A sztochasztika alapjai fizikusoknak

6. statisztikai alapfogalmak, maximum likelihood és momentum becslés

1. Egy játék fizikai terhelhetőségére elvégzett tesztek kg-ban a következő eredményeket adták: 40, 45, 40, 42, 36.

(a) A minta alapján adjuk meg a terhelhetőség empirikus eloszlásfüggvényét, mintaátlagát, korrigált/korrigálatlan empirikus szórásnégyzetét és a mediánt!

(b) Tegyük fel, hogy a háttérváltozó normális eloszlást követ 2 szórással. Adjunk 95%-os konfidenciaintervallumot a várható értékre!

2. Egy alkatrészekből álló sokaság 6 mintapéldányának következő volt a hónapokban mért teljes élettartama: 39, 45, 67, 50, 50, 60.

(a) Adjuk meg az empirikus eloszlásfüggvényt, a mintaátlagot, és a korrigált/korrigálatlan empirikus szórásnégyzetet!

(b) Tegyük fel, hogy az élettartam exponenciális eloszlást követ. Adjunk meg a paraméter maximum likelihood és momentumbecslését is! Adjunk 90%-os konfindenciaintervallumot az eloszlás paraméterére.

3. Egy almáskertben véletlenszerűen, egymástól függetlenül találhatók fertőzött fák. Tíz egyforma nagy, egyenként három sorból álló ültetvényben rendre 0, 3, 0, 1, 0, 0, 2, 1, 1, 2 beteg fát találtak.

(a) Adjuk meg az empirikus eloszlásfüggvényt, a mintaátlagot és az empirikus szórásnégyzetet!

(b) Tegyük fel, hogy a beteg fák száma Poisson-eloszlást követ. Adjunk maximum likelihood becslést és momentumbecslést az egy sorban található fák számának várható értékére!

4. Egy adott típusú izzó élettartamára öt mérés alapján a következő adataink vannak: 2,3, 4, 1,7, 3,2, 2,8.

(a) Adjuk meg az empirikus eloszlásfüggvényt, a mintaátlagot és az empirikus szórásnégyzetet!

(b) Tegyük fel, hogy a háttéreloszlás exponenciális ismeretlen paraméterrel. Adjunk ML és momentumbecslést a paraméterre!

5. Családok jövedelmét egy olyan skálán mérjük, ahol $X = 1$ a létminimumnak felel meg. Feltételezzük, hogy a jövedelem eloszlása $f(x) = \frac{\theta}{x^{\theta+1}}$, $x \geq 1$, sűrűségfüggvénnyel adható meg. Adjunk maximum likelihood becslést θ-ra, ha 10 véletlenszerűen választott család jövedelme: 1,53, 2,76, 19,65, 4,16, 7,31, 1,21, 254,2, 5,43, 1,12, 1,63.

6. Egy tóban a halakat egy betegség támadta meg, mely ismert p valószínűséggel pusztítja el az egyes egyedeket. A kifogott haltetemek k számából adjunk maximum likelihood becslés a betegség előtt a tóban élt halak számára!

7. Egy alkatrész élettartama exponenciális eloszlású θ/t várható értékkel, ha t hőmérsékleten működtetjük. Tegyük fel, hogy az n megfigyelést $t_1, \ldots, t_n$ hőmérsékleten végeztük és $x_1, \ldots, x_n$ élettartamokat figyeltünk meg. Adjunk ML becslést θ-ra!

8. Augusztusban 5 éjszakán át figyeltük meg a hullócsillagok számát. A következő mintát kaptuk: 4, 3, 7, 2, 4. A hullócsillagok száma egy este Poisson-eloszlású. Adjunk ML becslést az eloszlás paraméterére!

9. Legyenek $X_1, \ldots, X_n$ független, azonos eloszlású véletlen változók, melyek közös eloszlásfüggvénye F_θ, várható értéke $\mu(\theta)$, a szórásnégyzete σ^2. A szórásnégyzet ismert, a várható értéket becsüljük a mintaátlaggal. Legalább mekkora legyen n, hogy

$$\mathbf{P}_\theta \left(|\overline{X}_n - \mu| > 0,01 \right) \leq 0,01.$$

Használjuk a Csebisev-egyenlőtlenséget! És ha normális közelítést használunk?

10. Legyenek $X_1, \ldots, X_n$ független, Egyenletes$(0, \theta)$ eloszlású véletlen változók, ahol $\theta > 0$. Határozzuk meg $X_{n,n} = \max\{X_1, \ldots, X_n\}$ eloszlás-, és sűrűségfüggvényét, várható értékét és szórását!

Mutassuk meg, hogy

$$T_1(\mathbf{X}) = \frac{n+1}{2n} X_{n,n}$$

torzítatlan becslése $(\theta/2)$-nek, ami hatásosabb, mint a $\overline{X}_n$ mintaátlag! Igazoljuk, hogy T_1 gyengén konzisztens!

11. Még mindig egyenletes eloszlás. Legyenek $X_1, \ldots, X_{2n+1}$ független, Egyenletes$(0, \theta)$ eloszlású véletlen változók, ahol $\theta > 0$. Jelölje Y a nagyság szerint középső mintaelemet. Határozzuk meg Y eloszlását, várható értékét és szórását! Igazoljuk, hogy Y konzisztens becslése a $\theta/2$-nek!

12. Egy gyárban a termékek minőségét úgy ellenőrzik, hogy minden nap n terméket vizsgálnak meg. Az adott napi összes gyártmányt akkor fogadják el, ha minden megvizsgált gyártmány jó. Azt tapasztalták, hogy m nap alatt összesen x-szer fogadták el a napi gyártmányokat. Adjunk maximum likelihood becslést annak a valószínűségére, hogy egy termék selejtes!

13. Mendel törvényei szerint egy növény AA, AB, BB genotípusa rendre θ^2, $2\theta(1-\theta)$, $(1-\theta)^2$ arányban fordul elő. Egy területen a három gyakoriságra x_{AA}, x_{AB}, és x_{BB} adódott. Adjunk maximum likelihood becslést θ-ra! Mutassuk meg, hogy a kapott becslés torzítatlan!

14. A színtévesztés leggyakoribb fajtája az X kromoszómához kapcsoltan, nemhez kötötten öröklődik. Emiatt a férfiaknál a gyakoriság p, míg a nőknél csak p^2. Adjunk maximum likelihood becslést p-re, az alapján, hogy M férfiból m, N nőből pedig n volt színtévesztő!

15. Egy céllövő ismeretlen p valószínűséggel talál el egy célpontot. Adjunk ML-becslést p-re, ha az első sikeres lövés k-adikra következett be. Torzítatlan-e a kapott becslés? A második sikeres lövésre további ℓ lövésig kellett várni. Ezt figyelembe véve adjunk ML-becslést p-re!

16. Legyenek $X_1, \ldots, X_n$ független véletlen változók, $f(x) = \frac{2x}{3\theta^2}$, $\theta \leq x \leq 2\theta$, sűrűségfüggvénnyel. Adjunk becslést θ-ra momentum módszerrel és ML módszerrel is! Torzítatlan-e, ill. aszimptotikusan torzítatlan-e a kapott becslés?