Two-Way Analysis of Variance - no interaction

Example: Tests were conducted to assess the effects of two factors, engine type, and propellant type, on propellant burn rate in fired missiles. Three engine types and four propellant types were tested.

Twenty-four missiles were selected from a large production batch. The missiles were randomly split into three groups of size eight. The first group of eight had engine type 1 installed, the second group had engine type 2, and the third group received engine type 3.‘

Each group of eight was randomly divided into four groups of two. The first such group was assigned propellant type 1, the second group was assigned propellant type 2, and so on.

Data on burn rate were collected, as follows:

Engine	Propellant Type			
type	1	2	3	4
1	34.0	30.1	29.8	29.0
	32.7	32.8	26.7	28.9
2	32.0	30.2	28.7	27.6
	33.2	29.8	28.1	27.8
3	28.4	27.3	29.7	28.8
	29.3	28.9	27.3	29.1

We want to determine whether either **factor**, engine type (factor A) or propellant type (factor B), has a significant effect on burn rate.

Let Y_{ijk} denote the k'th observation at the i'th level of factor A and the j'th level of factor B.

The two factor model (without interaction) is:

$Y_{ijk} = \mu + \alpha_i + \beta_j + \epsilon_{ijk} \qquad i = 1, 2, 3, \qquad j = 1, 2, 3, 4, \qquad k = 1, 2,$ where

1. μ is the overall mean
2. $\sum_i \alpha_i = 0$
3. $\sum_j \beta_j = 0$
4. we assume ϵ_{ijk} are iid $N(0, \sigma^2)$
5. The mean of Y_{ijk} is

$$\mu_{ijk} = \mathbf{E}[Y_{ijk}] \ = \mu + \alpha_i + \beta_j$$

This model specifies that a plot of the mean against the levels of factor A consists of parallel lines for each different level of factor B, and a plot of the mean against the levels of factor B consists of parallel lines for each different level of factor A.

More generally, there will be I levels of factor A, J levels of factor B, and K replicates at each combination of levels of factors A and B.

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

$i = 1, 2, \ldots, I, \qquad j = 1, 2, \ldots, J, \qquad k = 1, 2, \ldots, K$

- In the example, $I = 3$, $J = 4$, $K = 2$, and there are $n = 24$ observations in total. There are $K = 2$ **replicates** at each level the factors A and B, and the experimental design is said to be **balanced**, because there are the same number of replicates in each cell.

The following table gives the cell means: $\bar{Y}_{ij.}$

Engine type	Propellant Type 1	2	3	4
1	33.35	31.45	28.25	28.95
2	32.60	30.00	28.40	27.70
3	28.85	28.10	28.50	28.95

- The estimated grand mean is: $\hat{\mu} = \bar{y}_{...} = 29.5917$

- the estimated factor A level means are:
  $\bar{y}_{1..} = (33.35 + 31.45 + 28.25 + 29.95)/4 = 30.50$
  $\bar{y}_{2..} = (32.6 + 30 + 28.4 + 27.7)/4 = 29.675$
  $\bar{y}_{3..} = (28.85 + 28.1 + 28.5 + 28.95)/4 = 28.60$

- and the estimated factor B level means are:
  $\bar{y}_{.1.} = (33.35 + 32.6 + 28.85)/3 = 31.60$
  $\bar{y}_{.2.} = (31.45 + 30 + 28.1)/3 = 29.85$
  $\bar{y}_{.3.} = (28.25 + 28.4 + 28.5)/3 = 28.383$
  $\bar{y}_{.4.} = (28.95 + 27.7 + 28.95)/3 = 28.533$

Estimation of Model Parameters

- $\hat{\mu} = \bar{y}_{...} = (30.5 + 29.675 + 28.6)/3 = (31.6 + 29.85 + 28.383 + 28.533)/4 = 29.5917$

- $\hat{\alpha}_i = \bar{y}_{i..} - \bar{y}_{...}$
  $\hat{\alpha}_1 = \bar{y}_{1..} - \bar{y}_{...} = 30.50 - 29.5917 = .908$
  $\hat{\alpha}_2 = \bar{y}_{2..} - \bar{y}_{...} = 29.675 - 29.5917 = .083$
  $\hat{\alpha}_3 = \bar{y}_{3..} - \bar{y}_{...} = 28.60 - 29.5917 = -.992$

  Note that $\hat{\alpha}_1 + \hat{\alpha}_2 + \hat{\alpha}_3 = 0$.

- $\hat{\beta}_j = \bar{y}_{.j.} - \bar{y}_{...}$
  $\hat{\beta}_1 = \bar{y}_{.1.} - \bar{y}_{...} = 31.60 - 29.5917 = 2.0083$
  $\hat{\beta}_2 = \bar{y}_{.2.} - \bar{y}_{...} = 29.85 - 29.5917 = .2583$
  $\hat{\beta}_3 = \bar{y}_{.3.} - \bar{y}_{...} = 28.383 - 29.5917 = -1.2087$
  $\hat{\beta}_4 = \bar{y}_{.4.} - \bar{y}_{...} = 28.533 - 29.5917 = -1.0587$

  Note that $\hat{\beta}_1 + \hat{\beta}_2 + \hat{\beta}_3 + \hat{\beta}_4 = 0$.

In the twoway model without interaction, the estimated cell means are:

$$\hat{\mu}_{ijk} = \hat{E}[Y_{ijk}] = \hat{\mu} + \hat{\alpha}_i + \hat{\beta}_j$$

The estimated means are as follows.

Engine type	Propellant type 1	2	3	4
1	32.508	30.758	29.291	29.441
2	31.683	29.933	28.466	28.616
3	30.608	28.858	27.391	27.541

The residuals are the differences between the observations and the estimated means ($Y_{ijk} - \hat{\mu}_{ijk}$). They are given in the following table:

Engine	Propellant Type			
type	1	2	3	4
1	1.492	-0.658	0.509	-0.441
	0.192	2.042	-2.591	-0.541
2	0.317	0.267	0.234	-1.016
	1.517	-0.133	-0.366	-0.816
3	-2.208	-1.558	2.309	1.259
	-1.308	0.042	-0.091	1.559

- The mean of the residuals is 0. This will always be the case.
- The sum of squares of the residuals is the error sum of squares (SSE) in the ANOVA table.

```
> resids
 1.492 -0.658  0.509 -0.441  0.192  2.042 -2.591 -0.541  0.317  0.267
 0.234 -1.016  1.517 -0.133 -0.366 -0.816 -2.208 -1.558  2.309  1.259
-1.308  0.042 -0.091  1.559
```

```
> mean(resids)
[1] 5e-04
```

The sum of squares of the residuals is 37.07.

```
> sum(resids^2)
[1] 37.07334
```

Formulas for Sums of Squares

$SSA = JK \sum_i (\bar{y}_{i..} - \bar{y}_{...})^2 = JK \sum_i \hat{\alpha}_i^2 = 4 \times 2 \times (.908^2 + .083^2 + (-.992)^2) = 14.52$

$SSB = IK \sum_j (\bar{y}_{.j.} - \bar{y}_{...})^2 = IK \sum_j \hat{\beta}_j^2 = 3 \times 2 \times (2.0083^2 + .2583^2 + (-1.2087)^2 + (-1.0587^2)) = 40.08$

$SSE = 37.07$

$SST = \sum_i \sum_j \sum_k (y_{ijk} - \bar{y}_{...})^2 = \sum_i \sum_j \sum_k (y_{ijk} - 29.5917)^2 = 91.68$

- Note the additivity relationsip, **SST=SSA+SSB+SSE**.
- If there are no replicates (only one observation per cell), then K will be equal to 1 in these formulas.
- The total degrees of freedom is the number of observations (24) minus 1, or 23. In general this will be $n - 1$.
- The degrees of freedom for factor A is the number of levels of A (3) minus 1, or 2. In general this will be $I - 1$.
- The degrees of freedom for factor B is the number of levels of B (4) minus 1, or 3. In general this will be $J - 1$.
- The degrees of freedom for error is the total number of degrees of freedom, minus the degrees of freedom for A, minus the degrees of freedom for B, or 23-2-3=18. In general this will be $(n-1) - (I-1) - (J-1) = n - I - J + 1$.
- This allows us the build the ANOVA table, as follows.

Source	DF	SS	MS	F	P
A	2	14.52	MSA=7.26	MSA/MSE=3.53	
B	3	40.08	MSB=13.36	MSB/MSE=6.49	
Error	18	37.07	MSE=2.06		
Total	23	91.68			

The hypotheses of interest are:

- $H_{0A} : \alpha_1 = \alpha_2 = \ldots = \alpha_I = 0$ (no effect of factor A)
- The observed test statistic for H_{0A} is $F_{obsA} = MSA/MSE$, and the p-value is $P(F_{I-1,n-I-J+1}) > F_{obsA}$, or in the present case, $P(F_{2,18}) > 3.53$. Referring to the F table, we see that the p-value is in (.05,.10).

- $H_{0B} : \beta_1 = \beta_2 = \ldots = \beta_I = 0$ (no effect of factor B)
- The observed test statistic for H_{0B} is $F_{obsB} = MSB/MSE$, and the p-value is $P(F_{J-1,n-I-J+1}) > F_{obsB}$, or in the present case, $P(F_{3,18}) > 6.49$.
  Calculating in R, the p-value is:

  ```
  > pf(6.49,3,18,lower.tail=F)
  [1] 0.003614655
  ```

  Calculating in minitab, the p-value is:

Here is the output from fitting an additive two way ANOVA in minitab.

```
MTB > set c1
DATA> 34 32.7 30.1 32.8 29.8 26.7 29 28.9
DATA> 32 33.2 30.2 29.8 28.7 28.1 27.6 27.8
DATA> 28.4 29.3 27.3 28.9 29.7 27.3 28.8 29.1
DATA> end
MTB > set c2
DATA> 8(1) 8(2) 8(3)
DATA> set c3
DATA> 1 1 2 2 3 3 4 4 1 1 2 2 3 3 4 4 1 1 2 2 3 3 4 4
DATA> end

MTB > twoway c1 c2 c3;
SUBC> additive.

Two-way ANOVA: C1 versus C2, C3

Source  DF       SS       MS     F      P
C2       2  14.5233   7.2617  3.53  0.051
C3       3  40.0817  13.3606  6.49  0.004
Error   18  37.0733   2.0596
Total   23  91.6783

S = 1.435   R-Sq = 59.56%   R-Sq(adj) = 48.33%


            Individual 95% CIs For Mean Based on
            Pooled StDev
C2    Mean    -+---------+---------+---------+--------
1   30.500                   (--------*--------)
2   29.675            (--------*--------)
3   28.600    (--------*--------)
              -+---------+---------+---------+--------
            27.6      28.8      30.0      31.2
```

```
               Individual 95% CIs For Mean Based on
               Pooled StDev
C3     Mean  ---------+---------+---------+---------+
1   31.6000                      (--------*-------)
2   29.8500            (-------*-------)
3   28.3833  (-------*-------)
4   28.5333   (-------*-------)
             ---------+---------+---------+---------+
                   28.5      30.0      31.5      33.0
```

Factor A (engine type), with a pvalue=.051 is marginally significant.

Factor B (propellant type), with a pvalue= .004 is highly significant.

By way of comparison, look what happens if we forget to add the second factor. Following are the 1 way ANOVAS for engine type and propellant separately. Note that the lines for "Total" and treatment factor (A or B) are unchanged. The error sums of squares and degrees of freedom are pooled values from the twoway ANOVA table. (eg. 77.16= 40.0817+37.0733; 21=18+3). The important thing to note is that factor A is now declared to be completely unimportant (p-value=.164)? What happened? By neglecting to include factor B, the SSE has more than doubled, while the error df has increased by only 3. The resulting estimate of the error variance (3.67) is nearly twice what it was in the two factor model, making the differences between engine types appear to be insignificant.

```
MTB >  oneway c1 c2
One-way ANOVA: C1 versus C2
Source  DF     SS    MS     F      P
C2       2  14.52  7.26  1.98  0.164
Error   21  77.16  3.67
Total   23  91.68

MTB > oneway c1 c3
One-way ANOVA: C1 versus C3
Source  DF     SS     MS     F      P
C3       3  40.08  13.36  5.18  0.008
Error   20  51.60   2.58
Total   23  91.68
```

Two-Way Analysis of Variance - with interaction

Let us go back to our two factor example on missile burn rate.

The two factor model without interaction was
$Y_{ijk} = \mu + \alpha_i + \beta_j + \epsilon_{ijk} \qquad i = 1, 2, 3, \qquad j = 1, 2, 3, 4, \qquad k = 1, 2,$

This model is rather restrictive in that it assumes that the difference in mean burn rate for two propellant types does not depend on the engine type that was being used, and the difference in mean burn rate for two engine types does not depend on the propellants which were being used. In fact, some propellants might work best with certain engine types, and vice versa, so we need to consider a more general model.

The **two factor model with interaction** is:
$Y_{ijk} = \mu+\alpha_i+\beta_j+\gamma_{ij}+\epsilon_{ijk} \qquad i = 1, 2, \ldots, I, \qquad j = 1, 2, \ldots, J, \qquad k = 1, 2, \ldots, K.$

where

- μ is the overall mean
- $\sum_{i=1}^{I} \alpha_i = 0$
- $\sum_{j=1}^{J} \beta_j = 0$
- $\sum_{i=1}^{I} \gamma_{ij} = 0$ for each $j = 1, 2, \ldots, J$
- $\sum_{j=1}^{J} \gamma_{ij} = 0$ for each $i = 1, 2, \ldots, I$
- we assume ϵ_{ijk} are iid $N(0, \sigma^2)$
- in this case the mean of Y_{ijk} is

$$\mu_{ijk} = \mathbf{E}[Y_{ijk}] \; = \mu + \alpha_i + \beta_j + \gamma_{ij}$$

The sum constraints ensure that there is a unique correspondence between the parameters (the α's, β's, γ's and μ) and the means of the random variables (the μ_{ijk}'s).

In the twoway model with interaction, the estimate of the mean μ_{ijk} is given by $\hat{\mu}_{ijk} = \bar{y}_{ij.}$. In the example, these were calculated as:

Engine	Propellant Type			
type	1	2	3	4
1	33.35	31.45	28.25	28.95
2	32.60	30.00	28.40	27.70
3	28.85	28.10	28.50	28.95

leading to the residuals:

Engine	Propellant Type			
type	1	2	3	4
1	0.65	-1.35	1.55	0.05
	-0.65	1.35	-1.55	-0.05
2	-0.60	0.20	0.30	-0.10
	0.60	-0.20	-0.30	0.10
3	-0.45	-0.80	1.20	-0.15
	0.45	0.80	-1.20	0.15

- As usual, the sum of the residuals equals 0.
- The sum of squares of the residuals is $SSE = 14.91$.

- The total sum of squares SST, sum of squares for engine type SSA and sum of squares for propellant type SSB are as before.
- The twoway model with interaction has a sum of squares for term for interaction

$$SS_{AB} = \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=1}^{K} (\bar{Y}_{ij.} - \bar{Y}_{i..} - \bar{Y}_{.j.} + \bar{Y}_{...})^2$$

- $SST = SSA + SSB + SS_{AB} + SSE$
- The degrees of freedom for interaction is $(I-1)(J-1)$ and the degrees of freedom for error is $IJ(K-1)$. The degrees of freedom for factors A and B and total are unchanged, and again, there is an additivity relationship for the degrees of freedom.

The ANOVA table for the model with interaction is:

Source	DF	SS	MS	F	P
A	I-1	SSA	MSA	MSA/MSE	
B	J-1	SSB	MSB	MSB/MSE	
AB	(I-1)(J-1)	SSAB	MSAB	MSAB/MSE	
Error	IJ(K-1)	SSE	MSE		
Total	IJK-1	SST			

In the example:

Source	DF	SS	MS	F	P
A	2	14.52	7.26	5.84	
B	3	40.08	13.36	10.75	
AB	6	22.17	3.70	2.97	
Error	12	14.91	1.243		
Total	23	91.68			

1. The p-value for the test for no interaction between propellant type and engine type, formally $\gamma_{ij} = 0$ for all i and j, is given by $pvalue = P(F_{(I-1)(J-1),IJ(K-1)} > MSAB/MSE$. For these data, the p-value is $P(F_{6,12} > 2.97) \in (.05, .1)$. In this case the test for interaction is not significant (ie we conclude there is no interaction between factors A and B), which indicates that the profile plots of the means are parallel. That is, the additive model is reasonable.

2. It only makes sense to test for the **main effects** of factors A and B if there is no interaction between the factors. In this case:

   (a) in testing for the main effect of engine type, the p-value is $p = P(F_{2,12} > 5.84) \in (.01, .05)$

   (b) in testing for the main effect of propellant, the p-value is $p = P(F_{3,12} > 10.75) < .01$.

   When testing at level .05, we conclude that there are significant differences between propellant types, and between engine types.

Here is the minitab output which verifies the calculations:

```
MTB > Twoway c1 c2 c3.

Two-way ANOVA: C1 versus C2, C3

Source       DF       SS       MS      F      P
C2            2  14.5233   7.2617   5.84  0.017
C3            3  40.0817  13.3606  10.75  0.001
Interaction   6  22.1633   3.6939   2.97  0.051
Error        12  14.9100   1.2425
Total        23  91.6783
```

Notes:

- If we had found that there was a significant interaction between propellant and engine type, we should not test for the main effects of those factors, as we know that there are effects of each factor, but that the effect of one factor will depend on the level of the other factor, because the mean profiles are NOT parallel.
- In the example, our conclusion was that there is no interaction, but that there are significant main effects of engine type AND propellant type. We can use a Bonferroni procedure to determine for which levels the effects of factor A are different, and for which levels the effects of factor B are different. The details are described in the notes on "Post-hoc comparisons".
- In one way analysis of variance there is one factor. In two way ANOVA there are two factors. In general there may be several factors, and the analysis of variance extends to that case, but we will limit our discussion in this course to one and two way ANOVA.

Blocking

Example: A new drug is being tested which is supposed to boost the average immune response to infection.

One scenario for a randomized controlled study is as follows:

- sample individuals from a population
- randomly assign individuals in the sample to the new treatment or to a control
- challenge the individuals with an antigen (eg a flu vaccine), and measure antibody levels 2 months later
- use a t-test to compare the mean antibody levels in the two groups. In the case there are several treatment groups (corresponding, say, to different doses of the treatment drug), use oneway ANOVA to compare the means of the associated groups.

A problem with this design is that it is known that a variety of factors affect the immune response. In particular, the immune response (more specifically, the production of antibodies) is known to be depressed in smokers, and so, even if the treatment is effective at increasing the average immune response, if the treatment group is dominated by smokers relative to the control group, the positive effect of the treatment will be masked.

Therefore, we would like to include smoker vs non-smoker as a second factor. The problem is that we have no way to randomly assign levels of this second factor to the experimental units (the subjects). Formally, groups of experimental units which are expected to be similar are referred to as **blocks**, and the associated factor is referred to as a **blocking factor**. By including the blocking factor in our experimental design we remove the variation (in the outcome variable) attributable to different values of the blocking factor.

In essence, we replace the anova table:

Source	DF	SS	MS	F
Treatments	I-1	SS_{tr}	MS_{tr}	MS_{tr}/MSE_1
Error1	n-I	SSE_1	MSE_1	
Total	n-1	SST		

by the table:

Source	DF	SS	MS	F
Treatments	I-1	SS_{tr}	MS_{tr}	$F_{obsTr} = MS_{tr}/MSE_2$
Blocks	J-1	SS_{blocks}	MS_{blocks}	$F_{obsBlocks} = MS_{blocks}/MSE_2$
Error2	n-I-J+1	SSE_2	MSE_2	
Total	n-1	SST		

where

- $SSE_1 = SSE_2 + SS_{blocks}$
- $n - i = (n - I - J + 1) + (j - 1)$

When we have included the blocking factor:

1. To test the hypothesis that there is no effect of the treatment, the p-value is $P(F_{i-1,n-I-J+1} > F_{obsTr})$.
2. Was the blocking effective? More formally, are the response means different across levels of the blocking factor. To test this, the p-value is $P(F_{J-1,n-I-J+1} > F_{obsBlocks})$.

Thus, inclusion of a blocking factor in a single factor experiment leads us to a two factor experiment, which we analyse using 2-way ANOVA, where one of the factors is of a special type **a blocking factor that cannot be randomly allocated to experimental units.**