ANOVA - Analysis of Multiple Factor Experiments

October 17, 2017

Randomized Block Design - Two Way ANOVA

Suppose in a random experiment we have a treatment that we are interested in and we have a blocking factor.

Example

Looking for differences in miles per gallon in 3 different gasoline treatments (treatment factor) and we want to block out differences in engine types (blocking factor). So we can assign each car to each gasoline and measure mpg.

Suppose there are k levels of factor A (treatment) and b levels of factor B (blocking).

Fixed Effect Model

Suppose we have a treatment effect, A with k levels and a blocking effect, B with b levels.

Model

$$Y_{ij} = \mu + \tau_i + \beta_j + \epsilon_{ij}$$

where

- Y_{ij} is the response in the ith treatment group of the jth block
- μ is the grand mean
- τ_i is the treatment effect in the ith treatment group
- β_j is the blocking effect in the jth block
- ϵ_{ij} is the error term which we assume to be iid $N(0, \sigma^2)$

ANOVA Table

As in the one-way ANOVA we can construct an ANOVA table to test for a treatment effect and for a blocking effect. Define means:

$$\bar{Y}_{i\bullet} = \frac{1}{b}\sum_{j=1}^{b} Y_{ij} \text{ average of } i\text{th treatment}$$

$$\bar{Y}_{\bullet j} = \frac{1}{k}\sum_{i=1}^{k} Y_{ij} \text{ average of } j\text{th block}$$

$$\bar{Y}_{\bullet\bullet} = \frac{1}{bk}\sum_{i=1}^{K}\sum_{j=1}^{b} Y_{ij} \text{ grand mean}$$

Sum of Squares

Sum of Squares	Measures
$SST = \sum_{i=1}^{k} \sum_{j=1}^{b} (Y_{ij} - \bar{Y}_{\bullet\bullet})^2$	total variation
$SSA = \sum_{i=1}^{k} \sum_{j=1}^{b} (\bar{Y}_{i\bullet} - \bar{Y}_{\bullet\bullet})^2$	variation due to treatment
$SSB = \sum_{i=1}^{k} \sum_{j=1}^{b} (\bar{Y}_{\bullet j} - \bar{Y}_{\bullet\bullet})^2$	variation due to the block
$SSE = SST - SSA - SSB$	variation due to the error term

ANOVA table

Source	**SS**	**df**	**MS**	**F**
Treatment	SSA	k-1	MSA = SSA/(k-1)	$F_A = MSA/MSE$
Block	SSB	b-1	MSB = SSB/(b-1)	$F_B = MSA/MSE$
Error	SSE	(k-1)(b-1)	MSE	
Total	SST	kb-1		

Test for difference in treatment effects

$H_0 : \tau_1 = \cdots = \tau_k$ Under H_0,

$$F_A = \frac{MSA}{MSE} \sim F_{k-1,(k-1)(b-1)}$$

Reject H_0 if F_A is too big.

Test for block effect

$H_0 : \beta_1 = \cdots = \beta_b$
Under H_0

$$F_B = \frac{MSB}{MSE} \sim F_{b-1,(k-1)(b-1)}$$

Reject H_0 if F is too big.

Example

A experiment was conducted to determine if oven position had an effect on drip loss in meat loaf. Three batches of loaves were made, each consisting of 8 loaves. The eight loaves were randomly assigned to one of the eight oven positions. Drip loss was measured for each loaf.

- Treamtent = oven position $k = 8$
- block = batches $b = 3$
- number of responses = $kb = 24$.

ANOVA table

Source	SS	df	MS	F	p-value
Oven Position	40.396	7	5.771	8.70	0.0003
Batch	16.259	2	8.130	12.24	0.0008
Error	9.290	14	0.664		
Total	65.945	23			

We can reject H_0 : no treatment effect (p-value $= 0.0003$).
We can reject H_0 : no block effect (p-value $= 0.0008$).
Conclude there is an effect of oven position and the meat loaf batch on the drip loss.

Random Effects Model

If both Treatment and Block are random samples from larger populations then we have a random effects model.

Model

$$Y_{ij} = \mu + \alpha_i + \beta_j + \epsilon_{ij}$$

- μ is the overall mean
- α_i is the treatment effect random variable assumed iid $N(0, \sigma_A^2)$
- β_j is the block effect random variable assumed iid $N(0, \sigma_B^2)$
- ϵ_{ij} is the error term assumed iid $N(0, \sigma^2)$

Further α, β and ϵ are assumed independent of each other. So $Var(Y_{ij}) = \sigma_A^2 + \sigma_B^2 + \sigma^2$.

Statistical Tests - Random Effects Model

To test for treatment effect we test $H_0 : \sigma_A^2 = 0$ and to test for block effect we test $H_0 : \sigma_B^2 = 0$ using the same ANOVA table as in the fixed effect case.

Mixed Effect Model

Sometimes the treatment(block) is a fixed effect while the block(treatment) is a random effect. Technically speaking this is the case in the meat loaf example, as the oven position was a fixed effect but the batch was a random effect.

Model

$$Y_{ij} = \mu + \tau_i + \beta_j + \epsilon_{ij}$$

- μ is the overall mean
- τ_i is the fixed treatment effect $i = 1, \ldots, k$
- β_j is the random batch effect $j = 1, \ldots, b$ assumed to be iid $N(0, \sigma_B^2)$
- ϵ_{ij} is the error term assumed to be iid $N(0, \sigma^2)$
- Further ϵ_{ij} are independent of β_j

Mixed model hypothesis tests

We can use the same ANOVA table to test

- $H_0 : \tau_1 = \tau_2 = \cdots = \tau_k = 0$ no treatment effect
- $H_0 : \sigma_B^2 = 0$ no batch effect

Balanced/Unbalanced designs

In the one-way model, the ANOVA table test (F test) is appropriate even when the design is not balanced ($n_1 \neq n_2 \neq \cdots n_K$).
The two way model is only appropriate if the design is balanced, that is when the number of Y's is the same for each treatment/block combination. When the design is modestly unbalanced, you can still use the ANOVA table test and the results will be approximately correct. If the design is badly unbalanced, you should not use ANOVA. In this case the unbalanced design is no longer orthogonal

$$SST \neq SSA + SSB + SSE$$

and so the F statistics no longer have an F distribution.
Instead you should use a regression model, assigning dummy variables for the treatment factor and batch factor.

Interaction effects

Suppose we have a two-way classification with two factors A with $a \geq 2$ levels and B with $b \geq 2$ levels. Both factors may effect the observed response variable Y. Further suppose we have n replications within each treatment combination resulting in a total of $N = nab$ total observations. In this case we are interested in whether each factor has an effect on response and in whether their interaction also has a significant effect.

model

$$Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \epsilon_{ijk}$$

- μ grand mean
- α_i is effect due to the ith level of treatment A
- β_j is effect due to the jth level of treatment B
- $(\alpha\beta)_{ij}$ is the effect of the interaction of the ith level of Treatment A with the jth level of Treatment B
- ϵ_{ijk} is the error term

Two-way model can be fixed effects, mixed effects or random effects model.

ANOVA Table for an Interaction model

Fixed Effects Model

Source	SS	df	MS	F	H_0
Treatment A	SSA	a-1	MSA	MSA/MSE	$H_0 : \alpha_i = 0$
Treatment B	SSB	b-1	MSB	MSB/MSE	$H_0 : \beta_j = 0$
Interaction AB	SSAB	(a-1)(b-1)	MSAB	MSAB/MSE	$H_0 : (\alpha\beta)_{ij} =$
Error	SSE	N-ab	MSE		
heightTotal	SST	N-1			

Interpretation of Results

Test the interaction effect first

- If the interaction effect is not significant (can't reject H_0) then there is no interaction effect and you proceed with the analysis of each factor individually
- If the interaction effect is significant then you must interpret your results in terms of the interaction effects.