6. 추정식

❍ 모평균(비율)

– 본 조사에서의 주요 추정대상은 모평균이나 모비율이며 모비율 추정 및 추정량의 분산은 모평균 추정방법과 동일

– 다만, 모비율 추정의 경우 조사변수 값이 어떤 특성을 갖고 있는가에 따라 1 또는 0의 값을 갖는다는 점이 차이임

– 본 조사에서 각종 모평균 추정을 위해서 사용된 가중치를 이용한 추정량은 다음과 같이 정의됨

$$\bar{y}=\frac{\sum_{h=1}^{L}\sum_{i=1}^{n_h}\sum_{j=1}^{m_{hi}}\sum_{l=1}^{k_{hij}} w_{hijl}y_{hijl}}{\sum_{h=1}^{L}\sum_{i=1}^{n_h}\sum_{j=1}^{m_{hi}}\sum_{l=1}^{k_{hij}} w_{hijl}}=\frac{\sum_{h=1}^{L}\sum_{i=1}^{n_h}\sum_{j=1}^{m_{hi}}\sum_{l=1}^{k_{hij}} w_{hijl}y_{hijl}}{w_{....}}$$

– 여기서, w_{hijl}은 각 응답자에 부여된 가중치이고, y_{hijl}은 각 응답결과로 모비율 추정의 경우는 특정 속성을 갖고 있는 경우는 1 아니면 0의 값을 갖음. L은 층의 수, n_h는 층 h에서의 1차 표본추출단위인 표본 조사구의 수, m_{hi}는 층 h내 i번째 표본 조사구의 표본가구 수, k_{hij}는 층 h내 i번째 표본 조사구의 j번 표본가구의 응답자 수

$w_{....}=\sum_{h=1}^{L}\sum_{i=1}^{n_h}\sum_{j=1}^{m_{hi}}\sum_{l=1}^{m_{hij}} w_{hijl}$은 전체 응답자에 대한 가중치의 합계임

❍ 모평균(비율)추정량의 분산

– 앞서 제시한 모평균 추정량에 대해서 층화와 3단 집락추출 등의 표본설계를 반영한 추정분산은 다음과 같이 추정

$$var(\bar{y})=\sum_{h=1}^{L}\frac{n_h(1-f_h)}{n_h-1}\sum_{i=1}^{n_h}(e_{hi..}-\overline{e_{h...}})^2$$

– 여기서, L은 층의 수, n_h는 층 h에서의 1차 추출단위(PSU)인 조사구 수, m_{hi}는 층 h내 i번째 표본조사구의 표본 가구수 임

$$f_h=n_h/N_h,\ e_{hi..}=\left(\sum_{j=1}^{m_{hi}}\sum_{l=1}^{k_{hij}} w_{hijl}(y_{hijl}-\bar{y})\right)/w_{....},\ \overline{e_{h...}}=\left(\sum_{i=1}^{n_h} e_{hi..}\right)/n_hm_{hi}$$

– 모평균 및 모비율 추정에 대한 표준오차(standard error), 상대표준오차(relative standard error), 95% 신뢰수준 오차의 한계는 다음과 같음

$$s.e(\bar{y})=\sqrt{var(\bar{y})},$$

$$rse(\bar{y})=\frac{s.e(\bar{y})}{\bar{y}}\times 100(\%),$$

$$\text{오차의 한계} = \pm 1.96\times\sqrt{var(\bar{y})}$$