13. Det randomiserte komplette blokk-forsøket

I ein-vegs variansanalyse, kan forskjellar i eksperimentelle einingar ~~kan~~ vere så store at det blir vanskeleg å detektere forskjellar mellom populasjonane. Ein måte å oppnå større likheit mellom einingane er ~~å~~ å gjere ei blokkdeling.

Ein annan måte å uttrykke det på er som følgjer. Når du veit eller mistenkjer at det er spesielle kjelder som forårsakar uønska forandringar i dei observerte verdiane, kan du redusere eller eliminere effekten av dei ved blokkdeling. Eksperimenta vil no bli randomiserte innan blokker.

Ein modell for det randomiserte blokk forsøket er

$$Y_{ij} = \mu_j + \varepsilon_{ij} \left\{ \begin{array}{l} N(0,\sigma^2) \text{ og uavh.} \\ j=1,2,\dots,k, \quad i=1,2,\dots,b. \end{array} \right.$$

der b er talet på blokker og k er talet på observasjonar innan kvar blokk.

Ei moglighet er å la $Y_{ij} = \mu_j + \beta_i + \varepsilon_{ij}$.

Men då blir: $E[\bar{Y}_{..}] = \dfrac{\sum_{i=1}^{b}\sum_{j=1}^{k}(\mu_j + \beta_i)}{kb} = \dfrac{\sum_{i=1}^{b}\left(\sum_{j=1}^{k}\mu_j + k\beta_i\right)}{kb}$

$$= \frac{b\sum_{j=1}^{k}\mu_j + k\sum_{i=1}^{b}\beta_i}{kb} = \frac{kb\bar{\mu} + kb\bar{\beta}}{kb} = \bar{\mu} + \bar{\beta}$$

der $\bar{\mu} = \frac{1}{k}\sum_{j=1}^{k}\mu_j$ og $\bar{\beta} = \frac{1}{b}\sum_{i=1}^{b}\beta_i$

Spesielt $E[\bar{Y}_{i\cdot} - \bar{Y}_{\cdot\cdot}] = \frac{1}{k}\sum_{j=1}^{k} E[Y_{ij}] - E[\bar{Y}_{\cdot\cdot}]$

$= \frac{1}{k}\sum_{j=1}^{k}(\mu_j + \beta_i) - E[\bar{Y}_{\cdot\cdot}] = \bar{\mu} + \beta_i - (\bar{\mu} + \bar{\beta}) = \beta_i - \bar{\beta}$

Ei bitre bilorcerming

$\mu_{ij} = \mu' + \alpha_j' + \beta_i'$

Vi har $\mu_{ij} = \mu' + \bar{\alpha}_\cdot' + \bar{\beta}_\cdot' + \alpha_j' - \bar{\alpha}_\cdot' + \beta_i' - \bar{\beta}_\cdot'$

der $\bar{\alpha}_\cdot' = \sum_{j=1}^{k} \frac{\alpha_j'}{k}$ og $\bar{\beta}_\cdot' = \sum_{i=1}^{b} \frac{\beta_i'}{b}$

Definer $\mu = \mu' + \bar{\alpha}_\cdot' + \bar{\beta}_\cdot'$ $\alpha_j = \alpha_j' - \bar{\alpha}_\cdot'$ og $\beta_i = \beta_i' - \bar{\beta}_\cdot'$

og vi får

$Y_{ij} = \overbrace{\mu + \alpha_j + \beta_i}^{\mu_{ij}} + \varepsilon_{ij}$ $\left\{ \begin{array}{l} N(0,\sigma^2) \text{ og uavh.} \\ j = 1, 2, \ldots, k, \quad i = 1, 2, \ldots, b \end{array} \right.$

der $\sum_{j=1}^{k} \alpha_j = \sum_{i=1}^{b} \beta_i = 0$

Oppsplitting av variasjon

$Y_{ij} = \bar{Y}_{\cdot\cdot} + \bar{Y}_{\cdot j} - \bar{Y}_{\cdot\cdot} + Y_{ij} - \bar{Y}_{\cdot j} = \bar{Y}_{\cdot\cdot} + \underbrace{\bar{Y}_{\cdot j} - \bar{Y}_{\cdot\cdot}}_{\hat{\alpha}_j} + \underbrace{\bar{Y}_{i\cdot} - \bar{Y}_{\cdot\cdot}}_{\hat{\beta}_i} + Y_{ij} - \bar{Y}_{\cdot j} - \bar{Y}_{i\cdot} + \bar{Y}_{\cdot\cdot}$

der $\bar{Y}_{\cdot j} = \frac{1}{b}\sum_{i=1}^{b} Y_{ij}$, $\bar{Y}_{i\cdot} = \frac{1}{k}\sum_{j=1}^{k} Y_{ij}$ og $\bar{Y}_{\cdot\cdot} = \frac{1}{kb}\sum_{i=1}^{b}\sum_{j=1}^{k} Y_{ij}$

$E[\bar{Y}_{\cdot\cdot}] = \frac{1}{kb}\sum_{i=1}^{k}\sum_{j=1}^{k}(\mu + \alpha_j + \beta_i) = \mu$

$E[\bar{Y}_{\cdot j} - \bar{Y}_{\cdot\cdot}] = \frac{1}{b}\sum_{i=1}^{b}[\mu + \alpha_j + \beta_i] - \mu = \mu + \alpha_j - \mu = \alpha_j$

Tilsvar. $E[\bar{Y}_{i\cdot} - \bar{Y}_{\cdot\cdot}] = \beta_i$

$\cancel{E[\bar{Y}_{\cdot j} - \bar{Y}_{\cdot\cdot}] = \alpha_j}$

Frå einvegs variansanalyse har vi

$$\sum_{i=1}^{b}\sum_{j=1}^{k}(Y_{ij}-\bar{Y}_{..})^2 = \sum_{i=1}^{b}\sum_{j=1}^{k}(\bar{Y}_{.j}-\bar{Y}_{..})^2 + \underbrace{\sum_{i=1}^{b}\sum_{j=1}^{k}(\bar{Y}_{i.}-\bar{Y}_{..}+Y_{ij}-\bar{Y}_{.j}-\bar{Y}_{i.}+\bar{Y}_{..})^2}_{SSE_{one-way}}$$

$$SSE_{one-way} = \sum_{i=1}^{b}\sum_{j=1}^{k}(\bar{Y}_{i.}-\bar{Y}_{..})^2 + \sum_{i=1}^{b}\sum_{j=1}^{k}(Y_{ij}-\bar{Y}_{.j}-\bar{Y}_{i.}+\bar{Y}_{..})^2$$

$$+\underbrace{2\sum_{i=1}^{b}\sum_{j=1}^{k}(\bar{Y}_{i.}-\bar{Y}_{..})(Y_{ij}-\bar{Y}_{.j}+\bar{Y}_{i.}+\bar{Y}_{..})}_{R}$$

$$\bar{Y}_{i.} = \frac{1}{k}\sum_{j=1}^{k} Y_{ij}$$

$$R = \underbrace{2\sum_{i=1}^{b}\sum_{j=1}^{k}(\bar{Y}_{i.}-\bar{Y}_{..})(Y_{ij}-\bar{Y}_{i.})}_{0} - \underbrace{2\sum_{i=1}^{b}\sum_{j=1}^{k}(\bar{Y}_{i.}-\bar{Y}_{..})(\bar{Y}_{.j}-\bar{Y}_{..})}_{0}$$

og vi får:

$$\sum_{i=1}^{b}\sum_{j=1}^{k}(Y_{ij}-\bar{Y}_{..})^2 = \sum_{i=1}^{b}\sum_{j=1}^{k}(\bar{Y}_{.j}-\bar{Y}_{..})^2 + \sum_{i=1}^{b}\sum_{j=1}^{k}(\bar{Y}_{i.}-\bar{Y}_{..})^2 + \sum_{i=1}^{b}\sum_{j=1}^{k}(Y_{ij}-\bar{Y}_{.j}-\bar{Y}_{i.}+\bar{Y}_{..})^2$$

$$SSTOT = SSTR + \underbrace{SSB + SSE}_{SSE_{one-way}} \qquad \boxed{\text{Teorem } 13.2.1 \text{ a}}$$

SSE

$$SSE = \sum_{i=1}^{b}\sum_{j=1}^{k} X_{ij}^2 \quad \text{der} \quad X_{ij} = Y_{ij} - \hat{\mu}_{ij}$$

$$Y_{ij}-\hat{\mu}_{ij} = Y_{ij} - \hat{\alpha}_j - \hat{\beta}_i - \hat{\mu} = Y_{ij} - \underbrace{(\bar{Y}_{.j}-\bar{Y}_{..})}_{\hat{\alpha}_j} - \underbrace{(\bar{Y}_{i.}-\bar{Y}_{..})}_{\hat{\beta}_i} - \underbrace{\bar{Y}_{..}}_{\hat{\mu}}$$

$$= Y_{ij} - \bar{Y}_{.j} - \bar{Y}_{i.} + \bar{Y}_{..}, \qquad E[Y_{ij}-\hat{\mu}_{ij}] = 0$$

Det trengs: $\overbrace{b-1 + k-1 + 1}^{\text{å estimere:}} = b+k-1$ parametrar for estimere $\hat{\mu}_{ij}$. Derfor har SSE $bk-b-k+1$ ~~fridomsgrader~~ $=(b-1)(k-1)$ fridomsgrader og $\frac{SSE}{\sigma^2} \sim \chi^2((b-1)(k-1))$ alltid.