4.3. Ortalama ve Toplam Tahminlerinin Varyanslarının Tahminleri

$$S^2 = \frac{\sum_{i=1}^{N}(y_i - \bar{Y})^2}{N-1}, s^2 = \frac{\sum_{i=1}^{n}(y_i - \bar{y})^2}{n-1}$$

Burada s^2, S^2'nin örneklem tahminidir. $\hat{Y}$ toplam ve $\bar{y}$ ortalama tahmin edicilerinin varyanslarının tahminine II. derece tahmin denir. Bu tahminler:

$$\hat{V}(\bar{y}) = (1-f)\frac{s^2}{n}$$

$$\hat{V}(\hat{Y}) = N^2(1-f)\frac{S^2}{n}, f = \frac{n}{N}$$

$$sh(\bar{y}) = s\sqrt{\frac{1-f}{n}}$$

$$sh(\hat{Y}) = N \times s\sqrt{\frac{1-f}{n}}$$

TEOREM 3:

BRÖ'de s^2, S^2'nin yansız bir tahmin edicisidir.

$$E(s^2) = S^2$$

İSPAT 3:

$$s^2 = \frac{\sum_{i=1}^{n}(y_i - \bar{y})^2}{n-1} = \frac{\sum_{i=1}^{n}[(y_i - \bar{Y}) - (\bar{y} - \bar{Y})]^2}{n-1}$$

$$= \frac{\sum_{i=1}^{n}(y_i - \bar{y})^2 - 2\sum_{i=1}^{n}(y_i - \bar{Y})(\bar{y} - \bar{Y}) + \sum_{i=1}^{n}(\bar{y} - \bar{Y})^2}{n-1}$$

$$\sum_{i=1}^{n}(y_i - \bar{Y})(\bar{y} - \bar{Y}) = (\bar{y} - \bar{Y})\sum_{i=1}^{n}(y_i - \bar{Y})$$

$$= (\bar{y} - \bar{Y})(n\bar{y} - n\bar{Y})$$

$$= n(\bar{y} - \bar{Y})^2$$

$$E(s^2) = \frac{E\sum_{i=1}^{n}(\bar{y} - \bar{Y})^2 - nE(\bar{y} - \bar{Y})^2}{n-1}$$

$$= \frac{1}{n-1}\left(E\sum_{i=1}^{n}(\bar{y} - \bar{Y})^2 - nE(\bar{y} - \bar{Y})^2 \right)$$

$$= \frac{1}{n-1}\left(E\sum_{i=1}^{n}(\bar{y} - \bar{Y})^2 - nV(\bar{y}) \right)$$

$$= \frac{1}{n-1}\left(\frac{n}{N}\sum_{i=1}^{n}(\bar{y} - \bar{Y})^2 - n\frac{N-n}{N}\frac{S^2}{n} \right)$$

$$= \frac{1}{n-1}\left(\frac{n}{N}S^2(N-1) - n\frac{N-n}{N}\frac{S^2}{n} \right)$$

$$= \frac{S^2}{N(n-1)}\left(\frac{n^2(N-1)}{n} - (N-n) \right)$$

$$= \frac{S^2}{N(n-1)}\frac{(n^2N - n^2 - nN + n^2)}{n} = \frac{S^2}{N(n-1)}\frac{nN(n-1)}{n}$$

$$E(s^2) = S^2$$

Örneklemede standart hata tahminleri 3 amaç için kullanılır:

1. Örnekleme yöntemlerini duyarlılık (varyans) yönünden birbiri ile karşılaştırmak için,
2. Bir araştırmanın planlama aşamasında örneklem genişliğini tahmin edebilmek için,
3. Tamamlanmış bir araştırmada duyarlılığı tam olarak elde edebilmek için.

Kitle Parametrelerinin güven aralığı:

θ kitle parametresi olmak üzere bunun tahmin edicisi $\hat{\theta}$'dır. $\hat{\theta}$'nın almış olduğu değere tahmin ya da nokta tahmini denir. Ayrıca θ belirli bir olasılıkla (güven düzeyi ile) tahmin edilir. Bu tahmine **aralık tahmini** denir ve şöyle ifade edilir.

$$P\left(\hat{\theta} - tablodeğeri \times sh(\hat{\theta}) \leq \theta \leq \hat{\theta} + tablodeğeri \times sh(\hat{\theta})\right) = 1 - \alpha$$

$sh(\hat{\theta})$, $\hat{\theta}$'nın standart hatasıdır. α 1. tip hata olmak üzere (ya da önem seviyesi) $1 - \alpha$'ya güven düzeyi denir. Tablo değeri ise ilgili teste ait iki yönlü tablo değeridir.

Kitle Ortalaması için güven arlığı:

$\theta = \hat{Y}$ ve $\hat{\theta} = \bar{y}$

$$P\left(\bar{y}-t_{\frac{\alpha}{2}}\sqrt{\frac{1-f}{n}}S \leq \bar{Y} \leq \bar{y}+t_{\frac{\alpha}{2}}\sqrt{\frac{1-f}{n}}S\right)=1-\alpha$$

$$\hat{V}(\bar{y})=(1-f)\frac{s^2}{n}$$

$$sh(\bar{y})=s\sqrt{\frac{1-f}{n}}$$

Kitle Toplamı için güven aralığı:

$\theta = Y$ ve $\hat{\theta} = \hat{Y}$

$$P\left(\hat{Y}-t_{\frac{\alpha}{2}}\sqrt{\frac{1-f}{n}}SN \leq Y \leq \hat{Y}+t_{\frac{\alpha}{2}}\sqrt{\frac{1-f}{n}}SN\right)=1-\alpha$$

$$P\left(N\bar{y}-t_{\frac{\alpha}{2}}\sqrt{\frac{1-f}{n}}SN \leq N\bar{Y} \leq N\bar{y}+t_{\frac{\alpha}{2}}\sqrt{\frac{1-f}{n}}SN\right)=1-\alpha$$

4.4. İki Değişkenin Birbirine Oranının Tahmini

Bazı araştırmalarda örneklem biriminden örneklem birimine değişen değerler alabilen iki değişkenin birbirine oranı tahmin edilmek isteniyor olabilir. Çoğu kez örneklem birimlerinin kümeler olduğu BRÖ'de bu tahmine gereksinim duyulabilir. Örneğin okulların örneklem birimi olarak alındığı araştırmada öğretmen ve öğrenci sayıları iki değişkendir. Öğretmen başına düşen öğrenci sayısı tahmin edilmek istenebilri. Değişkenlerden biri X_i diğeri Y_i ile gösterilirse

$$R=\frac{\sum_{i=1}^{N}\frac{y_i}{N}}{\sum_{i=1}^{N}\frac{x_i}{N}}=\frac{\bar{Y}}{\bar{X}}$$

$$\hat{R}=\frac{\sum_{i=1}^{n}\frac{y_i}{n}}{\sum_{i=1}^{n}\frac{x_i}{n}}=\frac{\bar{y}}{\bar{x}}$$

$\hat{R}$ tesadüfi değişkeninin dağılımı $\bar{y}$ tesadüfi değişkeninin dağılımına benzemez. $\hat{R}$ iki ortalamanın oranının dağılımı olduğu için normal dağılım göstermez. Küçük örneklemlerde dağılım pozitif yöne çarpıktır. Örneklem genişliği büyüdükçe dağılım normal dağılıma yakınlaşır. $\hat{R}$ tahmini bir miktar yanlıdır. Ancak örneklem genişliği büyüdükçe $(n > 60)$ yanlılık miktarı önemsizleşir.

Yanlılık miktarı:

$$E(\hat{R}) \neq R$$

$$Yanlılık\ Miktarı = E(\hat{R} - R) = E\left(\frac{\bar{y}}{\bar{x}} - R\right)$$

$$= E\left(\frac{\bar{y} - R\bar{x}}{\bar{x}}\right)$$

Bu beklenen değer sıfır ise o tahmin edici yansızdır. Yansızlık ise ancak $\bar{x}$ yerine $\bar{X}$ alınmasıyla sağlanabilir. O halde örneklem genişliği büyüdükçe $\bar{x}$, $\bar{X}$'ya yaklaşacağından bu beklenen değerde sıfıra yaklaşarak yanlılık miktarı önemsizleşecektir.

$$E\left(\frac{\bar{y} - R\bar{x}}{\bar{x}}\right) = \frac{1}{\bar{X}} E(\bar{y} - R\bar{x}) = \frac{1}{\bar{X}} E(\bar{y}) - \frac{1}{\bar{X}} RE(\bar{x})$$

$$= \frac{1}{\bar{X}} (\bar{Y} - R\bar{X})$$

$$= \frac{1}{\bar{X}} \left(\bar{Y} - \frac{\bar{Y}}{\bar{X}} \bar{X}\right) = 0$$