

Tabakalı Rastgele Örnekleme

Kitledeki birimlerin heterojen bir yapıya sahip olduğu durumlarda kullanılan bir örnekleme yöntemidir.

Kitle kendi içinde homojen kendi aralarında heterojen olan N_1, N_2, \dots, N_L büyüklüğündeki **tabaka (strata)** adı verilen ve

$$N = N_1 + N_2 + \dots + N_L = \sum_{h=1}^L N_h$$

koşulunu sağlayan alt gruplara ayrılır.

Tabakalardan örneklem çekerken farklı örnekleme yöntemleri kullanılabilir.

Tabakalardan basit rastgele örnekleme yöntemi kullanılarak örneklem çekiliyor ise bu örnekleme yöntemi **tabakalı rastgele örnekleme (stratified random sampling)** olarak adlandırılır.

Tabakalı rastgele örneklemede kitle hakkında çıkarımlar yapılabildiği gibi herbir tabaka hakkındada ayrı ayrı çıkarımlar yapılabilir.

Kitle Toplamı Y ve Kitle Ortalaması \bar{Y} 'nın Tahmini

Y_1, Y_2, \dots, Y_N , N çaplı bir kitle olsun. Bu durumda

$$N = N_1 + N_2 + \dots + N_L = \sum_{h=1}^L N_h$$

şeklinde ifade edilir. Burada,

N : Kitle çapı

N_h : h . tabakanın çapı

L : Tabaka sayısını

gösterir.

h . Tabaka Toplamı

$$Y_h = \sum_{i=1}^{N_h} Y_{hi}$$

olarak ifade edilir. Burada,

Y_{hi} : h . tabakadaki i . birimin almış olduğu değeri

gösterir.

Kitle Toplamı

$$Y = \sum_{h=1}^L Y_h = \sum_{h=1}^L \sum_{i=1}^{N_h} Y_{hi}$$

olarak ifade edilir. Bir başka deyişle, kitle toplamı tabakaların toplamına eşittir.

h. Tabaka Ortalaması

$$\bar{Y}_h = \frac{Y_h}{N_h} = \frac{\sum_{i=1}^{N_h} Y_{hi}}{N_h}$$

olarak ifade edilir.

Kitle Ortalaması

$$\bar{Y} = \frac{\sum_{h=1}^L Y_h}{N} = \frac{\sum_{h=1}^L \sum_{i=1}^{N_h} Y_{hi}}{N}$$

olarak ifade edilir.

N_h çaplı her bir tabakadan n_h çaplı örneklem çekildiğinde toplam örneklem çapı

$$n = n_1 + n_2 + \dots + n_L = \sum_{h=1}^L n_h$$

olarak ifade edilir.

Burada,

n_h : h. tabakadan seçilen örneklem çapını

L : Tabaka sayısını

gösterir.

h. Tabakanın Örneklem Toplamı

$$y_h = \sum_{i=1}^{n_h} y_{hi}$$

olarak ifade edilir.

Burada,

y_{hi} : h. tabakadaki örnekleme seçilen i. birimin almış olduğu değeri gösterir.

h. Tabakanın Örneklem Ortalaması

$$\bar{y}_h = \frac{y_h}{n_h} = \frac{\sum_{i=1}^{n_h} y_{hi}}{n_h}$$

olarak ifade edilir.

- h. tabakanın örneklem ortalaması \bar{y}_h , h. tabakanın ortalaması \bar{Y}_h 'nin yansız bir tahmin edicisidir. Bir başka deyişle,

$$E(\bar{y}_h) = \bar{Y}_h$$

dır.

h. Tabaka Toplamının Tahmin Edicisi

$$\hat{Y}_h = N_h \bar{y}_h$$

olarak ifade edilir.

- h. tabaka toplamının tahmin edicisi, h. tabaka toplamının yansız bir tahmin edicisidir. Bir başka deyişle,

$$E(\hat{Y}_h) = Y_h$$

dır.

Kitle Toplamı Y'in Tahmin Edicisi

$$\hat{Y}_{tb} = \sum_{h=1}^L \hat{Y}_h = \sum_{h=1}^L N_h \bar{y}_h$$

olarak ifade edilir. Bir başka deyişle, kitle toplamı Y'in tahmin edicisi tabaka toplamlarının tahmin edicilerinin toplamı olarak ifade edilir.

Kitle Ortalamasının Tahmin Edicisi

$$\bar{y}_{tb} = \frac{\hat{Y}_{tb}}{N} = \frac{\sum_{h=1}^L N_h \bar{y}_h}{N}$$

olarak ifade edilir.

Kitle ortalaması \bar{Y} 'nin tahmin edicisi, kitle ortalaması \bar{Y} 'nin yansız bir tahmin edicisidir. Bir başka deyişle,

$$E(\bar{y}_{tb}) = E\left[\frac{\sum_{h=1}^L N_h \bar{y}_h}{N}\right] = \frac{\sum_{h=1}^L N_h \bar{Y}_h}{N} = \bar{Y}$$

olarak bulunur.

Kitle Varyansı

Y_1, Y_2, \dots, Y_N , N çaplı bir kitle olsun. Bu kitlenin varyansı

$$\sigma^2 = \frac{1}{N} \sum_{h=1}^L \sum_{i=1}^{N_h} (Y_{hi} - \bar{Y})^2$$

yada

$$S^2 = \frac{1}{N-1} \sum_{h=1}^L \sum_{i=1}^{N_h} (Y_{hi} - \bar{Y})^2$$

olarak tanımlanır.

h. Tabakanın Varyansı

$$\sigma_h^2 = \frac{1}{N_h} \sum_{i=1}^{N_h} (Y_{hi} - \bar{Y}_h)^2$$

yada

$$S_h^2 = \frac{1}{N_h - 1} \sum_{i=1}^{N_h} (Y_{hi} - \bar{Y}_h)^2$$

olarak tanımlanır.

Not: $N \cong N - 1$ ise $\sigma_h^2 \cong S_h^2$ olarak alınabilir.

Y_1, Y_2, \dots, Y_N , N çaplı bir kitle olsun. Bu kitlenin varyansı

$$\begin{aligned} \sigma^2 &= \frac{1}{N} \sum_{h=1}^L \sum_{i=1}^{N_h} (Y_{hi} - \bar{Y})^2 \\ &= \frac{1}{N} \sum_{h=1}^L \sum_{i=1}^{N_h} (Y_{hi} - \bar{Y}_h)^2 + \frac{1}{N} \sum_{h=1}^L N_h (\bar{Y}_h - \bar{Y})^2 \\ &= \sigma_w^2 + \sigma_b^2 \end{aligned}$$

olarak elde edilir. Burada,

σ_w^2 : Tabaka içi varyans

σ_b^2 : Tabakalar arası varyans

olarak adlandırılır.