2. RELATIONSHIP BETWEEN A QUALITATIVE AND A QUANTITATIVE VARIABLE

Design and Data Analysis in Psychology II

Susana Sanduvete Chaves
SalvadorChacón Moscoso

1

1. INTRODUCTION

- You may examine gender differences in average salary or racial (white versus black) differences in average annual income.
- The variable (salary) to be tested should be interval or ratio (quantitative), whereas the gender or race variable should be binary (qualitative).

2

1. INTRODUCTION

- These analyses are used to compare group means (independent or dependent groups).
- The term *independent* is used because groups are conformed randomly (independent samples). Two separate sets of independent and identically distributed samples are obtained.

1. INTRODUCTION

- The term dependent is used because groups are paired with respect to some measure (dependent samples). Dependent samples (or "paired") *comparison* consist of a sample of matched pairs of similar units, or one group of units that has been tested twice.

1. INTRODUCTION

- A typical example of the repeated measures would be where subjects are tested prior to a treatment, say for high blood pressure, and the same subjects are tested again after treatment with a blood-pressure lowering medication.

1. INTRODUCTION

- Three steps to carry out:
  1. Testing statistical assumptions.
  2. Statistical tests of significance.
  3. Statistical power.

2. STEP 1. TESTING ASSUMPTIONS

- **Quantitative variable**: Interval or ratio variable.
- **Normal distribution**: the populations from which the samples are selected must be normal (Tests for Normality: Kolmogorov-Smirnov 'D' + Histogram of predicted 'Z' scores).

7

2. STEP 1. TESTING ASSUMPTIONS

The Central Limit Theorem says, however, that the distributions of y_1 *and* y_2 *are approximately normal when N is large.*

In practice, when $n_1 + n_2 \geq 30$, you do not need to worry too much about the normality assumption.

8

2. STEP 1. TESTING ASSUMPTIONS

- **Independence**: the observations within each sample must be independent (Durbin-Watson 'D' + scatter plot X-e).
- **Homoscedasticity** (homogeneity of variance): the populations from which the samples are selected must have equal variances. In this kind of studies, it is the most important assumption to take into account (Levene's test, F Max' + scatter plot y'-absolute errors).

9

2. STEP 1. TESTING ASSUMPTIONS

- To test homoscedasticity:
  - Levene's test (SPSS).
  - F-Max test:

$$\text{Fmax} = \frac{S^2 (largest)}{S^2 (smallest)}$$

  - $F \leq F_{(\alpha,\, k,\, n-1)}$ → Null hypothesis is accepted. The homogeneity of variance assumption has been satisfied.
  - $F > F_{(\alpha,\, k,\, n-1)}$ → Null hypothesis is rejected. The homogeneity of variance assumption has been violated.

    * k = number of groups (number of variances)

    n = number of participants in each group (same size is assumed)

10

2. STEP 1. TESTING ASSUMPTIONS: EXAMPLE

Test the homogeneity of variance assumption in the following data (A = nationality; Y = level of depression; α=0.05):

a_1: Spanish	a_2: Japanese
7	8
5	7
6	5
7	7
5	6
8	9
4	8
3	10
4	7
2	8

11

2. STEP 1. TESTING ASSUMPTIONS: EXAMPLE

a_1	a_2:	a^2_1	a^2_2
7	8	49	64
5	7	25	49
6	5	36	25
7	7	49	49
5	6	25	36
8	9	64	81
4	8	16	64
3	10	9	100
4	7	16	49
2	8	4	64
Σ=51	Σ=75	Σ=293	Σ=581

$$\overline{Y}_1 = \frac{\sum Y_1}{N} = \frac{51}{10} = 5.1$$

$$\overline{Y}_2 = \frac{\sum Y_2}{N} = \frac{75}{10} = 7.5$$

$$S^2_{Y_1} = \frac{\sum Y_1^2}{N} - \overline{Y}_1^2 = \frac{293}{10} - 5.1^2 = 3.29$$

$$S^2_{Y_2} = \frac{\sum Y_2^2}{N} - \overline{Y}_2^2 = \frac{581}{10} - 7.5^2 = 1.85$$

12

2. STEP 1. TESTING ASSUMPTIONS: EXAMPLE

$$\text{Fmax} = \frac{S^2(largest)}{S^2(smallest)} = \frac{3.29}{1.85} = 1.778$$

$F_{(\alpha,\, k,\, n-1)} = F_{(0.05,\, 2,\, 10-1)} = 4.43$

$F \leq F_{(\alpha,\, k,\, n-1)}$ → $1.778 \leq 4.43$ → Null hypothesis is accepted. The homogeneity of variance assumption has been satisfied.

13

3. Step 2. Statistical Tests of Significance

3.1. Step 2. Non-parametric Tests of Significance: assumptions are rejected

3.2. Step2. Parametric tests of significance: assumptions are accepted

14

3.1. STEP 2: NON-PARAMETRIC TESTS

3.1.1. Wilcoxon T test.

- When the population cannot be assumed to be normally distributed.
- Quantitative dependent variable.
- Two paired groups.
- Formula:

$$T_{+} = \sum R_i$$

$$T_{-} = \sum R_i$$

$$T = \text{the smallest of these two rank sums}$$

15

3.1.1. Wilcoxon T test

- Conclusions:
  - $T \geq T_{(\alpha, N)}$ → The null hypothesis is accepted. The variables are not related. There are not differences between groups.
  - $T < T_{(\alpha, N)}$ → The null hypothesis is rejected. The variables are related. There are differences between groups.

* N = number of pairs removing those which difference is 0.

16

3.1.1. Wilcoxon T test: example

In the table below, you can see the punctuation in social skills that a twin who went to the kindergarten (a_1) for a course, and the other twin who stayed at home (a_2). Are there statistical differences between both groups in social skills? (α=0.05).

a_1	a_2
82	63
69	42
73	74
43	37
58	51
56	43
76	80
65	82
73	53
66	66

3.1.1. Wilcoxon T test: example

a_1	a_2	d	Rank	Sign
82	63	19	7	+
69	42	27	9	+
73	74	-1	1	-
43	37	6	3	+
58	51	7	4	+
56	43	13	5	+
76	80	-4	2	-
65	82	-17	6	-
73	53	20	8	+
66	66	0		

3.1.1. Wilcoxon T test: example

$$T_+ = 36$$
$$T_- = \sum R_i = 9$$
$$T = 9$$

$T_{(\alpha, N)} = T_{(0.05, 9)} = 5$
$T > T_{(\alpha, N)} \rightarrow 9 > 5 \rightarrow$ The null hypothesis is accepted. The variables are not related. There are not statistical differences between groups in social skills.

3.1.2. Mann-Whitney U

- Quantitative or ordinal dependent variable.
- Two independent groups.
- Calculation:

$$U = n_1 n_2 + \frac{n_1(n_1+1)}{2} - \sum R_1$$
$$U = n_1 n_2 + \frac{n_2(n_2+1)}{2} - \sum R_2$$

U chosen = the smallest one

n_1 = smallest sample; n_2 = largest sample

3.1.2. Mann-Whitney U

- Conclusions:
  - $U \geq U_{(\alpha, n2, n1)}$ → The null hypothesis is accepted. The variables are not related. There are not statistical differences between groups.
  - $U < U_{(\alpha, n2, n1)}$ → The null hypothesis is rejected. The variables are related. There are statistical differences between groups.

21

3.1.2. Mann-Whitney U: example

We want to study if two different reading methods imply different results in reading qualifications (α=0.05).

Type of method	Qualifications				
Traditional (a_1)	80	85	25	70	90
New (a_2)	95	100	93	110	45

22

3.1.2. Mann-Whitney U: example

Type of method	Qualifications					ΣR
Traditional (a_1)	80 (4)	85 (5)	25 (1)	70 (3)	90 (6)	19
New (a_2)	95 (8)	100 (9)	93 (7)	110 (10)	45 (2)	36

*Values into brackets = ranks (orders)

$$U = n_1 n_2 + \frac{n_1(n_1+1)}{2} - \sum R_1 = 5*5 + \frac{5(5+1)}{2} - 19 = 21$$

$$U = n_1 n_2 + \frac{n_2(n_2+1)}{2} - \sum R_2 = 5*5 + \frac{5(5+1)}{2} - 36 = 4$$

U = 4

3.1.2. Mann-Whitney U: example

$U > U_{(0.05, 5, 5)}$ → 4 > 2 → The null hypothesis is accepted. The variables are not related. There are not statistical differences in reading qualifications depending on the reading method.

3.1.3. Kruskal Wallis H

- Quantitative or ordinal dependent variable.
- k independent groups.
- Formula:

$$H = \frac{12}{N(N+1)} \sum_{j=1}^{k} \frac{R_j^2}{n_j} - 3(N+1)$$

25

3.1.3. Kruskal Wallis H

- Conclusions:
  - $H \leq H_{(\alpha, k, n1, n2, n3)}$ → The null hypothesis is accepted. The variables are not related. There are not statistical differences between groups.
  - $H > H_{(\alpha, k, n1, n2, n3)}$ → The null hypothesis is rejected. The variables are related. There are statistical differences between groups.

26

3.1.3. Kruskal Wallis H: example

Do exist statistical differences in levels of authoritarianism between teachers from three different types of schools? (α=0.05)

State school (a_1)	Private school (a_2)	State-subsidized school (a_3)
96	82	115
128	124	149
83	132	166
61	135	147
101	109	

3.1.3. Kruskal Wallis H: example

State school (a_1)	Private school (a_2)	State-subsidized school (a_3)
96 (4)	82 (2)	115 (7)
128 (9)	124 (8)	149 (13)
83 (3)	132 (10)	166 (14)
61 (1)	135 (11)	147 (12)
101 (5)	109 (6)	
$\Sigma R_1 = 22$	$\Sigma R_2 = 37$	$\Sigma R_3 = 46$

*Values into brackets = ranks (orders)

3.1.3. Kruskal Wallis H: example

$$H = \frac{12}{N(N+1)} \sum_{j=1}^{k} \frac{R_j^2}{n_j} - 3(N+1) =$$

$$\frac{12}{14(14+1)} \left(\frac{(22)^2}{5} + \frac{(37)^2}{5} + \frac{(46)^2}{4} \right) - 3(14+1) = 6.4$$

$H > H_{(\alpha,\ k,\ n1,\ n2,\ n3)}$ → 6.4 > 5.666 → The null hypothesis is rejected. The variables are related. There are statistical differences in levels of authoritarianism between teachers from different schools.

29

3.2. STEP 2. PARAMETRIC TESTS OF SIGNIFICANCE

3.2.1. Linear regression analysis: qualitative independent variable with:

- Two groups: **lesson 3** (simple linear regression).
- More than two groups: **lesson 4** (multiple linear regression).

3.2.2. T-tests (independent and dependent samples).

3.2.3. One-way ANOVA ‘F’.

30

3.2.2. T-tests

- Two groups.
- Quantitative dependent variable (Interval or ratio variable).
- Different situations:
  - Independent two-sample t-test.
    - Equal sample sizes, equal variances in the distribution.
    - Equal sample sizes, unequal variances in the distribution.
    - Unequal sample sizes, equal variances in the distribution.
    - Unequal sample sizes, unequal variances in the distribution.
  - Dependent t-test for paired samples.

31

3.2.2. T-tests: significance in all cases

- $t > t_{(theoretical)}$ → Null hypothesis is rejected. The model is valid. The slope is statistically different from 0. There is, therefore, relationship between variables.

- $t \leq t_{(theoretical)}$ → Null hypothesis is accepted. The model is not valid. The slope is statistically equal to 0. There is not, therefore, relationship between variables.

32

3.2.2. T-tests: two independent samples

Equal sample sizes, equal variances

$$t = \frac{\overline{X}_1 - \overline{X}_2}{S_{X_1X_2}\sqrt{\frac{2}{n}}}$$

Where the grand standard deviation or pooled standard deviation is

$$S_{X_1X_2} = \sqrt{\frac{S_{X_1}^2 + S_{X_2}^2}{2}}$$

The degrees of freedom for this test is N – 2.

33

3.2.2. T-tests: two independent samples

Equal sample sizes, equal variances: example

A researcher wants to study the effect of depriving rats of water, and the time they spend in covering a labyrinth. 60 rats were assigned randomly to 12-hour deprivation group (sample A) and 24-hour deprivation group (sample B). The 30 rats in group A spent a mean of 5.9 minutes in covering the labyrinth, whereas the 30 rats in group B spent a mean of 6.4 minutes. The standard deviation were 0.29 and 0.3 respectively. Assuming that the two distributions have the same variance, did statistical differences between groups exist? (α=0.05)

34

3.2.2. T-tests: two independent samples

Equal sample sizes, equal variances: example

$$t = \frac{\overline{X}_1 - \overline{X}_2}{S_{X_1X_2}\sqrt{\frac{2}{n}}} = \frac{5.9 - 6.4}{0.295\sqrt{\frac{2}{30}}} = \frac{-0.5}{0.295\sqrt{0.067}} = \frac{-0.5}{0.295 * 0.259} = \frac{-0.5}{0.076} = 6.579$$

$$S_{X_1X_2} = \sqrt{\frac{S_{X_1}^2 + S_{X_2}^2}{2}} = \sqrt{\frac{0.29^2 + 0.3^2}{2}} = \sqrt{\frac{0.084 + 0.09}{2}} = \sqrt{\frac{0.174}{2}} = \sqrt{0.087} = 0.295$$

$$t > t_{(\alpha, N-2)} \rightarrow t > t_{(0.05, 58)} \rightarrow 6.579 > 2$$

Null hypothesis is rejected. There were statistical differences between groups

35

3.2.2. T-tests: two independent samples

Equal sample sizes, unequal variances

$$t = \frac{\overline{X}_1 - \overline{X}_2}{S_{X_1X_2}\sqrt{\frac{2}{n}}}$$

Where the grand standard deviation or pooled standard deviation is

$$S_{X_1X_2} = \sqrt{\frac{S_{X_1}^2 + S_{X_2}^2}{2}}$$

The degrees of freedom for this test is:

$$d.f. = \frac{\left(S_1^2 / n_1 + S_2^2 / n_2\right)^2}{(S_1^2 / n_1)^2 / (n_1 - 1) + (S_2^2 / n_2)^2 / (n_2 - 1)}$$

36

3.2.2. T-tests: two independent samples

Equal sample sizes, unequal variances: example

If in the previous example about rats, we assume that the variances of the two distributions were unequal, would statistical differences between variances exist? (α=0.05)

3.2.2. T-tests: two independent samples

Equal sample sizes, unequal variances: example

$$t = 6.579$$

$$d.f. = \frac{\left(S_1^2/n_1 + S_2^2/n_2\right)^2}{(S_1^2/n_1)^2/(n_1-1) + (S_2^2/n_2)^2/(n_2-1)} = \frac{\left(0.29^2/30 + 0.3^2/30\right)^2}{(0.29^2/30)^2/(30-1) + (0.3^2/30)^2/(30-1)} = 57.933$$

$$t > t_{(\alpha, d.f.)} \rightarrow t > t_{(0.05, 58)} \rightarrow 6.579 > 2$$

Null hypothesis is rejected. There were statistical differences between groups

3.2.2. T-tests (two independent samples)

Unequal sample sizes, equal variances

$$t = \frac{\overline{X}_1 - \overline{X}_2}{S_{X_1X_2}\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$$

where

$$S_{X_1X_2} = \sqrt{\frac{(n_1-1)S_{X_1}^2 + (n_2-1)S_{X_2}^2}{n_1+n_2-2}}$$

The degrees of freedom for this test is $n_1 + n_2 - 2$

39

3.2.2. T-tests (two independent samples)

Unequal sample sizes, equal variances: example

A researcher wanted to study the level of relax and the flexibility. 9 gymnasts were assigned randomly to two groups. The first group (experimental) participated in a 14-week training program with relax and flexibility exercises; the second group (control) participated in the same program, but without relax exercises. The measures in flexibility obtained were the following:

E: 130 127 142 145 135

C: 110 125 120 115

Variances are 58.7 in E and 41.67 in C. Populations present equal variances. Are there statistical differences between groups? (α=0.05)

40

3.2.2. T-tests (two independent samples)

Unequal sample sizes, equal variances

$$t = \frac{\overline{X}_1 - \overline{X}_2}{S_{X_1X_2}\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} = \frac{135.8-117.5}{7.169\sqrt{\frac{1}{5}+\frac{1}{4}}} = \frac{18.3}{7.169\sqrt{0.2+0.25}} = 3.805$$

$$\overline{X}_E = \frac{\sum X_E}{n_E} = \frac{679}{5} = 135.8$$

$$\overline{X}_C = \frac{\sum X_C}{n_C} = \frac{470}{4} = 117.5$$

3.2.2. T-tests (two independent samples)

Unequal sample sizes, equal variances

$$S_{X_1X_2} = \sqrt{\frac{(n_1-1)S_{X_1}^2+(n_2-1)S_{X_2}^2}{n_1+n_2-2}} = \sqrt{\frac{(5-1)58.7+(4-1)41.67}{5+4-2}}$$

$$S_{X_1X_2} = \sqrt{\frac{4*58.7+3*41.67}{7}} = \sqrt{\frac{234.8+125.01}{7}} = \sqrt{\frac{359.81}{7}} = \sqrt{51.401} = 7.169$$

$$t > t_{(\alpha,n_1+n_2-2)} \rightarrow t > t_{(0.05,7)} \rightarrow 3.805 > 2.365$$

Null hypothesis is rejected. There were statistical differences between groups.

3.2.2. T-tests (two independent samples)

Unequal sample sizes, unequal variances

$$t = \frac{\overline{X}_1 - \overline{X}_2}{S_{\overline{X}_1 - \overline{X}_2}}$$

Where the not pooled standard deviation is:

$$S_{\overline{X}_1 - \overline{X}_2} = \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$$

$$d.f. = \frac{\left(S_1^2 / n_1 + S_2^2 / n_2\right)^2}{(S_1^2 / n_1)^2 / (n_1 - 1) + (S_2^2 / n_2)^2 / (n_2 - 1)}$$

43

3.2.2. T-tests (two independent samples)

Unequal sample sizes, unequal variances: example

Supposing that in the previous example about gymnasts, the populations present unequal variances, are there differences between groups? (α=0.05)

44

3.2.2. T-tests (two independent samples)

Unequal sample sizes, unequal variances: example

$$t = \frac{\overline{X}_1 - \overline{X}_2}{S_{\overline{X}_1 - \overline{X}_2}} = \frac{135.8 - 117.5}{4.707} = \frac{18.3}{4.707} = 3.888$$

$$S_{\overline{X}_1 - \overline{X}_2} = \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}} = \sqrt{\frac{58.7}{5} + \frac{41.67}{4}}$$

$$S_{\overline{X}_1 - \overline{X}_2} = \sqrt{11.74 + 10.417} = \sqrt{22.157} = 4.707$$

45

3.2.2. T-tests (two independent samples)

Unequal sample sizes, unequal variances: example

$$d.f. = \frac{\left(S_1^2 / n_1 + S_2^2 / n_2\right)^2}{(S_1^2 / n_1)^2 / (n_1 - 1) + (S_2^2 / n_2)^2 / (n_2 - 1)}$$

$$d.f. = \frac{\left(58.7 / 5 + 41.67 / 4\right)^2}{(58.7 / 5)^2 / (5 - 1) + (41.67 / 4)^2 / (4 - 1)}$$

$$d.f. = \frac{(11.74 + 10.417)^2}{(11.74)^2 / 4 + (10.417)^2 / 3} = \frac{22.157^2}{137.828 / 4 + 108.514 / 3}$$

$$d.f. = \frac{490.933}{34.457 + 36.171} = \frac{490.933}{70.628} = 6.951$$

$$t > t_{(\alpha, d.f)} \rightarrow t > t_{(0.05, 6.951)} \rightarrow 3.388 > 2.365$$

Null hypothesis is rejected. There were statistical differences between groups.

46

3.2.2. T-tests (two dependent samples)

Dependent *t*-test for paired samples

$$t = \frac{\overline{D} - \mu_0}{S_D}$$

$$\overline{D} = \sum D_j / n$$

$$S_D = \sqrt{\frac{\sum \left(D_j - \overline{D}\right)^2}{n_1 + n_2 - 2}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

The constant μ_0 is non-zero if you want to test whether the average of the difference is significantly different from μ_0.

The degree of freedom used is $n - 1$.

47

3.2.2. T-tests (two dependent samples)

Dependent *t*-test for paired samples

A researcher wanted to study the level of relax and the flexibility. 9 gymnasts were measured before and after participating in a 14-week training program with relax and flexibility exercises. The measures in flexibility obtained were the following:

Before (B): 130 134 150 170 135

After (A) : 110 125 120 115 112

Are there statistically differences between the measures obtained before and after the program? (α=0.05)

48

3.2.2. T-tests (two dependent samples)

Dependent *t*-test for paired samples

B	A	D_j	$D_j - \overline{D}$	$(D_j - \overline{D})^2$
130	110	20	-7.4	54.76
134	125	9	-18.4	338.56
150	120	30	2.6	6.76
170	115	55	27.6	761.76
135	112	23	-4.4	19.36
719	**582**	**137**	**0**	**1181.2**

3.2.2. T-tests (two dependent samples)

Dependent *t*-test for paired samples

$$t = \frac{\overline{D} - \mu_0}{S_D} = \frac{27.4 - 0}{7.67} = 3.57$$

$$\overline{D} = \sum D_j / n = 137 / 5 = 27.4$$

$$S_D = \sqrt{\frac{\sum \left(D_j - \overline{D}\right)^2}{n_1 + n_2 - 2}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)} = \sqrt{\frac{1181.2}{5 + 5 - 2}\left(\frac{1}{5} + \frac{1}{5}\right)}$$

$$S_D = \sqrt{147.65 * 0.4} = \sqrt{59.06} = 7.67$$

$$t > t_{(\alpha, n-1)} \rightarrow t > t_{(0.05,4)} \rightarrow 3.57 > 2.776$$

Null hypothesis is rejected. There were statistical differences between groups.

3.2.3. One-way ANOVA 'F'

- While the t-test is limited to compare means of two groups, one-way ANOVA can compare more than two groups. Therefore, the t-test is considered a special case of one-way ANOVA.

51

3.2.3. One-way ANOVA 'F'

- 'Factor': Qualitative independent variable (number of values ≥ 2; number of conditions).
- Quantitative dependent variable (Interval or ratio variable).
- Dependent variable normally distributed.
- Independence of error effects.
- Homogeneity of variance.
- Sphericity (in longitudinal studies).

52

3.2.3. One-way ANOVA 'F'

Types of hypothesis:

- General hypothesis : The purpose is to test for significant differences between group means, and this is done by analyzing the variances. It is a general decision. There is no specification of which pairs of groups are significantly different (concrete conditions).

- Specific hypothesis: There are specifications of which pairs of groups are significantly different (concrete conditions):
  - Post hoc: firstly, the general hypothesis is tested; secondly, if null hypothesis is rejected, specific hypothesis are studied (section 3.2.3.1).
  - A priori: specific hypothesis are studied directly (section 3.2.3.2.).

One-tailed vs. two-tailed hypothesis

53

3.2.3. One-way ANOVA 'F'

Three steps to carry out:

1st. One-way ANOVA: (General hypothesis). ANOVA is an *omnibus* test statistic and cannot tell you which specific groups were significantly different from each other, only that at least two groups were. If null hypothesis is rejected:

2nd. *Post-hoc* test: (Specific hypothesis). To determine which specific groups differed from each other.

3rd. STEP 3. Statistical power: goodness of fit.

54

3.2.3. ANOVA table for one-way case

Sources of variation	Sum of Squares	Degrees of Freedom	Mean Square	F
Treatments Between groups	$n\sum_{j=1}^{a}(\bar{y}_{.j}-\bar{y}_{..})^2$	K-1	SS_b/df_b	MS_b/MS_w
Residuals/Error Within groups	$\sum_{i=1}^{n}\sum_{j=1}^{a}(y_{ij}-\bar{y}_{.j})^2$	K (n-1)	SS_W/df_W	
Total	$\sum_{i=1}^{n}\sum_{j=1}^{a}(y_{ij}-\bar{y}_{..})^2$	N -1		

55

3.2.3. Logic of ANOVA

$$SS_{\text{Total}} = SS_{\text{Error}} + SS_{\text{Treatments}}$$

$$\text{df}_{\text{Total}} = \text{df}_{\text{Error}} + \text{df}_{\text{Treatments}}$$

$$F - \frac{\text{variance between items}}{\text{variance within items}}$$

$F_{crit}(\alpha = 0.05$ K-1, K (n-1))

56

3.2.3. **One-way ANOVA: example**

Consider an experiment to study the effect of three different levels of some factor on a response (e.g. three types of fertilizer on plant growth). If we had 6 observations for each level, we could write the outcome of the experiment in a table like the following, where a_1, a_2, and a_3 are the three levels of the factor being studied. (α=0.05)

a_1	a_2	a_3
6	8	13
8	12	9
4	9	11
5	11	8
3	6	7
4	8	12

57

3.2.3. **One-way ANOVA example**

- The null hypothesis, denoted H_0, for the overall F-test for this experiment would be that all three levels of the factor produce the same response, on average. To calculate the F-ratio:

58

3.2.3. **One-way ANOVA example**

a_1	a_2	a_3
6	8	13
8	12	9
4	9	11
5	11	8
3	6	7
4	8	12
30	**54**	**60**

3.2.3. **One-way ANOVA example**

Step 1: Calculate the mean within each group:

$$\overline{Y}_1 = \frac{1}{6}\sum Y_{1i} = \frac{6+8+4+5+3+4}{6} = 5$$
$$\overline{Y}_2 = \frac{1}{6}\sum Y_{2i} = \frac{8+12+9+11+6+8}{6} = 9$$
$$\overline{Y}_3 = \frac{1}{6}\sum Y_{3i} = \frac{13+9+11+8+7+12}{6} = 10$$

3.2.3. **One-way ANOVA example**

Step 2: Calculate the overall mean:

$$\overline{Y} = \frac{\sum_i \overline{Y}_i}{a} = \frac{\overline{Y}_1 + \overline{Y}_2 + \overline{Y}_3}{a} = \frac{5+9+10}{3} = 8$$

where a is the number of groups.

Step 3: Calculate the "between-group" sum of squares:

$$SS_B = n(\overline{Y}_1 - \overline{Y})^2 + n(\overline{Y}_2 - \overline{Y})^2 + n(\overline{Y}_3 - \overline{Y})^2$$

$$= 6(5-8)^2 + 6(9-8)^2 + 6(10-8)^2 = 84$$

where n is the number of data values per group.

The between-group degrees of freedom is the number of groups minus one

$df_b = 3 - 1 = 2$

so the between-group mean square value is

$MS_B = 84 / 2 = 42$

3.2.3. One-way ANOVA example

Step 4: Calculate the "within-group" sum of squares. Begin by centering the data in each group

a_1	a_2	a_3
6 − 5 = 1	8 − 9 = -1	13 − 10 = 3
8 − 5 = 3	12 − 9 = 3	9 − 10 = -1
4 − 5 = -1	9 − 9 = 0	11 − 10 = 1
5 − 5 = 0	11 − 9 = 2	8 − 10 = -2
3 − 5 = -2	6 − 9 = -3	7 − 10 = -3
4 − 5 = -1	8 − 9 = -1	12 − 10 = 2

The within-group sum of squares is the sum of squares of all 18 values in this table

$SS_W = 1 + 9 + 1 + 0 + 4 + 1 + 1 + 9 + 0 + 4 + 9 + 1 + 9 + 1 + 1 + 4 + 9 + 4 = 68$

The within-group degrees of freedom is

$df_W = a(n - 1) = 3(6 - 1) = 15$

Thus the within-group mean square value is $MS_W = SS_W/df_W = 68/15 \approx 4.5$

3.2.3. One-way ANOVA example

SS_T=84+68=152
df_T=18-1=17

Step 5: The F-ratio is

$$F = \frac{MS_B}{MS_W} \approx 42/4.5 \approx 9.3$$

$F_{crit}(2,15) = 3.68$ at $\alpha = 0.05$.

Since $F = 9.3 > 3.68$, the results are significant at the 5% significance level. One would reject the null hypothesis, concluding that there is strong evidence that the expected values in the three groups differ.

63

3.2.3. One-way ANOVA example

SV	SS	df	MS	F
Between	84	2	42	9.3
Within	68	15	4.5	
Total	152	17		

64

3.2.3.1. 2nd step. Post-hoc test: (Specific hypothesis): To determine which specific groups differed from each other.

- Post hoc tests can only be used when the 'omnibus' ANOVA found a significant effect. If the F-value for a factor turns out non-significant, you cannot go further with the analysis. This 'protects' the post hoc test from being (ab)used too liberally. They are designed to avoid hence of type I error.
  - Tukey.
  - Scheffé.

65

3.2.3.1. Post-hoc test

- DHS Tukey:
  - Comparison of all pairs.
  - Limitation: You can only compare pairs.
  - HSD test: Tukey's Honestly Significant Difference.

$$\psi(HSD) = q_{\alpha,wdf,k}\sqrt{\frac{MS_w}{n}}$$

  - $(|\overline{y_i} - \overline{y_j}|) \geq HSD_{tukey}$ → There are statistically significant differences between groups.
  - $(|\overline{y_i} - \overline{y_j}|) < HSD_{tukey}$ →There are not statistically significant differences between groups.
  - Example: In the previous example, are there statistically significant differences between groups 1 and 3? (α=0.05)

66

3.2.3.1. Post-hoc test: example

$$\psi(HSD) = q_{k(n-1),k}\sqrt{\frac{MS_w}{n}} = q_{3(6-1),3}\sqrt{\frac{4.5}{6}} = q_{15,3}\sqrt{0.75} = 3.67 * 0.866 = 3.178$$

$$|\overline{y}_1 - \overline{y}_3| = |5-10| = 5$$

$$5 > 3.178$$

- There are statistically significant differences between groups 1 and 3.

67

3.2.3.1. One-way ANOVA 'F' Post-hoc tests

- Scheffé:
  - More conservative.
  - Recommended when few contrasts are developed.

$$|\overline{Y}_i - \overline{Y}_{i'}| \geq \sqrt{(k-1)F_{\alpha,k-1,(n-1)k}}\sqrt{MS_e\frac{\sum_1^p a_j^2}{n}}$$

There are statistically significant differences between groups.

$$|\overline{Y}_i - \overline{Y}_{i'}| < \sqrt{(k-1)F_{\alpha,k-1,(n-1)k}}\sqrt{MS_e\frac{\sum_1^p a_j^2}{n}}$$

There are not statistically significant differences between groups.

3.2.3.1. One-way ANOVA 'F' Post-hoc tests: example

Example: Are there statistical differences between groups 1-2 and group 3?

3.2.3.1. One-way ANOVA 'F' Post-hoc tests: example

$$(\overline{Y}_i - \overline{Y}_{i'}) = \sqrt{(k-1)F_{\alpha,k-1,(n-1)k}}\sqrt{MS_w \frac{\sum_1^p a_j^2}{n}}$$

$$\left|\overline{Y}_{1,2} - \overline{Y}_3\right| = \left|\frac{\overline{Y}_1 + \overline{Y}_2}{2} - \overline{Y}_3\right| = \left|\frac{9+5}{2} - 10\right| = 7 - 10 = 3$$

$$\sqrt{(k-1)F_{\alpha,k-1,(n-1)k}}\sqrt{MS_e \frac{\sum_1^p a_j^2}{n}} = \sqrt{(3-1)F_{0.05,3-1,(6-1)3}}\sqrt{4.5\frac{1^2+1^2+(-2)^2}{6}}$$

$$\sqrt{2F_{0.05,2,15}}\sqrt{4.5\frac{6}{6}} = \sqrt{2*3.68}\sqrt{4.5} = \sqrt{7.36}\sqrt{4.5} = 2.713*2.121 = 5.754$$

5.754>3 – There are not statistical differences between groups 1-2 and group 3

3.2.3. One-way ANOVA example

The p-value for this test is 0.002.

After performing the F-test, it is common to carry out some "post-hoc" analysis of the group means. In this case, the first two group means differ by **4** units, the first and third group means differ by **5** units, and the second and third group means differ by only **1** unit.

Thus the first group is strongly different from the other groups, as the mean difference is more times the standard error, so we can be highly confident that the population of the **first group differs** from the population means of the other groups. However there is no evidence that the second and third groups have different population means from each other, as their mean difference of one unit is comparable to the standard error.

Note $F(x, y)$ denotes an F-distribution with x **degrees of freedom** in the numerator and y degrees of freedom in the denominator.

71

3.2.3.2. A priori comparisons

- They are used when the general hypothesis is not of interest; only specific hypothesis are set out.
- Snedecor F is not calculated.
- Types of a priori comparisons:
  - **Non-orthogonal contrasts:**
    - When lot of contrasts are done, type I error (α =probability of being wrong when rejecting the null hypothesis) increases.
    - When c>k-1(c=number of contrasts), correction of α is required.
    - When c≤k-1, some authors recommend the correction and others consider that it is not necessary.
  - **Orthogonal contrasts**: α does not increase, so correction is not necessary (in post-hoc contrasts, the correction is not necessary either because Scheffé and Tukey formulas already control this increase).

72

3.2.3.2. A priori comparisons

- Formula to calculate SS_{bc}:

$$SS_{bc1} = \frac{n\left(\sum_{1}^{p} a_j \overline{y_j}\right)^2}{\sum_{1}^{p} a_j^2}$$

73

3.2.3.2. A priori comparisons A. Orthogonal contrasts

- Characteristics:
  - Their coefficients sum 0.
  - If we multiply the coefficients of two orthogonal contrasts and sum the result, we also obtain 0.
- Tendency contrasts could be carried out: e.g., contrasts could be linear, quadratic and cubic.
- A "F" is going to be calculated for each contrast.

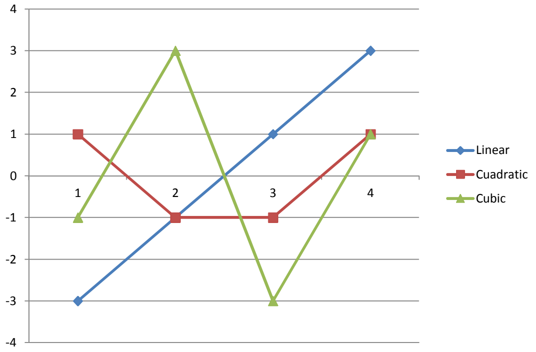
74

3.2.3.2. A priori comparisons
A. Orthogonal contrasts

- k-1 possible contrasts.
- Example of coefficients to study tendency:

	a_1	a_2	a_3	a_4
Linear	-3	-1	1	3
Quadratic	1	-1	-1	1
Cubic	-1	3	-3	1

3.2.3.2. A priori comparisons
A. Orthogonal contrasts



3.2.3.2. A priori comparisons
A. Orthogonal contrasts

Example 1: how to create orthogonal coefficients

	a_1	a_2	a_3	a_4	a_5
Contrast 1	+4	-1	-1	-1	-1
Contrast 2	0	-1	-1	+3	-1
Contrast 3					
Contrast 4					

77

3.2.3.2. A priori comparisons
A. Orthogonal contrasts

Example 1: how to create orthogonal coefficients

	a_1	a_2	a_3	a_4	a_5
Contrast 1	+4	-1	-1	-1	-1
Contrast 2	0	-1	-1	+3	-1
Contrast 3	0	+2	-1	0	-1
Contrast 4	0	0	+1	0	-1

78

3.2.3.2. A priori comparisons
A. Orthogonal contrasts

Example 2: how to create orthogonal coefficients

	a_1	a_2	a_3	a_4	a_5
Contrast 1	+3	+3	-2	-2	-2
Contrast 2	0	0	+1	-2	+1
Contrast 3					
Contrast 4					

79

3.2.3.2. A priori comparisons
A. Orthogonal contrasts

Example 2: how to create orthogonal coefficients

	a_1	a_2	a_3	a_4	a_5
Contrast 1	+3	+3	-2	-2	-2
Contrast 2	0	0	+1	-2	+1
Contrast 3	+1	-1	0	0	0
Contrast 4	0	0	+1	0	-1

80

3.2.3.2. A priori comparisons A. Orthogonal contrasts

- Example (obtained from López, J., Trigo, M. E. y Arias, M. A.(1999). *Diseños experimentales. Planificación y análisis*. Sevilla: Kronos):

a_1	a_2	a_3
14	9	7
12	12	5
18	9	1
9	4	3
10	5	6
15	9	2

Knowing that SS_W=128, do this data present linear or quadratic tendency? (α=0.05)

81

3.2.3.2. A priori comparisons A. Orthogonal contrasts

a_1	a_2	a_3
14	9	7
12	12	5
18	9	1
9	4	3
10	5	6
15	9	2
∑=78	∑=48	∑=24
$\bar{y}_1 = 78/6 = 13$	$\bar{y}_1 = 48/6 = 8$	$\bar{y}_1 = 24/6 = 4$

82

3.2.3.2. A priori comparisons
A. Orthogonal contrasts

$$SS_{b\,linear} = \frac{n\left(\sum_{1}^{p} a_j \overline{y_j}\right)^2}{\sum_{1}^{p} a_j^2} = \frac{6[(-1)(13)+(0)(8)+(+1)(4)]^2}{(-1)^2+(0)^2+(+1)^2} = \frac{486}{2} = 243$$

$$SS_{b\,cuadratic} = \frac{n\left(\sum_{1}^{p} a_j \overline{y_j}\right)^2}{\sum_{1}^{p} a_j^2} = \frac{6[(+1)(13)+(-2)(8)+(+1)(4)]^2}{(+1)^2+(-2)^2+(+1)^2} = \frac{6}{6} = 1$$

83

3.2.3.2. A priori comparisons
A. Orthogonal contrasts

SV	SS	df	MS	F
Between	244	2		
linear	243	1	243	28.48
quadratic	1	1	1	0.117
Within	128	15	8.53	
Total	372	17		

$F_{(\alpha,\, k-1,\, k(n-1))} = F_{(0.05,1,15)} = 4.54$

28.48>4.54 → There is linear tendency
0.117<4.54 → There is not quadratic tendency

84

3.2.3.2. A priori comparisons
B. Non-orthogonal contrasts

When lot of contrasts are done, type I error (α =probability of being wrong when rejecting the null hypothesis) increases for the group of contrasts :

Comparisons	Type I error by contrast	Type I error for the group of contrasts
C1	0.05	0.05
C2	0.05	0.05 x 2 = 0.1
C3	0.05	0.05 x 3 = 0.15

85

3.2.3.2. A priori comparisons
B. Non-orthogonal contrasts

Bonferroni: method for controlling the increase of α

$$\alpha_c = \frac{\alpha}{C}$$

Being:

- α = usually, 0.05 or 0.01
- C = number of contrasts

Comparisons	Type I error by contrast	Type I error for the group of contrasts
C1	0.05/3	0.05/3= 0.017
C2	0.05/3	(0.05/3) x 2 = 0.03
C3	0.05/3	(0.05/3) x 3 = 0.05

The highest type I error for the group of contrasts is α

86

3.2.3.3. One-way ANOVA 'F' in longitudinal studies

- In a longitudinal study, each participant presents at least a score for each condition.
- The most important assumption to take into account is the **sphericity**.
- Steps:
  - 1. Testing assumptions.
  - 2. One-way ANOVA F.
  - 3. Effect size.

3.2.3.3. One-way ANOVA 'F' in longitudinal studies

Step 1. Testing assumptions:

- **Normality:** it is not important because F is robust even when it is violated.
- **Independence of errors:** it is usually violated because the measurements of the same person tends to be similar.
- **Homoscedasticity:** it is usually violated because the scores obtained closer in time are more similar.

Although the two last assumptions are violated, F could be used if the relationship between variances and covariances are of **sphericity**.

3.2.3.3. One-way ANOVA 'F' in longitudinal studies

Step 2. One-way ANOVA F

SV	SS	df	MS	F
BETWEEN SUBJECTS	$k\sum_{i=1}^{n}\left(\bar{y}_{i.}-\bar{y}_{..}\right)^2$	n-1		
WITHIN SUBJECTS	$\sum_{i=1}\sum_{j=1}^{k}\left(y_{ij}-\bar{y}_{i.}\right)^2$	n(k-1)		
A	$n\sum_{i=1}^{k}\left(\bar{y}_{.j}-\bar{y}_{..}\right)^2$	k-1	SS_A/df_A	MS_A/MS_{SxA}
Subjects x A (error)	$\sum_{i=1}^{n}\sum_{j=1}^{k}\left(y_{ij}+\bar{y}_{..}-\bar{y}_{i.}-\bar{y}_{.j}\right)^2$	(n-1)(k-1)	SS_{SxA}/df_{SxA}	
TOTAL	$\sum_{i=1}^{n}\sum_{j=1}^{k}\left(y_{ij}-\bar{y}_{..}\right)^2$	nk-1		

89

3.2.3.3. One-way ANOVA 'F' in longitudinal studies

Step 2. One-way ANOVA F

- **Three stages test:**
  - It is used in order to try to avoid unnecessary calculations. **Adjusted ε** has to be calculated if we can not conclude based on the two first stages, and we have to carry out the third one.
  - ε (epsilon = measurement of the degree in which the assumption sphericity is violated) is going to be multiplied by the degrees of freedom of the theoretical F.

90

3.2.3.3. One-way ANOVA 'F' in longitudinal studies

Step 2. One-way ANOVA F

Stage 1: normal F ($\varepsilon=1$) (too easy to reject H_0).

- If H_o is accepted, we do not have to do anything else.
- If H_o is rejected, accepting this conclusion would imply to have a high level of possibility of committing type I error. We carry out stage 2.

Stage 2: conservative F; $\varepsilon=1/(k-1)$ (too difficult to reject H_0).

- If H_o is rejected, we do not have to do anything else.
- If H_o is accepted, we have to carry out stage 3.

Stage 3: F; $1/(k-1)< \varepsilon <1$.

91

3.2.3.3. One-way ANOVA 'F' in longitudinal studies

Example. We would like to study the effect that different revisions of a material (a_1=first revision; a_2=second revision; a_3=third revision) produce in the memorization.

Participants	a_1	a_2	a_3
1	14	18	22
2	16	22	27
3	14	20	30
4	18	23	24
5	22	26	27
6	24	23	26

92

3.2.3.3. One-way ANOVA 'F' in longitudinal studies

Participants	a_1	a_2	a_3	$\sum$	$\bar{y}_{i.}$
1	14	18	22	**54**	**18**
2	16	22	27	**65**	**21.67**
3	14	20	30	**64**	**21.33**
4	18	23	24	**65**	**21.67**
5	22	26	27	**75**	**25**
6	24	23	26	**73**	**24.33**
$\sum$	**108**	**132**	**156**		
$\bar{y}_{.j}$	**18**	**22**	**26**		

93

3.2.3.3. One-way ANOVA 'F' in longitudinal studies

$$\bar{y}_{..} = \frac{18+22+26}{3} = \frac{18+21.67+21.33+21.67+25+24.33}{6} = 22$$

$$SS_A = 6\left[(18-22)^2 + (22-22)^2 + (26-22)^2\right] =$$
$$= 6(16+0+16) = 32*6 = 192$$

$$SS_S = 3\begin{bmatrix}(18-22)^2 + (21.67-22)^2 + (21.33-22)^2 \\ +(21.67-22)^2 + (25-22)^2 + (24.33-22)^2\end{bmatrix} =$$
$$= 3(16+0.11+0.45+0.11+9+5.43) = 93.33$$

94

3.2.3.3. One-way ANOVA ‘F’ in longitudinal studies

$$SS_{AxS} = (14+22-18-18)^2 + (16+22-21.67-18)^2 + (14+22-21.33-18)^2 +$$
$$(18+22-21.67-18)^2 + (22+22-25-18)^2 + (24+22-24.33-18)^2 +$$
$$(18+22-18-22)^2 + (22+22-21.67-22)^2 + (20+22-21.33-22)^2 +$$
$$(23+22-21.67-22)^2 + (26+22-25-22)^2 + (23+22-24.33-22)^2 +$$
$$(22+22-18-26)^2 + (27+22-21.67-26)^2 + (30+22-21.33-26)^2 +$$
$$(24+22-21.67-26)^2 + (27+22-25-26)^2 + (26+22-24.33-26)^2 =$$
$$0+2.79+11.09+0.11+1+13.47+0+0.33+1.77+1.77+1+1.77+0+1.77$$
$$+21.81+2.79+4+5.43=70.67$$

95

3.2.3.3. One-way ANOVA ‘F’ in longitudinal studies

SV	SS	df	MS	F
BETWEEN SUBJECTS	93.33	5		
WITHIN SUBJECTS	262.67	12		
A	192	2	96	13.58
Subjects x A (error)	70.67	10	7.07	
TOTAL	356	17		

96

3.2.3.3. One-way ANOVA ‘F’ in longitudinal studies

$F_{(\alpha,\ k-1,\ (n-1)(k-1))}=F_{(0.05,\ 2,\ 10)}=4.1$

Three stages test:

Stage 1: normal F (ε=1).

$F_{(\alpha,\ (k-1)\varepsilon,\ (n-1)(k-1)\ \varepsilon}=F_{(0.05,\ 2*1,\ 10*1)}=4.1$

13.58>4.1→ H_o is rejected, so we have to carry out stage 2.

Stage 2: conservative F; ε=1/(k-1)=1/(3-1)=0.5.

$F_{(\alpha,\ (k-1)\varepsilon,\ (n-1)(k-1)\ \varepsilon}=F_{(0.05,\ 2*0.5,\ 10*0.5)}=F_{(0.05,\ 1,\ 5)}=6.61$

13.58>6.61→ H_o is rejected, so we do not have to carry out stage 3. This is the final decision.

97

4. STEP 3. Statistical power

- After analyzing the existence of significant relation between variables, we have to verify whether our **conclusions** are or not **erroneous** because of **statistical power**.

- The **power** of a statistical test is the probability that the test will reject a false null hypothesis (i.e. that it will not make a type II error). As power increases, the chances of a type II error decrease. The probability of a type II error is referred to as the false negative (β). Therefore, power is equal to 1 − β, which is equal to sensitivity.

98

4. STEP 3. Statistical power

	DECISION	
	Accept H_o	Reject H_o
H_o is true	1-α: Level of confidence	α : Type I error
H_o is false	β: Type II error	**1- β** **Power**

99

4. STEP 3. Statistical power

- Power analysis can be **used** to calculate the minimum **sample size** required to accept the outcome of a statistical test with a particular level of confidence. It can also be used to calculate the minimum **effect size** that is likely to be detected in a study using a given sample size.

- In statistics, an **effect size** is a measure of the strength of the relationship between two variables in a statistical population, or a sample-based estimation of that quantity.

100

4. STEP 3. Statistical power

- **Pearson's correlation**, often denoted 'r', is widely used as an effect size when **two quantitative variables** are available; for instance, if one was studying the relationship between birth weight and longevity.

- A related effect size is **the coefficient of determination** (**r-squared = R^2**). In the case of two variables, this is a measure of the proportion of variance shared by both, and varies from 0 to 1. An R^2 of 0.21 means that 21% of the variance of a variable is shared with the other variable.

101

4. STEP 3. Statistical power

- Effect size:

$$R^2 = \frac{SS_B}{SS_T}$$

- Low: ≤ 0.18
- Medium: [0.38-0.58]
- High: ≥ 0.67

102

4. STEP 3. STATISTICAL POWER. EFFECT SIZE. GOODNESS OF FIT

	Significant effect	Non-significant effect
High effect size	The effect probably exists	The non-significance can be due to low statistical power
Low effect size	The statistical significance can be due to an excessive high statistical power	The effect probably does not exist

4. STEP 3. STATISTICAL POWER. EFFECT SIZE. GOODNESS OF FIT: EXAMPLE

- Based on the example presented in slide 57 (first example of F), conclude about the statistical power.

SV	SS	df	MS	F
Between	84	2	42	9.3
Within	68	15	4.5	
Total	152	17		

4. STEP 3. STATISTICAL POWER. EFFECT SIZE. GOODNESS OF FIT: EXAMPLE

- Medium effect size:

$$R^2 = \frac{SS_B}{SS_T} = \frac{84}{152} = 0.553$$

- Significant: 9.3 > 3.68
- Conclusion: The effect probably exists

105