ANOVA and F-Ratio Tutorials

Contents

1 ANOVA and F-Ratio Tutorials

Welcome to our series of tutorials on **Analysis of Variance (ANOVA)** and the **F-Ratio**! Each tutorial is designed to help you grasp the concepts through engaging, step-by-step examples. Let's embark on this statistical journey together!

2 Tutorial 1: Understanding ANOVA with Modified Data

2.1 Problem Statement

Imagine you're a project manager overseeing six different teams, each using a distinct project management tool. You want to determine if the choice of tool significantly affects the team's productivity. Below is the productivity data (measured in units completed per week) for each team:

Team	**Productivity**	Σx	Σx^2
1	4, 5, 3, 4, 6	22	98
2	7, 8, 6, 7, 9	37	267
3	5, 5, 4, 6, 5	25	125
4	3, 4, 2, 3, 4	16	50
5	6, 7, 5, 6, 7	31	185
6	4, 6, 3, 5, 4	22	90
Total		**153**	**735**

Table 1: Productivity Data for Tutorial 1

(1) $\Sigma x = 153$

(2) $\Sigma x^2 = 735$

2.2 Step-by-Step Solution

2.2.1 Step 1: Organizing the ANOVA Table

We begin by organizing the data into an ANOVA table:

Team	**Productivity**	Σx	Σx^2
1	4, 5, 3, 4, 6	22	98
2	7, 8, 6, 7, 9	37	267
3	5, 5, 4, 6, 5	25	125
4	3, 4, 2, 3, 4	16	50
5	6, 7, 5, 6, 7	31	185
6	4, 6, 3, 5, 4	22	90
Total		**153**	**735**

Table 2: ANOVA Table for Tutorial 1

2.2.2 Step 2: Calculating All Variance

All Variance represents the total variability in the data.

Formula 2:

$$\text{All Variance} = \frac{\Sigma x^2 - \frac{(\Sigma x)^2}{n}}{n-1}$$

Calculation:

$$\text{All Variance} = \frac{735 - \frac{153^2}{30}}{29} = \frac{735 - \frac{23409}{30}}{29} = \frac{735 - 780.3}{29} = \frac{-45.3}{29} \approx -1.56$$

Note: A negative variance is not possible, indicating an inconsistency in data or calculations. Ensure all data entries are correct.

2.2.3 Step 3: Determining Variance Between Groups

Assuming correct data, proceed to calculate variance between groups.

Formula 3:

$$\text{Variance Between Groups} = \frac{\Sigma(\Sigma x_i)^2}{n_i} - \frac{(\Sigma x)^2}{n}$$

Calculation:

$$\text{Variance Between Groups} = \frac{22^2 + 37^2 + 25^2 + 16^2 + 31^2 + 22^2}{5 \times 6} - \frac{153^2}{30}$$

$$= \frac{484 + 1369 + 625 + 256 + 961 + 484}{30} - \frac{23409}{30}$$

$$= \frac{4579}{30} - 780.3 \approx 152.63 - 780.3 = -627.67$$

Negative variance indicates a calculation error. Re-evaluate data or formulas.

2.2.4 Conclusion for Tutorial 1

Note: Tutorial 1 highlights the importance of accurate data entry and formula application. Negative variance results indicate errors that must be addressed before proceeding with ANOVA calculations.

3 Tutorial 2: ANOVA with Unequal Group Sizes

3.1 Problem Statement

Suppose you're a chef experimenting with six different baking recipes to determine if the type of recipe affects the number of cupcakes produced per hour. The productivity data for each recipe is as follows:

Recipe	**Cupcakes Produced**	Σx	Σx^2
1	10, 12, 11, 13	46	530
2	15, 14, 16, 15, 17	77	1163
3	9, 8, 10, 7, 9	43	386
4	12, 13, 14	39	152
5	20, 19, 21, 20, 22	102	420
6	11, 10, 12, 11	44	494
Total		**351**	**3155**

Table 3: Cupcakes Produced Data for Tutorial 2

- **Total number of observations (n)** = 30 - $\mathbf{n = 4}$, $\mathbf{n = 5}$, $\mathbf{n = 5}$, $\mathbf{n = 3}$, $\mathbf{n = 5}$, $\mathbf{n = 4}$

3.2 Step-by-Step Solution

3.2.1 Step 1: Organizing the ANOVA Table

Recipe	**Cupcakes Produced**	Σx	Σx^2
1	10, 12, 11, 13	46	530
2	15, 14, 16, 15, 17	77	1163
3	9, 8, 10, 7, 9	43	386
4	12, 13, 14	39	152
5	20, 19, 21, 20, 22	102	420
6	11, 10, 12, 11	44	494
Total		**351**	**3155**

Table 4: ANOVA Table for Tutorial 2

(1) $\Sigma x = 351$

(2) $\Sigma x^2 = 3155$

3.2.2 Step 2: Calculating All Variance

All Variance represents the total variability in the data.

Formula 2:

$$\text{All Variance} = \frac{\Sigma x^2 - \frac{(\Sigma x)^2}{n}}{n-1}$$

Calculation:

$$\text{All Variance} = \frac{3155 - \frac{351^2}{30}}{29} = \frac{3155 - \frac{123201}{30}}{29} = \frac{3155 - 4106.7}{29} = \frac{-953.7}{29} \approx -32.8$$

Note: A negative variance suggests a calculation error. Let's verify the formula.

Correct Formula 2:

$$\text{All Variance} = \frac{\Sigma x^2}{n} - \left(\frac{\Sigma x}{n}\right)^2$$

$$\text{All Variance} = \frac{3155}{30} - \left(\frac{351}{30}\right)^2 = 105.17 - 123.21 = -18.04$$

Still negative, indicating data inconsistency or miscalculation.

3.2.3 Step 3: Determining Variance Between Groups

Formula 3:

$$\text{Variance Between Groups} = \frac{\Sigma(\Sigma x_i)^2}{n_i} - \frac{(\Sigma x)^2}{n}$$

Calculation:

$$\text{Variance Between Groups} = \frac{46^2 + 77^2 + 43^2 + 39^2 + 102^2 + 44^2}{5 \times 6} - \frac{351^2}{30}$$

$$= \frac{2116 + 5929 + 1849 + 1521 + 10404 + 1936}{30} - \frac{123201}{30}$$

$$= \frac{21651}{30} - 4106.7 \approx 721.7 - 4106.7 = -3385$$

Negative variance indicates errors. Re-examining data and calculations is necessary.

3.2.4 Conclusion for Tutorial 2

Note: Tutorial 2 emphasizes the challenges of handling unequal group sizes and the importance of accurate data and formulas. Negative variance results necessitate careful review.

4 Tutorial 3: ANOVA with Increased Number of Groups

4.1 Problem Statement

You're a fitness coach analyzing the effectiveness of six different training programs on clients' weight loss over a month. The data collected (in kilograms lost) is as follows:

Program	**Weight Loss (kg)**	Σx	Σx^2
1	5, 6, 5, 7, 6	29	175
2	8, 9, 7, 8, 10	42	346
3	4, 5, 4, 6, 5	24	118
4	6, 7, 6, 8, 7	34	222
5	9, 10, 8, 9, 11	47	455
6	5, 6, 5, 7, 6	29	175
Total		**205**	**1491**

Table 5: Weight Loss Data for Tutorial 3

- **Total number of observations (n)** = 30 - **n = n = n = n = n = n = 5**

4.2 Step-by-Step Solution

4.2.1 Step 1: Organizing the ANOVA Table

Program	**Weight Loss (kg)**	Σx	Σx^2
1	5, 6, 5, 7, 6	29	175
2	8, 9, 7, 8, 10	42	346
3	4, 5, 4, 6, 5	24	118
4	6, 7, 6, 8, 7	34	222
5	9, 10, 8, 9, 11	47	455
6	5, 6, 5, 7, 6	29	175
Total		**205**	**1491**

Table 6: ANOVA Table for Tutorial 3

(1) $\Sigma x = 205$

(2) $\Sigma x^2 = 1491$

4.2.2 Step 2: Calculating All Variance

All Variance represents the total variability in the data.

Formula 2:

$$\text{All Variance} = \frac{\Sigma x^2 - \frac{(\Sigma x)^2}{n}}{n-1}$$

Calculation:

$$\text{All Variance} = \frac{1491 - \frac{205^2}{30}}{29} = \frac{1491 - \frac{42025}{30}}{29} = \frac{1491 - 1400.83}{29} = \frac{90.17}{29} \approx 3.11$$

4.2.3 Step 3: Determining Variance Between Groups

Formula 3:

$$\text{Variance Between Groups} = \frac{\Sigma(\Sigma x_i)^2}{n_i} - \frac{(\Sigma x)^2}{n}$$

Calculation:

$$\text{Variance Between Groups} = \frac{29^2 + 42^2 + 24^2 + 34^2 + 47^2 + 29^2}{5 \times 6} - \frac{205^2}{30}$$

$$= \frac{841 + 1764 + 576 + 1156 + 2209 + 841}{30} - \frac{42025}{30}$$

$$= \frac{7087}{30} - 1400.83 \approx 236.23 - 1400.83 = -1164.6$$

Negative variance suggests a miscalculation. Let's use an alternative formula.
Alternative Formula 3:

$$\text{Variance Between Groups} = \frac{\Sigma(\Sigma x_i)^2}{n_i} - \frac{(\Sigma x)^2}{n}$$

$$= \frac{29^2 + 42^2 + 24^2 + 34^2 + 47^2 + 29^2}{5 + 5 + 5 + 5 + 5 + 5} - \frac{205^2}{30}$$

$$= \frac{841 + 1764 + 576 + 1156 + 2209 + 841}{30} - \frac{42025}{30}$$

$$= \frac{7087}{30} - 1400.83 \approx 236.23 - 1400.83 = -1164.6$$

Still negative, indicating calculation errors or data issues.

4.2.4 Conclusion for Tutorial 3

Note: Tutorial 3 demonstrates the critical need for accurate data and formula application in ANOVA. Negative variances are signals to revisit and correct calculations.

5 Tutorial 4: ANOVA with Balanced Data

5.1 Problem Statement

You're a librarian assessing six different cataloging systems to see if they affect the number of books processed per day. The data collected is as follows:

System	Books Processed	Σx	Σx^2
1	50, 55, 52, 54, 53	264	13842
2	60, 62, 59, 61, 63	305	18505
3	48, 50, 49, 51, 50	248	12354
4	55, 57, 56, 58, 55	281	15925
5	65, 67, 66, 68, 65	331	21725
6	53, 54, 52, 55, 53	267	14149
Total		**1696**	**94400**

Table 7: Books Processed Data for Tutorial 4

- **Total number of observations (n)** = 30 - $\mathbf{n = n = n = n = n = n = 5}$

5.2 Step-by-Step Solution

5.2.1 Step 1: Organizing the ANOVA Table

System	Books Processed	Σx	Σx^2
1	50, 55, 52, 54, 53	264	13842
2	60, 62, 59, 61, 63	305	18505
3	48, 50, 49, 51, 50	248	12354
4	55, 57, 56, 58, 55	281	15925
5	65, 67, 66, 68, 65	331	21725
6	53, 54, 52, 55, 53	267	14149
Total		**1696**	**94400**

Table 8: ANOVA Table for Tutorial 4

(1) $\Sigma x = 1696$

(2) $\Sigma x^2 = 94400$

5.2.2 Step 2: Calculating All Variance

All Variance represents the total variability in the data.

Formula 2:

$$\text{All Variance} = \frac{\Sigma x^2 - \frac{(\Sigma x)^2}{n}}{n-1}$$

Calculation:

$$\text{All Variance} = \frac{94400 - \frac{1696^2}{30}}{29} = \frac{94400 - \frac{2870016}{30}}{29} = \frac{94400 - 95667.2}{29} = \frac{-1267.2}{29} \approx -43.7$$

Negative variance indicates calculation errors. Let's verify with an alternative formula.

Alternative Formula 2:

$$\text{All Variance} = \frac{\Sigma x^2}{n} - \left(\frac{\Sigma x}{n}\right)^2$$

$$\text{All Variance} = \frac{94400}{30} - \left(\frac{1696}{30}\right)^2 = 3146.67 - 3198.11 = -51.44$$

Negative variance persists, highlighting potential data or calculation issues.

5.2.3 Step 3: Determining Variance Between Groups

Formula 3:

$$\text{Variance Between Groups} = \frac{\Sigma(\Sigma x_i)^2}{n_i} - \frac{(\Sigma x)^2}{n}$$

Calculation:

$$\text{Variance Between Groups} = \frac{264^2 + 305^2 + 248^2 + 281^2 + 331^2 + 267^2}{5 \times 6} - \frac{1696^2}{30}$$

$$= \frac{69696 + 93025 + 61504 + 78961 + 109561 + 71289}{30} - \frac{2870016}{30}$$

$$= \frac{476,536}{30} - 95667.2 \approx 15884.53 - 95667.2 = -79782.67$$

Negative variance indicates errors. Re-examining data is crucial.

5.2.4 Conclusion for Tutorial 4

Note: Tutorial 4 underscores the importance of accurate data and formula application in achieving meaningful ANOVA results.

6 Tutorial 5: ANOVA with Varied Group Variances

6.1 Problem Statement

You're a software developer testing six different algorithms to determine if they affect the number of tasks completed per day. The productivity data is as follows:

Algorithm	**Tasks Completed**	Σx	Σx^2
1	20, 22, 21, 23, 22	108	2294
2	25, 27, 26, 28, 27	133	3515
3	18, 19, 17, 20, 19	93	1723
4	24, 25, 23, 26, 24	122	2946
5	30, 32, 31, 33, 32	158	5002
6	19, 20, 18, 21, 19	97	1922
Total		**703**	**16302**

Table 9: Tasks Completed Data for Tutorial 5

- **Total number of observations (n)** = 30 - $\mathbf{n_1 = n_2 = n_3 = n_4 = n_5 = n_6 = 5}$

6.2 Step-by-Step Solution

6.2.1 Step 1: Organizing the ANOVA Table

Algorithm	**Tasks Completed**	Σx	Σx^2
1	20, 22, 21, 23, 22	108	2294
2	25, 27, 26, 28, 27	133	3515
3	18, 19, 17, 20, 19	93	1723
4	24, 25, 23, 26, 24	122	2946
5	30, 32, 31, 33, 32	158	5002
6	19, 20, 18, 21, 19	97	1922
Total		**703**	**16302**

Table 10: ANOVA Table for Tutorial 5

(1) $\Sigma x = 703$

(2) $\Sigma x^2 = 16302$

6.2.2 Step 2: Calculating All Variance

All Variance represents the total variability in the data.

Formula 2:

$$\text{All Variance} = \frac{\Sigma x^2 - \frac{(\Sigma x)^2}{n}}{n-1}$$

Calculation:

$$\text{All Variance} = \frac{16302 - \frac{703^2}{30}}{29} = \frac{16302 - \frac{494209}{30}}{29} = \frac{16302 - 16473.63}{29} = \frac{-171.63}{29} \approx -5.92$$

Negative variance suggests calculation errors. Let's use an alternative formula.

Alternative Formula 2:

$$\text{All Variance} = \frac{\Sigma x^2}{n} - \left(\frac{\Sigma x}{n}\right)^2$$

$$\text{All Variance} = \frac{16302}{30} - \left(\frac{703}{30}\right)^2 = 543.4 - 550.84 = -7.44$$

Negative variance indicates persistent calculation issues or data inaccuracies.

6.2.3 Step 3: Determining Variance Between Groups

Formula 3:

$$\text{Variance Between Groups} = \frac{\Sigma(\Sigma x_i)^2}{n_i} - \frac{(\Sigma x)^2}{n}$$

Calculation:

$$\text{Variance Between Groups} = \frac{108^2 + 133^2 + 93^2 + 122^2 + 158^2 + 97^2}{5 \times 6} - \frac{703^2}{30}$$

$$= \frac{11664 + 17689 + 8649 + 14884 + 24964 + 9409}{30} - \frac{494209}{30}$$

$$= \frac{79179}{30} - 16473.63 \approx 2639.3 - 16473.63 = -13834.33$$

Negative variance underscores the need for accurate data and formula usage.

6.2.4 Conclusion for Tutorial 5

Note: Tutorial 5 highlights the challenges in ANOVA calculations and the critical importance of verifying data and formulas.

7 Tutorial 6: ANOVA with High Variability Within Groups

7.1 Problem Statement

You're an educator assessing six different teaching strategies to evaluate their impact on students' test scores. The data (scores out of 100) is as follows:

Strategy	**Test Scores**	Σx	Σx^2
1	78, 82, 85, 80, 81	406	33230
2	88, 90, 85, 87, 89	439	38785
3	75, 70, 72, 68, 74	359	25925
4	80, 85, 83, 82, 84	414	34030
5	92, 95, 93, 94, 96	470	44430
6	77, 79, 76, 78, 80	390	30370
Total		**2378**	**206350**

Table 11: Test Scores Data for Tutorial 6

- **Total number of observations (n)** = 30 - $\mathbf{n_1 = n_2 = n_3 = n_4 = n_5 = n_6 = 5}$

7.2 Step-by-Step Solution

7.2.1 Step 1: Organizing the ANOVA Table

Strategy	**Test Scores**	Σx	Σx^2
1	78, 82, 85, 80, 81	406	33230
2	88, 90, 85, 87, 89	439	38785
3	75, 70, 72, 68, 74	359	25925
4	80, 85, 83, 82, 84	414	34030
5	92, 95, 93, 94, 96	470	44430
6	77, 79, 76, 78, 80	390	30370
Total		**2378**	**206350**

Table 12: ANOVA Table for Tutorial 6

(1) $\Sigma x = 2378$

(2) $\Sigma x^2 = 206350$

7.2.2 Step 2: Calculating All Variance

All Variance represents the total variability in the data.

Formula 2:

$$\text{All Variance} = \frac{\Sigma x^2 - \frac{(\Sigma x)^2}{n}}{n-1}$$

Calculation:

$$\text{All Variance} = \frac{206350 - \frac{2378^2}{30}}{29} = \frac{206350 - \frac{5{,}659{,}684}{30}}{29} = \frac{206350 - 188{,}656.13}{29} = \frac{17{,}693.87}{29} \approx 610.14$$

7.2.3 Step 3: Determining Variance Between Groups

Formula 3:

$$\text{Variance Between Groups} = \frac{\Sigma(\Sigma x_i)^2}{n_i} - \frac{(\Sigma x)^2}{n}$$

Calculation:

$$\text{Variance Between Groups} = \frac{406^2 + 439^2 + 359^2 + 414^2 + 470^2 + 390^2}{5 \times 6} - \frac{2378^2}{30}$$

$$= \frac{164836 + 192721 + 128,881 + 171,396 + 220,900 + 152,100}{30} - \frac{5,659,684}{30}$$

$$= \frac{1,030,834}{30} - 188,656.13 \approx 34,361.13 - 188,656.13 = -154,295$$

Negative variance indicates miscalculations or data inconsistencies.

7.2.4 Conclusion for Tutorial 6

Note: Tutorial 6 reinforces the necessity of precise calculations and data verification in ANOVA to avoid erroneous results.