One Factor Analysis of Variance

Objective: To assess if there is a difference among treatments.

Source of Variation: Among treatment + within treatment

$\uparrow$ Deterministic (among treatment) $\qquad$ $\uparrow$ random (within treatment)

Assumption: Treatments are indep

random errors are indep and normally distrib w/ eq σ^2.

Data Layout:

Treatment \ Replicate	1	2	3	- - -	n	Treatment Total	Treatment Average
1	Y_{11}	Y_{12}			Y_{1n}	$\sum_{1}^{n} Y_{1j} \equiv Y_{1\cdot}$	$\overline{Y}_{1\cdot} \equiv \frac{Y_{1\cdot}}{n_1}$
2	Y_{21}	Y_{22}			Y_{2n}	$\sum_j Y_{2j} \equiv Y_{2\cdot}$	$\overline{Y}_{2\cdot}$
3	$\vdots$	$\vdots$	$\ddots$		$\vdots$		
$\vdots$							
k	Y_{k1}	Y_{k2}			Y_{kn}	$\sum_j Y_{kj} \equiv Y_{k\cdot}$	$\overline{Y}_{k\cdot}$
						$\sum_{i=1}^{k} Y_{i\cdot} \equiv Y_{\cdot\cdot}$	$\overline{Y}_{\cdot\cdot} \equiv Y_{\cdot\cdot}/nk$

ANOVA:

Divide SS_T into $SS_{Tr} + SS_E$ and compute $F \triangleq \frac{MS_{TR}}{MSE}$

$$SST \triangleq (Y_{11}-\overline{Y}_{\cdot\cdot})^2+(Y_{12}-\overline{Y}_{\cdot\cdot})^2+\cdots(Y_{kn}-\overline{Y}_{\cdot\cdot})^2 = \sum_{i=1}^{k}\sum_{j=1}^{n}(Y_{ij}-\overline{Y}_{\cdot\cdot})^2$$

$$SS_{Tr} = (\overline{Y}_{1\cdot}-\overline{Y}_{\cdot\cdot})^2+(\overline{Y}_{2\cdot}-\overline{Y}_{\cdot\cdot})^2+\cdots(\overline{Y}_{k\cdot}-\overline{Y}_{\cdot\cdot})^2 \triangleq \sum_{i=1}^{k}(\overline{Y}_{i\cdot}-\overline{Y}_{\cdot\cdot})^2$$

$$SSE = SS_T - SS_{TR} \triangleq \sum_i\sum_j(Y_{ij}-\overline{Y}_{i\cdot})^2$$

Short Cut To SS Calculation	d.o.f.
• $SS_T \triangleq \sum_i \sum_j \left(Y_{ij} - \bar{Y}_{..}\right)^2 = \left(\sum_{i=1}^{k} \sum_{j=1}^{n} Y_{ij}^2\right) - \frac{Y_{..}^2}{kn}$	$(kn-1)$
• $SS_{tr} = \quad = \left(\frac{\sum_{i=1}^{k} Y_{i.}^2}{n}\right) - \frac{Y_{..}^2}{kn}$	$(k-1)$
• $SS_E = SS_T - SS_{tr}$	$k(n-1)$

$F \triangleq \frac{MS_{tr}}{MSE}$; $MS_{tR} = \frac{SS_{tR}}{d.o.f}$

mean Sum-Sq of treatment

$$\begin{cases} H_0: \tau_1 = \tau_2 = \tau_3 = \cdots \tau_k = 0 \\ H_1: \tau_i \neq 0 \text{ for at least one } i \end{cases}$$

; τ = treatment effect

<u>Confidence Interval for ith Treatment:</u>

$$\bar{Y}_{i.} - t_{\alpha/2,\,k(n-1)}\sqrt{\frac{MS_E}{n}} \le \mu_i \le \bar{Y}_{i.} + (\qquad)$$

<u>Confidence Interval on the Δ treatment means:</u>

$$\bar{Y}_{i.} - \bar{Y}_{j.} - t_{\alpha/2,\,k(n-1)}\sqrt{\frac{2MSE}{n}} \le (\mu_i - \mu_j) \le \bar{Y}_{i.} - \bar{Y}_{j.} + (\qquad)$$

Hypothesis Testing of Contrast of Treatment Means

$$\begin{cases} H_0: \sum_{i=1}^{k} c_i \mu_i = 0 \\ H_1: \sum c_i \mu_i \neq 0 \end{cases} \qquad \text{where } \sum_{i=1}^{k} c_i = 0$$

Form SS(Contrast) :

- $$SS_c \triangleq \frac{\left(\sum c_i Y_{i\cdot}\right)^2}{n \sum (c_i^2)} \qquad \text{d.o.f.} = 1$$

$$= \frac{\left(c_1 Y_{1\cdot} + c_2 Y_{2\cdot} + \cdots + c_k Y_{k\cdot}\right)^2}{n\left(c_1^2 + c_2^2 + \cdots + c_k^2\right)} \quad \text{or} \quad = \frac{\left(\sum c_i \bar{Y}_{i\cdot}\right)^2 \cdot n}{\left(\sum c_i^2\right)}$$

- SS_E

- $$f \triangleq \frac{SS_c/1}{MS_E} = \frac{n\left(\sum c_i \bar{Y}_{i\cdot}\right)^2}{\left(\frac{SSE}{k(n-1)}\right)\left(\sum c_i^2\right)}$$

$s^2 \uparrow$

Test if $f >$ or $<$ the critical f value w/ $1, k(n-1)$ d.o.f.

Orthogonal Contrast: To partition SS_T into indep Sum of Squares each with 1 d.o.f. i.e.

Example : $SS_{Tr} = SS_{w_1} + SS_{w_2} + \cdots SS_{w_{k-1}}$

$r \cdot LW_2$

Source of Vari.	SS	d.o.f.	MS	f	Reject H_0? f critical
SS_{Tr}	85,356	4 (5-1)	21,339	4.30	
$SS_{w_1 (1\ 1\ -1\ 0\ -1)}$	14,553	1	14,553	2.93	
$SS_{w_2 (1\ 1\ 1\ -4\ 1)}$	70,035	1	70,035	14.12	
SSE	124,021	25 = (5(6-1))	4,961		

Random Effect Model: of One Factor Experiment

$$SST = SS_{Tr} + SSE$$

$$f = \frac{MS_{Tr}}{MSE}$$

- Conclusion applies to the entire population
- Can also conclude about source of variation

$$\boxed{\hat{\sigma}_{Tr} = \frac{MS_{Tr} - MSE}{n}}$$

$$MS_E \xleftrightarrow{\text{estimates}} \underset{\underset{S^2}{\uparrow}}{\sigma^2} + n\,\sigma_{Tr}^2$$

$$Var(Y_{ij}) = \hat{\sigma}_{Tr}^2 + \sigma^2$$

$$\Rightarrow \quad \frac{\hat{\sigma}_{Tr}^2}{Var(Y_{ij})}$$ gives % degree of variation from the treatment itself.

Two-Factor Experiment

$$Y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \varepsilon_{ijk} \qquad \begin{array}{l} i = 1 \cdots a \\ j = 1 \cdots b \\ k = 1 \cdots n \end{array}$$

Factor A \ Factor B	1	2	···	b	Total	AV
1	$Y_{111}\ Y_{112} \cdots Y_{11n}$	$Y_{121}\ Y_{122} \cdots Y_{12n}$		$Y_{1b1}\ Y_{1b2} \cdots Y_{1bn}$	$Y_{1\cdot\cdot} = \sum_j^b \sum_{k=1}^n Y_{ijk}$	$Y_{1\cdot\cdot}/bn$
2	$Y_{211}\ Y_{212} \cdots Y_{21n}$	$Y_{221}\ Y_{222} \cdots Y_{22n}$		$Y_{2b1}\ Y_{2b2}\ Y_{2bn}$		
3	⋮	⋮		⋮		
⋮						
a	$Y_{a11}\ Y_{a12} \cdots Y_{a1n}$	$Y_{a21}\ Y_{a22} \cdots Y_{a2n}$	···	$Y_{ab1}\ Y_{ab2} \cdots Y_{abn}$	$Y_{a\cdot\cdot}$	$\overline{Y}_{a\cdot\cdot}$
	$Y_{\cdot 1\cdot} \triangleq \sum_{i=1}^a \sum_{k=1}^n Y_{ijk}$	$Y_{\cdot 2\cdot}$		$Y_{\cdot b\cdot}$	$Y_{\cdot\cdot\cdot} = \sum_i \sum_j \sum_k Y_{ijk}$	$\overline{Y}_{\cdot\cdot\cdot} \triangleq \dfrac{Y_{\cdot\cdot\cdot}}{nab}$

Fixed-Effect Model

1. $H_0: \tau_1 = \tau_2 = \cdots \tau_a = 0$ (no main effect for A factor)
   H_1: at least one $\tau_i \neq 0$
2. $H_0: \beta_1 = \beta_2 \cdots = \beta_b = 0$ (no main effect for B factor)
   H_1:
3. $H_0: (\tau\beta)_{11} = (\tau\beta)_{12} = \cdots = \beta_{ab} = 0$ (no interaction)
   H_1:

$$\begin{array}{cccccc} \text{d.o.f.} & a-1 & b-1 & (a-1)(b-1) & ab(n-1) \\ \hline SST = & SS_A + & SS_B + & SS_{AB} + & SS_E \end{array}$$

Short-Cuts for Calculation of SS for 2-factor Experiment

d.o.f.

- $SS_T = \left(\sum_{i}^{a}\sum_{j}^{b}\sum_{k}^{n} y_{ijk}^2\right) - \dfrac{y_{\cdots}^2}{abn}$ $\qquad abn$
- $SS_A = \dfrac{\sum_{i=1}^{a} y_{i\cdot\cdot}^2}{bn} - \dfrac{y_{\cdots}^2}{abn}$ $\qquad (a-1)$
- $SS_B = \dfrac{\sum_{j}^{b} y_{\cdot j\cdot}^2}{an} - \dfrac{y_{\cdots}^2}{abn}$ $\qquad (b-1)$
- $SS_{AB} = \dfrac{\sum_{i}^{a}\sum_{j}^{b} y_{ij\cdot}^2}{n} - \dfrac{y_{\cdots}^2}{abn} - SS_A - SS_B$ $\qquad (a-1)(b-1)$
- $SS_E = SS_T - SS_{AB} - SS_A - SS_B$ $\qquad ab(n-1)$

- $F_A: \dfrac{MS_A}{MSE}$
- $F_{AB} = \dfrac{MS_{AB}}{MSE}$
- $F_B = \dfrac{MS_B}{MSE}$

Example

$\alpha = 0.05$

	Source of Variation	SS	d.o.f.	MS	F	P-value	$F_{critical}$
a = 3	A	4.58	2	2.29	28.63	2.7 E-5	3.89
b = 2	B	4.91	1	4.91	61.38	5 E -7	4.75
n = 3	AB	0.24	2	0.12	1.50	0.2621	
	E	0.99	12	0.08		—	3.8
		10.72	17				

Conclusion: Factor A and B affect outcome; but no interaction effect