Felsőbb Matematika Informatikusoknak - Sztochasztika
2. ZH minta

2015 ősz

A ZH-n 5 feladat lesz. Ebbe nem fér bele minden feladattípusból $1-1$, ezért az 5 feladat véletlenszerűen lesz kiválasztva az alábbi típuspéldákhoz nagyon hasonló feladatok közül, valamint esetleg a házi feladatokhoz nagyon hasonló feladatok közül. (A két csoport között jelentős az átfedés, de nincs tartalmazás.)

1. Pistike e-mail címére munkahelyi levelek, magánlevelek és spam érkezik. Munkahelyi levélből átlagosan kétóránként egy, magánlevélből kétóránként három, spamből pedig óránként egy.

   a.) Mennyi a valószínűsége, hogy 10:00 és 18:00 között pont 3 munkahelyi emailt kap?

   b.) Mennyi a valószínűsége, hogy 10:00 és 12:00 között a háromféle levélből összesen 4-et kap?

   c.) Mennyi a valószínűsége, hogy 10:00 és 12:00 között összesen 4 levelet kap, de ebből egy se spam?

2. Legyenek $X_1, X_2, \dots$ független valószínűségi változók, melyek eloszlása egyenletes a $[0,1]$ intervallumon. Legyen
$$M_n = \max\{X_1, \dots, X_n\}$$
és legyen F_n az M_n eloszlásfüggvénye, vagyis $F_n(x) := \mathbb{P}(M_n < x)$.

   a.) Számoljuk ki az F_n eloszlásfüggvényt. *Segítség: az $\{X_1, \dots, X_n\}$ maximuma persze pontosan akkor kisebb x-nél, ha minden X_i külön-külön kisebb x-nél.*

   b.) Legyen $Y_n = n(1 - M_n)$ és legyen G_n az Y_n eloszlásfüggvénye. Számoljuk ki a G_n eloszlásfüggvényt. *Segítség:* $\mathbb{P}(Y_n < y) = \mathbb{P}(n(1 - M_n) < y) = \mathbb{P}(M_n > 1 - \frac{y}{n})$.

   c.) Számoljuk ki a $G(y) := \lim_{n\to\infty} G_n(y)$ határértéket.

   d.) A valószínűségi változók gyenge konvergenciájának eloszlásfüggvényes definíciója alapján mutassuk meg, hogy az Y_n sorozat gyengén konvergens, és nevezzük meg a határeloszlást.

3. a.) Legyen az X valószínűségi változó pesszimista geometriai eloszlású $p \in (0,1)$ paraméterrel, vagyis $\mathbb{P}(X = k) = q^k p$ $(k = 0, 1, \dots)$, ahol $q := 1 - p$. Számoljuk ki X generátorfüggvényét – vagyis a $g_X(z) := \sum_{k=0}^{\infty} \mathbb{P}(X = k) z^k$ függvényt. *(Segítség: $\sum_{k=0}^{\infty} x^k = \frac{1}{1-x}$, ha $-1 < x < 1$.)*

   b.) Egy Y nemnegatív egész értékű valószínűségi változó generátorfüggvénye $g_Y(z) = \frac{2+3z+4z^3+z^7}{10}$. Határozzuk meg ebből Y eloszlását, vagyis hogy melyik $k \in \mathbb{N}$ értéket mekkora valószínűséggel veszi fel. *(Segítség: ezt a generátorfüggvényt könnyen fel lehet írni $g_Y(z) = \sum_k p_k z^k$ alakban.)*

4. Egy X valószínűségi változó generátorfüggvénye $g(z) = \left(\frac{1}{2-z}\right)^3 e^{3(z-1)}$.

   a.) Mennyi X várható értéke?

   b.) Mennyi X szórása?

5. Móricka programot írt, de nem tudja lefordítani, mert van benne egy hiba, így a fordító kiír egy hibaüzenetet. Ezért Móricka módosít valamit a programon, de sajnos nem érti a hibaüzenetet. Ezért a módosítás csak $\frac{4}{10}$ valószínűséggel vezet a hiba megszűnéséhez, $\frac{6}{10}$ valószínűséggel viszont egy újabb hibát hoz létre (miközben a régi hiba is megmarad).

Ezután Móricka újra megpróbálja lefordítani a programot, és ha van hiba, akkor minden hiba után kap egy hibaüzenetet. Ekkor mindegyik hibát megpróbálja kijavítani, majd újra megpróbálja lefordítani, stb. A folyamat során Móricka két fordítási kísérlet között mindig minden hibát megpróbál kijavítani, és a javítási kísérletek egymástól függetlenül mindig $\frac{4}{10}$ valószínűséggel sikeresek, illetve $\frac{6}{10}$ valószínűséggel vezetnek újabb hibához.

a.) Mennyi a valószínűsége, hogy Mórickának legfeljebb 4 próbálkozásból sikerül lefordítania a programot (a legelső, sikertelen kísérletet nem beleértve)?

b.) Mennyi a (legelső sikertelent követő) negyedik fordítási kísérlet során talált hibák számának várható értéke?

c.) Mennyi a valószínűsége, hogy a program valaha is lefordítható lesz?

6. Egy kis vezeték nélküli internet-szolgáltató háromféle díjcsomagot kínál. Az első csomag legfeljebb 2 Mbit/sec (letöltési) sebességet biztosít. Erre 100-an fizetnek elő, és csúcsidőben átlagosan 1 Mbit/sec adatot töltenek le. A második csomag legfeljebb 3 Mbit/sec sebességet biztosít, erre 300-an fizetnek elő, és csúcsidőben átlagosan 2 Mbit/sec adatot töltenek le. A harmadik csomag legfeljebb 6 Mbit/sec sebességet biztosít, erre 150-an fizetnek elő, és csúcsidőben átlagosan 3 Mbit/sec adatot töltenek le. Az egyes felhasználók igénye véletlen és független. A szolgáltató eszközei összesen legfeljebb 1300 Mbit/sec sávszélességet tudnak biztosítani az előfizetőknek. Adjunk nagy eltérés becslést annak valószínűségére, hogy ez (a csúcsidő egy adott pillanatában) kevésnek bizonyul.

7. Egy hamis érmén a fej valószínűsége $\frac{4}{10}$, de Móricka azt állítja, hogy ő mégis képes vele 1000 dobásból legalább 500-szor fejet dobni. Pistike a centrális határeloszlás tétel segítségével akarja meggyőzni Mórickát, hogy ennek a valószínűsége rettentő kicsi. Igen ám, de legfeljebb mennnyi lehet a Pistike által számolt becslés hibája a Berry-Esseen tétel szerint? (A tételbeli konstans választható 0.48-nak.)

8. Móricka elhatározza, hogy addig dobál egy dobókockát, amíg 1000-szer ki nem jön neki a hatos. (Persze nem feltétlenül kell az 1000 hatosnak egymás után kijönni.) Valamelyik nagy eltérés tétel segítségével becsüljük meg annak a valószínűségét, hogy ez legfeljebb 5000 dobásból sikerül neki.

(Segítség: a p paraméterű Bernoulli eloszlás Cramer féle rátafüggvénye

$$I(x) = x \ln\left(\frac{x}{1-x}\frac{1-p}{p}\right) + \ln\left(\frac{1-x}{1-p}\right)$$

(ha $0 < x < 1$). A p paraméterű (optimista) geometriai eloszlás Cramer féle rátafüggvénye

$$I(x) = x \ln\left(\frac{x-1}{x}\frac{1}{1-p}\right) + \ln\left(\frac{1}{p}\frac{1-p}{x-1}\right)$$

(ha $x > 1$).)