6. TABAKALI ÖRNEKLEME

N genişliğinde bir kitle, kendi içlerinde homojen, aralarında da heterojen olan $N_1, N_2, \dots, N_L$ genişliklerinde alt kitlelerden oluşan $(\sum_{n=1}^{L} N_h = N)$ ve her bir alt kitleye uygulanan örnekleme yöntemine tabakalı örnekleme denir.

Her tabakaya aynı örnekleme yöntemi uygulanabileceği gibi farklı örnekleme yöntemleri de uygulanabilir. Örneğin her bir tabakaya Basit Rastgele Örnekleme yöntemi uygulanabileceği gibi farklı örnekleme yöntemi de uygulanabilir. Örneğin her tabakaya BRÖ uygulanırsa tabakalı rastgele örnekleme yöntemi olur. Sistematik örnekleme yöntemi uygulanırsa, tabakalı sistematik örnekleme yöntemi olur. BU bağlamda tabakalı örnekleme bir örnekleme yöntemi değil bilinen yöntemlerin alt kümelere uygulanmasıdır. Tabakalı örnekleme yapmadaki amaç varyansı küçültmektir. Varyansın küçülmesi ise kitle parametresinin daha dar bir alanda tahmin edilmesi anlamına gelir. Ancak tabaka sayısı arttıkça varyansta aynı oranda azalmaz. Çünkü her bir tabakaya düşen örneklem genişliği azalacağından istenen küçültme doğru orantılı olarak sağlanamaz. Bu bağlamda yönetsel ve idari zorluklar dışında tabaka olmalıdır. Örneğin ülke, bölge veya şehirler soyo-ekonomik yapıya göre fakir, orta ve zengin gibi 3 tabakaya ayrılabileceği gibi kamuoyu araştırmalarında ise kırsal ve kentsel kesim gibi iki tabakaya ayırmak mümkün olacaktır.

6.1. h-ıncı Tabakaya Ait Parametreler ve Tahmin Ediciler

L tabakadan oluşan bir kitlede h tabaka indisini, i de birim indisini gösteriyor olsun. h-ıncı tabaka için gösterimler, parametreler ve tahmin ediciler aşağıdadır.

Gösterimler:

Y_{h_i}: h-ıncı tabakadaki i-nci birim değeri

N_h: h-ıncı tabaka (kitle için) genişliği

n_h: h-ıncı tabaka (örneklem için) genişliği

$W_h = N_n/N$: h-ıncı tabaka ağırlığı

$f_h = n_h/n$: h-ıncı tabakanın örnekleme oranı

Parametreler:

- h-ıncı alt kitle ortalaması ve varyansı sırasıyla şöyledir:

$$\bar{Y}_h = \frac{\sum_{i=1}^{N_h} y_{h_i}}{N_h}$$

$$S_h^2 = \frac{\sum_{i=1}^{N_h}\left(y_{h_i} - \bar{Y}_h\right)^2}{N_h - 1}$$

- h-ıncı alt kitleye ait belli özelliğe sahip birimler oranı ve varyansı şöyledir:

$$P_h = \frac{A_h}{N_h}$$

$$S_h^2 = \frac{N_h P_h Q_h}{N_h - 1}$$

Tahmin ediciler:

- h-ıncı alt örneklem için ortalama ve varyans sırasıyla şöyledir:

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

$$s_h^2 = \frac{\sum_{i=1}^{n_h}\left(y_{h_i} - \bar{y}_h\right)^2}{n_h - 1}$$

- h-ıncı alt örnekleme için belli özelliğe sahip birimler oranı ve varyansı şöyledir:

$$p_h = \frac{a_h}{n_h}$$

$$s_h^2 = \frac{n_h p_h q_h}{n_h - 1}$$

6.2. Kitle Ortalaması ve Toplamının Tahmini

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

$$Y = \sum_{i=1}^{N_h} Y_h$$

Bir kitle tabakalara ayrılmış ise gerçek değeri $\bar{Y}_{tb}$, tahmini değeri $\bar{y}_{tb}$ ile gösterilsin.

$$\bar{Y}_{tb} = \frac{Y}{N} = \frac{\sum_{h=1}^{L} N_h \bar{Y}_h}{N} = \frac{1}{N}\sum_{h=1}^{L}\sum_{i=1}^{N_h} y_{h_i}$$

$$\bar{Y}_{tb} = \sum_{h=1}^{L} W_h \bar{Y}_h$$

$\bar{Y}_h$ yerine örneklemden $\bar{y}_{tb}$'ın kullanılmasıyla

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

$$= \sum_{h=1}^{L} W_h \bar{y}_h$$

Kitle toplamının tahmini ise

$$\hat{Y}_{tb} = N\bar{y}_{tb} \; ya \; da$$

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

NOT: Eğer her tabakada $\bar{y}_h$, $\bar{Y}_h$'ın yansız bir tahmini ise $\bar{y}_{tb}$'da $\bar{Y}_{tb}$'nın yansız bir tahmin edicisidir. Yani;

$$E(\bar{y}_{tb}) = \bar{Y}_{tb}$$

6.3. Kitle Ortalaması ve Toplamını Tahmin Edicilerini Varyanslarının Tahmin Edicileri

Teorem:

$$V(\bar{y}_{tb}) = \sum_{h=1}^{L} W_h^2 V(\bar{y}_h)$$

$$V(\bar{y}_h) = E(\bar{y}_h - \bar{Y}_h)^2$$

Bu teoremden aşağıdaki sonuçlar veriliyor:

Sonuç 1: Tabakalı örneklemede kitle toplamı tahmin edicisinin varyansı:

$$\hat{Y}_{tb} = N\bar{y}_{tb}$$

$$V(\bar{y}_{tb}) = N^2 V(\bar{y}_{tb})$$

$$V(\bar{y}_{tb}) = \sum_{h=1}^{L} W_h^2 V(\bar{y}_h)$$

Sonuç 2: Tabakalı örneklemede kitle ortalaması ve toplam tahmin edicisinin varyansının tahmin edicisi:

$$\hat{V}(\bar{y}_{tb}) = \sum_{h=1}^{L} W_h^2 \hat{V}(\bar{y}_h)$$

$$\hat{V}\left(\hat{Y}_{tb}\right) = \sum_{h=1}^{L} N_h^2 \hat{V}(\bar{y}_h)$$

Burada $\hat{V}(\bar{y}_h)$, h-ıncı tabakaya uygulanan örnekleme yöntemine göre ortalamanın varyans tahminidir.