

Example:

A new method to determine the amount of low-calorie sweetener in different food samples has been introduced by a company. The company wants to apply this method on four food samples. The company has four labs. So the tests that involve the application of this new method to each of the food samples will be carried out in each of the four labs. Each of the labs have reported the mean recovery percentages of the amount of low-calorie sweetener they could detect on each of the food samples. The data are given below.

Lab	Food samples			
	1	2	3	4
1	99.5	83.0	96.5	96.8
2	105.0	105.5	104.0	108.0
3	95.4	81.9	87.4	86.3
4	93.7	80.8	84.5	70.3

Table 1: Mean recovery percentages for a sweetener in four food samples

It seems that different labs have different results for each sample. What we have to ascertain is if any of these results have occurred due to chance variation. To establish this, we go for two-way ANOVA without replication. Why ‘without replication’? Because from each lab we have only one value for percentage recovery (which is the mean value). If there had been more than one value for percentage recovery from each lab, for each food sample, we would have to go for two-way ANOVA with replication.

Using two-way ANOVA without replication, we are going to calculate the F statistic for both the groups (i.e. food samples/columns) and the blocks (i.e. labs/rows).

The main equations for two-way ANOVA without replication are given below with the expanded meaning of each term.

The main equations for two-way ANOVA without replication are given below with the expanded meaning of each term.

$$SST = SSG + SSB + SSE \quad (1)$$

$$F_{groups} = \frac{\left(\frac{SSG}{df_{groups}} \right)}{\left(\frac{SSE}{df_{error}} \right)} \quad (2)$$

$$F_{blocks} = \frac{\left(\frac{SSB}{df_{blocks}} \right)}{\left(\frac{SSE}{df_{error}} \right)} \quad (3)$$

SST = Sum of squares total

SSG = Sum of squares groups (i.e. columns)

SSB = Sum of squares blocks (i.e. rows)

SSE = Sum of squares error

df_{groups} = degrees of freedom groups

df_{blocks} = degrees of freedom blocks

df_{error} = degrees of freedom error

Calculating F_{groups} and F_{blocks} manually

1. Add all the 16 mean recovery values and divide by 16 to get the overall mean, $\mu_{TOT} = 92.4$
2. Find the difference between each of the 16 recovery values and the

Calculating F_{groups} and F_{blocks} manually

1. Add all the 16 mean recovery values and divide by 16 to get the overall mean, $\mu_{TOT} = 92.4$
2. Find the difference between each of the 16 recovery values and the overall mean, square the differences, and add them up to obtain SST

Lab	Sample	Recovery	Overall mean	Difference	Squared difference
1	1	99.5	92.4	7.1	50.4
	2	83.0	92.4	-9.4	88.4
	3	96.5	92.4	4.1	16.8
	4	96.8	92.4	4.4	19.4
2	1	105.0	92.4	12.6	158.8
	2	105.5	92.4	13.1	171.6
	3	104.0	92.4	11.6	134.6
	4	108.0	92.4	15.6	243.4
3	1	95.4	92.4	3.0	9.0
	2	81.9	92.4	-10.5	110.3
	3	87.4	92.4	-5.0	25.0
	4	86.3	92.4	-6.1	37.2
4	1	93.7	92.4	1.3	1.7
	2	80.8	92.4	-11.6	134.6
	3	84.5	92.4	-7.9	62.4
	4	70.3	92.4	-22.1	488.4
			Sum of squares total (SST)		1751.8

3. Find the difference between each group mean and the overall mean, square the differences, add them up, and multiply by the number of items in each group to obtain SSG

Lab	Food samples			
	1	2	3	4
1	99.5	83.0	96.5	96.8
2	105.0	105.5	104.0	108.0
3	95.4	81.9	87.4	86.3
4	93.7	80.8	84.5	70.3
	98.4	87.8	93.1	90.4

$$SSG = 4 \times \left\{ (98.4 - 92.4)^2 + (87.8 - 92.4)^2 + (93.1 - 92.4)^2 + (90.4 - 92.4)^2 \right\} = 247.4$$

4. Find the difference between each block mean and the overall mean, square the differences, add them up, and multiply by the number of items in each group to obtain SSB

Lab	Food samples				
	1	2	3	4	
1	99.5	83.0	96.5	96.8	94.0
2	105.0	105.5	104.0	108.0	105.6
3	95.4	81.9	87.4	86.3	87.8
4	93.7	80.8	84.5	70.3	82.3

$$SSB = 4 * \left\{ (94.0 - 92.4)^2 + (105.6 - 92.4)^2 + (87.8 - 92.4)^2 + (82.3 - 92.4)^2 \right\} = \\ 1201.7$$

5. Since we have already calculated SST , SSG , and SSB , we can calculate SSE using equation (1)

$$SSE = SST - SSG - SSB = 302.7$$

6. Calculating degrees of freedom

$$df_{groups} = \text{number of groups} - 1 = 4 - 1 = 3$$

$$df_{blocks} = \text{number of blocks} - 1 = 4 - 1 = 3$$

$$df_{error} = df_{groups} \times df_{blocks} = 3 \times 3 = 9$$

7. Calculating F_{groups} using equation (2)

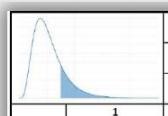
$$F_{groups} = \frac{\left(\frac{SSG}{df_{groups}} \right)}{\left(\frac{SSE}{df_{error}} \right)} = \frac{\left(\frac{247.4}{3} \right)}{\left(\frac{302.7}{9} \right)} = 2.452$$

$$F_{groups} = \frac{\left(\frac{SSG}{df_{groups}}\right)}{\left(\frac{SSE}{df_{error}}\right)} = \frac{\left(\frac{247.4}{3}\right)}{\left(\frac{302.7}{9}\right)} = 2.452$$

8. Calculating F_{blocks} using equation (3)

$$F_{blocks} = \frac{\left(\frac{SSB}{df_{blocks}}\right)}{\left(\frac{SSE}{df_{error}}\right)} = \frac{\left(\frac{1201.7}{3}\right)}{\left(\frac{302.7}{9}\right)} = 11.91$$

Since in our example, number of groups and number of blocks are 4 each, F table lookup for numerator (groups/blocks) degrees of freedom = 3 and denominator (error) degrees of freedom = 9, at alpha = 0.05 is shown below as $F_{critical} = 3.86$



		F-Table Upper Tail Area of 0.05																	
		Numerator df																	
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Denominator df	1	161.45	199.50	215.71	224.58	230.16	233.99	236.77	238.88	240.54	241.88	242.98	243.91	244.69	245.36	245.95	246.46	246.92	247.30
	2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40	19.41	19.42	19.43	19.43	19.44	19.44	19.45	19.46
	3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.76	8.74	8.73	8.71	8.70	8.69	8.68	8.67
	4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.94	5.91	5.89	5.87	5.86	5.84	5.83	5.82
	5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.70	4.68	4.66	4.64	4.62	4.60	4.59	4.58
	6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.03	4.00	3.98	3.96	3.94	3.92	3.91	3.90
	7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.60	3.57	3.55	3.53	3.51	3.49	3.48	3.47
	8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.31	3.28	3.26	3.24	3.22	3.20	3.19	3.18
	9	5.12	4.29	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.10	3.07	3.05	3.03	3.01	2.99	2.97	2.96
	10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.94	2.91	2.89	2.86	2.85	2.83	2.81	2.80
	11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.82	2.79	2.76	2.74	2.72	2.70	2.69	2.68
	12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.72	2.69	2.66	2.64	2.62	2.60	2.58	2.57
	13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.63	2.60	2.58	2.55	2.53	2.51	2.50	2.49
	14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.57	2.53	2.51	2.48	2.46	2.44	2.43	2.42
	15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.51	2.48	2.45	2.42	2.40	2.38	2.37	2.36
	16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.46	2.42	2.40	2.37	2.35	2.33	2.32	2.31
	17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.41	2.38	2.35	2.33	2.31	2.29	2.27	2.26
	18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.37	2.34	2.31	2.29	2.27	2.25	2.23	2.22
	19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.34	2.31	2.28	2.26	2.23	2.21	2.20	2.19
	20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.31	2.28	2.25	2.22	2.20	2.18	2.16	2.14
	21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.28	2.25	2.22	2.20	2.18	2.16	2.14	2.13
	22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.26	2.23	2.20	2.17	2.15	2.13	2.11	2.10
	23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27	2.24	2.20	2.18	2.15	2.13	2.11	2.09	2.08
	24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.22	2.18	2.15	2.13	2.11	2.09	2.07	2.06
	25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24	2.20	2.16	2.14	2.11	2.09	2.07	2.05	2.04
	26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22	2.18	2.15	2.12	2.09	2.07	2.05	2.03	2.02
	27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25	2.20	2.17	2.13	2.10	2.08	2.06	2.04	2.02	2.01
	28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19	2.15	2.12	2.09	2.06	2.04	2.02	2.00	1.99

Conclusion

Null hypothesis for groups/food samples: $H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$

But $F_{groups} = 2.452$ is smaller than $F_{critical} = 3.86$. Therefore we cannot reject the null hypothesis which means that there is insufficient evidence to conclude that the recovery depends on the sample type.

Null hypothesis for blocks/labs: $H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$

$F_{blocks} = 11.91$ far exceeds $F_{critical} = 3.86$. Therefore we reject the null hypothesis for the blocks. This means that recoveries for different labs are significantly different at the 95% level of confidence.