Least-square solution

$$\hat{\beta} = (X^T X)^{-1} X^T Y.$$

$$X^T X = \begin{pmatrix} n_1 & & & 0 \\ & n_2 & & \\ & & \ddots & \\ 0 & & & n_r \end{pmatrix}$$

$$\hat{\beta} = \begin{pmatrix} \bar{y}_{1\cdot} \\ \vdots \\ \bar{y}_{r\cdot} \end{pmatrix} = \begin{pmatrix} \frac{1}{n_1} \sum_{j=1}^{n_1} y_{1j} \\ \vdots \\ \frac{1}{n_r} \sum_{j=1}^{n_r} y_{rj} \end{pmatrix}$$

Denote $y_{i\cdot} = \sum_{j=1}^{n_i} y_{ij}$

$$y_{\cdot\cdot} = \sum_{i=1}^{r} y_{i\cdot} = \sum_i \sum_j y_{ij}$$

$$\bar{y}_{i\cdot} = \frac{1}{n_i} y_{i\cdot}$$

$$\bar{y}_{\cdot\cdot} = \frac{1}{n_T} y_{\cdot\cdot}$$

Least squares: $\underset{\{\mu_i\}}{\text{minimize}} \sum_i \sum_j (y_{ij} - \mu_i)^2$

$\Leftrightarrow$ minimize $\sum_j (y_{ij} - \mu_i)^2$ for each $i = 1, \dots r$ seperately

$$\Rightarrow \hat{\mu}_i = \bar{y}_{i\cdot}$$

$$\hat{y}_{ij} = \bar{y}_{i\cdot}$$

MLE : Maximize $\mathcal{L}(\mu_1, \cdots \mu_r, \sigma^2) \Rightarrow$ same result.

Residuals : $e_{ij} = y_{ij} - \hat{y}_{ij} = y_{ij} - \bar{y}_{i\cdot}$

Analysis of variance

$$SSTO = \text{total SS} = \sum_i \sum_j (y_{ij} - \bar{y}_{\cdot\cdot})^2$$

$$SSTR = \text{treatment SS} = \sum_i \sum_{j=1}^{n_i} (\bar{y}_{i\cdot} - \bar{y}_{\cdot\cdot})^2$$

(SSB)
↓
between

$$= \sum_i n_i (\bar{y}_{i\cdot} - \bar{y}_{\cdot\cdot})^2$$

$$SSE = \text{Error SS} = \sum_i \sum_j (y_{ij} - \bar{y}_{i\cdot})^2$$

(SSW)
↓
within

$$SSTO = SSTR + SSE$$

Degree of freedom

df. TO $= n_T - 1$, because there are n_T deviations $(y_{ij} - \bar{y}_{..})$, but $\sum_i \sum_j (y_{ij} - \bar{y}_{..}) = 0$

df. Error $= n_T - r$, because each $\sum_j (y_{ij} - \bar{y}_{i.})$ has $(n_i - 1)$ d.f. $\Rightarrow \sum_i (n_i - 1) = n_T - r$

df. Treatment $= r - 1$, because there are r estimated treatment deviation $(\bar{y}_{i.} - \bar{y}_{..})$, but $\sum_i n_i (\bar{y}_{i.} - \bar{y}_{..}) = 0$

Mean squares: $MS = \dfrac{SS}{df}$

Under assumption $\varepsilon_{ij} \sim \text{Normal}(0, \sigma^2)$, iid

$$E(MSE) = E\{\sum_i \sum_j (y_{ij} - \bar{y}_{i.})^2\} / (n_T - r)$$

$$= E\{\sum_i (n_i - 1) S_i^2\} / n_T - r$$

$$S_i^2 = \frac{1}{n_i - 1} \sum_j (y_{ij} - \bar{y}_{i.})^2$$

$$= \sum_i (n_i - 1)\, E\, S_i^2 \Big/ n_T - r$$

$$= \frac{\sum_i (n_i - 1)\, \sigma^2}{n_T - r} = \sigma^2$$

$$E(MSTR) = \frac{E \sum_i n_i (\bar{y}_{i\cdot} - \bar{y}_{\cdot\cdot})^2}{r-1}$$

where
$$\bar{y}_{i\cdot} = \mu_i + \bar{\varepsilon}_{i\cdot}$$
$$\bar{y}_{\cdot\cdot} = \bar{\mu}_\cdot + \bar{\varepsilon}_{\cdot\cdot}$$
$$\bar{\mu}_\cdot = \frac{\sum_{i=1}^{r} n_i \mu_i}{n_T}$$

$$= \frac{1}{r-1} E\left\{ \sum_i n_i (\mu_i + \bar{\varepsilon}_{i\cdot} - \bar{\mu}_\cdot - \bar{\varepsilon}_{\cdot\cdot})^2 \right\}$$

$$= \frac{1}{r-1}\Big\{ E \sum_i n_i (\mu_i - \bar{\mu}_\cdot)^2 + E \sum_i n_i (\bar{\varepsilon}_{i\cdot} - \bar{\varepsilon}_{\cdot\cdot})^2$$
$$+ \underbrace{2 E \sum_i n_i (\mu_i - \bar{\mu}_\cdot)(\bar{\varepsilon}_{i\cdot} - \bar{\varepsilon}_{\cdot\cdot})}_{=0} \Big\}$$

$$= \frac{1}{r-1} \sum_i n_i (\mu_i - \bar{\mu}_\cdot)^2 + \frac{1}{r-1} \sum_i n_i E(\bar{\varepsilon}_{i\cdot}^2 - 2\bar{\varepsilon}_{i\cdot}\bar{\varepsilon}_{\cdot\cdot} + \bar{\varepsilon}_{\cdot\cdot}^2)$$

$$= \qquad = \qquad + \frac{1}{r-1} \sum_i n_i \Big\{ Var(\bar{\varepsilon}_{i\cdot}) + Var(\bar{\varepsilon}_{\cdot\cdot})$$
$$- 2 E\Big(\bar{\varepsilon}_{i\cdot} \sum_{i=1}^{r} n_i \bar{\varepsilon}_{i\cdot} \Big/ n_T\Big) \Big\}$$

28

$$= \quad = \quad + \frac{1}{r-1}\sum_{i=1}^{r} n_i \left\{ \frac{\sigma^2}{n_i} + \frac{\sigma^2}{n_T} - 2\frac{n_i}{n_T}\frac{\sigma^2}{n_i} \right\}$$

$$E(\bar{\varepsilon}_{1\cdot}\, \bar{\varepsilon}_{2\cdot}) = E\bar{\varepsilon}_{1\cdot}\, E\bar{\varepsilon}_{2\cdot} = 0$$

$$= \quad = \quad + \frac{\sigma^2}{r-1}\sum_{i=1}^{r} n_i \left\{ \frac{1}{n_i} - \frac{1}{n_T} \right\}$$

$$= \quad = \quad + \frac{\sigma^2}{r-1}(r-1)$$

$$= \frac{1}{r-1}\sum_i n_i (\mu_i - \bar{\mu}_\cdot)^2 + \sigma^2$$

So $E(MSTR) = \sigma^2 + \dfrac{\sum n_i (\mu_i - \bar{\mu}_\cdot)^2}{r-1}$

$E(MSE) = \sigma^2$

ANOVA F-test.

$H_0: \mu_1 = \mu_2 = \cdots \mu_r$ $\quad H_A:$ H_0 is not true

Under H_0, $E(MSTR) = E(MSE)$

$$F = \frac{MSTR}{MSE} = \frac{MSTR/\sigma^2}{MSE/\sigma^2}$$

Under H_0, F is a ratio of two independent χ^2 variables.

- Cochran's thm.

If $y_1, \dots y_n \sim$ iid $N(\mu, \sigma^2)$

and $SSTO = \sum_{i=1}^{r} SS_i$ is decomposed into r sum of squares each with df_i degree of freedom

then $\frac{SS_i}{\sigma^2}$ are independent χ^2 r.v.s

if $\sum_{i=1}^{r} df_i = n-1$

$$SSTO = SSTR + SSE$$

$$n_T - 1 = r-1 + n_T - r$$

$\Rightarrow$ under H_0, $F = \dfrac{MSTR}{MSE} \sim F_{r-1,\, n_T - r}$.

always a right-tailed test.

under H_A: $F \sim$ noncentral F-distr.

SSTR and SSE are still indep.

Remark: extra SS principle.

Full model: $y_{ij} = \mu_i + \varepsilon_{ij}$

Reduced model: $y_{ij} = \mu + \varepsilon_{ij}$

$$SSE(F) = \sum_i \sum_j (y_{ij} - \bar{y}_{i\cdot})^2$$

$$SSE(R) = \sum_i \sum_j (y_{ij} - \bar{y}_{\cdot\cdot})^2$$

$$F = \frac{(SSE(R) - SSE(F)) / (df_R - df_F)}{SSE(F) / df_F}$$

$$= \frac{MSTR}{MSE} \Leftrightarrow \text{ANOVA } F \text{ test}$$

Factor - effects model.

Instead of $y_{ij} = \mu_i + \varepsilon_{ij}$

We can write $y_{ij} = \mu + \tau_i + \varepsilon_{ij}$

$\tau_i = \mu_i - \mu$

where μ is defined as

(1) $\mu = \bar{\mu}_{\cdot} = \frac{\sum \mu_i}{r}$, then $\sum_{i=1}^{r} \tau_i = 0$

(2) $\mu = \sum_{i=1}^{r} w_i \mu_i$, $\sum w_i = 1$

$\downarrow$

weighted mean

then $\sum_i w_i \tau_i = 0$

Test of $H_0 : \mu_1 = \mu_2 = \cdots \mu_r$

$\Leftrightarrow H_0 : \tau_1 = \tau_2 = \cdots \tau_r$

Regression approach to ANOVA.

$y_{ij} = \mu + \tau_i + \varepsilon_{ij}$

$\sum_{i=1}^{r} \tau_i = 0 \Rightarrow \tau_r = -\tau_1 - \cdots \tau_{r-1}$

(32)

So $\beta = \begin{pmatrix} \mu \\ \tau_1 \\ \vdots \\ \tau_{r-1} \end{pmatrix}$ $\quad X = \begin{array}{c} \begin{pmatrix} 1 & 1 & & \\ 1 & 1 & & \\ \vdots & \vdots & 1 & \\ \vdots & \vdots & 1 & 1 \\ 1 & -1 & -1 & -1 \\ 1 & -1 & -1 & -1 \end{pmatrix} \end{array} \begin{array}{l} \to \text{treatment } 1 \\ \to \text{treatment } 2 \\ \\ \to \text{treatment } r \end{array}$

$$y_{ij} = \mu + \tau_1 1_{i=1} + \cdots \tau_{r-1} 1_{i=r-1} - (\tau_1 + \cdots \tau_{r-1}) 1_{i=r} + \varepsilon_{ij}$$

or $$y_{ij} = \mu + \tau_1 X_{ij1} + \cdots + \tau_{r-1} X_{ijr-1} + \varepsilon_{ij}$$

$$X_{ijk} = \begin{cases} 1 & \text{if treatment } k \\ 0 & \text{if } \cancel{\text{treatment}} \text{ not } k \quad 1 \le k \le r-1 \\ -1 & \text{if treatment } r \end{cases}$$

· Randomization Tests

model : $y_{ij} = \mu + \tau_i + \varepsilon_{ij}$

where ε_{ij} are fixed, associated with experimental unit (i, j)

Treatment are randomized (assigned at random)

τ_i random.

all assignment are equally likely.

EX: Medicine A

5 , 7

Medicine B

3 , 4

$$F = \frac{MSTR}{MSE}$$

5	7	3	4	prob
A	A	B	B	
5	7	3	4	$\frac{1}{6}$
5	4	3	7	$\frac{1}{6}$
5	3	4	7	$\frac{1}{6}$
3	7	4	5	$\frac{1}{6}$
3	4	5	7	$\frac{1}{6}$
4	7	5	3	$\frac{1}{6}$

$\binom{4}{2} = 6$

$$F = \frac{MSTR}{MSE} \qquad \bar{Y}_{\cdot\cdot} = \frac{19}{4}$$

$$\frac{2\cdot\left[\left(\frac{5+7}{2} - \frac{19}{4}\right)^2 + \left(\frac{3+4}{2} - \frac{19}{4}\right)^2\right]/2}{\left[\left(5 - \frac{5+7}{2}\right)^2 + (7-6)^2 + (3-3.5)^2 + (4-3.5)^2\right]/2} = F_{obs}$$

$F = ?$

$\vdots$

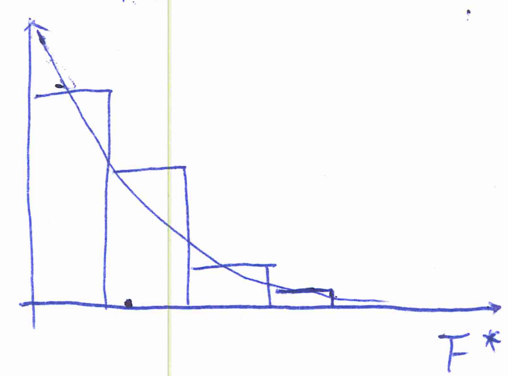
P-value $= P(F \geq F_{obs}) = \frac{1}{6}$

obs.	1	2	3	4	5	6
	25	20	24	23	22	21

$$\binom{6}{3} = 20$$

permutation						Prob.	F
1_{25}	2_{20}	3_{24}	4_{23}	5_{22}	6_{21}	$\frac{1}{20}$	$F^* = F_{obs}$
1_{25}	2_{20}	4_{23}	3	5	6	$\frac{1}{20}$	F^*
1	2	5	3	4	6	⋮	⋮
1	2	6	3	4	5	⋮	

$$\text{P-value} = \mathbb{P}(F \geq F_{obs}) = \frac{\#\{\text{permutations} \mid F^* \geq F_{obs}\}}{20}$$



Choise of sample size

· power approach

$$\text{power} = P(\text{Reject } H_0 \mid H_A)$$

$$= P(F^* > F_{1-\alpha, r-1, n_T - r} ; \varphi)$$

φ → noncentrality parameter.

where $\varphi = \frac{1}{\sigma}\sqrt{\frac{\sum_{i=1}^{r} n_i (\mu_i - \mu)^2}{r}}$

how to choose n

$\hookrightarrow$ sample size of each treatment.

Δ = range = $\max(\mu_i) - \min(\mu_i)$

based on desired Δ, α, β, σ (or $\frac{\Delta}{\sigma}$).

α → significance, β → type II

Find n from Table B12

(it gives sample size for each group)

In R.

pwr.anova.test(f = , r , n , sig.level , power)

$\downarrow$ effect size

$$= \frac{1}{\sigma}\sqrt{\sum_{i=1}^{r} \frac{n_i}{n_T}(\mu_i - \mu)^2}$$

ex: pwr.anova.test(f = .3, r = 3, n = 10, sig.level = .05)

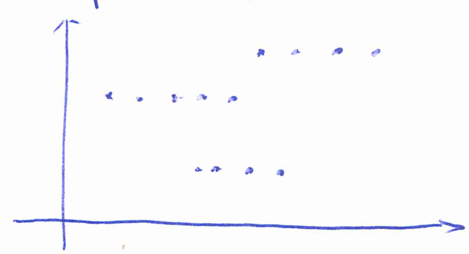
power =

pwr.anova.test(f = .3, r = 3, power = .8, sig.level = .05)

n = ?

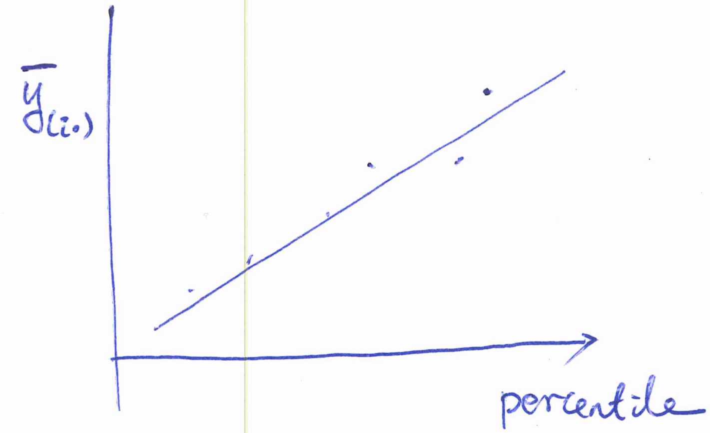
Analysis of Factor level effect
(treatment effects)

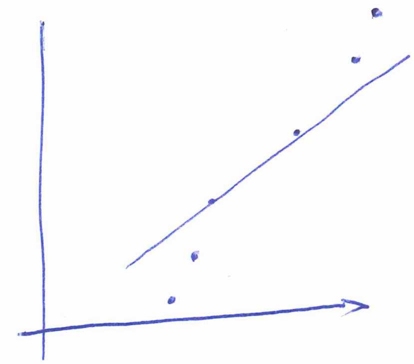
(1) plots. (.) line plot



(..) Normal prob. plot (only if $n_1 = n_2 = \cdots n$)

if $\cancel{\bar{y}_{(i\cdot)}}$ $\mu_1 = \mu_2 = \cdots \mu_r$ if not





if $\mu_1 = \mu_2 = \ldots \mu_r = \mu$

for $\bar{y}_{(1\cdot)}, \ldots \bar{y}_{(r\cdot)}$

$$p(t_1, \ldots t_r) = \frac{r!}{\sqrt{\sigma}^r (2\pi)^{r/2}} e^{-\frac{\sum_{i=1}^{r}(t_i - \mu)^2}{2\sigma^2}} \quad \text{if } t_1 < t_2 < \ldots t_r$$

$$E(\bar{y}_{(i\cdot)}) \approx \mu + \sigma \Phi^{-1}\left(\frac{i - .375}{r + .25}\right)$$

$$\hat{E}(\bar{y}_{(i\cdot)}) = \bar{Y}_{\cdot\cdot} + \hat{\sigma} \Phi^{-1}\left(\frac{i - .375}{r + .25}\right)$$

- Inference for sigle factor level mean.
  - Estimation and testing.

$$\hat{\mu}_i = \bar{y}_{i\cdot} , \quad \bar{y}_{i\cdot} \sim Normal(\mu_i , \frac{\sigma^2}{n_i})$$

$$S^2(\bar{y}_{i\cdot}) = \frac{MSE}{n_i}$$

$$\frac{\bar{y}_{i\cdot} - \mu_i}{S(\bar{y}_{i\cdot})} \sim t_{n_T - r}$$

C.I. for μ_i $\quad \bar{y}_{i\cdot} \pm t_{1-\frac{\alpha}{2}, n_T - r} S(\bar{y}_{i\cdot})$

Test: $H_0 : \mu_i = c$, by $t = \dfrac{\bar{y}_{i\cdot} - c}{S(\bar{y}_{i\cdot})} \sim t_{n_T - r}$

- Inference for difference between 2 factor level means

$D = \mu_i - \mu_{i'}$, $i, i' = 1, \cdots r$

  - Estimation: $\hat{D} = \bar{y}_{i\cdot} - \bar{y}_{i'\cdot} \sim Normal(\mu_i - \mu_{i'} , \sigma^2(\frac{1}{n_i} + \frac{1}{n_{i'}}))$

$$S^2(\hat{D}) = MSE(\frac{1}{n_i} + \frac{1}{n_{i'}})$$

$H_0 : \mu_i = \mu_{i'}$, $t = \dfrac{\hat{D}}{S(\hat{D})} \sim t_{n_T - r}$