Felsőbb Matematika Villamosmérnököknek - Sztochasztika
2. ZH megoldások

2016 ősz, 2016.11.29 18:00

Munkaidő: 90 perc.

1. Egy radioaktív mintában másodpercenkent átlagosan 3 kis energiájú és 1 nagy energiájú alfa-részecske keletkezik. A detektorunk a nagy energiájú részecskéket 90% valószínűséggel észleli, a kis energiájúakat viszont csak 20% valószínűséggel (a többi részecskétől függetlenül). Mennyi a valószínűsége, hogy egy két másodperc hosszú időintervallumban legalább 4 részecskét észlel?

   Megoldás: A nagyenergiájú alfa-részecskék Poisson-folyamat szerint keletkeznek, 1 rátával. (Az időt mérjük másodpercben.) Az *észlelt* nagyenergiájú részecskék folyamata ennek egy ritkítása, ezért szintén Poisson-folyamat, $1 \cdot 0.9 = 0.9$ rátával. Ugyanígy, a kisenergiájú alfa-részecskék is Poisson-folyamat szerint keletkeznek 3 rátával, tehát az *észlelt* kisenergiájú részecskék folyamata is Poisson-folyamat, $3 \cdot 0.2 = 0.6$ rátával. A kétféle észlelt részecske folyamata független, ezért az uniójuk is Poisson-folyamat – és a ráták összeadódnak.

   Összefoglalva a lényeg: az *összes észlelt* részecske folyamata Poisson-folyamat, $0.9+0.6 = 1.5$ rátával.

   Legyen X a két másodperc alatt észlelt részecskék száma. Így $X \sim Poi(\lambda)$, ahol $\lambda = \mathbb{E}X = 2 \cdot 1.5 = 3$. Vagyis $k = 0, 1, 2, \ldots$-re $\mathbb{P}(X = k) = e^{-\lambda}\frac{\lambda^k}{k!} = e^{-3}\frac{3^k}{k!}$. A kérdésre a válasz:

$$\mathbb{P}(X \geq 4) = 1 - \sum_{k=0}^{3} \mathbb{P}(X = k) = 1 - e^{-3}\left(1 + 3 + \frac{3^2}{2} + \frac{3^3}{6}\right) = 1 - 13e^{-3} \approx 0.3528 \approx 35\%.$$

2. Az X valószínűségi változó generátorfüggvénye $g(z) = c(z+z^2+z^4+z^5+z^8+z^{11}+z^{13}+z^{15})$.

   a.) (1 pont) Mennyi a c konstans értéke?

   b.) (4 pont) Mennyi X várható értéke?

   c.) (4 pont) Mennyi a $\mathbb{P}(X = 8)$ valószínűség?

   Megoldás:

   a.) $g(1) = 1$ mindig, vagyis $c = \frac{1}{8}$.

   b.) $\mathbb{E}X = g'(1) = \frac{1}{8}(1 + 2z + 4z^3 + 5z^4 + 8z^7 + 11z^{10} + 13z^{12} + 15z^{14}) \,|_{z=1} = \frac{59}{8} = 7.375$.
   Avagy: a $g(z) = \sum_{k=0}^{\infty} p_k z^k$ generátorfüggvényből leolvasható az X eloszlása:

k	1	2	4	5	8	11	13	15
$p_k = \mathbb{P}(X = k)$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$

   amiből szintén könnyen $\mathbb{E}X = \frac{1+2+4+5+8+11+13+15}{8} = \frac{59}{8} = 7.375$.

   c.) $\mathbb{P}(X = 8) = p_8$ a z^8 együtthatója a $g(z) = \sum_{k=0}^{\infty} p_k z^k$ sorfejtésben, vagyis esetünkben $\mathbb{P}(X = 8) = \frac{1}{8}$.

3. Móricka az egyetemi órák látogatásának egészségügyi kockázatairól ír egy kamu lánclevelet, és elküldi 10 ismerősének a nulladik napon. A levélben benne van, hogy a címzett küldje tovább újabb 10 embernek. A levelet a címzettek egymástól függetlenül 90% valószínűséggel olvasatlanul törlik, ám a maradék 10% valószínűséggel tényleg továbbküldik 10 embernek, a következő napon.

   a.) Várhatóan hányan küldenek levelet a harmadik napon?

b.) Mennyi a valószínűsége annak, hogy előbb-utóbb senki nem küldi tovább a levelet?

c.) Mennyi a levelet továbbküldő emberek számának várható értéke?

Megoldás: Legyen Z_n az n-edik napon levelet küldő emberek száma. *(Vagyis nem az az érdekes, hogy hány levelet küld – aki küld, az úgyis pont* 10*-et küld, – hanem, hogy hány embert fertőz meg a lánclevél-küldés.)* Ez a Z_n Galton-Watson elágazó folyamat. Mindenki $n = 10$ embert próbál megfertőzni, és minden próbálkozás $p = \frac{1}{10}$ valószínűséggel sikeres, így a közvetlen utódok számának várható értéke $m = 1$.

(Konkrétan az egylépéses utódszámeloszlás $X \sim Bin(n,p) = Bin(10, \frac{1}{10})$, de ez most nekünk csak annyiban kell, hogy $m = np = 1$.)

Ezért

a.) $\mathbb{E}Z_3 = m^3 = 1$.

b.) $m = 1$, vagyis a folyamat kritikus (és nem elfajult). Ezért a kihalás valószínűsége 1.

c.) A folyamat kritikus, ezért a teljes létszám várható értéke $\mathbb{E}N = \infty$.

4. Egy béka a számegyenesen ugrál. A nullából indul, majd minden másodpercben urgik egyet: $\frac{1}{3}$ valószínűséggel helyben, $\frac{1}{3}$ valószínűséggel egy egységnyit jobbra, $\frac{1}{3}$ valószínűséggel pedig egy egységnyit balra – az előzményektől függetlenül. Móricka a centrális határeloszlás tétel segítségével próbálja megbecsülni annak valószínűségét, hogy a béka 150 ugrás után legalább 10 egységnyivel jutott jobbra. Legfeljebb mennyi lesz Móricka becslésének hibája a Berry-Esseen tétel szerint? *(A tételben szereplő konstanst vehetjük* 0.4748*-nak.)*

Megoldás: Legyen X_i a béka i-edik ugrása. A szöveg szerint az X_i-k függetlenek és egyenletesek a $\{-1, 0-1\}$ halmazon. Legyen $n = 150$ és $S_n = X_1 + \cdots + X_n$ a béka helye 150 ugrás után. Móricka a $\mathbb{P}(S_n \geq 10)$ valószínűséget akarja CHT-vel becsülni. A Berry-Esseen tétel szerint a becslés hibájára

$$\text{hiba} \leq \frac{C\delta}{\sqrt{n}\sigma^3},$$

ahol $C = 0.4748$, σ az X_i-k (közös) szórása és $\delta = \mathbb{E}(|X_i - m|^3)$, ahol $m = \mathbb{E}X_i$. Esetünkben

$$\begin{aligned}
m &= \frac{-1+0+1}{3} = 0 \\
\sigma^2 &= \mathbb{E}X_i^2 - m^2 = \mathbb{E}X_i^2 = \frac{(-1)^2 + 0^2 + 1^2}{3} = \frac{2}{3} \\
\delta &= \mathbb{E}(|X_i - m|^3) = \mathbb{E}(|X_i|^3) = \frac{|-1|^3 + |0|^3 + |1|^3}{3} = \frac{2}{3}.
\end{aligned}$$

Ezeket visszhelyettesítve

$$\text{hiba} \leq \frac{0.4748 \cdot \frac{2}{3}}{\sqrt{150}\sqrt{\frac{2}{3}}^3} = \frac{0.4748}{\sqrt{100}} = 0.04748 \approx 4.7\%.$$

5. Egy 45 pontos ZH-n a hallgatók hosszú évek tapasztalata szerint átlagosan 29 pontot szoktak elérni. Adjunk nagy eltérés becslést annak valószínűségére, hogy idén a 130 hallgató átlaga legfeljebb 20 pont lesz. *(Tegyük fel, hogy a feladatsor ugyanolyan nehéz, mint máskor, és a hallgatók is ugyanolyan felkészültek, mint máskor. Az egyes hallgatók eredményei függetlenek. Negatív pontszámot nem lehet elérni.)*

Megoldás: Legyen $n = 130$ és legyen $i = 1, 2, \ldots, n$-re X_i az i-edik hallgató pontszáma. Így $S_n := X_1 + \cdots + X_n$ az évfolyam összpontszáma. Feladat a $\mathbb{P}(\frac{S_n}{n} \le 20) = \mathbb{P}(S_n \le 2600)$ valószínűség becslése.

A feladat szerint az X_i-k függetlenek és korlátosak: $0 = a_i \le X_i \le b_i = 45$. Egyébként az eloszlásukról semmit nem tudunk – nincs okunk pl. feltételezni, hogy azonos eloszlásúak lennének. Így a Cramér féle nagy eltérés tétel nem használható.

Szerencsére tudjuk viszont az *összeg* várható értékét. Pontosabban, a szöveg szerint $\mathbb{E}\frac{S_n}{n} = 29$, amiből $\mathbb{E}S_n = 29n = 3770$. A Hoeffding egyenlőtlenség alkalmazásához ez éppen elég: Hoeffding szerint $t \ge 0$-ra

$$\mathbb{P}(S_n \le \mathbb{E}S_n - t) \le \exp\left\{-\frac{2t^2}{\sum_{i=1}^n (b_i - a_i)^2}\right\}.$$

Esetünkben $\sum_{i=1}^n (b_i - a_i)^2 = 130(45-0)^2 = 263250$, így $t := 3770 - 2600 = 1170$ választással

$$\mathbb{P}(S_n \le 2600) = \mathbb{P}(S_n \le \mathbb{E}S_n - t) \le \exp\left\{-\frac{2 \cdot 1170^2}{263250}\right\} = e^{-10.4} \approx 3 \cdot 10^{-5}.$$