**Matematika A4** *(Valószínűségszámítás),* **1. zárthelyi,** *2009. október 19. 8:00,* **Vetier András kurzusa**
*Munkaidő: 60 perc. A megoldásokhoz adjon magyarázatot! A végeredményeket elég numerikus képlettel megadni.*

# Feladatok

1. Egy bizonyos betegséggel kapcsolatban egy orvosi teszt 0,9 valószínűséggel helyes diagnózist ad, ha az illető beteg, de 0,05 valószínűséggel téves diagnózist ad, ha az illető egészséges. A lakosság egy ezreléke szenved ebben a betegségben. Feltéve, hogy a teszt szerint beteg valaki, mi a valószínűsége annak, hogy mégis egészséges?

2. Felmérést végzünk a múlt heti ötös lottó nyertesei (azaz legalább két találatot elért játékosai) között. Csak azokat kérdezzük, akik rendszeresen játszanak hetente egy szelvénnyel, és ezt megelőzően is nyertek már (sok ilyen van). A kérdés az, hogy ezt megelőzően hány hete nyertek utoljára. Mit tippel, hány heti várakozást fogunk kapni a leggyakrabban? Átlagosan hány hetet kellett várni egy ilyen lottózónak a két nyerés között? Indokolja meg az alkalmazott eloszlás jogosságát!

3. Adja meg egy 0 és 2 között egyenletes eloszlású véletlen szám négyzetének $F(x)$ eloszlás- és $f(x)$ sűrűségfüggvényének a képletét minden $x$ valós számra.

4. Egy 4 tagú társaság tagjai egymástól függetlenül mennek el esténként a kártyaklubba ultizni. Tegyük fel, hogy megjelenésük valószínűségei: $0,9; 0,8; 0,7; 0,6$. Az, hogy hányan jelennek meg egy este, egy $X$ valószínűségi változót jelent.

   a) Hogyan szimulálná $X$-et Excel-lel? Adjon $X$-re egy definíciót egy Excel táblázat egyetlen sorában! (Adja meg - szépen tálalva - a sor celláiba kerülő utasításokat!)

   b) Ha ezek után megismétli a definíciót, mondjuk, 1000 sorban, akkor az 1000 szimulációs eredményből hogyan közelítené $X$ eloszlását grafikusan, illetve $X$ várható értékét numerikusan? (Világosan írja le a választ - szépen olvasható - mondatokban!)

# Megoldások

1. Az $A := \{\text{egeszséges}\}$ $B := \{\text{teszt szerint beteg}\}$ jelöléssel a Bayes tételből

   $$\mathbb{P}(A|B) = \frac{\mathbb{P}(B|A)\mathbb{P}(A)}{\mathbb{P}(B|A)\mathbb{P}(A) + \mathbb{P}(B|\bar{A})\mathbb{P}(\bar{A})},$$

   ahol $\bar{A}$ az $A$ esemény komplementerét jelöli. **(2 pont)**

   A megadott adatokból kapjuk, hogy a fenti kifejezés számlálója $0,05 * 0,999$, **(3 pont)**

   nevezője pedig $0,05 * 0,999 + 0,9 * 0,001$. **(5 pont)**

2. Először kiszámoljuk annak a valószínűségét ($p$), hogy egy adott szelvény adott héten nyer (azaz legalább két találatos):

   $$p = \mathbb{P}(\text{nyer}) = 1 - \mathbb{P}(\text{0 vagy 1 találat}) = 1 - (\mathbb{P}(\text{0 találat}) + \mathbb{P}(\text{1 találat})) = 1 - \frac{\binom{85}{5} + 5\binom{85}{4}}{\binom{90}{5}} \quad (\approx 0,023)$$

   **(4 pont)**

   Mivel az egyes heteken történő nyerések egymástól függetlenek, és mindegyik $p$ valószínűséggel következik be, ezért az $X =$ „ezt megelőzően hány hete nyertek utoljára” valószínűségi változó (optimista) geometriai eloszlást követ az előzőleg számolt $p$ paraméterrel. **(2 pont)**

A leggyakrabban kapott várakozási időre (hetekben) a geometriai eloszlás móduszával tippelhetünk, ami minden esetben 1. (Ez onnan látható, hogy $p_k = (1-p)^k p$ szigorúan monoton csökkenő sorozat, ami $k = 1$ -nél veszi fel a maximumát.) Tehát a tipp: 1 hét. **(2 pont)**

A két nyertes hét közötti átlagos várakozási idő megegyezik a geometriai eloszlású $X$ várhatóértékével, ami

$$\mathbb{E}(X) = \frac{1}{p} \ \ (\approx 42,9 \text{ hét}).$$

**(2 pont)**

3. Legyen $U$ a $[0,2]$ intervallumon egyenletes eloszlású valószínűségi változó (ezzel ekvivalens: $U = 2 * RND$). Nyilván $U^2$ egy valószínűséggel 0 és 4 között vesz fel értékeket, ezért a sűrűségfüggvényre $f(x) = 0$, ha $x \notin [0,4]$, az eloszlásfüggvényre pedig $F(x) = 0$, ha $x < 0$, és $F(x) = 1$, ha $x > 4$ **(3 pont)**

   Legyen ezek után $x \in [0,4]$. Ekkor

   $$F(x) = \mathbb{P}(U^2 \leq x) = \mathbb{P}(U \leq \sqrt{x}) = \sqrt{x}/2,$$

   hiszen $\sqrt{x} \in [0,2]$, és $U$ $[0,2]$-n egyenletes eloszlású. **(5 pont)**

   Már csak a sűrűségfüggvényt kell meghatároznunk $x \in [0,4]$ esetén. Ez pedig az eloszlásfüggvény deriváltja, azaz

   $$f(x) = \left(\frac{\sqrt{x}}{2}\right)' = \frac{1}{4\sqrt{x}}$$

   . **(2 pont)**

   A derivált a $[0,4]$ intervallum végpontjaiban nem létezik, hiszen az $x = 0$ és az $x = 4$ helyeken az eloszlásfüggvénynek törése van, de mint tudjuk, a sűrűségfüggvény véges sok pontban tetszőlegesen változtatható, így mindegy, hogy $f(0)$-t és $f(4)$-et mennyinek definiáljuk.

4. A helyes megoldás az alábbi Excel fájlból kiolvasható:

   `zh1_2009_10_19_vetier_4-ik_feladat_megoldasa.xls`

   A megoldás pontozása:

   A valószínűségi változó helyes definiálása IF és SUM utasításokkal (zöld cellák): **(2+2 pont)**

   A gyakoriságok helyes meghatározása a FREQUENCY utasítással (világoskék cellák): **(2 pont)**

   A relatív gyakoriságok meghatározása OSZTÁS-sal (sötétkék cellák): **(1 pont)**

   Utalás az ÁBRA helyes elkészítésére: **(1 pont)**

   A várható érték közelítése az átlaggal, amit az AVERAGE utasítás hajt végre (lila cella): **(2 pont)**