Recap

- a paired t-test
- Bayesian inference : model selection
  Bootstrap inference : create samples by allowing for replacement

- ANOVA (Analysis of variance) ← alternative to t-test especially in the case of more than two levels in your modifications.

- $y_{ij} = \alpha + \beta_i + e_{ij}$

  we have g - groups of levels $i = 1, 2, \dots, g$
  and at each 'i' level we have n_i samples
  $j = 1, 2, \dots, n_i$

- $\beta_i \neq 0$ y_{ij} follows a multiple mean model
  (atleast one group level means are different)

  $\beta_i = 0 \Rightarrow$ all means are equal
  (a single global mean)

- Times

  $g=4$ $\begin{cases} B_1 & [\, y_{1,1}, y_{1,2}, \dots y_{1,n_1}] \to \text{\# samples in group} \\ B_2 & \text{''} \\ B_3 & \text{''} \\ B_4 & \text{''} \end{cases}$

  $\alpha \to$ mean of $\sum_{i=1}^{g} n_i$ samples $= \bar{y}_{\cdot,\cdot}$

  $\beta_i \to$ mean of $\sum_{j=1}^{n_i} (y_{i,j} - \bar{y}_{\cdot,\cdot})$

- Single mean model (vs) multiple means model

  $\mu_i = \alpha + \beta_i$ ← m.m.m
  $\mu = \alpha$ ← s.m.m

- all means equal is Null hypothes → Single mean
  atleast one μ_i is diff Alter. hyp → multiple mean

- $y_{ij} = \alpha + \beta_i + e_{ij}$

  $\mu_i + e_{ij} \sim N(\mu_i, \sigma^2)$
  ↓
  $N(0, \sigma^2)$

  $= \mu + e_{ij} \sim N(\mu, \sigma^2)$

- Occam's Razor - prefer simple model consistent w/ the data only move to complex when necessary

  (Bayesian model selection)

- $y_{i,j} = \alpha + \beta_i + e_{ij}$

  $r_{i,j} = y_{ij} - \bar{y}_{\cdot,\cdot}$ (single mean model)
  ↑
  global mean of the data

  $r_{i,j} = y_{ij} - \bar{y}_{i,\cdot}$ (multiple mean model)

- least squares approach : Sum of squares of residuals (SS)

- $r_{ij} = y_{ij} - \bar{y}_{\cdot,\cdot}$

  $= (\bar{y}_{i,\cdot} - \bar{y}_{\cdot,\cdot}) + (y_{ij} - \bar{y}_{i,\cdot})$

  $= \beta_i + e_{ij}$

  $$\sum_{i}^{g} \sum_{j}^{n_i} (y_{ij} - \bar{y}_{\cdot,\cdot})^2 = \sum_{i}^{g} \sum_{j}^{n_i} \beta_i^2 + \sum_{i}^{g} \sum_{j}^{n_i} e_{ij}^2 + \underbrace{2 \sum \sum \beta_i e_{ij}}_{0}$$

  $$\boxed{SS_{Total} = SS_{level} + SS_{error}}$$

  $SS_{level} = \sum_{j}^{g} n_i \beta_i^2$

- Total samples of 40 ($g = 4$, $n_i = 10$) → fixed SS_{Total}

  smaller SS_{level} → single mean model
  $\beta_i = \bar{y}_{i,\cdot} - y_{\cdot,\cdot}$
  ⇑
  (between groups) explained variance by the group/level

  (within group) e_{ij} → unexplained by the grouping

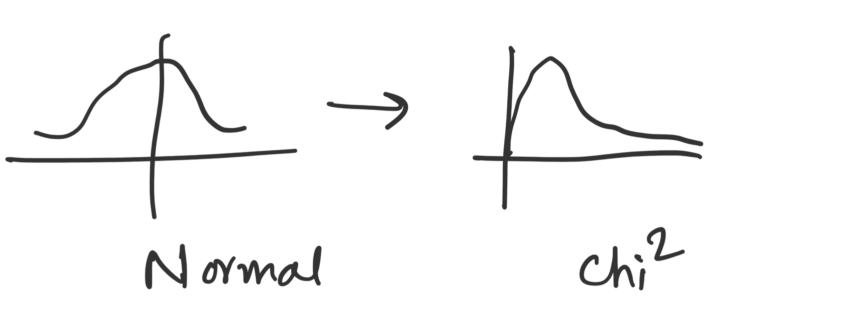
- $MS_{level} = \dfrac{SS_{level}}{g-1}$ $MS_{error} = \dfrac{SS_{error}}{N-g}$

- $f_0 = \dfrac{MS_{level}}{MS_{error}}$ $N = \sum_{i}^{g} n_i$

  $f_0 > 1$ multiplemean model

  $f_0 \approx 1$ variation in each group = var from random sampling
  (single mean model)

- F- distribution (



  $f = \dfrac{var(chi^2)}{var(chi^2)}$

  $Pr(f > f_0, f \approx 1) = p > 0.05$

  ⇑ How extreme the value f_0 is under the Null model
  ⇑ single mean model Null