TAHMİN EDİCİLERDE ARANAN ÖZELLİKLER

Tahmin edici elde etme yöntemlerini kullanarak elde ettiğimiz tahmin edicilerin hangisinin daha iyi olduğu, tahmin yaparken hangisinin tercih edileceği merak konusudur. Tahmin edicilerde aranan özellikler aşağıdaki gibi verilebilir:

- Yansızlık (Unbiased)
- Etkinlik (Efficiency)
- Tutarlılık (Consistency)

YANSIZLIK

Tanım: $\mathrm{T}(X_1, X_2, \dots, X_n)$, θ parametresi için bir tahmin edici olmak üzere her θ için,

$$E_\theta(\mathrm{T}(X_1, X_2, \dots, X_n)) = \theta$$

oluyorsa , $\mathrm{T}(X_1, X_2, \dots, X_n)$ tahmin edicisine θ parametresi için yansız bir tahmin edici denir. O zaman

$$Bias_\theta\,(\mathrm{T}(X_1, X_2, \dots, X_n)) = E_\theta(\mathrm{T}(X_1, X_2, \dots, X_n)) - \theta$$

değerine, T tahmin edicisinin yan değeri (veya yanlılığı) denir.

ETKİNLİK

Tanım 1: T_1 *ve* T_2 , θ parametresinin *yansız* iki tahmin edicisi olsun.

$$E_\theta T_1^2 < \infty \ \text{ ve } \ E_\theta T_2^2 < \infty$$

olduğunu varsayalım. $T_1^{'}in$ T_2 ‘ye göre göreceli etkinliği,

$$eff_\theta\left(\frac{T_1}{T_2}\right) = \frac{Var_\theta(T_1)}{Var_\theta(T_2)}$$

ile tanımlanır ve eğer

$$eff_\theta\left(\frac{T_1}{T_2}\right) < 1$$

ise T_1, T_2 ‘den daha etkindir denir.

Tanım 2:

n örnek ölçümünü gösterdiğine göre

$$\lim_{n\to\infty} eff_\theta(T_1) = 1$$

ise T_1 tahmininin asimptotik olarak (en) etkin olduğu söylenir ve T_1 için

$$\lim_{n\to\infty} E_\theta(T_1) = \theta$$

ise T_1asimtotik olarak yansızdır.

Örnek: (Momentler metodu)

$X_1, X_2, \dots, X_n \sim$ Düzgün $(0, \theta)$ ise

a) Momentler metodunu kullanarak θ ‘ nu tahmin edicisini bulunuz?

Çözüm: Düzgün dağılım için kitle momenti;

$$\mathrm{E}(x) = \frac{\theta}{2}$$

dir. 1. kitle momentini 1.örneklem momentine eşitlersek Momentler tahmin edicisi

$$\frac{1}{n}\sum_{i=1}^{n} x_i = \frac{\theta}{2} \Longrightarrow \bar{X} = \frac{\theta}{2} \Longrightarrow \tilde{\theta} = 2\bar{X}$$

bulunur.

b) $X_1, \ldots, X_n \sim$ Düzgün $(0, \theta)$ ise Momentler tahmin edicisini mi yoksa En Çok Olabilirlik tahmin edicisini mi tercih edersiniz? Neden?

Momentler yöntemi ile elde edilen tahmin edici $\tilde{\theta} = 2\bar{X}$ ise

$$E\left(\tilde{\theta}\right) = E\left(2\bar{X}\right) = 2\, E\left(\bar{X}\right) = 2 \,.\, \frac{\theta}{2} = \theta$$

$$V\left(\tilde{\theta}\right) = V\left(2\bar{X}\right) = 4\, V\left(\bar{X}\right) = 4 \,.\, \frac{V\,(X)}{n} = 4 \,.\, \frac{1}{n} \,.\, \frac{\theta^2}{12} = \frac{\theta^2}{3n}$$

En Çok Olabilirlik yöntemi ile elde edilen tahmin edici $\hat{\theta} = X_{(n)}$ ise

$$E\left(X_{(n)}\right) = \int_0^{\theta} x \,.\, \frac{nX^{n-1}}{\theta^n}\, dx = \frac{n}{\theta^n} \int_0^{\theta} X^n .\, dx = \frac{n}{n+1}\, \theta$$

$$E\left(X_{(n)}^2\right) = \int_0^{\theta} x^2\, \frac{n.x^{n-1}}{\theta^n} .dx = \frac{n}{\theta^n} \int_0^{\theta} x^{n+1} .dx$$

$$= \frac{n}{\theta^n} \,.\, \frac{x^{n+2}}{n+2}\Big|_0^{\theta} = \frac{n}{\theta^n} \,.\, \frac{\theta^{n+2}}{n+2}$$

$$= \frac{n}{\theta^n} \,.\, \frac{\theta^n . \theta^2}{n+2} = \frac{n}{n+2} \,.\, \theta^2$$

$$\Rightarrow V\left(X_{(n)}\right) = E\left(X_{(n)}^2\right) - \left[E\left(X_{(n)}\right)\right]^2$$

$$V\left(X_{(n)}\right) = \frac{\theta^2 n}{(n+1)^2(n+2)}$$

$$V\left(\tilde{\theta}\right) = \frac{\theta^2}{3n} \geq V\left(\hat{\theta}\right) = \frac{\theta^2 n}{(n+2)(n+1)^2}$$

En Çok olabilirlik tahmin edicisi yanlı, fakat asimptotik olarak yansızdır. Momentler tahmin edicisi yansızdır. Varyanslar incelendiğinde En Çok olabilirlik tahmin edicisinin varyansı, Momentler tahmin edicisinin varyansından daha küçüktür.

"Bu durumda ne yapacağız?"

Bu durumda Hata Kareler Ortalamasına (MSE) bakılır.

$$\text{MSE} = Yan^2 + \text{Varyans}$$

TUTARLILIK:

Tanım:

T_n , θ' nın tahmin edicilerinin bir dizisi olsun. Bu tahmin ediciler eğer her ε > 0 için

$$\lim_{n\to\infty} P\,[\,|T_n - \theta| < \varepsilon] = 1 \quad , \forall\, \theta \text{ için}$$

ise tutarlı tahmin ediciler olarak adlandırılırlar.

Teorem:

T_n , θ ' nın tahmin edicilerinin bir dizisi olsun. Eğer

- ❖ $\lim_{n\to\infty} Var_\theta(T_n) = 0$ (MSE tutarlılık)

ve

- ❖ $\lim_{n\to\infty} E_\theta(T_n) = \theta$ (Asimptotik yansızlık)

koşulları sağlanırsa, T_n tahmin edicilerinin dizisi θ parametresi için tutarlıdır.

Örnek:

$X_1, X_2, \dots, X_n$ beklenen değeri μ , varyansı 1 olan normal dağılımdan bir örneklem olsun.

$$T_n = \frac{X_1 + X_2 + \cdots + X_n}{n}$$

tahmin edicilerinin dizisini göz önüne alalım. $T_n \sim \mathrm{N}\left(\mu, \frac{1}{n}\right)$ olduğunu biliyoruz. O zaman, her ε > 0 için

$$P\,(|T_n - \mu| < \varepsilon) = P\,(-\varepsilon < T_n - \mu < \varepsilon)$$
$$= P\,(\mu - \varepsilon < T_n < \mu + \varepsilon)$$

$$= \int_{\mu-\varepsilon}^{\mu+\varepsilon} \frac{1}{\sqrt{2\pi\left(\frac{1}{n}\right)}} . e^{-\frac{1}{2\left(\frac{1}{n}\right)}(x-\mu)^2} . dx$$

$$= \int_{t=-\varepsilon\sqrt{n}}^{\varepsilon\sqrt{n}} \left(\frac{1}{2\pi}\right)^{\frac{1}{2}} . e^{-\frac{t^2}{2}} . \mathrm{dt}$$

$$\mathrm{P}\left(-\varepsilon\sqrt{n} < z < \varepsilon\sqrt{n}\right) \xrightarrow{n \to \infty} 1$$

T_n tahmin edicilerinin dizisi μ ‘ nün tutarlı bir tahmin edicisidir.

Kaynaklar

(1) Akdi, Y. (2010) Matematiksel İstatistiğe Giriş, Gazi Kitabevi, Ankara.

(2) Hogg, R. V. And Craing, A. T. (1989). Introduction to Mathematical Statistics. 4th Ed., New York: Macmillan Publishing Co.

(3) Mendenhall, W., Wackerly, D. D. and Scheaffer, R. (1990). Mathematical Statistics with Applications. 4th Ed., Boston: PWS-Kent Publishing Company.

(4) Hogg, R. V. And Tanis, E. A. (1993) Probability and Statistical Inference. 4th Ed., New York: Macmillan Publishing Co.

(5) Larson, H. J. (1982). Introduction to Probability Theory and Statistical Inference. 3rd Ed., New York: John Wiley ve Sons.

(6) Öztürk, F., Akdi, Y., Aydoğdu, H. Ve Karabulut, İ. (2006). Parametre Tahmini ve Hipotez Testi, Bıçaklar Kitabevi, Ankara.

(7) Casella, G. ve Berger, R.L. (2002). Statistical Inference, Second Edition, Duxbury.