The Design and Analysis of Experiments

Completely Randomised Design (CRD)

Aim: To find if there is a significant difference between the population means of c "treatments" ($c \geq 3$).

Data Pattern: n observations are taken on each of c treatments and the resulting data can be tabulated as follows:

	Population mean	Observations	Total	Sample mean
Treatment 1	μ_1	$x_{11},\ x_{12},\ \cdots,\ x_{1n}$	T_1	$\bar{x}_1$
Treatment 2	μ_2	$x_{21},\ x_{22},\ \cdots,\ x_{2n}$	T_2	$\bar{x}_2$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
Treatment c	μ_c	$x_{c1},\ x_{c2},\ \cdots,\ x_{cn}$	T_c	$\bar{x}_c$

where

$$T_i = \sum_{j=1}^{n} x_{ij}, \quad \bar{x}_i = \frac{T_i}{n}.$$

Analysis of Variance (ANOVA): Compute

$$\begin{aligned}
T &= \sum_{j=1}^{c} T_j, \\
SS_{total} &= \sum_{i=1}^{c}\sum_{j=1}^{n} x_{ij}^2 - \frac{T^2}{cn}, \\
SS_{treat} &= \frac{1}{n}\sum_{i=1}^{c} T_i^2 - \frac{T^2}{cn}, \\
SS_{error} &= SS_{total} - SS_{treat}, \\
MS_{treat} &= \frac{SS_{treat}}{c-1}, \\
MS_{error} &= \frac{SS_{error}}{c(n-1)}.
\end{aligned}$$

Arrange them as the one-way ANOVA table:

Source of variation	Sum of squares	d.f.	Mean square	F-ratio
Treatments	SS_{treat}	$c-1$	MS_{treat}	$\frac{MS_{treat}}{MS_{error}}$
Error	SS_{error}	$c(n-1)$	MS_{error}	
Total variation	SS_{total}	$cn-1$		

Hypothesis Test:

H_0: $\mu_1 = \mu_2 = \cdots = \mu_c$ (i.e. treatment means are the same).
H_1: μ_i differ (i.e. treatment means are different).

Given the significance level α, find $F_{\alpha,c-1,c(n-1)}$ from the F-distribution table with $\nu_1 = c-1$ and $\nu_2 = c(n-1)$. If

$$F = \frac{MS_{treat}}{MS_{error}} > F_{\alpha,c-1,c(n-1)}$$

reject H_0 and conclude that the treatment means are significantly different at α level of significance; otherwise do not reject H_0 and conclude that the treatment means are not significantly different at α level of significance.

Confidence Intervals: Compute

$$s = \sqrt{MS_{error}}.$$

The $(1-\alpha)100\%$ confidence interval for μ_i, the population mean of treatment i is given by

$$\bar{x}_i \pm t_{\alpha/2,c(n-1)} s\sqrt{\frac{1}{n}}.$$

The $(1-\alpha)100\%$ confidence interval for $\mu_i - \mu_j$, the difference between the population means of treatments i and j is given by

$$\bar{x}_i - \bar{x}_j \pm t_{\alpha/2,c(n-1)} s\sqrt{\frac{2}{n}}.$$

Randomised Block Design (RBD)

Aim: To find (i) if there is a significant difference between the means of c "treatments", and (ii) if there is a significant difference between the means of r "blocks".

Data Pattern:

	Treat 1	$\cdots$	Treat c	Block total	Block mean
Block 1	x_{11}	$\cdots$	x_{1c}	B_1	$\bar{B}_1$
Block 2	x_{21}	$\cdots$	x_{2c}	B_2	$\bar{B}_2$
$\vdots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$
Block r	x_{r1}	$\cdots$	x_{rc}	B_r	$\bar{B}_r$
Treat total	T_1	$\cdots$	T_c		
Treat mean	$\bar{T}_1$	$\cdots$	$\bar{T}_c$		

where

$$B_i = \sum_{j=1}^{c} x_{ij}, \quad \bar{B}_i = \frac{B_i}{c},$$

$$T_j = \sum_{i=1}^{r} x_{ij}, \quad \bar{x}_j = \frac{T_j}{r}.$$

Analysis of Variance (ANOVA): Compute

$$
\begin{aligned}
G &= \sum_{i=1}^{c} T_i, \quad \bar{G} = \frac{G}{rc}, \\
SS_{total} &= \sum_{i=1}^{r}\sum_{j=1}^{c} x_{ij}^2 - \frac{G^2}{rc}, \\
SS_{treat} &= \frac{1}{r}\sum_{j=1}^{c} T_j^2 - \frac{G^2}{rc}, \\
SS_{block} &= \frac{1}{c}\sum_{i=1}^{r} B_i^2 - \frac{G^2}{rc}, \\
SS_{error} &= SS_{total} - SS_{treat} - SS_{block}, \\
MS_{treat} &= \frac{SS_{treat}}{c-1}, \\
MS_{block} &= \frac{SS_{block}}{r-1}, \\
MS_{error} &= \frac{SS_{error}}{(r-1)(c-1)}.
\end{aligned}
$$

Arrange them as the ANOVA table:

Source of variation	Sum of squares	d.f.	Mean square	F-ratio
Treatments	SS_{treat}	$c-1$	MS_{treat}	$\frac{MS_{treat}}{MS_{error}}$
Blocks	SS_{block}	$r-1$	MS_{block}	$\frac{MS_{block}}{MS_{error}}$
Error	SS_{error}	$(r-1)(c-1)$	MS_{error}	
Total variation	SS_{total}	$rc-1$		

Hypothesis Tests:

H_0: treatment means are the same.

H_1: treatment means are different.

Given the significance level α, find $F_{\alpha,c-1,(r-1)(c-1)}$ from the F-distribution table with $\nu_1 = c-1$ and $\nu_2 = (r-1)(c-1)$. If

$$
F = \frac{MS_{treat}}{MS_{error}} > F_{\alpha,c-1,(r-1)(c-1)}
$$

reject H_0 and conclude that the treatment means are significantly different at α level of significance; otherwise do not reject H_0 and conclude that the treatment means are not significantly different at α level of significance.

H_0: block means are the same.

H_1: block means are different.

Given the significance level α, find $F_{\alpha,r-1,(r-1)(c-1)}$ from the F-distribution table with $\nu_1 = r-1$ and $\nu_2 = (r-1)(c-1)$. If

$$F = \frac{MS_{block}}{MS_{error}} > F_{\alpha,r-1,(r-1)(c-1)}$$

reject H_0 and conclude that the block means are significantly different at α level of significance; otherwise do not reject H_0 and conclude that the block means are not significantly different at α level of significance.

Confidence Intervals: Compute

$$s = \sqrt{MS_{error}}.$$

The $(1-\alpha)100\%$ confidence interval for μ_j, the population mean of treatment j, is given by

$$\bar{T}_j \pm t_{\alpha/2,(r-1)(c-1)} s \sqrt{\frac{1}{r}}.$$

The $(1-\alpha)100\%$ confidence interval for $\mu_j - \mu_k$, the difference between the population means of treatments j and k, is given by

$$\bar{T}_j - \bar{T}_k \pm t_{\alpha/2,(r-1)(c-1)} s \sqrt{\frac{2}{r}}.$$