Consider the following anova program and its output. Note that the least significant difference test has more power than the t-test.

```
data ;
input trt $ y @@ ;
datalines ;
A 6    A 0    A 2    A 8   A 11
A 4    A 13   A 1    A 8   A 0
B 0    B 2    B 3    B 1   B 18
B 4    B 14   B 9    B 1   B 9
C 13   C 10   C 18   C 5   C 23
C 12   C 5    C 16   C 1   C 20
;

proc anova ;
model y = trt ;
means trt / lsd ttest ;
```

```
                              Analysis of Variance

    Source     DF     Sum of Squares      Mean Square     F Value     Pr > F
    Model       2       293.60000000     146.80000000        3.98     0.0305
    Error      27       995.10000000      36.85555556
    Total      29      1288.70000000

               R-Square     Coeff Var      Root MSE       Y Mean
               0.227826     76.846553      6.070878     7.900000

     Source     DF         Anova SS      Mean Square     F Value     Pr > F
     TRT         2     293.60000000     146.80000000        3.98     0.0305

                                 Mean Response

             TRT      N         Mean Y     95% CI MIN     95% CI MAX
             A       10       5.300000       1.360937       9.239063
             B       10       6.100000       2.160937      10.039063
             C       10      12.300000       8.360937      16.239063

                        Least Significant Difference Test

  TRT    TRT      Delta Y    95% CI MIN     95% CI MAX    t Value    Pr > |t|
  A      B      -0.800000     -6.370677       4.770677      -0.29      0.7705
  A      C      -7.000000    -12.570677      -1.429323      -2.58      0.0157 *
  B      A       0.800000     -4.770677       6.370677       0.29      0.7705
  B      C      -6.200000    -11.770677      -0.629323      -2.28      0.0305 *
  C      A       7.000000      1.429323      12.570677       2.58      0.0157 *
  C      B       6.200000      0.629323      11.770677       2.28      0.0305 *

                              Two Sample t-Test

  TRT    TRT      Delta Y    95% CI MIN     95% CI MAX    t Value    Pr > |t|
  A      B      -0.800000     -5.922307       4.322307      -0.33      0.7466
  A      C      -7.000000    -12.664270      -1.335730      -2.60      0.0182 *
  B      A       0.800000     -4.322307       5.922307       0.33      0.7466
  B      C      -6.200000    -12.467653       0.067653      -2.08      0.0523
  C      A       7.000000      1.335730      12.664270       2.60      0.0182 *
  C      B       6.200000     -0.067653      12.467653       2.08      0.0523
```

Let us take a closer look at the analysis of variance table.

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	293.60000000	146.80000000	3.98	0.0305
Error	27	995.10000000	36.85555556		
Total	29	1288.70000000			

This is how the table values are computed where n is the number of observations and p is the number of model parameters.

Source	DF	Sum of Squares	Mean Square	F-value	p-value
Model	$p-1$	SSR	MSR = SSR/$(p-1)$	$F^* = \text{MSR}/\text{MSE}$	$1 - F(F^*, p-1, n-p)$
Error	$n-p$	SSE	MSE = SSE/$(n-p)$		
Total	$n-1$	SST			

For the following sum of squares calculations, y are observed values and $\hat{y}$ are predicted values.

$$\begin{aligned}
\text{SSR} &= \sum(\hat{y}_i - \bar{y})^2 \\
\text{SSE} &= \sum(y_i - \hat{y}_i)^2 \\
\text{SST} &= \sum(y_i - \bar{y})^2
\end{aligned}$$

The p-value in the anova table is used for checking that the regression model is better than the mean $\bar{y}$. The null hypothesis is that the model is no better than the mean, that is

$$H_0 : \text{SSE} = \text{SST}$$

Under H_0 we have SSR $= 0$ hence MSR $= 0$ and

$$H_0 : F^* = 0$$

Recall that the p-value is (loosely) the probability that H_0 is true. Hence for small p-values, reject H_0 and conclude that the regression model is better than the mean.

Let us take a closer look at the mean response table.

Mean Response

TRT	N	Mean Y	95% CI MIN	95% CI MAX
A	10	5.300000	1.360937	9.239063
B	10	6.100000	2.160937	10.039063
C	10	12.300000	8.360937	16.239063

Recall that the confidence interval for a treatment mean is

$$\bar{y} \pm t(1-\alpha/2, \mathrm{dfe}) \times \mathrm{SE}, \quad \mathrm{SE} = \sqrt{\frac{\mathrm{MSE}}{n}}$$

where SE is standard error and MSE (mean square error) is estimated model variance. From the analysis of variance table at the top of the output we have

Source	DF	Sum of Squares	Mean Square
Error	27	995.10000000	36.85555556

Hence

$$\mathrm{dfe} = 27, \quad \mathrm{MSE} = 36.85555556$$

The confidence interval for the mean of treatment A can be checked by typing the following into R.

```
ybar = 5.3
n = 10
MSE = 36.85555556
dfe = 27
alpha = 0.05
SE = sqrt(MSE / n)
t = qt(1 - alpha/2, dfe) * SE
ybar - t
ybar + t
```

R prints the following results.

```
[1] 1.360937
[1] 9.239063
```

The R results match the mean response table for treatment A.

TRT	N	Mean Y	95% CI MIN	95% CI MAX
A	10	5.300000	1.360937	9.239063

Let us take a closer look at the first line of the least significant difference table.

```
                      Least Significant Difference Test

  TRT    TRT      Delta Y    95% CI MIN    95% CI MAX    t Value    Pr > |t|
  A      B      -0.800000     -6.370677      4.770677      -0.29      0.7705
```

The least significant difference of two treatment means $\bar{y}_A$ and $\bar{y}_B$ is

$$\mathrm{LSD} = t(1-\alpha/2, \mathrm{dfe}) \times \mathrm{SE}, \quad \mathrm{SE} = \sqrt{\mathrm{MSE} \times \left(\frac{1}{n_A} + \frac{1}{n_B}\right)}$$

The corresponding confidence interval is

$$(\bar{y}_A - \bar{y}_B) \pm \mathrm{LSD}$$

The confidence interval in the LSD table can be checked by typing the following into R.

```
ybarA = 5.3
ybarB = 6.1
nA = 10
nB = 10
MSE = 36.85555556
dfe = 27
alpha = 0.05
SE = sqrt(MSE * (1/nA + 1/nB))
LSD = qt(1 - alpha/2, dfe) * SE
ybarA - ybarB - LSD
ybarA - ybarB + LSD
```

R prints the following results.

```
[1] -6.370677
[1] 4.770677
```

The R results match the confidence interval in the LSD table.


```
  TRT    TRT      Delta Y    95% CI MIN    95% CI MAX    t Value    Pr > |t|
  A      B      -0.800000     -6.370677      4.770677      -0.29      0.7705
```

Let us take a closer look at the first line of the t-test table.

```
                          Two Sample t-Test

 TRT   TRT     Delta Y    95% CI MIN    95% CI MAX    t Value    Pr > |t|
 A     B     -0.800000     -5.922307      4.322307      -0.33      0.7466
```

The t-test confidence interval is

$$(\bar{y}_A - \bar{y}_B) \pm t(1 - \alpha/2, \mathrm{dfe}) \times \mathrm{SE}$$

where

$$\mathrm{SE} = \sqrt{\frac{\mathrm{SSE}}{\mathrm{dfe}} \times \left(\frac{1}{n_A} + \frac{1}{n_B}\right)}, \quad \mathrm{SSE} = \sum (y_A - \bar{y}_A)^2 + \sum (y_B - \bar{y}_B)^2$$

and

$$\mathrm{dfe} = n_A + n_B - 2$$

The confidence interval can be checked by typing the following into R.

```
yA = c(6,0,2,8,11,4,13,1,8,0)
yB = c(0,2,3,1,18,4,14,9,1,9)
nA = length(yA)
nB = length(yB)
dfe = nA + nB - 2
SSE = var(yA) * (nA - 1) + var(yB) * (nB - 1)
MSE = SSE / dfe
SE = sqrt(MSE * (1/nA + 1/nB))
alpha = 0.05
t = qt(1 - alpha/2, dfe) * SE
mean(yA) - mean(yB) - t
mean(yA) - mean(yB) + t
```

R prints the following result which matches the above t-test table.

```
[1] -5.922307
[1] 4.322307
```

R's t-test function gives the same result.

```
t.test(yA,yB,var.equal=TRUE)

Two Sample t-test

data:  yA and yB
t = -0.32812, df = 18, p-value = 0.7466
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -5.922307  4.322307
```