ANOVA基本概念

- 四大檢定中，$H_0: \mu_1 = \mu_2 \Rightarrow$ 用 t 檢定

  統計量為 $t = \dfrac{\overline{x}_1 - \overline{x}_2}{S\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t_{(n_1+n_2-2)}$

  由此延伸至3個以上 populations

  $H_0: \mu_1 = \mu_2 = \mu_3 \Rightarrow$ 該用什麼檢定？

  Answer：用 F 檢定

  統計量 $F = \dfrac{\frac{SSB}{k-1}}{\frac{SSE}{n-k}} \sim F_{(k-1,\, n-k)}$

  組數(3)

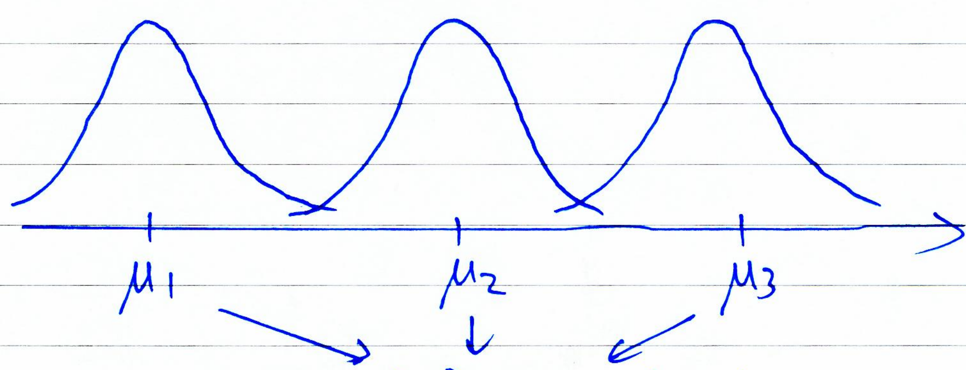
  Idea：

  比較"變異"以對"平均值"進行推論

  ANOVA (變異數分析)

- ANOVA

  $H_0: \mu_1 = \mu_2 = \mu_3$ (前提 $\sigma_1^2 = \sigma_2^2 = \sigma_3^2$)



Q：此3者是否足夠接近？

"變異" → 平方和 (SS = sum of squares)

① 總變異 (SST) T = total

$$SST = \sum_{i=1}^{k} \sum_{j=1}^{n_i} (x_{ij} - \underset{\text{全体總平均}}{\bar{x}})^2$$

② 組間變異 (SSB) B = between

$$SSB = \sum_{i=1}^{k} n_i (\underset{\text{組平均}}{\bar{x}_i} - \underset{\text{總平均}}{\bar{x}})^2$$

③ 組內變異 (SSE) E = error

$$SSE = \sum_{i=1}^{k} \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)^2$$

數學上可証明

$$\underset{(\text{總變異})}{SST} = \underset{(\text{組間變異})}{SSB} + \underset{(\text{組內變異})}{SSE}$$

ANOVA table

變異(SS)	df	均方(MS)	F
SSB (組間)	$k-1$	$MSB = \frac{SSB}{k-1}$	相除 (比較)
SSE (組內)	$n-k$	$MSE = \frac{SSE}{n-k}$	
SST (總)	$n-1$		

$$F = \frac{\overset{\text{組間}}{MSB}}{\underset{\text{組內}}{MSE}} \sim F_{(k-1,\,n-k)}$$

← F↑, 表組間差異相对組內差異較大
⇒ 傾向拒絕 H_0