TWO WAY ANALYSIS OF VARIANCE MODEL

Consider a completely randomized design for an experiment with two treatment factors A and B. We will assume that every level of A is observed with every level of B (so the factors are *crossed*).

Notation:

A has a levels coded 1, 2, … , a
B has b levels coded 1,2, … , b
v = total number of treatments (= ab)

Example: In the battery example, we had two factors:

A with levels 1 = alkaline, 2 = heavy duty
B with levels 1 = name brand, 2 = store brand
So a = 2, b = 2, v = 4.

When designing the experiment, we could imagine various scenarios. Four possibilities (not a complete list):

I) The means of LPUC for levels of B don' t depend on the level of A, and are higher for level 1 of B than for level 2 of B.

II) Treatment makes no difference in LPUC.

III) When A is at level 1, the mean LPUC is higher for level 1 of B than for level 2 of B, but when A is at level 2, the means for both levels of B are the same.

IV) For level 1 of A, the mean LPUC if higher for level 1 of B than for level 2 of B, but for level 2 of A, the mean of LPUC is lower for level 1 of B than for level 2 of B.

Such scenarios can be illustrated with *interaction plots*. Here is an example for case (I).

```
I)              |
                |                                  Brand: Name = 1, Store = 2
        Group   |     1--------1
         mean   |
                |     2--------2
                |
                |___________________
                      1        2      Type battery: Alkaline = 1, Heavy duty = 2
```

Exercise: Sketch the appropriate interaction plots for cases II - IV.

II)

Brand: Name = 1, Store = 2

Group mean

1 2 Type battery: Alkaline = 1, Heavy duty = 2

III)

Brand: Name = 1, Store = 2

Group mean

1 2 Type battery: Alkaline = 1, Heavy duty = 2

IV)

Brand: Name = 1, Store = 2

Group mean

1 2 Type battery: Alkaline = 1, Heavy duty = 2

Possible Models

Let Y_{ijt} denote the random variable giving the response for observation t of the treatment at level i of A and level j of B. (r_{ij} = number of observations at level i of A and level j of B.)

1. The *cell-means model*:

$Y_{ijt} = \mu + \tau_{ij} + \varepsilon_{ijt}$.

The ε_{ijt} are independent random variables.

Each $\varepsilon_{ijt} \sim N(0, \sigma^2)$

Exercise: What can you say about the τ_{ij}'s in each scenario (I) - (IV) above when using the cell-means model?

2. The *main effects model* (also known as the *two-way additive model*):

$Y_{ijt} = \mu + \alpha_i + \beta_j + \varepsilon_{ijt}$.

The ε_{ijt} are independent random variables.

Each $\varepsilon_{ijt} \sim N(0, \sigma^2)$

Exercise:

a. Can each of scenarios I - IV be modeled by the main effects model? If not, which ones can and which ones can't? In any of the scenarios that can be modeled by the main effect model, what is the connection between α_i, β_j of this model and τ_{ij} of the cell-means model?

b. For the battery situation, find $E(Y_{11} - Y_{12})$ and $E(Y_{21} - Y_{22})$. What are the implications of what you find?

3. The *two-way analysis of variance model* (also known as the *two-way complete model*):

$Y_{ijt} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijt}$.

The ε_{ijt} are independent random variables.

Each $\varepsilon_{ijt} \sim N(0, \sigma^2)$

Note that:

- The two-way ANOVA model is equivalent to the cell-means model: $\tau_{ij} = \alpha_i + \beta_j + (\alpha\beta)_{ij}$
- The main effects model is a submodel of the two-way ANOVA model -- it is the special case of the two-way model when all $(\alpha\beta)_{ij} = 0$.

The terms $(\alpha\beta)_{ij}$ are called *interaction terms*. The following exercise illustrates why.

Exercise: In each scenario I - IV above, give plausible values of the parameters α_i, β_j , and $(\alpha\beta)_{ij}$.