$H_0: \sigma_\tau^2 = 0$ in LNp.4-63 cf.

One-way Random Effects Model: ANOVA

check LNp.5

- balance: In the following, assume $n_1 = \cdots = n_k = n$.
- The null hypothesis for the FEM:

  same null dist.

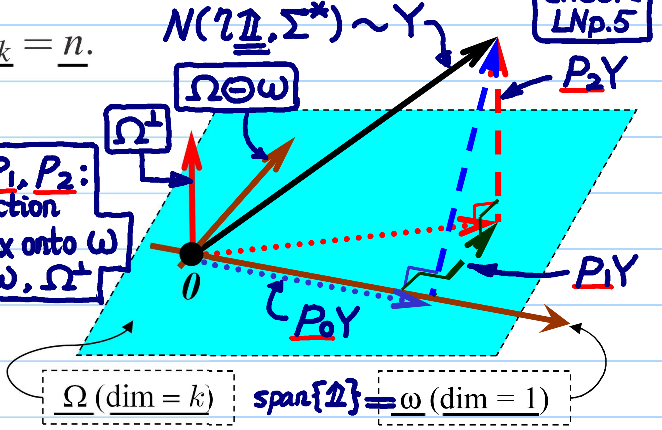
  $$H_0: \tau_1 = \cdots = \tau_k$$

  $Y \sim N(\eta \mathbb{1}, \sigma^2 I)$

  should be replaced by (meaning?)

  $$H_0^*: \sigma_\tau^2 = 0.$$

  Under H_0^*, the F-test and the ANOVA table in LNp. 3-6 still holds.



- Reason: under H_0^*,

  (under H_0^*) $SSTr \sim \sigma^2 \chi^2_{k-1}$,

  and

  (under H_0^* & H_A^*) $SSE \sim \sigma^2 \chi^2_{N-k}$,

  and they are independent.

  Therefore the F-test has the distribution $F_{k-1,N-k}$ under H_0^*.

  $\dfrac{SSTr/(k-1)}{SSE/(N-k)}$

$E(P_0Y) = P_0\eta\mathbb{1} = \eta\mathbb{1}$, $E(P_1Y) = E(P_2Y) = 0$

$\Omega^\perp$, Ω : different eigenspace of Σ^*

σ^2, $n\sigma_\tau^2 + \sigma^2$: eigenvalue (exercise, use the vectors in LNp.3-4)

(0) $SSTr = \|P_1Y\|^2 = \sum_i n(\bar{y}_{i\cdot} - \bar{y}_{\cdot\cdot})^2$ (LNp.3-5)

$= \sum_i (\sqrt{n}\,\bar{y}_{i\cdot} - \sqrt{n}\,\bar{y}_{\cdot\cdot})^2 \sim (n\sigma_\tau^2 + \sigma^2)\chi^2_{k-1}$

$\sqrt{n}\,\bar{y}_{i\cdot} = \sqrt{n}(\eta + \tau_i + \bar{\varepsilon}_{i\cdot}) \overset{iid}{\sim} N(\sqrt{n}\,\eta, n\sigma_\tau^2 + \sigma^2)$

$\bar{y}_{\cdot\cdot} = \bar{\bar{y}}_{i\cdot}$ (check LNp.36)

$SSE = \|P_2Y\|^2 = \sum_i \sum_j (y_{ij} - \bar{y}_{i\cdot})^2$ (LNp.3-5) $\sim \sigma^2 \chi^2_{N-k}$, $k(n-1)$

$y_{ij} = \eta + \tau_i + \varepsilon_{ij}$

$= \sum_j (\varepsilon_{ij} - \bar{\varepsilon}_{i\cdot})^2 \sim \sigma^2 \chi^2_{n-1}$

$cov(P_1Y, P_2Y) = P_1\,cov(Y)\,P_2^T$
$= P_1 \Sigma^* P_2^T = 0$

columns of P_1 & P_2 are eigenvectors of Σ^* & $P_1P_2 = 0$

ANOVA Tables ($n_i = n$)

- We can apply the *same* ANOVA and F-test in the fixed effects case for analyzing data.

  same test statistic, same null dist

different (cf. E(MS) in LNp.3-6)

ANOVA table (FEM) in LNp.3-6 cf.

Source	d.f.	SS	MS	E(MS)
treatment	$k-1$	$SSTr$	$MSTr = \frac{SSTr}{k-1}$	$\sigma^2 + n\sigma_\tau^2$
residual	$N-k$	SSE	$MSE = \frac{SSE}{N-k}$	σ^2
total	$N-1$			

Under $H_0^* \cup H_A^*$

Under H_0^*, $\sigma_\tau^2 = 0$, $E(MSTr) = \sigma^2$

Pulp Experiment

ANOVA result (FEM) in LNp.3-7 cf.

Source	d.f.	SS	MS	E(MS)
treatment	3	1.34	0.447	$\sigma^2 + 5\sigma_\tau^2$
residual	16	1.70	0.106	σ^2
total	19	3.04		

- However, we need to compute the expected mean squares under the alternative of $\sigma_\tau^2 > 0$,

  (i) for sample size determination, and

  (ii) to estimate the variance components. (σ_τ^2 & σ^2)

Expected Mean Squares for Treatments

SSE/N-k : an unbiased estimator of σ^2

- Equation (1) holds independent of σ_τ^2,

(LNp.3-33) $\sigma^2\chi^2_{N-k} \sim$

$$E(MSE) = E\left(\frac{SSE}{N-k}\right) = \sigma^2. \quad (1)$$

SSE only contains information of error var. component σ^2

- Under the alternative: $\sigma_\tau^2 > 0$, and for $n_i = n$,

(LNp.3-33) $(n\sigma_\tau^2+\sigma^2)\chi^2_{k-1} \sim$

$$E(MSTr) = E\left(\frac{SSTr}{k-1}\right) = \sigma^2 + n\sigma_\tau^2. \quad (2)$$

$$E\left(\underbrace{\frac{\frac{SSTr}{k-1} - \frac{SSE}{N-k}}{n}}_{\text{an unbiased estimator of } \sigma_\tau^2}\right) = \sigma_\tau^2$$

SSTr contains information about factor var. component σ_τ^2 & error var. component σ^2 (cf. E(SSTr) of FEM in LNp.3-6)

- For unequal n_i's, n in (2) is replaced by

$$n' = \frac{1}{k-1}\left[\sum_{i=1}^k n_i - \frac{\sum_{i=1}^k n_i^2}{\sum_{i=1}^k n_i}\right].$$

(exercise) use (o) in LNp.3-33

Proof of (2)

$y_{ij} = \eta + \tau_i + \varepsilon_{ij}$

$z_1, \ldots, z_k \overset{\text{i.i.d.}}{\sim} N(\mu, \theta^2)$

$\frac{\sum_{i=1}^k (z_i - \bar{z})^2}{\theta^2} \sim \chi^2_{k-1}$

$\Rightarrow \sum_{i=1}^k (z_i-\bar{z})^2 \sim \theta^2\chi^2_{k-1}$

$$\underbrace{\bar{y}_{i.}}_{\eta+\tau_i+\bar{\varepsilon}_{i.}} - \underbrace{\bar{y}_{..}}_{\eta+\bar{\tau}_.+\bar{\varepsilon}_{..}} = (\tau_i - \bar{\tau}_.) + (\bar{\varepsilon}_{i.} - \bar{\varepsilon}_{..})$$

LNp.3-5

$$SSTr = \sum_{i=1}^k n(\bar{y}_{i.} - \bar{y}_{..})^2$$

$$= n\left\{\sum_{i=1}^k (\tau_i - \bar{\tau}_.)^2 + \sum_{i=1}^k (\bar{\varepsilon}_{i.} - \bar{\varepsilon}_{..})^2 + 2\sum_{i=1}^k \underbrace{(\bar{\varepsilon}_{i.} - \bar{\varepsilon}_{..})(\tau_i - \bar{\tau}_.)}_{\text{independent}}\right\}.$$

E E E mean=0

The cross product term has mean 0 (because τ and ε are independent). It can be shown that

$$E\left(\underbrace{\sum_{i=1}^k (\tau_i - \bar{\tau}_.)^2}_{\sim \sigma_\tau^2\chi^2_{k-1}}\right) = (k-1)\sigma_\tau^2 \quad \text{and} \quad E\left(\underbrace{\sum_{i=1}^k (\bar{\varepsilon}_{i.} - \bar{\varepsilon}_{..})^2}_{\sim \frac{\sigma^2}{n}\chi^2_{k-1}}\right) = \frac{(k-1)\sigma^2}{n}.$$

iid $N(0, \sigma_\tau^2)$ average of τ_i's; iid $N(0, \sigma^2/n)$ average of $\bar{\varepsilon}_{i.}$'s

Therefore

$$E(SSTr) = n(k-1)\sigma_\tau^2 + (k-1)\sigma^2,$$

$$E(MSTr) = E\left(\frac{SSTr}{k-1}\right) = \sigma^2 + n\sigma_\tau^2.$$

cf. In FEM (LNp.3-6) $E(SSTr) = n\cdot\sum_i(\tau_i - \bar{\tau})^2 + (k-1)\sigma^2$

Variance components: estimation of σ^2 and σ_τ^2

- From equations (1) and (2) in LNp. 3-35, we obtain the following unbiased estimates of the variance components:

Can this be always ≥0? (Note. $\sigma_\tau^2 \geq 0$)

same as the $\hat{\sigma}^2$ in FEM

$$\hat{\sigma}^2 = MSE \quad \text{and} \quad \hat{\sigma}_\tau^2 = \frac{MSTr - MSE}{n}.$$

check LNp3-35

$MSTr/MSE$ ‖

Note that $\hat{\sigma}_\tau^2 \geq 0$ if and only if $MSTr \geq MSE$, which is equivalent to $F \geq 1$. Therefore a *negative* variance estimate $\hat{\sigma}_\tau^2$ occurs only if the value of the F statistic is less than 1. Obviously the null hypothesis H_0 is not rejected when $F \leq 1$. Since variance cannot be negative, a negative variance estimate is replaced by 0. This does not mean that σ_τ^2 is zero. It simply means that there is not enough information in the data to get a good estimate of σ_τ^2.

$E(F_{n_1,n_2}) = \frac{n_2}{n_2-2}$

$H_0: \sigma_\tau^2 = 0$

not "accept H_0"

- For the pulp experiment, $n = 5$, $\hat{\sigma}^2 = 0.106$, $\hat{\sigma}_\tau^2 = (0.447 - 0.106)/5 = 0.068$, i.e., sheet-to-sheet variance (within same operator) is 0.106, which is about 50% higher than operator-to-operator variance 0.068.

cf. ↕ v

a property of operator population

Implications on process improvement: try to reduce the two sources of variation, also considering costs.

Estimation of Overall Mean η

← the only fixed effect in REM

check graph in LNp3-31

the intercept parameter in FEM is usually of no interest

- In REM, η, the population mean, is often of interest. cf.

From $E(y_{ij}) = \eta$, we use the estimate

In FEM, $E(y_{ij}) = \mu_i = \eta + \tau_i$ cf.

For balanced data, GLS=OLS

$$\hat{\eta} = \bar{y}_{\cdot\cdot}$$

same as the $\hat{\eta}$ in FEM under sum coding (LNp3-8), but in the case of FEM $\eta = (\mu_1 + \cdots + \mu_k)/k$

- $Var(\hat{\eta}) = Var(\bar{\tau}_\cdot + \bar{\varepsilon}_{\cdot\cdot}) = \frac{\sigma_\tau^2}{k} + \frac{\sigma^2}{N}$, where $N = \sum_{i=1}^k n_i$.

$\hat{\eta} = \bar{y}_{\cdot\cdot} = \eta + \bar{\tau}_\cdot + \bar{\varepsilon}_{\cdot\cdot} \sim N(\eta, \sigma_\tau^2/k + \sigma^2/N)$

$= E(MSTr)$

pivotal quantity $\frac{\bar{y}_{\cdot\cdot} - \eta}{se(\hat{\eta})}$

For $n_i = n$, $Var(\hat{\eta}) = \frac{\sigma_\tau^2}{k} + \frac{\sigma^2}{nk} = \frac{1}{nk}(\sigma^2 + n\sigma_\tau^2)$.

$s.e.(\hat{\eta}) = \sqrt{\frac{MSTr}{nk}}$

Using (2) in LNp.3-35, $\frac{MSTr}{nk}$ is an unbiased estimate of $Var(\hat{\eta})$.

Confidence interval for η: $\bar{y}_{\cdot\cdot}$ and $MSTr$ are indep. (LNp.33)

$SSTr \sim (\sigma^2 + n\sigma_\tau^2)\chi^2_{k-1}$

estimate ± (critical value) × s.e.(estimate)

$$\hat{\eta} \pm t_{k-1,\frac{\alpha}{2}} \sqrt{\frac{MSTr}{nk}}$$

cf. N

In FEM (under sum coding) C.I. for η: $\bar{y}_{\cdot\cdot} \pm t_{N-k,\frac{\alpha}{2}} \sqrt{\frac{MSE}{N}}$

$SSE \sim \sigma^2 \chi^2_{N-k}$

- In the pulp experiment, $\hat{\eta} = 60.40$, $MSTr = 0.447$, and the 95% confidence interval for η is

$$60.40 \pm 3.182\sqrt{\frac{0.447}{5 \times 4}} = [59.92, 60.88].$$

compare REM (LNp3-32) & Split-plot design (LNp.4-45~66, future lecture)

❖ **Reading**: textbook, 2.5