Example: ANOVA for determining the best level of fertilizer application (Bus Q600)

Let's consider the fertilizer level choice problem. We will start with the data set and then perform the ANOVA calculations.

The problem is,

$$H_0 \ : \ \mu_1 = \mu_2 = \mu_3 \text{ (i.e., all three levels have the same effectiveness)}$$
$$H_a \ : \ \text{at least two levels differ in effectiveness.}$$

Data Set

First, recall that we write the sample variance as

$$s^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}.$$

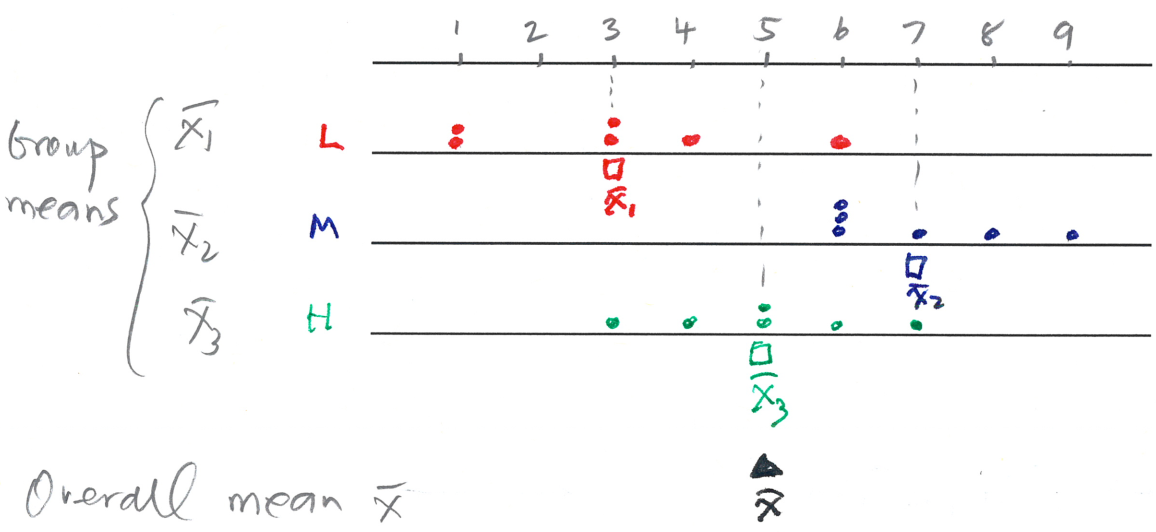
An alternative notation for this is

$$MS = \frac{SS}{n-1}$$

where

MS (Mean Square)	is the new notation for	s^2
SS (Sum of Squares)	is the new notation for	$\sum_{i=1}^{n}(x_i - \bar{x})^2$.

Here is the dotplot for the data values followed by a summary table.

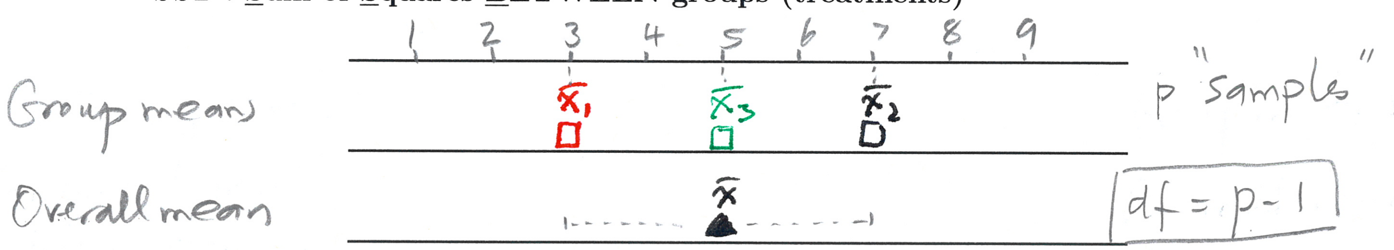


Experiment j	Level i (L, M or H)		
	L $(i=1)$	M $(i=2)$	H $(i=3)$
$j=1$	$x_{11}=6$	$x_{21}=6$	$x_{31}=5$
$j=2$	$x_{12}=4$	$x_{22}=6$	$x_{32}=4$
$j=3$	$x_{13}=3$	$x_{23}=9$	$x_{33}=3$
$j=4$	$x_{14}=3$	$x_{24}=8$	$x_{34}=6$
$j=5$	$x_{15}=1$	$x_{25}=7$	$x_{35}=5$
$j=6$	$x_{16}=1$	$x_{26}=6$	$x_{36}=7$
Sum	$\sum_{j=1}^{6} x_{1j}=18$	$\sum_{j=1}^{6} x_{2j}=42$	$\sum_{j=1}^{6} x_{3j}=30$
n_i :	$n_1=6$	$n_2=6$	$n_3=6$
Group means → $\bar{x}_i$:	$\bar{x}_1=3$	$\bar{x}_2=7$	$\bar{x}_3=5$
Overall mean → $\bar{x}=$		$\frac{18+42+30}{6+6+6}=5$	
$n=$		$n_1+n_2+n_3=18$	
$p=$		3	

SS : **<u>S</u>um of <u>S</u>quares (useful in variance-type calculations for ANOVA)**

Here we calculate the SSB and SSE values and add them to find the SST.

SSB : **<u>S</u>um of <u>S</u>quares <u>B</u>ETWEEN groups (treatments)**



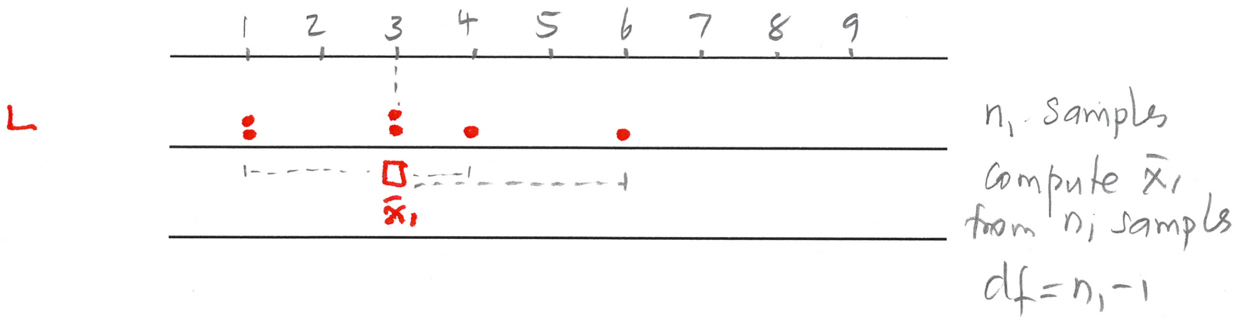
$$
\begin{aligned}
SSB &= n_1(\bar{x}_1-\bar{x})^2+n_2(\bar{x}_2-\bar{x})^2+n_3(\bar{x}_3-\bar{x})^2 \\
&= 6\cdot(3-5)^2+6\cdot(7-5)^2+6\cdot(5-5)^2 \\
&= 6\cdot 4+6\cdot 4+0 \\
SSB &= 48
\end{aligned}
$$

SSE : **Sum of Squares for ERRORS (within groups)**

$$SSE = \underbrace{\sum_{j=1}^{n_1}(x_{1j} - \bar{x}_1)^2}_{\text{For Level } i=1} + \underbrace{\sum_{j=1}^{n_2}(x_{2j} - \bar{x}_2)^2}_{\text{For Level } i=2} + \underbrace{\sum_{j=1}^{n_3}(x_{3j} - \bar{x}_3)^2}_{\text{For Level } i=3}$$

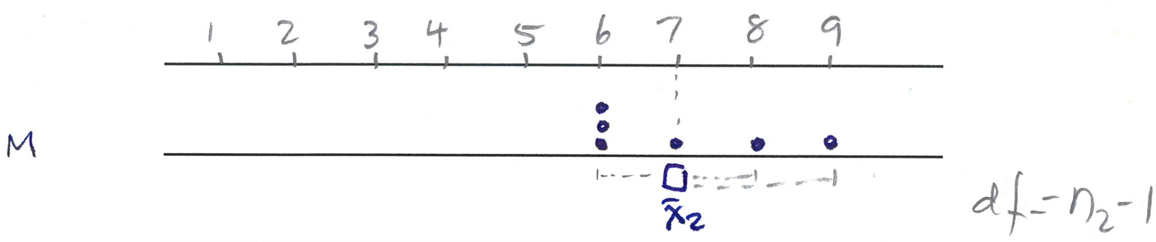
Let's consider each term separately now.

Level $i = 1$



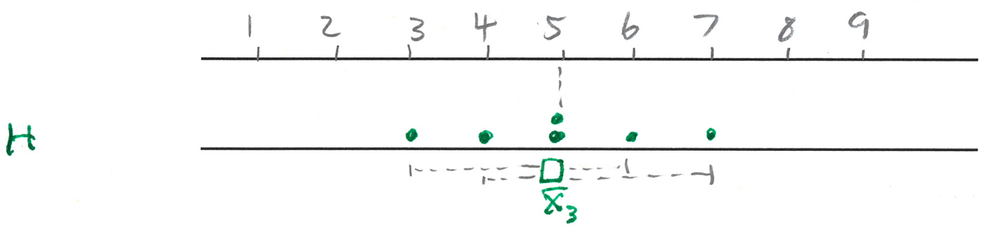
$$\begin{aligned}\sum_{j=1}^{n_1}(x_{1j} - \bar{x}_1)^2 &= (x_{11} - \bar{x}_1)^2 + (x_{12} - \bar{x}_1)^2 + (x_{13} - \bar{x}_1)^2 + (x_{14} - \bar{x}_1)^2 \\ &\quad +(x_{15} - \bar{x}_1)^2 + (x_{16} - \bar{x}_1)^2 \\ &= (6-3)^2 + (4-3)^2 + (3-3)^2 + (3-3)^2 + (1-3)^2 + (1-3)^2 \\ &= 9 + 1 + 0 + 0 + 4 + 4 \\ &= 18\end{aligned}$$

Level $i = 2$



$$\begin{aligned}\sum_{j=1}^{n_2}(x_{2j} - \bar{x}_2)^2 &= 1 + 1 + 4 + 1 + 0 + 1 \\ &= 8\end{aligned}$$

Level $i = 3$



$df = n_3 - 1$

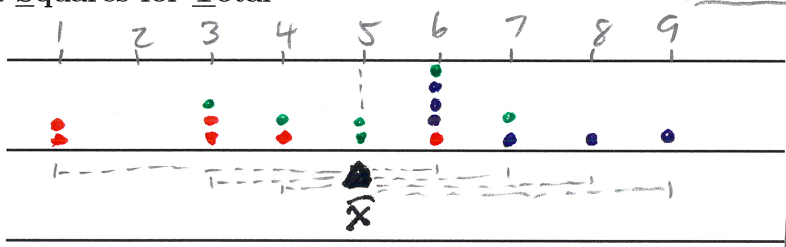
$$\sum_{j=1}^{n_3} (x_3 - \bar{x}_3)^2 = 0 + 1 + 4 + 1 + 0 + 4$$
$$= 10.$$

So, we have

$$SSE = 18 + 8 + 10$$
$$= 36.$$

$df = (n_1 - 1) + (n_2 - 1) + (n_3 - 1)$
$= n_1 + n_2 + n_3 - 3$
$df = n - p$

SST : **S**um **of** **S**quares **for** **T**otal



$n = n_1 + n_2 + n_3$ samples

$df = n - 1$

$df = (p-1) + (n-p)$
$= n - 1$

$$SST = SSB + SSE$$
$$= 48 + 36$$
$$= 84$$

Question What is the main source of variation? Is it between, or within, groups?

To answer this question, we calculate the measures that are similar to variance.

Mean Square (MS)
(similar to s^2)

MS Between (treatments)	MS Error (within groups)
$\downarrow$	$\downarrow$
$MSB = \dfrac{SSB}{p-1} = \dfrac{48}{3-1} = 24$	$MSE = \dfrac{SSE}{n-p} = \dfrac{36}{18-3} = 2.4$

Now, recall that

- if MSB is large and MSE is small, we may reject H_0,
- if MSB is small and MSE is large, we may not reject (i.e., "accept") H_0.

Let's find the test statistic F for this problem:

$$F = \frac{MSB}{MSE}, \text{ with } df = (p-1, n-p).$$

If this number is large, then variations between groups is large and we may reject H_0.

- In our case, $\alpha = 0.05$, $p-1 = 3-1 = 2$, $n-p = 18-3 = 15$ and so, $F_\alpha^{(2,15)} = 3.68$ from Table A.7.
- But we have, for the test statistics, $F = 24/2.4 = 10$, so we reject H_0.

Here is the summary table.

Source	df	SS	MS	F	p-value
Between (treatments)	$p-1=2$	$SSB=48$	$MSB = \frac{SSB}{p-1} = 24$	$F = \frac{MSB}{MSE} = 10$	.0017
Error (within groups)	$n-p=15$	$SSE=36$	$MSE = \frac{SSE}{n-p} = 2.4$		
Total	$n-1=17$	$SST=84$