REPLICATED DESIGN

Determination of standard deviation of the average effect $\overline{E}_j$:

Notation: * $Var(\overline{E}_j)$ = Variance of the average effect $\overline{E}_j$.

* σ_i^2 = Variance of the outcome of exper. i $= Var(Y_i)$

* a_{ij} = level (+1 or −1) of factor j in experiment i.

$$Var(\overline{E}_j) = Var\left(\sum_{i=1}^{N} \frac{a_{ij} Y_i}{N/2}\right) = \frac{4}{N^2} Var\left(\sum_{i=1}^{N} a_{ij} Y_i\right)$$

$$= \frac{4}{N^2} \sum_{i=1}^{N} Var(Y_i) = \frac{4}{N^2} \sum_{i=1}^{N} \sigma_i^2$$

$$\Rightarrow \sigma_{\overline{E}_j} = \sqrt{Var(\overline{E}_j)} = \frac{2}{N} \sqrt{\sum_{i=1}^{N} \sigma_i^2}$$

If variance across all experiments is the same, i.e., $\sigma_i = \sigma$ for all $i = 1, \dots, N = 2^k$:

$$\Rightarrow \sigma_{\overline{E}_j} = \frac{2}{N} \sqrt{\sum_{i=1}^{N} \sigma_i^2} = \frac{2}{N} \sqrt{N \sigma^2} = \frac{2}{\sqrt{N}} \cdot \sigma$$

Using average results $\bar{y}_i$ for experiment $i = 1, \dots, N \ (= 2^k)$, to estimate $\sigma_{\bar{E}_j}$:

Consider n_i replications of experiment i, $i = 1, \dots, N$.

Experiment	Result	$\bar{y}_i = \frac{\sum_{j=1}^{n_i} y_{ij}}{n_i}$	$s_i^2 = \frac{\sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2}{n_i - 1}$
1	$y_{11}, y_{12}, \dots, y_{1n_1}$	$\bar{y}_1$	s_1^2
2	$y_{21}, y_{22}, \dots, y_{2n_2}$	$\bar{y}_2$	s_2^2
$\vdots$	$\vdots \quad \ddots \quad \vdots$	$\vdots$	$\vdots$
N	$y_{N1}, y_{N2}, \dots, y_{Nn_2}$	$\bar{y}_N$	s_N^2

We obtain:

$$\mathrm{Var}(\bar{E}_j) = \mathrm{Var}\left(\sum_{i=1}^{N} \frac{a_{ij} \bar{Y}_i}{N/2} \right) = \dots = \frac{4}{N^2} \sum_{i=1}^{N} \mathrm{Var}(\bar{Y}_i)$$

$$= \frac{4}{N^2} \sum_{i=1}^{N} \sigma_{\bar{Y}_i}^2 \qquad \text{where} \quad \sigma_{\bar{Y}_i}^2 = \frac{\sigma_i^2}{n_i}$$

$$\Rightarrow \sigma_{\bar{E}_j} = \sqrt{\frac{4}{N^2} \sum_{i=1}^{N} \sigma_{\bar{Y}_i}^2} = \frac{2}{N} \sqrt{\sum_{i=1}^{N} \sigma_i^2 / n_i}$$

* If $\sigma_i = \sigma$ and $n_i = n \Rightarrow \sigma_{\bar{Y}_i}^2 = \sigma_{\bar{Y}}^2 \quad \forall i = 1, 2, \dots, N.$

$$\Rightarrow \boxed{\sigma_{\bar{E}_j} = \frac{2}{N} \sqrt{\frac{N \sigma^2}{n}} = \frac{2\sigma}{\sqrt{N \cdot n}} \quad \left(= \frac{2\sigma}{\sqrt{2^k n}} \right)}$$

Estimation of σ (in $\sigma_{\bar{E}_j} = \frac{2\sigma}{\sqrt{2^k n}}$):

Pooled variance:

$$S^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2 + \cdots + (n_N-1)S_N^2}{(n_1-1)+(n_2-1)+\cdots+(n_N-1)}$$

$= \{$ Assume that $n_i = n$ for all i $\}$

$$= \frac{\sum_{i=1}^{2^k}\sum_{j=1}^{n}(y_{ij}-\bar{y}_i)^2}{2^k(n-1)}.$$

$$\left(\text{or: } \frac{\sum_{i=1}^{N} s_i^2}{N} \quad \text{where } N = 2^k. \right)$$

Confidence intervals on the effects:

Note that: $T_0 = \dfrac{\overline{E}_j}{\sigma_{\overline{E}_j}} \in t\big(\underbrace{2^k(n-1)}_{}\big)$

$= \text{degrees of freedom}$

$= \underbrace{(n-1)+(n-1)+\dots+(n-1)}_{\#\, 2^k}$

$= 2^k(n-1)$

$= \text{total nbr of runs} - \text{nbr of model parameters}$

$= 2^k n - 2^k$

This gives confidence intervals:

$$\boxed{\overline{E}_j \pm t_{\alpha/2,\, 2^k(n-1)} \cdot \sigma_{\overline{E}_j}}$$

So, if zero is included in the interval, we cannot reject H_0. (H_0: The effect $\overline{E}_j$ is not significant, i.e., $\mu_j = 0$)