Likelihood függvény

Likelihood-függvény

Ha $X_1, ..., X_n$ i.i.d. véletlen változók, és $f(x|\theta)$ eloszlás függvény (PDF). Ekkor a **likelihood függvény**t úgy definiáljuk, hogy

$$\mathcal{L}(\theta) = \prod_{i=1}^{n} f(X_i|\theta).$$

A **log-likelihood függvény**t pedig a következő módon írjuk fel:
$\ell(\theta) = log\mathcal{L}(\theta)$

Maximum likelihood becslő

A **maximum likelihood** becslő (MLE), jelölése $\hat{\theta}_n$, nem más, mint θ azon értéke, amely mellett a likelihood függvény ($\mathcal{L}(\theta)$) a legnagyobb.

Likelihood függvény: megjegyzés

Mivel a logaritmus egy szigorúan monotone növekvő transzformáció, ezért a log-likelihood és a likelihood függvények ugyanazon θ értékek mellett veszik fel a szélső értékeiket. Ebből következik, hogy a log-likelihood függvény maximumának megkeresése ugyanarra a válaszra vezet. A gyakorlatban gyakrabban alkalmazzuk a log-likelihoodot, mert könnyebben kiszámítható.

Ha $\mathcal{L}(\theta)$-t megszorozzuk bármely olyan tetszőleg pozitív c számmal, amely nem függ θ-tól, akkor a MLE nem változik. Emiatt a likelihood függvényben lévő konstans értékeket gyakran el is hagyjuk.

MLE tulajdonságok

Bizonyos regularitási feltételek teljesülése esetén az MLE elvű becslő olyan jó tulajdonságokkal rendelkezik, amelyek népszerű becslési módszerré teszik.

Tulajdonságok

1. Invariancia: Ha $\hat{\theta}_{MLE}$ a legjobb becslésünk θ-ra, akkor, $g(\hat{\theta}_{MLE})$ a a legjobb becslés $g(\theta)$-ra
2. Konzisztencia: $\hat{\theta} \xrightarrow{p} \theta$
3. Aszimptotikus normalitás: $\frac{\hat{\theta}-\theta}{se} \xrightarrow{d} N(0,1)$
4. Aszimptotikus hatékonyság

Regresszió

A hipotézis vizsgálat során, annak érdekében, hogy következtetni tudjunk a becsült paraméterekről, feltevéseket tettünk a hibatagok eloszlására vonatkozóan. A Maximum Likelihood (MLE) elvű becslés során ezt a feltevést a becslési folyamat során is felhasználjuk.

Ha feltesszük, hogy a hibatagok:

1. normális eloszlásúak
2. homoszkedasztikusak
3. feltételesen korrelálatlanok

akkor a likelihood függvény:

$$\mathcal{L}(\beta, \sigma^2) = (2\pi\sigma^2)^{-\frac{n}{2}} \exp - \frac{(Y - X\beta)^T (Y - X\beta)}{2\sigma^2}.$$

A log-likelihood függvény:

$$\ell(\beta, \sigma^2) = -\frac{n}{2} log(2\pi) - \frac{n}{2} log(\sigma^2) - \frac{(Y - X\beta)^T (Y - X\beta)}{2\sigma^2}$$

Elsőrendű feltételek

A log-likelihood függvény maximuma az ismeretlen paraméterek függvényében úgy számolható ki, hogy vesszük azok parciális deriváltjait:

$$\frac{\partial \ell(\beta, \sigma^2)}{\partial \beta} = \frac{X^T(Y - X\hat{\beta})}{\sigma^2} = 0$$
$$\frac{\partial \ell(\beta, \sigma^2)}{\partial \sigma^2} = -\frac{n}{2\hat{\sigma}^2} + \frac{(Y - X\hat{\beta})^T(Y - X\hat{\beta})}{2\hat{\sigma}^4} = 0.$$

Ezen egyenleteket felhasznála arra a megoldásra jutunk, hogy:

$$\hat{\beta}_{MLE} = (X^T X)^{-1} X^T Y$$
$$\hat{\sigma}^2_{MLE} = n^{-1}(Y - X\hat{\beta})^T(Y - X\hat{\beta}) = n^{-1}\epsilon^T\epsilon$$

Ahogyan látható, a MLE elven becsült regressziós koefficiensek **megegyeznek** az OLS elven becsült koeffíciensekkel, azonban a varianciák különböznek. Az MLE a korrigálatlan minta varianciát adja vissza. (Emiatt lesz csak aszimptotikusan torzítatlan a becslés)

Hatékonyság

Legkisebb varianciájú torzítatlan becslő

Amennyiben **feltesszük**, hogy a Gauss-Markov feltevések teljesültek, és a hibatag normális eloszlású, akkor

$$\hat{\beta} = \hat{\beta}_{MLE}$$

a legkisebb varianciájú β torzítatlan becslői körében.

A Gauss-Markov tételhez képest ezen tétel kis különbsége annál lényegesebb. A MLE elvű becslés esetén - a feltételek teljesülése mellett - már nem csak a lineáris, hanem a nem-lineáris becslők körében sem találunk kisebb varianciájut.