Küme Örneklemesi

Ne zaman kullanılır?

Kitlenin çok büyük olduğu dolayısıyla tüm birimlerin tam listesine ulaşılamadığı durumlarda uygulama maliyetini azaltmak için tercih edilen bir yöntemdir.

Küme örneklemesinin adımları

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

- Bu kitle kendi içerisinde heterojen özellik gösteren birimler bir arada olacak şekilde M tane kümeye bölünür.
- Küme çapları

  $$N = N_1 + N_2 + \cdots + N_M = \sum_{i=1}^{M} N_i$$

  eşitliği sağlanacak şekilde

  $$N_1, N_2, \dots, N_M$$

  olarak gösterilir.

- Bu kümelerin kendi aralarında homojen olması istenir.
- Bu kümelerden m tanesi rastgele olarak seçilir.
- Seçilen m küme içerisinden

  $$n = n_1 + n_2 + \cdots + n_m = \sum_{i=1}^{m} n_i$$

  eşitliği sağlanacak şekilde

  $$n_1, n_2, \dots, n_m$$

  çaplı örneklemler seçilir.
- M kümeye *birincil örnekleme birimi (böb)* adı verilir.
- M küme içerisinden rastgele olarak seçilen m kümeye *ikincil örnekleme birimi (iöb)* adı verilir.

Küme örneklemesinde kullanılan semboller

M: Küme çapı
m: Örnekleme seçilen kümelerin çapı
N: Kitle çapı
N_i: *i*. Kümenin çapı
n: Toplam örneklem çapı
n_i: i. Kümeden örnekleme seçilen birimlerin çapı

olarak ifade edilir.

Küme örneklemesinin uygulamada en yaygın kullanılan hali

- İki-aşamalı (**two-stage**) örnekleme

olarak adlandırılır.

Not: Küme örneklemesinin üç ve üçten fazla aşamalı türleri de vardır.

Küme Örneklemesinin Avantajları:

- İlgilendiğimiz kitlenin listesini oluşturmak yerine seçtiğimiz kümelerin listesini oluşturmak yeterlidir.
- Büyük çaplı kitlelerde uygulama maliyeti ve uygulama süresi bakımından avantaj sağlar.

Küme Örneklemesinin Dezavantajları

- Basit rastgele örnekleme ve tabakalı örneklemeye göre daha düşük bir etkinliğe sahiptir.
- Örnekleme hatası yüksek olabilir.

Kitle Toplamının Tahmini

- Örnekleme birimleri iki aşamada seçilir.
- Birinci aşamada, M tane küme içerisinden m tane küme $\binom{M}{m}$ farklı şekilde seçilir.
- İkinci aşamada, örnekleme seçilen m kümeden sırasıyla $n_1, n_2, \ldots, n_m$ çaplı örneklemler rastgele olarak seçilir.
- i. küme toplamı

$$y_i = y_{i1} + y_{i2} + \cdots + y_{in_i} = \sum_{j=1}^{n_i} y_{ij}$$

olarak bulunur.

- i. küme ortalaması

$$\bar{\bar{y}}_i = \frac{1}{n_i}\left(y_{i1} + y_{i2} + \cdots + y_{in_i}\right) = \frac{1}{n_i}\sum_{j=1}^{n_i} y_{ij}$$

olarak bulunur.

Burada, i. küme ortalaması, i. kümedeki birim başına ortalama için yansız bir tahmin edicidir. Bir başka deyişle,

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

olur.

- i. kümenin toplamının tahmin edicisi

$$\hat{Y}_i = N_i\bar{\bar{y}}_i$$
$$= \frac{N_i}{n_i}\left(y_{i1} + y_{i2} + \cdots + y_{in_i}\right)$$

olarak tanımlanır.

- i. kümenin toplamının tahmin edicisi, i. kümenin toplamı için yansızdır. Bir başka deyişle,

$$E\left(\hat{Y}_i\right) = Y_i$$

olarak ifade edilir.

m küme toplamının tahmin edicisi

$$\hat{y} = \sum_{i=1}^{m} \hat{Y}_i$$
$$= \sum_{i=1}^{m} N_i\bar{\bar{y}}_i$$
$$= \sum_{i=1}^{m} \frac{N_i}{n_i}\left(y_{i1} + y_{i2} + \cdots + y_{in_i}\right)$$

olarak ifade edilir.

Küme başına ortalamanın tahmin edicisi

$$\hat{\bar{\bar{Y}}} = \frac{\hat{y}}{m}$$
$$= \frac{1}{m}\sum_{i=1}^{m} \hat{Y}_i$$
$$= \frac{1}{m}\sum_{i=1}^{m} N_i\bar{\bar{y}}_i$$
$$= \frac{1}{m}\sum_{i=1}^{m} N_i\frac{y_i}{n_i}$$

$$= \frac{1}{m} \sum_{i=1}^{m} \frac{N_i}{n_i} \left(y_{i1} + y_{i2} + \cdots + y_{in_i} \right)$$

olarak ifade edilir.

Burada, küme başına ortalamanın tahmin edicisi, kitledeki küme başına ortalama için yansız bir tahmin edicidir. Bir başka deyişle,

$$E\left(\hat{\bar{\bar{Y}}}\right) = \bar{Y}$$

olarak ifade edilir.

Kitle Toplamının Tahmin Edicisi

Kitle toplamı,

$$\hat{Y} = M\hat{\bar{\bar{Y}}}$$

$$= M \frac{\hat{y}}{m}$$

$$= \frac{M}{m} \sum_{i=1}^{m} \hat{Y}_i$$

$$= \frac{M}{m} \sum_{i=1}^{m} N_i \bar{\bar{y}}_i$$

$$= \frac{M}{m} \sum_{i=1}^{m} \frac{N_i}{n_i} \left(y_{i1} + y_{i2} + \cdots + y_{in_i} \right)$$

olarak ifade edilir.

Burada, kitle toplamının tahmin edicisi, kitle toplamı için yansız bir tahmin edicidir. Bir başka deyişle,

$$E\left(\hat{Y}\right) = Y$$

olarak ifade edilir.

Kitlede, birim başına ortalamanın tahmin edicisi

$$\hat{\bar{\bar{Y}}} = \frac{\hat{Y}}{N}$$

$$= \frac{1}{N} \frac{M}{m} \sum_{i=1}^{m} \frac{N_i}{n_i} \left(y_{i1} + y_{i2} + \cdots + y_{in_i} \right)$$

olarak ifade edilir.

Kitlede birim başına ortalamanın tahmin edicisi, kitle ortalaması için yansız bir tahmin edicidir. Bir başka deyişle,

$$E\left(\hat{\bar{\bar{Y}}}\right) = \bar{\bar{Y}}$$

olarak ifade edilir.