■ 공분산분석(Analysis of Covariance)

- 요인(factor)과 공변량(covariate)을 설명변수로 갖는 선형모형

  ↳ concomitant variable

- 공변량 x가 반응변수를 예측하는데 어느 정도 기여할 수 있거나 확률화를 했으나 각 처리에서 x가 균형을 맞추지 못한 경우

  ⇨ 모형에 공변량을 포함시킴으로써 실험의 정도와 검정력을 높을 수 있음

- 확률화가 적용되지 않은 관측연구(observational study)의 경우 x가 골고루 섞여 있을 보장이 없음 ⇨ 모형에 공변량을 포함시킴으로써 편의(bias)를 줄임

○ 완전임의배치법

요인			
1	2	...	k
(y_{11}, x_{11}) (y_{12}, x_{12}) $\vdots$ (y_{1n_1}, x_{1n_1})	(y_{21}, x_{21}) (y_{22}, x_{22}) $\vdots$ (y_{2n_2}, x_{2n_2})	$\cdots$ $\cdots$ $\ddots$ $\cdots$	(y_{k1}, x_{k1}) (y_{k2}, x_{k2}) $\vdots$ (y_{kn_k}, x_{kn_k})

$$Y_{ij} = \alpha + \tau_i + \beta x_{ij} + \varepsilon_{ij} = \mu + \tau_i + \beta(x_{ij} - \overline{x}_{..}) + \varepsilon_{ij}$$

- 가정: $\varepsilon_{ij} \sim$ iid $N(0, \sigma^2)$, $\sum \tau_i = 0$
- $\alpha = \mu - \beta \overline{x}_{..}$

- $Y_{ij} \sim N(\mu_{ij}, \sigma^2)$
  - $E(Y_{ij}) = \mu_{ij} = \mu + \tau_i + \beta(x_{ij} - \overline{x}_{..}) \quad \Rightarrow \quad \sum_i \sum_j \mu_{ij} = n\mu, \qquad (n = n_1 + \cdots + n_k)$
  - $E(\overline{Y}_{i.}) = \mu + \tau_i + \beta(\overline{x}_{i.} - \overline{x}_{..})$, $\quad E(\overline{Y}_{..}) = \mu$

- 공분산분석 모형에 대한 최소제곱추정량
  - $\hat{\mu} = \overline{Y}_{..}$
  - $\hat{\beta} = \dfrac{\sum_i \sum_j (x_{ij} - \overline{x}_{i.})(Y_{ij} - \overline{Y}_{i.})}{\sum_i \sum_j (x_{ij} - \overline{x}_{i.})^2}$
  - $\hat{\tau}_i = \overline{Y}_{i.} - \overline{Y}_{..} - \hat{\beta}(\overline{x}_{i.} - \overline{x}_{..})$
  - $\hat{\sigma}^2 = \dfrac{1}{n-k-1} \sum_i \sum_j \left(Y_{ij} - \overline{Y}_{i.} - \hat{\beta}(x_{ij} - \overline{x}_{i.}) \right)^2$

- 처리효과검정 ($H_0 : \tau_1 = \tau_2 = \cdots = \tau_k = 0$)
  - $SSE(R(\tau))$: 귀무가설(Reduced model) 하에서의 SSE
  - $SSE(F)$: 대립가설(Full model) 하에서의 SSE
  - 검정통계량

$$F^* = \frac{SSE(R(\tau)) - SSE(F)}{(n-2)-(n-k-1)} \Big/ \frac{SSE(F)}{n-k-1} \sim F_{k-1,n-k-1}$$

- 기울기 β에 대한 검정 ($H_0 : \beta = 0$)
  - 검정통계량

$$F^* = \frac{SSE(R(\beta)) - SSE(F)}{(n-k)-(n-k-1)} \Big/ \frac{SSE(F)}{n-k-1} \sim F_{1,n-k-1}$$

- μ_{ij}에 대한 추정: $E(Y_{ij}) = \mu_{ij}$ $\Leftarrow$ $\hat{Y}_{ij} = \hat{\alpha} + \hat{\tau}_i + \hat{\beta} x_{ij}$

- 기울기의 동일성(test of parallel slopes)

$$Y_{ij} = \alpha + \tau_i + \beta_i x_{ij} + \varepsilon_{ij}$$

  - 기울기의 동일성: $H_0 : \beta_1 = \beta_2 = \ \cdots \ = \beta_k$
  - $Y_{ij} = \alpha + \tau_i + \beta + (\tau\beta)_i + \varepsilon_{ij}$ 에서 상호작용 $(\tau\beta)_i$ 에 대한 유의성 검정
  - Full 모형과 Reduced 모형의 SSE 비교

$$F^* = \frac{SSE(R) - SSE(F)}{k-1} / \frac{SSE(F)}{n-2k} \sim F_{k-1,n-2k}$$

  - 기울기의 동일성 가정이 성립 ⇨ 공분산분석
  - 기울기의 동일성 가정 할 수 없음 ⇨ 일반선형모형을 통해 각 처리수준에서의 기울기 추정하고 기울기에 대한 차이 및 전반적 현상을 해석