Statisztika I - paraméterbecslés

Sztochasztika

Horváth Illés

2022/11/29

1. Definíciók, példák
2. Statisztikák általános tulajdonságai
3. Paraméterbecslés I - momentumbecslés
4. Likelihood függvény
5. Paraméterbecslés II - maximum likelihood becslés

Motiváció

"Egy alacsony forgalmú úton percenként átlagosan $1,8$ autó halad el."

Motiváció

"Egy alacsony forgalmú úton percenként átlagosan $1,8$ autó halad el."

Hogyan kapták a fenti információt?

"Egy alacsony forgalmú úton percenként átlagosan $1,8$ autó halad el."

Hogyan kapták a fenti információt?

Általában megfigyeléseken alapuló becslés révén.

Motiváció

"Egy alacsony forgalmú úton percenként átlagosan $1,8$ autó halad el."

Hogyan kapták a fenti információt?

Általában megfigyeléseken alapuló becslés révén.

Példa. 5 diszjunkt 1 perces intervallumban megszámoljuk az autókat, és az 1, 4, 0, 3, 1 mintát kapjuk. Próbáljunk ez alapján egy naiv becslést adni az autók érkezési rátájára.

Minta, háttéreloszlás

Az általános keret a következő. Egy *minta* az $X_1, X_2, \dots, X_n$ fae valószínűségi változók egy ismeretlen háttéreloszlásból. n a minta mérete.

Szokás jelölni a mintát $X_1, X_2, \dots, X_n$-nel, ha azt akarjuk hangsúlyozni, hogy véletlenszerű, és szokás jelölni $x_1, x_2, \dots, x_n$-nel is, ha a konkrét realizációra vagyunk kíváncsiak.

Minta, háttéreloszlás

Az általános keret a következő. Egy *minta* az $X_1, X_2, \dots, X_n$ fae valószínűségi változók egy ismeretlen háttéreloszlásból. n a minta mérete.

Szokás jelölni a mintát $X_1, X_2, \dots, X_n$-nel, ha azt akarjuk hangsúlyozni, hogy véletlenszerű, és szokás jelölni $x_1, x_2, \dots, x_n$-nel is, ha a konkrét realizációra vagyunk kíváncsiak.

Általában a háttéreloszlás valamilyen paraméteres eloszlás-családból származik, amit $\mathbb{P}_\theta(.)$-val jelölünk (vagy $f_\theta(.)$ sűrűségfüggvénnyel a folytonos esetben).
θ a paraméter, ami adott tartományból vehet fel értékeket (a legtipikusabb tartományok $\mathbb{R}, \mathbb{R}^+$ és $\mathbb{Z}^+$). Több paraméteres eloszlás-család is lehetséges.

Minta, háttéreloszlás

Az általános keret a következő. Egy *minta* az $X_1, X_2, \ldots, X_n$ fae valószínűségi változók egy ismeretlen háttéreloszlásból. n a minta mérete.

Szokás jelölni a mintát $X_1, X_2, \ldots, X_n$-nel, ha azt akarjuk hangsúlyozni, hogy véletlenszerű, és szokás jelölni $x_1, x_2, \ldots, x_n$-nel is, ha a konkrét realizációra vagyunk kíváncsiak.

Általában a háttéreloszlás valamilyen paraméteres eloszlás-családból származik, amit $\mathbb{P}_\theta(.)$-val jelölünk (vagy $f_\theta(.)$ sűrűségfüggvénnyel a folytonos esetben).
θ a paraméter, ami adott tartományból vehet fel értékeket (a legtipikusabb tartományok $\mathbb{R}, \mathbb{R}^+$ és $\mathbb{Z}^+$). Több paraméteres eloszlás-család is lehetséges.

Az előző, autós példában a háttéreloszlás POI(λ), ahol $\lambda > 0$ ismeretlen.

Statisztikák

Egy *statisztika* a minta egy $T = T(x_1, \ldots, x_n)$ függvénye.

Statisztikák

Egy *statisztika* a minta egy $T = T(x_1, \ldots, x_n)$ függvénye.

Nevezetes statisztikák:

- mintaátlag:
$$\bar{x} = \frac{x_1 + \cdots + x_n}{n}.$$

Statisztikák

Egy *statisztika* a minta egy $T = T(x_1, \dots, x_n)$ függvénye.

Nevezetes statisztikák:

- mintaátlag:
$$\bar{x} = \frac{x_1 + \dots + x_n}{n}.$$
- medián: a mintát növekvő sorba rendezve a középső elem (páros sok elem esetén a két középső elem átlaga). Az elemek fele nagyobb vagy egyenlő, fele kisebb vagy egyenlő a mediánnál.

Statisztikák

Egy *statisztika* a minta egy $T = T(x_1, \ldots, x_n)$ függvénye.

Nevezetes statisztikák:

- mintaátlag:
$$\bar{x} = \frac{x_1 + \cdots + x_n}{n}.$$
- medián: a mintát növekvő sorba rendezve a középső elem (páros sok elem esetén a két középső elem átlaga). Az elemek fele nagyobb vagy egyenlő, fele kisebb vagy egyenlő a mediánnál.

Az átlag és a medián is a minta tipikus viselkedését akarják jellemezni egyetlen számmal (de az értékük általában eltérő).

Statisztikák

- minta minimuma és maximuma:

$$x_{\min} = \min(x_1, \dots, x_n), \qquad x_{\max} = \max(x_1, \dots, x_n).$$

- minta terjedelme:

$$x_{\max} - x_{\min}.$$

Statisztikák

- minta minimuma és maximuma:

$$x_{\min} = \min(x_1, \ldots, x_n), \qquad x_{\max} = \max(x_1, \ldots, x_n).$$

- minta terjedelme:

$$x_{\max} - x_{\min}.$$

- minta varianciája (vagy empirikus variancia):

$$s_n^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2.$$

- minta szórása (vagy empirikus szórás):

$$s_n = \sqrt{s_n^2}.$$

Statisztikák

- minta minimuma és maximuma:

$$x_{\min} = \min(x_1, \ldots, x_n), \qquad x_{\max} = \max(x_1, \ldots, x_n).$$

- minta terjedelme:

$$x_{\max} - x_{\min}.$$

- minta varianciája (vagy empirikus variancia):

$$s_n^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2.$$

- minta szórása (vagy empirikus szórás):

$$s_n = \sqrt{s_n^2}.$$

A minta terjedelme és az empirikus szórás is azt jellemzik, hogy a minta mennyire szóródik.

Becslésnek tulajdonságai

Pontbecslésnek vagy becslésnek egy olyan T statisztikát nevezünk, amit arra használunk, hogy az ismeretlen θ paramétert becsüljük meg (esetleg annak egy $f(\theta)$ függvényét).

Becslésnek tulajdonságai

Pontbecslésnek vagy becslésnek egy olyan T statisztikát nevezünk, amit arra használunk, hogy az ismeretlen θ paramétert becsüljük meg (esetleg annak egy $f(\theta)$ függvényét).

Becslések tulajdonságai:

- egy T statisztika *torzítatlan* becslés θ-ra (vagy $f(\theta)$-ra), ha

$$\mathbb{E}_\theta(T(X_1, \ldots, X_n)) = \theta \quad (\text{vagy } f(\theta)).$$

Példa. Ha a háttéreloszlás POI(λ), akkor az $\bar{X}$ mintaátlag torzítatlan becslés λ-ra:

$$\mathbb{E}_\lambda(\bar{X}) = \mathbb{E}_\lambda\left(\frac{X_1 + \cdots + X_n}{n}\right) = \frac{\lambda + \cdots + \lambda}{n} = \lambda.$$

Statisztikák tulajdonságai

Lemma

Tetszőleges $\mathbb{P}_\theta(.)$ *háttéreloszlásra* $\bar{X}$ *torzítatlan becslés az*

$$f(\theta) = \mathbb{E}_\theta(X_1)$$

függvényre, és

$$s_n^{*2} = \frac{n}{n-1} s_n^2$$

torzítatlan becslés a

$$g(\theta) = \mathbb{D}_\theta^2(X_1)$$

függvényre.

Nem biz.

Lemma

Tetszőleges $\mathbb{P}_\theta(.)$ *háttéreloszlásra* $\bar{X}$ *torzítatlan becslés az*

$$f(\theta) = \mathbb{E}_\theta(X_1)$$

függvényre, és

$$s_n^{*2} = \frac{n}{n-1} s_n^2$$

torzítatlan becslés a

$$g(\theta) = \mathbb{D}_\theta^2(X_1)$$

függvényre.

Nem biz.

s_n^{*2} a *korrigált empirikus variancia*. Bessel-korrekció néven is ismert.

Statisztikák tulajdonságai

Sokféle torzítatlan becslés lehetséges; például X_1 is torzítatlan becslés $f(\theta) = \mathbb{E}_\theta(X_1)$-re. Persze intuitívan $\bar{X}$ jobb becslés kell, hogy legyen; ezt formalizáljuk következőnek.

Statisztikák tulajdonságai

Sokféle torzítatlan becslés lehetséges; például X_1 is torzítatlan becslés $f(\theta) = \mathbb{E}_\theta(X_1)$-re. Persze intuitívan $\bar{X}$ jobb becslés kell, hogy legyen; ezt formalizáljuk következőnek.

- Ha T_1 és T_2 mindkettő torzítatlan becslés θ-ra, akkor T_1-et hatásosabbnak nevezzük T_2-nél, ha θ minden értékére

$$\mathbb{D}_\theta^2(T_1) \leq \mathbb{D}_\theta^2(T_2).$$

- Egy torzítatlan becslés hatásos, ha nincs nála hatásosabb torzítatlan becslés.

Statisztikák tulajdonságai

Sokféle torzítatlan becslés lehetséges; például X_1 is torzítatlan becslés $f(\theta) = \mathbb{E}_\theta(X_1)$-re. Persze intuitívan $\bar{X}$ jobb becslés kell, hogy legyen; ezt formalizáljuk következőnek.

- Ha T_1 és T_2 mindkettő torzítatlan becslés θ-ra, akkor T_1-et hatásosabbnak nevezzük T_2-nél, ha θ minden értékére

$$\mathbb{D}_\theta^2(T_1) \leq \mathbb{D}_\theta^2(T_2).$$

- Egy torzítatlan becslés hatásos, ha nincs nála hatásosabb torzítatlan becslés.

Példa. Ha a háttéreloszlás POI(λ), akkor

$$\mathbb{D}_\lambda(\bar{X}) = \frac{\sqrt{\lambda}}{\sqrt{n}}, \qquad \mathbb{D}_\lambda(X_1) = \sqrt{\lambda},$$

tehát $\bar{X}$ hatásosabb becslés λ-ra, mint X_1.

Statisztikák tulajdonságai

- Legyen T_n egy statisztika n mintaméretre. A $(T_n)_{n=1,2,\ldots}$ statisztika-sorozat konzisztens becslés θ-ra, ha

$$\lim_{n\to\infty} \mathbb{P}(|T_n - \theta| > \varepsilon) = 0 \qquad \forall \varepsilon > 0,$$

illetve konzisztens becslés $f(\theta)$-ra, ha

$$\lim_{n\to\infty} \mathbb{P}(|T_n - f(\theta)| > \varepsilon) = 0 \qquad \forall \varepsilon > 0.$$

Statisztikák tulajdonságai

- Legyen T_n egy statisztika n mintaméretre. A $(T_n)_{n=1,2,\dots}$ statisztika-sorozat konzisztens becslés θ-ra, ha

$$\lim_{n\to\infty} \mathbb{P}(|T_n - \theta| > \varepsilon) = 0 \qquad \forall \varepsilon > 0,$$

illetve konzisztens becslés $f(\theta)$-ra, ha

$$\lim_{n\to\infty} \mathbb{P}(|T_n - f(\theta)| > \varepsilon) = 0 \qquad \forall \varepsilon > 0.$$

A NSzT garantálja, hogy $\bar{X}$ konzisztens becslés $f(\theta) = \mathbb{E}_\theta(X_1)$-ra.

- Legyen T_n egy statisztika n mintaméretre. A $(T_n)_{n=1,2,\ldots}$ statisztika-sorozat konzisztens becslés θ-ra, ha

$$\lim_{n\to\infty} \mathbb{P}(|T_n - \theta| > \varepsilon) = 0 \qquad \forall\varepsilon > 0,$$

illetve konzisztens becslés $f(\theta)$-ra, ha

$$\lim_{n\to\infty} \mathbb{P}(|T_n - f(\theta)| > \varepsilon) = 0 \qquad \forall\varepsilon > 0.$$

A NSzT garantálja, hogy $\bar{X}$ konzisztens becslés $f(\theta) = \mathbb{E}_\theta(X_1)$-ra.

Például a POI(λ) háttéreloszlás esetén $\bar{X}$ konzisztens becslés λ-ra.

Paraméter becslés I - momentum-becslés

Egy általános módszer háttéreloszlás paraméterének becslésére a momentum-becslés.
Az alapötlet az, hogy θ-t úgy becsüljük, hogy

$$\mathbb{E}_\theta(X_1)) = \bar{x}$$

teljesüljön.

Egy általános módszer háttéreloszlás paraméterének becslésére a momentum-becslés.
Az alapötlet az, hogy θ-t úgy becsüljük, hogy

$$\mathbb{E}_\theta(X_1)) = \bar{x}$$

teljesüljön. Formálisan: ha a

$$g(\theta) = \mathbb{E}_\theta(X_1)$$

függvény invertálható, akkor θ momentum-becslése

$$\hat{\theta} = g^{-1}(\bar{X}).$$

Egy általános módszer háttéreloszlás paraméterének becslésére a momentum-becslés.
Az alapötlet az, hogy θ-t úgy becsüljük, hogy

$$\mathbb{E}_\theta(X_1)) = \bar{x}$$

teljesüljön. Formálisan: ha a

$$g(\theta) = \mathbb{E}_\theta(X_1)$$

függvény invertálható, akkor θ momentum-becslése

$$\hat{\theta} = g^{-1}(\bar{X}).$$

Ha a g függvény nem invertálható, akkor nem tudunk momentum-becslést adni.

Példa. A korábbi autós példában a minta 5 diszjunkt 1 perces intervallumra 1, 4, 0, 3, 1.

Példa. A korábbi autós példában a minta 5 diszjunkt 1 perces intervallumra 1, 4, 0, 3, 1.

A háttéreloszlás $X_1 \sim \mathrm{POI}(\lambda)$, és

$$g(\lambda) = \mathbb{E}_\lambda(X_1) = \lambda,$$

így g az identitásfüggvény, és a momentum-becslés

$$\hat{\lambda} = g^{-1}(\bar{X}) = \bar{X};$$

Paraméter becslés I - momentum-becslés

Példa. A korábbi autós példában a minta 5 diszjunkt 1 perces intervallumra 1, 4, 0, 3, 1.

A háttéreloszlás $X_1 \sim \mathrm{POI}(\lambda)$, és

$$g(\lambda) = \mathbb{E}_\lambda(X_1) = \lambda,$$

így g az identitásfüggvény, és a momentum-becslés

$$\hat{\lambda} = g^{-1}(\bar{X}) = \bar{X};$$

a konkrét realizációra a momentum-becslés

$$\hat{\lambda} = \bar{x} = \frac{1+4+0+3+1}{5} = 1.8.$$

Ha a háttéreloszlásnak 2 paramétere van: $\mathbb{P}_{\theta_1,\theta_2}(.)$, akkor vegyük a következő $\mathbb{R}^2 \to \mathbb{R}^2$ függvényt:

$$g(\theta_1, \theta_2) = (\mathbb{E}_{\theta_1,\theta_2}(X_1), \mathbb{D}^2_{\theta_1,\theta_2}(X_1)).$$

Momentum-becslés 2 paraméter esetén

Ha a háttéreloszlásnak 2 paramétere van: $\mathbb{P}_{\theta_1,\theta_2}(.)$, akkor vegyük a következő $\mathbb{R}^2 \to \mathbb{R}^2$ függvényt:

$$g(\theta_1, \theta_2) = (\mathbb{E}_{\theta_1,\theta_2}(X_1), \mathbb{D}^2_{\theta_1,\theta_2}(X_1)).$$

Ha g invertálható mint $\mathbb{R}^2 \to \mathbb{R}^2$ függvény, akkor a momentum-becslés

$$(\hat{\theta}_1, \hat{\theta}_2) = g^{-1}(\bar{x}, s_n^2)$$

ahol $\bar{x}$ a mintaátlag és s_n^2 az empirikus variancia.

Momentum-becslés 2 paraméter esetén

Ha a háttéreloszlásnak 2 paramétere van: $\mathbb{P}_{\theta_1,\theta_2}(.)$, akkor vegyük a következő $\mathbb{R}^2 \to \mathbb{R}^2$ függvényt:

$$g(\theta_1, \theta_2) = (\mathbb{E}_{\theta_1,\theta_2}(X_1), \mathbb{D}^2_{\theta_1,\theta_2}(X_1)).$$

Ha g invertálható mint $\mathbb{R}^2 \to \mathbb{R}^2$ függvény, akkor a momentum-becslés

$$(\hat{\theta}_1, \hat{\theta}_2) = g^{-1}(\bar{x}, s_n^2)$$

ahol $\bar{x}$ a mintaátlag és s_n^2 az empirikus variancia.

Még több paraméter esetén annyi magasabb momentumot $(\mathbb{E}(X^3), \mathbb{E}(X^4)$ stb.) tekintünk, ahány paraméter van.

Likelihood függvény

Egy $x_1, \dots, x_n$ minta *likelihood függvénye*

$$L_x(\theta) = \prod_{i=1}^{n} \mathbb{P}_\theta(X_i = x_i)$$

ha a háttéreloszlás diszkrét, és

$$L_x(\theta) = \prod_{i=1}^{n} f_\theta(X_i = x_i)$$

ha a háttéreloszlás folytonos $f_\theta(.)$ sűrűségfüggvénnyel.

Likelihood függvény

Egy $x_1, \ldots, x_n$ minta *likelihood függvénye*

$$L_x(\theta) = \prod_{i=1}^{n} \mathbb{P}_\theta(X_i = x_i)$$

ha a háttéreloszlás diszkrét, és

$$L_x(\theta) = \prod_{i=1}^{n} f_\theta(X_i = x_i)$$

ha a háttéreloszlás folytonos $f_\theta(.)$ sűrűségfüggvénnyel.

A likelihood függvény megegyezik a minta valószínűségével (vagy folytonos esetben sűrűségével), viszont a θ paraméter függvényében tekintjük.

A maximum likelihood (ML) becslés a következő: adott mintára legyen

$$\hat{\theta} = \arg\max_\theta \{L_x(\theta)\},$$

azaz azt a θ paramétert választjuk, amelyre $L_x(\theta)$ értéke maximális az adott mintára.

Maximum likelihood (ML) becslés

A maximum likelihood (ML) becslés a következő: adott mintára legyen

$$\hat{\theta} = \arg\max_{\theta}\{L_x(\theta)\},$$

azaz azt a θ paramétert választjuk, amelyre $L_x(\theta)$ értéke maximális az adott mintára.

A ML becslés eltér a Bayes-tételtől! Most θ értéke nem véletlenszerű (csak ismeretlen).

Maximum likelihood (ML) becslés

A ML becsléshez az $L_x(\theta)$ függvény maximumát kell meghatároznunk. Ha θ folytonos tartományból vehet fel értékeket, akkor ez megkapható a

$$\frac{\mathrm{d}}{\mathrm{d}\theta} L_x(\theta) = 0,$$

egyenlet megoldásával, majd a megoldások közül a maximumhely kiválasztásával.

Maximum likelihood (ML) becslés

A ML becsléshez az $L_x(\theta)$ függvény maximumát kell meghatároznunk. Ha θ folytonos tartományból vehet fel értékeket, akkor ez megkapható a

$$\frac{\mathrm{d}}{\mathrm{d}\theta} L_x(\theta) = 0,$$

egyenlet megoldásával, majd a megoldások közül a maximumhely kiválasztásával.

Van egy gyakorlati trükk, ami a fenti számítást általában egyszerűsíti. Egy $x_1, \ldots, x_n$ minta *log-likelihood függvénye*

$$\ell_x(\theta) = \log(L_x(\theta)) = \begin{cases} \sum_{i=1}^n \log(\mathbb{P}_\theta(X_i = x_i)) \\ \sum_{i=1}^n \log(f_\theta(X_i = x_i)) \end{cases}$$

Maximum likelihood (ML) becslés

Mivel a logaritmus függvény növő, ezért $\ell_x(\theta)$-nak ugyanott van a maximumhelye, mint $L_x(\theta)$-nak, és

$$\frac{\mathrm{d}}{\mathrm{d}\theta}\ell_x(\theta) = 0,$$

általában könnyebben megoldható.

Maximum likelihood (ML) becslés

Mivel a logaritmus függvény növő, ezért $\ell_x(\theta)$-nak ugyanott van a maximumhelye, mint $L_x(\theta)$-nak, és

$$\frac{\mathrm{d}}{\mathrm{d}\theta}\ell_x(\theta) = 0,$$

általában könnyebben megoldható.

Van egy olyan előnye is $\ell_x(\theta)$ használatának, hogy azon minimumhelyek, ahol $L_x(\theta) = 0$, $\frac{\mathrm{d}}{\mathrm{d}\theta}L_x(\theta) = 0$-nak megoldásai, de $\frac{\mathrm{d}}{\mathrm{d}\theta}\ell_x(\theta) = 0$-nak nem.

Maximum likelihood (ML) becslés

Mivel a logaritmus függvény növő, ezért $\ell_x(\theta)$-nak ugyanott van a maximumhelye, mint $L_x(\theta)$-nak, és

$$\frac{\mathrm{d}}{\mathrm{d}\theta}\ell_x(\theta) = 0,$$

általában könnyebben megoldható.

Van egy olyan előnye is $\ell_x(\theta)$ használatának, hogy azon minimumhelyek, ahol $L_x(\theta) = 0$, $\frac{\mathrm{d}}{\mathrm{d}\theta}L_x(\theta) = 0$-nak megoldásai, de $\frac{\mathrm{d}}{\mathrm{d}\theta}\ell_x(\theta) = 0$-nak nem.

Ökölszabály: ha a

$$\frac{\mathrm{d}}{\mathrm{d}\theta}\ell_x(\theta) = 0$$

egyenletnek egy megoldása van, és az nem a paramétertartomány szélén van, akkor az a megoldás a maximumhely és így a ML becslés θ-ra.

Maximum likelihood (ML) becslés

Ha θ csak egész értékeket vehet fel, akkor

$$\frac{\mathrm{d}}{\mathrm{d}\theta}\ell_x(\theta) = 0$$

helyett az

$$\frac{L_x(\theta+1)}{L_x(\theta)} = 1$$

egyenletet oldjuk meg.

Maximum likelihood (ML) becslés

Ha θ csak egész értékeket vehet fel, akkor

$$\frac{\mathrm{d}}{\mathrm{d}\theta}\ell_x(\theta) = 0$$

helyett az

$$\frac{L_x(\theta+1)}{L_x(\theta)} = 1$$

egyenletet oldjuk meg.

Az ötlet az, hogy egy bizonyos θ értékig $\frac{L_x(\theta+1)}{L_x(\theta)} > 1$ teljesül, és ahol az egyenlőtlenség megfordul $\frac{L_x(\theta+1)}{L_x(\theta)} < 1$-re, ott van a maximumhely.

Maximum likelihood (ML) becslés

A ML becslés akkor is működik, ha a háttéreloszlásnak 2 (vagy több) paramétere van, csak akkor többváltozós függvény maximumát keressük; folytonos paraméter esetén a szokásos eljárás:

$$\frac{\partial l_{\theta_1}}{\partial \theta_1}(\theta_1, \theta_2) = 0,$$
$$\frac{\partial l_{\theta_2}}{\partial \theta_2}(\theta_1, \theta_2) = 0$$

egyenletrendszert megoldani, majd a másodrendű parciális deriváltakból összeállított 2×2-es Hesse-mátrixról ellenőrizni, hogy negatív definit-e. (Illetve a korábbi ökölszabály itt is él: ha csak egy megoldás van, és az nem a paramétertartomány szélén van, akkor az a maximumhely.)

1. feladat

Egy pénzérme nem szabályos, p valószínűséggel a fej lesz felül $(0 < p < 1)$, de p értéke ismeretlen.

(a) A FIFFFIFFIF sorozatot kapjuk. Ez alapján adjunk maximum-likelihood-becslést p-re.

(b) Feldobjuk az érmét 10-szer, és azt tapasztaljuk, hogy a fejek száma 7 lett. Ez alapján adjunk maximum-likelihood-becslést p-re. Adjunk momentum-becslést is p-re.

(c) Feljegyezzük, hogy két szomszédos fej között mennyi írást dobtunk; a következőket kaptuk: 0, 1, 0, 0, 1, 0, 1. Ez alapján is adjunk maximum-likelihood-becslést és momentum-becslést p-re.

(d) Feljegyezzük, hogy két szomszédos írás között mennyi fejet dobtunk; a következőket kaptuk: 1, 3, 2, 1. Ez alapján is adjunk maximum-likelihood-becslést és momentum-becslést p-re. Magyarázzuk is meg a kapott eredményt.

Megoldás.

(a) A FIFFFIFFIF minta likelihood-függvénye

$$L_x(p) = p\cdot(1-p)\cdot p\cdot p\cdot p\cdot(1-p)\cdot p\cdot p\cdot(1-p)\cdot p = p^7(1-p)^3,$$

és a log-likelihood függvény

$$\ell_x(p) = \log(L_x(p)) = 7\log p + 3\log(1-p).$$

A ML becsléshez meg kell oldanunk az

$$\frac{\mathrm{d}}{\mathrm{d}p}\ell_x(p) = 0$$

egyenletet.

(a)

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}p}(7\log p + 3\log(1-p)) &= 0 \\
\frac{7}{p} - \frac{3}{1-p} &= 0 \\
7(1-p) - 3p &= 0 \\
7 - 10p &= 0 \\
p &= \frac{7}{10},
\end{aligned}$$

tehát a ML becslés

$$\hat{p} = \frac{7}{10}.$$

1. feladat

(a) Az érdekesség kedvéért vessük ezt össze azzal, mint ha az

$$\frac{\mathrm{d}}{\mathrm{d}p} L_x(p) = 0$$

1. feladat

(a) Az érdekesség kedvéért vessük ezt össze azzal, mint ha az

$$\frac{\mathrm{d}}{\mathrm{d}p} L_x(p) = 0$$

egyenletet oldjuk meg.

(a) Az érdekesség kedvéért vessük ezt össze azzal, mint ha az

$$\frac{\mathrm{d}}{\mathrm{d}p} L_x(p) = 0$$

egyenletet oldjuk meg. Innen

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}p}\left(p^7(1-p)^3\right) &= 0 \\ 7p^6(1-p)^3 + p^7 \cdot 3(1-p)^2 \cdot (-1) &= 0 \\ p^6(1-p)^2\,(7(1-p) - 3p) &= 0. \end{aligned}$$

(a) Az érdekesség kedvéért vessük ezt össze azzal, mint ha az

$$\frac{\mathrm{d}}{\mathrm{d}p} L_x(p) = 0$$

egyenletet oldjuk meg. Innen

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}p}\left(p^7(1-p)^3\right) &= 0 \\
7p^6(1-p)^3 + p^7 \cdot 3(1-p)^2 \cdot (-1) &= 0 \\
p^6(1-p)^2\left(7(1-p) - 3p\right) &= 0.
\end{aligned}$$

Ennek 3 megoldása van: $p = 0$, $p = 1$ és $p = \frac{7}{10}$, és el kell döntenünk, hogy melyik a maximumhely. $p = 0$-ra és $p = 1$-re $L_x(p) = 0$, tehát ezek minimumhelyek, és

$$\hat{p} = \frac{7}{10}.$$

(b) A minta ezúttal $n = 1$ elemű; $X_1 \sim \mathrm{BIN}(10, p)$ és $x_1 = 7$.

(b) A minta ezúttal $n = 1$ elemű; $X_1 \sim \mathrm{BIN}(10, p)$ és $x_1 = 7$.

A likelihood függvény ezúttal

$$L_x(p) = \mathbb{P}_p(X_1 = 7) = \binom{10}{7} p^7 (1-p)^3.$$

(b) A minta ezúttal $n = 1$ elemű; $X_1 \sim \mathrm{BIN}(10, p)$ és $x_1 = 7$.

A likelihood függvény ezúttal

$$L_x(p) = \mathbb{P}_p(X_1 = 7) = \binom{10}{7} p^7 (1-p)^3.$$

Vegyük észre, hogy ez csak egy konstans szorzóban különbözik az (a) részbeli likelihood-függvénytől, tehát a maximumhelyük szükségszerűen ugyanaz. Ez a további számításokból is kiderül.

(b) A log-likelihood függvény

$$\ell_x(p) = \log(L_x(p)) = \log \binom{10}{7} + 7 \log p + 3 \log(1-p),$$

és meg kell oldanunk a

$$\frac{\mathrm{d}}{\mathrm{d}p}\ell_x(p) = 0$$

egyenletet,

(b) A log-likelihood függvény

$$\ell_x(p) = \log(L_x(p)) = \log \binom{10}{7} + 7 \log p + 3 \log(1-p),$$

és meg kell oldanunk a

$$\frac{\mathrm{d}}{\mathrm{d}p} \ell_x(p) = 0$$

egyenletet, ahonnan

$$0 + \frac{7}{1-p} - \frac{3}{p} = 0.$$

Ezen a ponton az egyenlet már pontosan ugyanaz, mint az (a) részben, és a ML becslés is ugyanaz:

$$\hat{p} = \frac{7}{10}.$$

(b) Továbbra is $n = 1$, $X_1 \sim \mathrm{BIN}(10, p)$ és $x_1 = 7$. A momentum-becsléshez

$$g(p) = \mathbb{E}_p(X_1) = 10p$$

(b) Továbbra is $n = 1$, $X_1 \sim \mathrm{BIN}(10, p)$ és $x_1 = 7$. A momentum-becsléshez

$$g(p) = \mathbb{E}_p(X_1) = 10p$$

és így

$$g^{-1}(x) = x/10.$$

(b) Továbbra is $n = 1$, $X_1 \sim \mathrm{BIN}(10, p)$ és $x_1 = 7$. A momentum-becsléshez

$$g(p) = \mathbb{E}_p(X_1) = 10p$$

és így

$$g^{-1}(x) = x/10.$$

A momentum-becslés

$$\hat{p} = g^{-1}(\bar{x}) = \bar{x}/10 = 7/10.$$

(c) Két szomszédos fej között az írások száma $Y_i \sim \mathrm{PGEO}(p)$. A minta $n = 7$ elemű:

$$y_1 = 0,\ y_2 = 1,\ y_3 = 0,\ y_4 = 0,\ y_5 = 1,\ y_6 = 0,\ y_7 = 1.$$

(c) Két szomszédos fej között az írások száma $Y_i \sim \mathrm{PGEO}(p)$. A minta $n = 7$ elemű:

$$y_1 = 0,\ y_2 = 1,\ y_3 = 0,\ y_4 = 0,\ y_5 = 1,\ y_6 = 0,\ y_7 = 1.$$

A minta likelihood-függvénye

$$L_y(p) = \prod_{i=1}^{n} \mathbb{P}_p(Y_i = y_i) = \prod_{i=1}^{n} p(1-p)^{y_i} =$$

$$p \cdot p(1-p)^1 \cdot p \cdot p \cdot p(1-p)^1 \cdot p \cdot p(1-p)^1 \cdot p = p^7(1-p)^3.$$

(c) Két szomszédos fej között az írások száma $Y_i \sim \mathrm{PGEO}(p)$. A minta $n = 7$ elemű:

$$y_1 = 0,\ y_2 = 1,\ y_3 = 0,\ y_4 = 0,\ y_5 = 1,\ y_6 = 0,\ y_7 = 1.$$

A minta likelihood-függvénye

$$L_y(p) = \prod_{i=1}^{n} \mathbb{P}_p(Y_i = y_i) = \prod_{i=1}^{n} p(1-p)^{y_i} =$$

$$p \cdot p(1-p)^1 \cdot p \cdot p \cdot p(1-p)^1 \cdot p \cdot p(1-p)^1 \cdot p = p^7(1-p)^3.$$

Ez ugyanaz, mint az (a) részben, úgyhogy a ML-becslés ezúttal is

$$\hat{p} = \frac{7}{10}.$$

(c) A momentum-becsléshez

$$g(p) = \mathbb{E}_p(Y_1) = \frac{1}{p} - 1,$$

(c) A momentum-becsléshez

$$g(p) = \mathbb{E}_p(Y_1) = \frac{1}{p} - 1,$$

és

$$g^{-1}(y) = \frac{1}{y+1}.$$

Továbbá

$$\bar{y} = \frac{3}{7},$$

így a momentum-becslés

$$\hat{p} = g^{-1}(\bar{y}) = \frac{1}{1+\frac{3}{7}} = \frac{1}{\frac{10}{7}} = \frac{7}{10},$$

ugyanúgy, mint korábban.

1. feladat

(d) Két szomszédos írás között a fejek száma $Z_i \sim \mathrm{PGEO}(1-p)$.
A minta $n = 4$ elemű:

$$z_1 = 1,\ z_2 = 2,\ z_3 = 3,\ z_4 = 1.$$

1. feladat

(d) Két szomszédos írás között a fejek száma $Z_i \sim \mathrm{PGEO}(1-p)$. A minta $n = 4$ elemű:

$$z_1 = 1,\ z_2 = 2,\ z_3 = 3,\ z_4 = 1.$$

A minta likelihood-függvénye

$$L_z(p) = \prod_{i=1}^{n} \mathbb{P}_p(Z_i = z_i) = \prod_{i=1}^{n} (1-p)p^{z_i} =$$

$$(1-p)p^1 \cdot (1-p)p^2 \cdot (1-p)p^3 \cdot (1-p)p^1 = (1-p)^4 p^7.$$

1. feladat

(d) Két szomszédos írás között a fejek száma $Z_i \sim \mathrm{PGEO}(1-p)$. A minta $n = 4$ elemű:

$$z_1 = 1,\ z_2 = 2,\ z_3 = 3,\ z_4 = 1.$$

A minta likelihood-függvénye

$$L_z(p) = \prod_{i=1}^{n} \mathbb{P}_p(Z_i = z_i) = \prod_{i=1}^{n} (1-p)p^{z_i} =$$
$$(1-p)p^1 \cdot (1-p)p^2 \cdot (1-p)p^3 \cdot (1-p)p^1 = (1-p)^4 p^7.$$

Ez más, mint az (a) részben, a ML-becslés most

$$\hat{p} = \frac{7}{11}\text{-nek}$$

adódik.

1. feladat

(d) A momentum-becsléshez

$$g(p) = \mathbb{E}_p(Z_1) = \frac{1}{1-p} - 1,$$

(d) A momentum-becsléshez

$$g(p) = \mathbb{E}_p(Z_1) = \frac{1}{1-p} - 1,$$

és

$$g^{-1}(z) = 1 - \frac{1}{z+1}.$$

Továbbá

$$\bar{z} = \frac{1+2+3+1}{4} = \frac{7}{4},$$

így a momentum-becslés

$$\hat{p} = g^{-1}(\bar{z}) = 1 - \frac{1}{\frac{7}{4}+1} = 1 - \frac{1}{\frac{11}{4}} = 1 - \frac{4}{11} = \frac{7}{11}.$$

(d) A momentum-becsléshez

$$g(p) = \mathbb{E}_p(Z_1) = \frac{1}{1-p} - 1,$$

és

$$g^{-1}(z) = 1 - \frac{1}{z+1}.$$

Továbbá

$$\bar{z} = \frac{1+2+3+1}{4} = \frac{7}{4},$$

így a momentum-becslés

$$\hat{p} = g^{-1}(\bar{z}) = 1 - \frac{1}{\frac{7}{4}+1} = 1 - \frac{1}{\frac{11}{4}} = 1 - \frac{4}{11} = \frac{7}{11}.$$

Tehát az ugyanebből a mintából kiszámolt ML-becsléssel megegyezik, de a korábbiakkal nem.

1. feladat

(d) Az eltérés oka az, hogy a FIFFFIFFIF mintában a szomszédos fejek száma ugyan rendre 1, 2, 3, 1, azonban a minta utolsó eleme nem írás, így lehetséges, hogy további fejekkel folytatódna, és akkor az 1, 2, 3, 1 minta utolsó eleme is nagyobb lenne.

(d) Az eltérés oka az, hogy a FIFFFIFFIF mintában a szomszédos fejek száma ugyan rendre 1, 2, 3, 1, azonban a minta utolsó eleme nem írás, így lehetséges, hogy további fejekkel folytatódna, és akkor az 1, 2, 3, 1 minta utolsó eleme is nagyobb lenne.

Az 1, 2, 3, 1 minta valójában az FIFFFIFFIFI eredeti mintának felelne meg, ami 7 írást és 4 fejet tartalmaz, így arra az (a) és (b) rész becslése is $\frac{7}{11}$ lenne.

(d) Az eltérés oka az, hogy a FIFFFIFFIF mintában a szomszédos fejek száma ugyan rendre 1, 2, 3, 1, azonban a minta utolsó eleme nem írás, így lehetséges, hogy további fejekkel folytatódna, és akkor az 1, 2, 3, 1 minta utolsó eleme is nagyobb lenne.

Az 1, 2, 3, 1 minta valójában az FIFFFIFFIFI eredeti mintának felelne meg, ami 7 írást és 4 fejet tartalmaz, így arra az (a) és (b) rész becslése is $\frac{7}{11}$ lenne.

Összességében a (d) típusú felírással egyáltalán nem alkalmazható olyan mintára, aminek az utolsó eleme fej. (Ahogy egyébként a (c) típusú felírás sem alkalmazható olyan mintára, ahol az utolsó elem írás, de ez az eredeti mintánál nem volt gond.)

6. feladat

Az $X_1, \ldots, X_n$ minta az $U[0, a]$ háttéreloszlásból származik.

(a) Adjunk momentum-becslést a értékére a $0.38, 0.78, 2.22, 1.91, 1.71$ minta alapján.

(b) Adjunk maximum-likelihood becslést a értékére a $0.38, 0.78, 2.22, 1.91, 1.71$ minta alapján.

(c) Adjunk momentum becslést a értékére a $0.38, 0.20, 2.22, 0.16, 0.24$ minta alapján. Mit tapasztalunk?

Megoldás.

(b) A momentum-becsléshez szükségünk van $X_1 \sim U([0, a])$ várható értékére:

$$g(a) = \mathbb{E}_a(X_1) = \frac{a}{2},$$

így

$$g^{-1}(x) = 2x,$$

és a momentum-becslés a-ra

$$\hat{a} = 2 \cdot \bar{x} = 2 \cdot \frac{0.38 + 0.78 + 2.22 + 1.91 + 1.71}{5} = 2.8.$$

(b) U$(0, a)$ sűrűségfüggvénye

$$f_a(x) = \begin{cases} \frac{1}{a} & \text{ha } x \in (0, a) \\ 0 & \text{ha } x \notin (0, a) \end{cases}$$

és

$$L_x(a) = \prod_{i=1}^{5} f_a(x_i).$$

(b) U$(0, a)$ sűrűségfüggvénye

$$f_a(x) = \begin{cases} \frac{1}{a} & \text{ha } x \in (0, a) \\ 0 & \text{ha } x \notin (0, a) \end{cases}$$

és

$$L_x(a) = \prod_{i=1}^{5} f_a(x_i).$$

Hogy néz ki az $L_x(a)$ függvény? Számítsuk ki az értékét pl. az $a = 0.5$ pontban.

(b) $U(0, a)$ sűrűségfüggvénye

$$f_a(x) = \begin{cases} \frac{1}{a} & \text{ha } x \in (0, a) \\ 0 & \text{ha } x \notin (0, a) \end{cases}$$

és

$$L_x(a) = \prod_{i=1}^{5} f_a(x_i).$$

Hogy néz ki az $L_x(a)$ függvény? Számítsuk ki az értékét pl. az $a = 0.5$ pontban.

$$f_a(x_1) = f_{0.5}(0.38) = \frac{1}{0.5},$$

mivel $0.38 \in (0, 0.5)$.

6. feladat

(b)

$$f_a(x_2) = f_{0.5}(0.78) = 0,$$

mivel $0.78 \notin (0, 0.5)$.

(b)

$$f_a(x_2) = f_{0.5}(0.78) = 0,$$

mivel $0.78 \notin (0, 0.5)$.

Innen már mindenképpen

$$L_x(0.5) = 0.$$

(b)

$$f_a(x_2) = f_{0.5}(0.78) = 0,$$

mivel $0.78 \notin (0, 0.5)$.

Innen már mindenképpen

$$L_x(0.5) = 0.$$

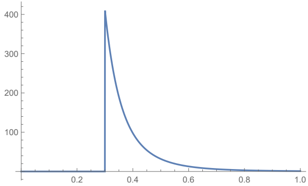
Sőt, $L_x(a) = 0$ teljesül minden $a < 2.22$ értékre. Másrészt viszont $a \geq 2.22$ értékre

$$f_a(x_i) = \frac{1}{a} \qquad i = 1, \ldots, 5,$$

és így

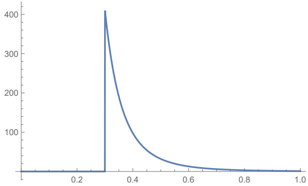
$$L_x(a) = \begin{cases} \frac{1}{a^5} & \text{ha } a \geq 2.22 \\ 0 & \text{ha } x < 2.22 \end{cases}$$

(b) $L_x(a)$ grafikonja így néz ki:



6. feladat

(b) $L_x(a)$ grafikonja így néz ki:

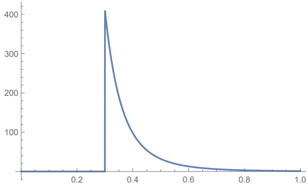


Nem folytonos, úgyhogy a maximumot nem deriválással határozzuk meg. Viszont a grafikon alapján világos, hogy a maximumhely, és így a ML becslés

$$\hat{a} = 2.22.$$

6. feladat

(b) $L_x(a)$ grafikonja így néz ki:



Nem folytonos, úgyhogy a maximumot nem deriválással határozzuk meg. Viszont a grafikon alapján világos, hogy a maximumhely, és így a ML becslés

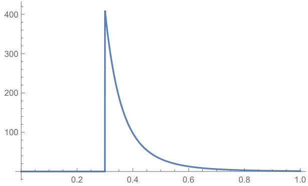
$$\hat{a} = 2.22.$$

Általában az $\mathrm{U}(0, a)$ háttéreloszlásra a ML becslés

$$\hat{a} = \max(X_i).$$

6. feladat

(b) $L_x(a)$ grafikonja így néz ki:



Nem folytonos, úgyhogy a maximumot nem deriválással határozzuk meg. Viszont a grafikon alapján világos, hogy a maximumhely, és így a ML becslés

$$\hat{a} = 2.22.$$

Általában az $\mathrm{U}(0, a)$ háttéreloszlásra a ML becslés

$$\hat{a} = \max(X_i).$$

Egyébként ez a becslés nem torzítatlan,

$$\mathbb{E}_a(\max(X_i)) = \frac{n}{n+1} a.$$

6. feladat

(c) Ugyanazt csináljuk, mint az (a) részben, csak a $0.38, 0.20, 2.22, 0.16, 0.24$ mintára, és

$$\hat{a} = 2 \cdot \bar{x} = 2 \cdot \frac{0.38 + 0.20 + 2.22 + 0.16 + 0.24}{5} = 1.28.$$

(c) Ugyanazt csináljuk, mint az (a) részben, csak a $0.38, 0.20, 2.22, 0.16, 0.24$ mintára, és

$$\hat{a} = 2 \cdot \bar{x} = 2 \cdot \frac{0.38 + 0.20 + 2.22 + 0.16 + 0.24}{5} = 1.28.$$

Ezzel az a gond, hogy $a = 1.28$ nem is lehetséges, mivel a mintaelemek között 2.22 is szerepel.

6. feladat

(c) Ugyanazt csináljuk, mint az (a) részben, csak a $0.38, 0.20, 2.22, 0.16, 0.24$ mintára, és

$$\hat{a} = 2 \cdot \bar{x} = 2 \cdot \frac{0.38 + 0.20 + 2.22 + 0.16 + 0.24}{5} = 1.28.$$

Ezzel az a gond, hogy $a = 1.28$ nem is lehetséges, mivel a mintaelemek között 2.22 is szerepel.

A momentum-becslésnél sajnos ez előfordulhat: nincs garancia arra, hogy a kapott paraméter ténylegesen lehetséges is.

(c) Ugyanazt csináljuk, mint az (a) részben, csak a $0.38, 0.20, 2.22, 0.16, 0.24$ mintára, és

$$\hat{a} = 2 \cdot \bar{x} = 2 \cdot \frac{0.38 + 0.20 + 2.22 + 0.16 + 0.24}{5} = 1.28.$$

Ezzel az a gond, hogy $a = 1.28$ nem is lehetséges, mivel a mintaelemek között 2.22 is szerepel.

A momentum-becslésnél sajnos ez előfordulhat: nincs garancia arra, hogy a kapott paraméter ténylegesen lehetséges is.

A ML-becsléssel ilyen gond nincs, ott a kapott paraméter mindig lehetséges.

7. feladat

Egy vizsgán a hallgatóknak átlagosan 60%-a megy át. Az előző félévben 14 hallgató ment át a vizsgán N-ből, N azonban ismeretlen. Adjunk ML-becslést N-re. Adjunk momentum-becslést N-re.

7. feladat

Egy vizsgán a hallgatóknak átlagosan 60%-a megy át. Az előző félévben 14 hallgató ment át a vizsgán N-ből, N azonban ismeretlen. Adjunk ML-becslést N-re. Adjunk momentum-becslést N-re.

Megoldás. A minta mérete $n = 1$; $X_1 \sim \mathrm{BIN}(N, 0.6)$, és

$$L_x(N) = \binom{N}{14} 0.6^{14} \cdot 0.4^{N-14}.$$

A maximumhely meghatározásához a következő egyenletet oldjuk meg:

$$\frac{L_x(N+1)}{L_x(N)} = 1,$$

mivel N egész értékű paraméter.

$$\frac{L_x(N+1)}{L_x(N)} = \frac{\binom{N+1}{14} 0.6^{14} \cdot 0.4^{N+1-14}}{\binom{N}{14} 0.6^{14} \cdot 0.4^{N-14}} =$$

$$\frac{\frac{(N+1)!}{14!(N+1-14)!} 0.6^{14} \cdot 0.4^{N+1-14}}{\frac{N!}{14!(N-14)!} 0.6^{14} \cdot 0.4^{N-14}} =$$

$$\frac{(N+1) \cdot 0.4}{(N-13)} = 1.$$

A megoldás $N \approx 22.33$.

7. feladat

Ez azt jelenti, hogy $14 \leq N \leq 22$ esetén

$$\frac{L_x(N+1)}{L_x(N)} > 1,$$

tehát $L_x(N)$ növekvő, és $23 \leq N$ esetén

$$\frac{L_x(N+1)}{L_x(N)} < 1,$$

azaz $L_x(N)$ csökkenő,

7. feladat

Ez azt jelenti, hogy $14 \leq N \leq 22$ esetén

$$\frac{L_x(N+1)}{L_x(N)} > 1,$$

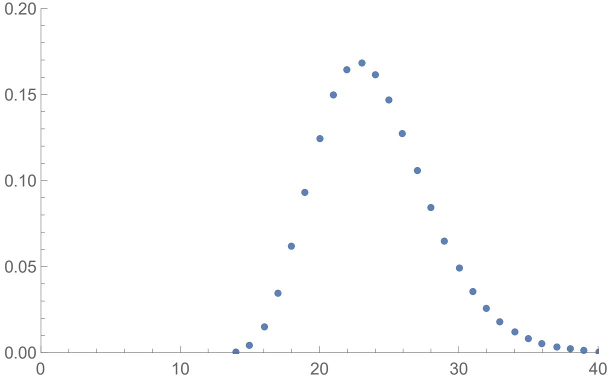
tehát $L_x(N)$ növekvő, és $23 \leq N$ esetén

$$\frac{L_x(N+1)}{L_x(N)} < 1,$$

azaz $L_x(N)$ csökkenő, a maximuma $N = 23$-ban van, és a ML becslés

$$\hat{N} = 23.$$

$L_x(N)$ grafikonja egyébként így néz ki:



7. feladat

A momentum-becsléshez

$$g(N) = \mathbb{E}_N(X_1) = 0.6N,$$

és így

$$g^{-1}(\bar{x}) = \frac{1}{0.6} \cdot 14 \approx 23.33.$$

7. feladat

A momentum-becsléshez

$$g(N) = \mathbb{E}_N(X_1) = 0.6N,$$

és így

$$g^{-1}(\bar{x}) = \frac{1}{0.6} \cdot 14 \approx 23.33.$$

Persze N csak egész lehet, ami a momentum-becslés esetén nem garantált. Jobb híján kerekíthetjük a legközelebbi egészre, és akkor a momentum-becslés is

$$\hat{N} = 23.$$

5. feladat

Egy M/M/1 szerver terheltségét ($\rho = \lambda/\mu$, ahol λ az érkezési, μ a kiszolgálási ráta) a bufferben sorban álló igények számából akarjuk megbecsülni. Tudjuk, hogy ha a terheltség ρ ($0 < \rho < 1$), akkor a bufferben lévő igények száma PGEO($1 - \rho$) eloszlású (beleértve az éppen kiszolgálás alatt álló igényt is).

5 távoli időpontban ránézve a szerverre, a bufferben lévő igények számára a következő adódott: 2, 0, 4, 1, 1. Adjunk ML-becslést ρ értékére a minta alapján. Adjunk momentum-becslést is ρ értékére.

5. feladat

Egy M/M/1 szerver terheltségét ($\rho = \lambda/\mu$, ahol λ az érkezési, μ a kiszolgálási ráta) a bufferben sorban álló igények számából akarjuk megbecsülni. Tudjuk, hogy ha a terheltség ρ ($0 < \rho < 1$), akkor a bufferben lévő igények száma PGEO($1 - \rho$) eloszlású (beleértve az éppen kiszolgálás alatt álló igényt is).

5 távoli időpontban ránézve a szerverre, a bufferben lévő igények számára a következő adódott: 2, 0, 4, 1, 1. Adjunk ML-becslést ρ értékére a minta alapján. Adjunk momentum-becslést is ρ értékére.

(Miért fontos, hogy távoli időpontok legyenek?)

5. feladat

Megoldás. $X_i \sim \mathrm{PGEO}(1-\rho)$ esetén

$$g(\rho) = \mathbb{E}_\rho(X_1) = \frac{\rho}{1-\rho} = \frac{1}{1-\rho} - 1,$$

5. feladat

Megoldás. $X_i \sim \mathrm{PGEO}(1-\rho)$ esetén

$$g(\rho) = \mathbb{E}_\rho(X_1) = \frac{\rho}{1-\rho} = \frac{1}{1-\rho} - 1,$$

aminek az inverze

$$g^{-1}(x) = 1 - \frac{1}{1+x},$$

5. feladat

Megoldás. $X_i \sim \mathrm{PGEO}(1-\rho)$ esetén

$$g(\rho) = \mathbb{E}_\rho(X_1) = \frac{\rho}{1-\rho} = \frac{1}{1-\rho} - 1,$$

aminek az inverze

$$g^{-1}(x) = 1 - \frac{1}{1+x},$$

és így a momentum-becslés

$$\hat{\rho} = g^{-1}(\bar{x}) = 1 - \frac{1}{1+\bar{x}} = 1 - \frac{1}{1+8/5} = \frac{8}{13}.$$

5. feladat

A likelihood és log-likelihood függvény

$$\begin{aligned} L_x(\rho) &= (1-\rho)\rho^2 \cdot (1-\rho)\rho^0 \cdot (1-\rho)\rho^4 \cdot (1-\rho)\rho^1 \cdot (1-\rho)\rho^1 = \\ &\qquad \rho^8(1-\rho)^5, \\ \ell_x(\rho) &= 8\log(\rho) + 5\log(1-\rho). \end{aligned}$$

A likelihood és log-likelihood függvény

$$L_x(\rho) = (1-\rho)\rho^2 \cdot (1-\rho)\rho^0 \cdot (1-\rho)\rho^4 \cdot (1-\rho)\rho^1 \cdot (1-\rho)\rho^1 = \rho^8(1-\rho)^5,$$
$$\ell_x(\rho) = 8\log(\rho) + 5\log(1-\rho).$$

$\frac{\mathrm{d}}{\mathrm{d}\rho}\ell_x(\rho) = 0$ megoldása alapján a ML-becslés

$$\hat{\rho} = \frac{8}{13},$$

tehát ebben az esetben megegyezik a momentum-becsléssel.

8. feladat

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

8. feladat

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

Megoldás. Ugyan a hőmérséklet miatt nem azonos eloszlású a minta, de azért a likelihood-függvényt fel tudjuk írni.
$X_i \sim \mathrm{EXP}(t_i/\theta)$ a feltétel szerint, és így

$$L_x(\theta) = \prod_{i=1}^{n} \frac{t_i}{\theta} e^{-\frac{t_i}{\theta} x_i}.$$

8. feladat

A log-likelihood-függvény

$$\ell_x(\theta) = \sum_{i=1}^{n} \log(t_i) - n\log(\theta) - \sum_{i=1}^{n} \frac{t_i x_i}{\theta},$$

8. feladat

A log-likelihood-függvény

$$\ell_x(\theta) = \sum_{i=1}^{n} \log(t_i) - n\log(\theta) - \sum_{i=1}^{n} \frac{t_i x_i}{\theta},$$

és

$$\frac{\mathrm{d}}{\mathrm{d}\theta}\ell_x(\theta) = 0 - \frac{n}{\theta} + \sum_{i=1}^{n} \frac{t_i x_i}{\theta^2} = 0$$

megoldása

$$\hat{\theta} = \frac{\sum_{i=1}^{n} t_i x_i}{n}.$$