Matematikai statisztika
Témakör 5: Bayes-statisztika alapjai

Elek Péter

2015 november

Motiváció: a klasszikus statisztika "problémái"

- A klasszikus statisztikában nem vesszük figyelembe a becslés előtt esetlegesen már meglevő információkat a paraméter hozzávetőleges értékéről (*a priori* információ).
  - Pl. korábbi, vagy más területen végzett becslések alapján – lásd pl. a korábbi feladatot a hierarchikus modellről
- Hasonlóan, a pontbecslések összehasonlítása problémás lehet, mert nem feltétlenül van egyértelmű rendezés két becslés között
  - a paramétertartománytól függhet, hogy melyik becslés jobb pl. az MSE-kritérium szerint

A priori és a posteriori eloszlás

- *a priori* eloszlás: a θ paraméter eloszlása a kísérlet megkezdése előtt rendelkezésre álló információk alapján
- Legyen $q(t)$ az a priori sűrűségfüggvény.
- Legyen a minta $\mathbf{x} = (x_1, \ldots, x_n)$ és a minta sűrűségfüggvénye a θ paramétertől függően $f_\theta(\mathbf{x})$.
- *a posteriori* eloszlás: a θ paraméter mintavétel utáni eloszlása.
- Ez a Bayes-tételből következően

  $$q(t|\mathbf{x}) = \frac{f_t(\mathbf{x}) \cdot q(t)}{f(\mathbf{x})},$$

  ahol

  $$f(\mathbf{x}) = \int_t f_t(\mathbf{x})\, q(t)\, dt.$$

- Megjegyzés: a nevező nem függ t-től, ezért azt nem feltétlenül kell explicit módon kiszámítani.

Becslések összehasonlítása Bayes-i keretben

- Korábban definiáltuk a négyzetes veszteségfüggvényt a becslések összehasonlítására:
$$MSE_\theta\left(\hat{\theta}\right) = E_\theta\left[\left(\hat{\theta}-\theta\right)^2\right],$$
- ami függhet θ értékétől, tehát nem feltétlenül teszi egyértelműen összehasonlíthatóvá a becslőfüggvényeket.
- Ha θ maga is valószínűségi változó, akkor definiálhatjuk az *a priori rizikó*t a fenti veszteségfüggvénynek az a priori eloszlás szerinti várható értékeként:
$$E\left[\left(\hat{\theta}-\theta\right)^2\right] = \int_t E_\theta\left[\left(\hat{\theta}-\theta\right)^2\right] q(t)\, dt.$$
- Így két becslőfüggvény egyértelműen összehasonlítható.
  - Az összehasonlítás eredménye persze függ $q(t)$-től.

Bayes-becslés

- A θ paraméter *Bayes-becslés*e az a posteriori eloszlás várható értéke:
$$\hat{\theta} = E(\theta|\mathbf{X}=\mathbf{x}) = \int_t t\cdot q(t|\mathbf{x})\, dt = \frac{\int_t t\cdot f_t(\mathbf{x})\cdot q(t)\, dt}{\int_t f_t(\mathbf{x})\cdot q(t)\, dt}.$$
- Megmutatható, hogy a Bayes-becslés minimalizálja az a priori rizikót.
- Tehát adott $q(t)$ a priori sűrűségfüggvény esetén ez az optimális becslőfüggvény.

Példa: Bayes-becslés normális eloszlású minta esetén

- Legyen $X_1,\ldots,X_n$ f.a.e. minta $N\left(\theta,\sigma^2\right)$ eloszlásból, ahol σ^2 az egyszerűség kedvéért ismert és θ a becsülendő paraméter.
- Legyen továbbá $\theta\sim N\left(\mu,\eta^2\right)$ az a priori eloszlás.
- Határozzuk meg θ Bayes-becslését!
- Megjegyzés: a példa kapcsolódhat a korábbi, hierarchikus modellre vonatkozó feladathoz.
  - Az a priori eloszlás jöhet az egységek összességére vonatkozó információból.
  - Empirikus Bayes-i módszerek: olyan módszerek, amikor az a priori eloszlást (más) adatok alapján becsüljük meg

Megoldás

- Az a posteriori sűrűségfüggvény (ahol a K_i konstansokat nem szükséges kiszámítanunk):

$$
\begin{aligned}
q(t|\mathbf{x}) &= K_1 f_t(\mathbf{x})\, q(t) = \\
&= K_2 \exp\left[-\frac{\sum_{i=1}^n (x_i - t)^2}{2\sigma^2}\right] \cdot \exp\left[-\frac{(t-\mu)^2}{2\eta^2}\right] = \\
&= K_3 \exp\left[-\frac{t^2}{2}\left(\frac{n}{\sigma^2} + \frac{1}{\eta^2}\right) + t\left(\frac{n\overline{x}}{\sigma^2} + \frac{\mu}{\eta^2}\right)\right] = \\
&= K_4 \exp\left[-\frac{1}{2}\left(\frac{n}{\sigma^2} + \frac{1}{\eta^2}\right)\left(t - \frac{n\overline{x}/\sigma^2 + \mu/\eta^2}{n/\sigma^2 + 1/\eta^2}\right)^2\right].
\end{aligned}
$$

Megoldás, folyt.

- Tehát az a posteriori eloszlás:

$$\theta|\mathbf{x} \sim N\left(\frac{n\overline{x}/\sigma^2 + \mu/\eta^2}{n/\sigma^2 + 1/\eta^2}, \frac{1}{n/\sigma^2 + 1/\eta^2}\right).$$

- θ Bayes-becslése ennek várható értéke:

$$\hat{\theta} = (1-\lambda)\cdot \overline{x} + \lambda\mu,$$

- ahol

$$\lambda = \frac{1/\eta^2}{n/\sigma^2 + 1/\eta^2}.$$

- A Bayes-becslés tehát a mintaátlag és az a priori várható érték súlyozott átlaga, a szórásoktól és mintaelemszámtól függő súllyal.

Bayes-becslések néhány tulajdonsága

- A Bayes-becslések (triviális esetektől eltekintve) nem torzítatlanok, de (általános regularitási feltételek mellett) konzisztensek és aszimptotikusan normálisak.
- Láttuk, hogy adott a priori eloszlás esetén minimalizálják az a priori rizikót.
- Bayes-becslések kiszámítása:
  - A számítás sokszor nem lehetséges analitikusan, csak szimulációs módszerekkel (pl. Markov-lánc Monte Carlo [MCMC])
  - Konjugált prior eloszlás: olyan a priori eloszlás, amikor az a posteriori eloszlás is ugyanazon eloszláscsaládba tartozik. Ilyenkor a számítások könnyebben végezhetők el. Példák:
    - normális eloszlás normális likelihood esetén (ld. korábban)
    - béta-eloszlás binomiális likelihood esetén (ld. később)

Feladat: Bayes-becslés binomiális minta esetén

- Legyen $X_1, \ldots, X_n$ f.a.e. minta p paraméterű Bernoulli-eloszlásból.
- Legyen továbbá p a priori eloszlása béta-eloszlás (α, β) paraméterrel.
- A béta-eloszlás tulajdonságai:
  - Sűrűségfüggvény: $f(x) = K \cdot x^{\alpha-1} (1-x)^{\beta-1}$ ha $0 < x < 1$ és 0 egyébként, ahol K megfelelő konstans (az ún. béta-függvény reciproka). Feltesszük, hogy $\alpha > 0$ és $\beta > 0$.
  - Várható érték: $\alpha / (\alpha + \beta)$
  - Segítségével rugalmasan lehet a $[0, 1]$ intervallum valószínűségi változóit modellezni.
  - Speciális eset: egyenletes eloszlás ha $\alpha = \beta = 1$
- Határozzuk meg p Bayes-becslését!