k!}e^{-\lambda}\\
&=(1-p)^{\ell}\lambda^{\ell}e^{-\lambda}\frac{1}{\ell!}\sum_{k=\ell}^{\infty}\frac{1}{(k-\ell)!}(\lambda p)^{k-\ell}\\
&=[\lambda(1-p)]^{\ell}e^{-\lambda}\frac{1}{\ell!}e^{\lambda p}\\
&=\frac{[\lambda(1-p)]^{\ell}}{\ell!}e^{-(1-p)\lambda}.
\end{aligned}$$

Az utolsó előtti egyenlőségnél használtuk, hogy az exponenciális függvény Taylor-sora jelent meg ($e^x=\sum_{k=0}^{\infty}x^k/k!$). Látjuk, hogy ez éppen egy $(1-p)\lambda$ paraméterű Poisson-eloszlás.

Tehát az első technikus után maradt hibás magok száma is Poisson. Azaz a második technikus is Poisson darab hibás magot kap, csak a paraméter $(1-p_1)\lambda$. A fentiek szerint a második technikus után maradt hibás magok száma is Poisson, a paraméter pedig $(1-p_1)(1-p_2)\lambda$, végül a harmadik után maradt hibás magok száma is Poisson-eloszlású, $(1-p_1)(1-p_2)(1-p_3)\lambda$ paraméterrel.