Experimental Designs

Basic Concept/Model, Part II

Yu-Ling Tseng
Dept. of Applied Math, NDHU

Basic/ideal design CRD

Completely Randomized Design

All the E.U.'s are allocated at random among all the treatments.

$$\begin{array}{ccccccc}
 & & \text{Group 1 } (n_1) & \longrightarrow & \text{Trt 1} & & \\
\text{Random} & \nearrow & & & & \searrow & \text{Compare} \\
(\text{e.u.s}) & \ldots & \ldots & \ldots & & \ldots & \\
\text{assignment} & \searrow & & & & \nearrow & \text{responses} \\
 & & \text{Group k } (n_k) & \longrightarrow & \text{Trt k} & &
\end{array}$$

Hence, groups of e.u.s that should be similar in all respects before we apply the trts.

The statistical model: $j = 1, \ldots k;\ i = 1, \ldots, n_j$

$$Y_{ij} = \mu + \tau_j + \epsilon_{ij}; \qquad (1)$$

- ✓ Y_{ij} : the ith obs on the jth trt
- ✓ μ : a common effect for the whole experiment

- τ_j : the effect of the jth trt
- ϵ_{ij} : the random error (experimental errors.)

  w/ $\epsilon_{ij} \overset{iid}{\sim} N(0, \sigma_\epsilon^2)$

Note
$\mu_j = \mu + \tau_j$:
the true mean of the response at the jth trt.

Objectives of most experiments:

Are there differences among trts' effects?
$\longrightarrow$ Testing
How large are the differences?
$\longrightarrow$ (pt, interval) est.

Let $\theta = \sum_{j=1}^{k} \tau_j^2$: overall measure of the trt effects.

Of interest: How big is θ related to σ_ϵ^2?

1 factor experiment

The jth trt = The jth level of the factor

- √ The fixed (effect) model
  Levels of the factor are set at specific values
  $\longleftarrow$ only interested in these levels
  $\longrightarrow$ where the inference can make . . .
- √ The random (effect) model
  Levels of the factor are randomly selected
  $\longleftarrow$ interest focus on the response variability caused by the differences among the levels of the factor
  $\longrightarrow$ wider inference

$*$ The fixed (effect) model

e.g. Meat storage experiment data

Factor: packaging method at 4 levels (cpw, vp, coo2n, co2)

$**$ Add $\tau_1, \ldots, \tau_k$ are assumed to be fixed parameters with $\sum \tau_j = 0$ (or $\tau_1 = 0$: baseline) to model (1). $**$

One factor fixed effect model:

$$Y_{ij} = \mu + \tau_j + \epsilon_{ij}; \qquad (2)$$

w/ $\epsilon_{ij} \overset{iid}{\sim} N(0, \sigma_\epsilon^2)$, $\tau_1, \ldots, \tau_k$ are fixed parameters and $\sum \tau_j = 0$ (or $\tau_1 = 0$: baseline)

Objectives:

- √ Test: $H_0 : \tau_1 = \tau_2 = \ldots = \tau_k = 0$
- √ pt est's/conf. intervals for μ, τ_j's, σ_ϵ^2 or ft's of these.

$*$ The random (effect) model

Levels of the factor are randomly selected

e.g. Casting of high temp. alloys

e.g. operators, days, batches of material, . . .

$**$ Add $\tau_1, \ldots, \tau_k$ are assumed to be r.v.'s with $\tau_j \overset{iid}{\sim} N(0, \sigma_\tau^2)$ to model (1). $**$

One factor random effect model:

$$Y_{ij} = \mu + \tau_j + \epsilon_{ij}; \tag{3}$$

w/ $\epsilon_{ij} \overset{iid}{\sim} N(0, \sigma_\epsilon^2) \perp \tau_j \overset{iid}{\sim} N(0, \sigma_\tau^2)$

Objectives:

- √ Test: $H_0 : \sigma_\tau^2 = 0$ ($\Longleftrightarrow$ equal trts' means)
- √ pt est's/conf. intervals for σ_τ^2 or $\sigma_\tau^2 / \sigma_\epsilon^2, \ldots$

Analysis

One-way ANOVA

Model (1): $j = 1, \ldots k;\ i = 1, \ldots, n_j$

$$Y_{ij} = \mu + \tau_j + \epsilon_{ij};$$

w/ $\epsilon_{ij} \overset{iid}{\sim} N(0, \sigma_\epsilon^2)$

τ_j's : fixed or random.

Note:

$$
\begin{aligned}
& Y_{ij} = \mu + \tau_j + \epsilon_{ij} \\
\text{c.f.} \quad & Y_{ij} = Y_{ij} \\
\Longleftrightarrow \quad & Y_{ij} = \mu + (\mu_j - \mu) + (Y_{ij} - \mu_j) \\
\Longleftrightarrow \quad & Y_{ij} - \mu = (\mu_j - \mu) + (Y_{ij} - \mu_j) \qquad (4)
\end{aligned}
$$

$\mu \hookrightarrow \bar{Y}_{..}$: sample mean of $N = \sum_j n_j$ obs.'s

$\mu_j \hookrightarrow \bar{Y}_{.j}$: sample mean of n_j obs.'s taken at trt j

(4) $\searrow$

$$
Y_{ij} - \bar{Y}_{..} = (\bar{Y}_{.j} - \bar{Y}_{..}) + (Y_{ij} - \bar{Y}_{.j}) \qquad (5)
$$

Square (5) and sum over i and j

$$\sum_{j=1}^{k}\sum_{i=1}^{n_j}(Y_{ij} - \bar{Y}_{..})^2 \tag{6}$$

$$= \sum_{j=1}^{k}\sum_{i=1}^{n_j}(\bar{Y}_{.j} - \bar{Y}_{..})^2 + \sum_{j=1}^{k}\sum_{i=1}^{n_j}(Y_{ij} - \bar{Y}_{.j})^2$$

$$\mathrm{SS}_{\mathrm{total}} = \mathrm{SS}_{\mathrm{treatment}} + \mathrm{SS}_{\mathrm{error}}$$

(6): fundamental equation of analysis of variance

ANOVA table for testing

H_0 : **no trt effects** ($\tau_1 = \tau_2 = \ldots = \tau_k = 0$ **or** $\sigma^2_\tau = 0$)

Source	df	SS	MS	F	P value
Trt	$\nu_1 = k - 1$	SS_{trt}	MS_{trt}	$f = \frac{\text{MS}_{\text{trt}}}{\text{MS}_{\text{err}}}$	$P(F_{(\nu_1,\nu_2)} \geq f)$
Error	$\nu_2 = N - k$	SS_{err}	MS_{err}		
Total	$N - 1$	SS_{total}			

Note: when $n_j = n, \forall j$ **:**
$E(MS_{err}) = \sigma^2_\epsilon$, $E(MS_{trt}) = \sigma^2_\epsilon + n\theta^2_\tau(\sigma^2_\tau)$,
w/ $\theta^2_\tau = \sum_{j=1}^{k} \tau_j^2 / (k-1)$

Note

$$\begin{aligned} SS_{\text{total}} &= \sum_{j=1}^{k}\sum_{i=1}^{n_j}(Y_{ij} - \bar{Y}_{..})^2 \\ SS_{\text{treatment}} &= \sum_{j=1}^{k}\sum_{i=1}^{n_j}(\bar{Y}_{.j} - \bar{Y}_{..})^2 \\ SS_{\text{error}} &= \sum_{j=1}^{k}\sum_{i=1}^{n_j}(Y_{ij} - \bar{Y}_{.j})^2 \end{aligned}$$

regardless τ is fixed or random.

2-factor experiment

Factor A with $a > 1$ levels
Factor B with $b > 1$ levels
$\longrightarrow$ # of treatments = $ab = k$

Model (1): $j = 1, \ldots k;\ i = 1, \ldots, n_j$

$$Y_{ij} = \mu + \tau_j + \epsilon_{ij} = \mu_j + \epsilon_{ij};$$

- ✓ Y_{ij} : the ith obs on the jth trt
- ✓ μ_j : the mean of the jth trt

jth trt, $j = 1, \ldots, k$

$\hookrightarrow (i, j)$ trt, $i = 1, \ldots, a, j = 1, \ldots b$

(1) becomes $i = 1, \ldots a;\ j = 1, \ldots, b;\ k = 1, \ldots, n_{ij}$

$$Y_{ijk} = \mu_{ij} + \epsilon_{(ij)k} = \mu + \tau_{ij} + +\epsilon_{(ij)k}$$

w/Y_{ijk} : the kth obs on trt (i, j) $\quad \epsilon_{(ij)k} \overset{iid}{\sim} N(0, \sigma^2_\epsilon)$

Let μ_i: mean of response w/ A at level i;
μ_j: mean of response w/ B at level j

$$Y_{ijk} = \mu_{ij} + \epsilon_{(ij)k} \quad \text{c.f.} \quad Y_{ijk} = Y_{ijk}$$

$$\begin{aligned}
\Longleftrightarrow \quad Y_{ijk} - \mu &= (\mu_{ij} - \mu) + (Y_{ijk} - \mu_{ij}) && (7)\\
\Longleftrightarrow \quad Y_{ijk} - \mu &= (\mu_i - \mu) + (\mu_j - \mu) \\
+ (Y_{ijk} - \mu_{ij}) &+ (\mu_{ij} - \mu_i - \mu_j + \mu) && (8)\\
&\Downarrow \quad \text{model} \\
Y_{ijk} &= \mu + A_i + B_j + AB_{ij} + \epsilon_{(ij)k}
\end{aligned}$$

$\hookrightarrow \mu_{ij} = \mu + A_i + B_j + AB_{ij} =$ common mean + main effects + interaction

$\mu \hookrightarrow \bar{Y}_{...}$: sample mean of $N = \sum_{ij} n_{ij}$ obs.'s

$\mu_i \hookrightarrow \bar{Y}_{i..}$: sample mean of $\sum_j n_{ij}$ obs.'s taken at level i of A

$\mu_j \hookrightarrow \bar{Y}_{.j.}$: sample mean of $\sum_i n_{ij}$ obs.'s taken at level j of B

$\mu_{ij} \hookrightarrow \bar{Y}_{ij.}$: sample mean of n_{ij} obs.'s taken at trt (i, j) in (7),(8)

$\hookrightarrow$

$$
\begin{aligned}
Y_{ijk} - \bar{Y}_{...} &= (\bar{Y}_{ij.} - \bar{Y}_{...}) + (Y_{ijk} - \bar{Y}_{ij.}) \\
&= (\bar{Y}_{i..} - \bar{Y}_{...}) + (\bar{Y}_{.j.} - \bar{Y}_{...}) \\
&\quad + (\bar{Y}_{ij.} - \bar{Y}_{i..} - \bar{Y}_{.j.} + \bar{Y}_{...}) + (Y_{ijk} - \bar{Y}_{ij.})
\end{aligned}
$$

Square and sum over i, j, k

$$\sum_i^a \sum_j^b \sum_k^{n_{ij}} (Y_{ijk} - \bar{Y}_{\cdots})^2$$

$$= \sum_i^a \sum_j^b \sum_k^{n_{ij}} (\bar{Y}_{ij\cdot} - \bar{Y}_{\cdots})^2 + \sum_i^a \sum_j^b \sum_k^{n_{ij}} (Y_{ijk} - \bar{Y}_{ij\cdot})^2$$

$$\mathrm{SS}_{\mathrm{total}} = \mathrm{SS}_{\mathrm{trt}} + \mathrm{SS}_{\mathrm{error}}$$

$$\sum_{i}^{a}\sum_{j}^{b}\sum_{k}^{n_{ij}}(\bar{Y}_{ij.} - \bar{Y}_{...})^2 + \sum_{i}^{a}\sum_{j}^{b}\sum_{k}^{n_{ij}}(Y_{ijk} - \bar{Y}_{ij.})^2$$

$$= \sum_{i}^{a}\sum_{j}^{b}\sum_{k}^{n_{ij}}(\bar{Y}_{i..} - \bar{Y}_{...})^2 + \sum_{i}^{a}\sum_{j}^{b}\sum_{k}^{n_{ij}}(\bar{Y}_{.j.} - \bar{Y}_{...})^2$$

$$+ \sum_{i}^{a}\sum_{j}^{b}\sum_{k}^{n_{ij}}(\bar{Y}_{ij.} - \bar{Y}_{i..} - \bar{Y}_{.j.} + \bar{Y}_{...})^2$$

$$+ \sum_{i}^{a}\sum_{j}^{b}\sum_{k}^{n_{ij}}(Y_{ijk} - \bar{Y}_{ij.})^2$$

$$
\begin{aligned}
\mathrm{SS}_{\mathrm{total}} &= \mathrm{SS}_{\mathrm{trt}} + \mathrm{SS}_{\mathrm{error}} \\
&= \mathrm{SS}_A + \mathrm{SS}_B + \mathrm{SS}_{AB} + \mathrm{SS}_{\mathrm{error}}
\end{aligned}
$$

The fundamental equation of analysis of variance, again

ANOVA table for testing H_0 : no trt effects
H_0 : no A main effects ; H_0 : no B main effects
H_0 : no A, B interaction:

Source	df	SS		MS	F
Trt	$\nu_1 = ab - 1$	SS_{trt}		MS_{trt}	$\frac{MS_{trt}}{MS_{err}}$
A	$a - 1$		SS_A	MS_A	?
B	$b - 1$		SS_B	MS_B	?
AB	$(a-1)(b-1)$		SS_{AB}	MS_{AB}	?
Error	$\nu_2 = N - ab$	SS_{err}		MS_{err}	
Total	$N - 1$	SS_{total}			

Multi-factor model

jth trt, $j = 1, \ldots, k = abc$ (3 factors)

$\hookrightarrow (i, j, k)$ trt, $i = 1, \ldots, a, j = 1, \ldots b, k = 1, \ldots c$

model:
$i = 1, \ldots a;\ j = 1, \ldots, b;\ k = 1, \ldots, c;\ l = 1, \ldots n_{ijk}$

$$Y_{ijkl} = \mu_{ijk} + \epsilon_{(ijk)l} = \mu + \tau_{ijk} + +\epsilon_{(ijk)l}$$

$\hookrightarrow$ multi-factor model . . . . . .

You get the idea, right?