One-way Random Effects Model

Tirthankar Dasgupta

Pulp Experiment Revisited

Reflectance data in pulp experiment: each of four operators made five pulp sheets; reflectance was read for each sheet using a brightness tester.

Randomization : assignment of 20 containers of pulp to operators and order of reading.

Table: Reflectance Data, Pulp Experiment

Operator A	Operator B	Operator C	Operator D
59.8	59.8	60.7	61.0
60.0	60.2	60.7	60.8
60.8	60.4	60.5	60.6
60.8	59.9	60.9	60.5
59.8	60.0	60.3	60.5

Objective : determine if there are differences among operators in making sheets and reading brightness.

Fixed versus Random Effects

- In the pulp experiment the effects τ_i are called *fixed* effects because the interest was in comparing the four *specific* operators in the study.
- If these four operators were chosen randomly from the population of operators in the plant, the interest would usually be in the variation among all operators in the population.
- Because the observed data are from operators randomly selected from the population, the variation among operators in the *population* is referred to as *random* effects.

One-way Random Effects Model

- Model:

$$y_{ij} = \mu + \tau_i + \epsilon_{ij},$$

  ϵ_{ij}'s are independent error terms with $N(0, \sigma^2)$, τ_i are independent $N(0, \sigma_\tau^2)$, and τ_i and ϵ_{ij} are independent.
- Here σ^2 and σ_τ^2 are the two *variance components* of the model. The variance among operators in the population is measured by σ_τ^2.

ANOVA decomposition and hypothesis testing

- The null hypothesis for the fixed effects model: $\tau_1 = \cdots = \tau_k$ should be replaced by

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

Under H_0, the F test and the ANOVA decomposition described earlier in the context of fixed effects model will still hold.

Reason: under H_0, $SSTr \sim \sigma^2 \chi^2_{k-1}$ and $SSE \sim \sigma^2 \chi^2_{N-k}$. Therefore the F-test has the distribution $F_{k-1,N-k}$ under H_0.

ANOVA decomposition and hypothesis testing (contd.)

Source	Degrees of Freedom (*df*)	Sum of Squares	Mean Squares	F
operator	3	1.34	0.447	4.20
residual	16	1.70	0.106	
total	19	3.04		

- We can apply the *same* ANOVA and F test in the fixed effects case for analyzing data. However, we need to compute the expected mean squares under the alternative of $\sigma_\tau^2 > 0$ to estimate the variance components.

Expected Mean Squares

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

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

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

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

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

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

ANOVA Tables ($n_i = n$)

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	$N-1$			

Pulp Experiment

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		

Estimation of σ^2 and σ_τ^2

- From equations (1) and (2), we obtain the following unbiased estimates of the variance components:

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

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.

Estimation of σ^2 and σ_τ^2

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

Estimation of Overall Mean μ

- In random effects model, μ, the population mean, is often of interest. From $E(y_{ij}) = \mu$, we use the estimate $\hat{\mu} = \bar{y}_{..}$.
- $Var(\hat{\mu}) = Var(\bar{\tau} + \bar{\epsilon}_{..}) = \frac{\sigma_\tau^2}{k} + \frac{\sigma^2}{N}$, where $N = \sum_{i=1}^{k} n_i$.
- For $n_i = n$, $Var(\hat{\mu}) = \frac{\sigma_\tau^2}{k} + \frac{\sigma^2}{nk} = \frac{1}{nk}\left(\sigma^2 + n\sigma_\tau^2\right)$.
- Using (2), $\frac{MSTr}{nk}$ is an unbiased estimate of $Var(\hat{\mu})$.

Confidence Interval for μ

- A $100(1-\alpha)\%$ Confidence interval for μ:

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

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

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