# Random Effects Model

( Eisenhart ~ 1947 )

fixed vs random effects in Anova

**Example:** Sodium content in Beer (USA data set)

**Experiment**
- Six brands
- Each brand 8 12 ounce bottles were sampled & Sodium content in mg were measured.

**Questions:**

1. Expected value of sodium content for a given brand ?
2. Does sodium content differ significantly between brands?
3. What is expected value of sodium content in Beer?
4. What is the variability of Sodium content between different Beer brands?

Data - (2005) Kutner et al

## Random effects model

$\boxed{\text{One way Anova}}$

$Y_{ij}$ - $j^{th}$ observation in factor level $i$

$$Y_{ij} = \mu + \alpha_i + \varepsilon_{ij} \quad j = 1, .., n_i \; ; \; i = 1, .. I$$

$\alpha_i$ - fixed effect

$\varepsilon_{ij} \sim N(0, \sigma^2)$ independent

Random effects model

$Y_{ij}$ - $j^{th}$ observation in factor level $i$

$$Y_{ij} = \mu + \alpha_i + \varepsilon_{ij} \quad j=1,..,n_i \; ; \; i=1,..I$$

$\varepsilon_{ij} \sim N(0, \sigma^2)$

$\alpha_i \sim N(0, \sigma_\alpha^2)$

independent

In R we used library(lme4) and applied lmer function :

$\hat{\mu} = 17.629$

$\hat{\sigma}^2 = 0.716$

$\hat{\sigma}_\alpha^2 = 26.274$

<u>REML</u> : restricted maximum likelihood

- Predict random effects
- Estimate fixed effects & variance parameters
- Parameters : $\mu$, $\sigma^2$, $\sigma_\alpha^2$

<u>Usage</u>:- - Model has less parameters
- (preferred) # of brands is large $(I >>)$
  # of repetitions is small $(n_i <<)$

i.e. if # $\alpha_i$ >>> and # $n_i$ <<< .. insufficient data to estimate $\alpha_i$

# Differences between fixed effect & Random effect

one way Anova

- $E[Y_{ij}] = \mu + \alpha_i$
- $Cov[Y_{ij}, Y_{ik}]$

$$= Cov(\mu + \alpha_i + \varepsilon_{ij}, \mu + \alpha_i + \varepsilon_{ik})$$

$$= Cov(\varepsilon_{ij}, \varepsilon_{ik}) = 0$$

All responses are independent
(∵ normality)

Random effects model

- $E[Y_{ij}] = \mu$
- $Var[Y_{ij}] = \sigma_\alpha^2 + \sigma^2$
- $Cov(Y_{ij}, Y_{ik})$

$$\overset{Ex}{=} \sigma_\alpha^2$$

(observations within group/factor are not statistically independent)

## Intra class Correlation

$$\rho = \frac{Cov(Y_{ik}, Y_{ij})}{\sqrt{Var(Y_{ik})\, var(Y_{ij})}} \overset{Ex}{=} \frac{\sigma_\alpha^2}{\sigma_\alpha^2 + \sigma^2}$$

## Hypothesis Testing:

$H_0: \sigma_\alpha^2 = 0$ (no variability in the factor levels)

$\Downarrow$

$H_A: \sigma_\alpha^2 > 0$ $\quad \alpha_i = 0, \; E[\alpha_i] = 0 \;\; \forall i$

### Use F-test:

(Recall: same as one way Anova F-test)

$$\frac{SS_\alpha / I-1}{RSS / n-I} \sim F_{I-1, n-I}$$

- $E[\bar{Y}_{..}] = \mu$ $\qquad \hat{\mu} := \bar{Y}_{..}$ is unbiased estimate for $\mu$

- $Var[\bar{Y}_{..}] \underset{Ex}{=} \frac{1}{n}\sigma^2 + \frac{\left(\sum_{i=1}^{I} n_i^2\right)}{n^2}\sigma_\alpha^2$

$\frac{SS_\alpha}{(I-1)n}$ is an estimate of $var(\bar{Y}_{..})$

- Confidence Interval $100(1-\alpha)\%$ for $\mu$ is

$$\left(\bar{Y}_{..} - t_{I-1}\left(1-\frac{\alpha}{2}\right)\sqrt{\frac{SS_A}{(I-1)n}} \;,\; \bar{Y}_{..} + t_{I-1}\left(1-\frac{\alpha}{2}\right)\sqrt{\frac{SS_A}{(I-1)n}}\right)$$

- $\hat{\sigma}^2 = \frac{RSS}{n-I}$

- $\widehat{n^2\sigma^2 + \sum_{i=1}^{I} n_i^2 \sigma_\alpha^2} = \frac{SS_\alpha}{(I-1)}$