Házi feladat II.

Elek Péter

Határidő: két hét múlva óra eleje

1. Feladat. Tegyük fel, hogy $\Pr(X = j) = 1/3$ ha $j = 1, 2, 3$.

a. Határozzuk meg $E(X)$ és $Var(X)$ értékét!

b. Legyen X_1, X_2, X_3 három elemű független minta ebből az eloszlásból. Soroljuk fel az összes lehetséges mintát azok valószínűségeivel együtt!

c. Határozzuk meg az $\overline{X}$ mintaátlag, a $\hat{m}$ tapasztalati medián, az X_1 első mintaelem, az s^{*2} korrigált tapasztalati variancia és az s^* korrigált tapasztalati szórás mintavételi eloszlását!

d. Határozzuk meg a fenti becslőfüggvények várható értékét és varianciáját! (Megjegyzés: egyes esetekben általános tételekre is lehet hivatkozni.)

e. Torzítatlanul becsüli-e s^* a sokasági szórást ebben az esetben? Torzítatlanul becsüli-e s^{*2} a sokasági varianciát ebben az esetben? Magyarázzuk meg a kapott eredményt! (Megjegyzés: ha $E\left(\widehat{\theta}\right) = \theta$, akkor általában nem teljesül $E\left(\sqrt{\widehat{\theta}}\right) = \sqrt{\theta}$.)

2. Feladat. Legyen X_1, X_2 és X_3 f.a.e. Bernoulli(p) eloszlású minta, és definiáljuk p két becslőfüggvényét a következőképpen: $\hat{p}_1 = X_1 (1 + X_2 - X_3)$ és $\hat{p}_2 = X_1 + X_2 - X_3$.

a. Mely becslőfüggvény(ek) torzítatlan(ok)?

b. Határozzuk meg az átlagos négyzetes eltérésüket!

c. Melyik becslőfüggvényt választanánk?

3. Feladat. A magyarországi felnőtt férfiak magassága közelítően normális eloszlású $\mu = 176$ cm várható értékkel és $\sigma = 7$ cm szórással. Ebből a sokaságból veszünk $n = 20$ elemű f.a.e. mintát. Határozzuk meg annak valószínűségét, hogy a mintaátlag 174 cm és 180 cm közé esik!

4. Feladat. A kurzus során bizonyítás nélkül közöltük, hogy normális eloszlású f.a.e. minta esetén $\frac{ns^2}{\sigma^2} \sim \chi^2_{n-1}$. Bizonyítsuk be ezt az állítást $n = 2$ esetén!

5. Feladat. *(Becslőfüggvények összehasonlítása a mintavételi eloszlásuk szimulálásával.)* Legyen X a $[0,1]$ intervallumon egyenletes eloszlású valószínűségi változó és vegyünk ebből $n = 3$ elemű f.a.e. mintát: X_1, X_2, X_3. Vizsgáljuk az $\overline{X}$ mintaátlag és a $\hat{m} = \text{Med}\,(X_1, X_2, X_3)$ tapasztalati medián mintavételi eloszlását. (Megjegyzés: a mintavételi eloszlások tulajdonságai lineáris transzformációval könnyen átvihetők tetszőleges, $[\theta_1, \theta_2]$ intervallumon egyenletes eloszlás esetére.)

a. Számítsuk ki az $E(\overline{X})$, $Var(\overline{X})$ és $E(\hat{m})$ értékeket!

b. Az R programcsomag segítségével szimuláljunk 5000 különböző, $n = 3$ elemű mintát, és minden esetben számítsuk ki $\overline{X}$ és $\hat{m}$ értékét. Ezek után közelítsük $\overline{X}$ and $\hat{m}$ várható értékét és varianciáját az 5000 szimulációból számított átlaggal és tapasztalati varianciával! Hasonlítsuk össze ezeket a közelítéseket a fent meghatározott elméleti értékekkel!

c. A mintaátlag vagy a tapasztalati medián a jobb becslőfüggvénye a várható értéknek egyenletes eloszlású minta esetén? Miért?

6. Feladat. *(A standardizált mintaátlag mintavételi eloszlása normális eloszlású minta esetén.)* Tudjuk, hogy normális eloszlású f.a.e. minta esetén $\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim N\,(0,1)$. A gyakorlatban azonban általában a $\frac{\overline{X}-\mu}{s^*/\sqrt{n}}$ változóval dolgozunk, amely már t-eloszlású.

Ennek illusztrálására generáljunk 5000 darab, $n = 4$ elemű mintát egy $\mu = 1$ és $\sigma = 2$ paraméterű normális eloszlásból, és számítsuk ki a $\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}$, valamint a $\frac{\overline{X}-\mu}{s^*/\sqrt{n}}$ mintastatiszikákat. Ábrázoljuk egy ábrán mindkettő mintavételi eloszlását, és értelmezzük a különbségeiket!

7. Feladat. Legyen $X_1, X_2, \ldots, X_n$ f.a.e. minta $E(X_i) = \mu \neq 0$ várható értékkel és $Var(X_i) = \sigma^2$ varianciával. Határozzuk meg $1/\overline{X}$ sztochasztikus határértékét és aszimptotikus eloszlását, ahol $\overline{X}$ a mintaátlag! Minden lépésben pontosan közöljük, hogy milyen állítást használunk!

8. Feladat. Tegyük fel, hogy egy pártra a választók ismeretlen p aránya szavazna. Milyen nagynak kell lennie egy kérdőíves felmérés véletlen mintájának ahhoz, hogy a pártra szavazók mintabeli aránya legalább 0,95 valószínűséggel a p-től legfeljebb 0,025-ra legyen?

9. Feladat. *(Lásd még a kismintás feladatsor 3. feladatát!)*

Legyen n elemű f.a.e. mintánk a $[0, 2\theta]$ intervallumon egyenletes eloszlásból. Két becslőfüggvényünk a θ ismeretlen várható értékre: az $\overline{X}_n$ mintaátlag, illetve a mintaelemek maximumának fele: $Z_n = M_n/2$. Vizsgáljuk szimulációval a két becslőfüggvény nagymintás tulajdonságait! Mivel θ az egyenletes eloszlás skálaparamétere, elegendő például a $\theta = 1/2$ esetet vizsgálni.

a. Illusztráljuk szimulációval (különböző n-ek választásával és megfelelő ábrák készítésével), hogy $\overline{X}_n$ a θ-t konzisztensen becsüli, aszimptotikusan normális eloszlású és szórása $n^{-1/2}$ nagyságrendben tart zérushoz!

b. Illusztráljuk szimulációval (különböző n-ek választásával és megfelelő ábrák készítésével), hogy Z_n konzisztensen becsüli θ-t, de $\sqrt{n}\,(Z_n - \theta)$ nem tart eloszlásban a normális eloszláshoz!

c. A kismintás feladatsor 3. feladatában kiszámítottuk, hogy $Var\,(Z_n) \sim c/n^2$, ha $n \to \infty$. Ezért azt várhatjuk, hogy a $W_n = n\,(\theta - Z_n)$ valószínűségi változóknak lesz (konstanstól különböző) határeloszlása. Illusztráljuk szimulációval (különböző n-ek választásával és megfelelő ábrák készítésével), hogy tényleg ez a helyzet! Milyen alakú a határeloszlás?

d. Bizonyítsuk be formálisan, hogy Z_n konzisztensen becsüli θ-t, és W_n aszimptotikus eloszlása exponenciális! Milyen paraméterű ez az exponenciális eloszlás?

   (Útmutatás: Határozzuk meg közvetlenül a $\Pr\,(W_n < t)$ valószínűségeket minden $t > 0$ esetén, és bizonyítsuk be, hogy az eloszlásfüggvény az exponenciális eloszlásfüggvényhez tart!)