Quantitative Methods

Parametric and Non-Parametric Tests of Independence

Intro and Exam Focus

- Parametric test of correlation
- Nonparametric tests of relationships
  - Rank correlation
  - Contingency tables for categorical data

Parametric Test of Correlation

- Test of whether population correlation coefficient equals zero

$$\text{t-stat} = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} \text{ with } n-2 \text{ degrees of freedom}$$

where: r = sample correlation coefficient

n = sample size

Parametric Test of Correlation: CFA Institute Example

The exhibit below shows the sample correlations between the monthly returns for four different mutual funds and the S&P 500. The correlations are based on 36 monthly observations.

	Fund 1	Fund 2	Fund 3	Fund 4	S&P500
Fund 1	1				
Fund 2	0.9231	1			
Fund 3	0.4771	0.4156	1		
Fund 4	0.7111	0.7238	0.3102	1	
S&P 500	0.8277	0.8223	0.5791	0.7515	1

Test the null hypothesis that the correlation of Fund 3 and Fund 4 is equal to zero against the alternative hypothesis that it is not equal to zero. Use a 5 percent significance level and critical *t*-values of ±2.032.

Parametric Test of Correlation: Solution

State the hypotheses:

H_0: vs. H_a:

Calculate test statistic:

t-stat = =

Make decision: test stat critical *t*-value (2.032) →
null

-3

Nonparametric Test: Rank Correlation

- **Spearman rank correlation test:** whether two sets of ranks are correlated
- Rank correlation: $r = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n(n^2 - 1)}$
- where: n = sample size
  d_i = difference between two ranks
- If n > 30, may use *t*-table with n – 2 degrees of freedom

Rank Correlation: Example

Time Period	Return of A	Return of B	Rank of A	Rank of B	Difference in Ranks (d_i)	d_i^2
1	12%	5%	3	10	–7	49
2	–2%	8%	36	12	24	576
…	…	…	…	…	…	…
52	0%	3%	28	18	10	100
					Sum	6533

Perform a test of independence on the data with 5% significance:

$$r = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n(n^2 - 1)} = \quad = \quad =$$

-1

Rank Correlation: Solution

Perform a test of independence on the data with 5% significance:

df	One-Tailed Probabilities (p)		
	p = 0.10	p = 0.05	p = 0.025
50	1.30	1.68	2.01

$$\text{t-stat} = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} = \quad = \quad =$$

T-stat critical value hence, null hypothesis that ρ = 0

-4

Contingency Tables

Earnings Growth	Dividend Yield			
	Low	Medium	High	Total
Low	28	53	42	123
Medium	42	32	39	113
High	49	25	14	88
Total	119	110	95	324

Degrees of freedom = (r – 1)(c – 1)

$$\text{Expected if independent} = \frac{\text{total for row } i \times \text{total for column } j}{\text{total for all columns and rows}}$$

Observed frequency in cell i,j

$$\text{Test statistic: } \chi^2 = \text{sum of } \frac{(O_{i,j} - E_{i,j})^2}{E_{i,j}}$$

Expected frequency in cell i,j

-2

Contingency Tables: CFAI Example

Consider the contingency table below, which classifies 500 randomly selected companies on the basis of two environmental, social, and governance (ESG) rating dimensions: environmental rating and governance rating.

	Governance Rating			
Environmental Rating	Progressive	Average	Poor	Total
Progressive	35	40	5	80
Average	80	130	50	260
Poor	40	60	60	160
Total	155	230	115	500

Using a 5 percent level of significance, determine whether these two ESG rating dimensions are independent of one another.

Contingency Tables: Solution

Expected frequency for cell (1,1) = =

Expected frequency for cell (1,2) = =

Expected frequency for cell (2,1) = =

Completing the table . . .

	Governance Rating		
Environmental Rating	Progressive	Average	Poor
Progressive			18.4
Average		119.6	59.8
Poor	49.6	73.6	36.8

Contingency Tables: Solution

Scaled squared deviation for cell (1,1) = =

Expected frequency for cell (1,2) = =

Expected frequency for cell (2,1) = =

Completing the table . . .

Environmental Rating	Governance Rating		
	Progressive	Average	Poor
Progressive			9.759
Average		0.904	1.606
Poor	1.858	2.513	14.626

Total scaled squared deviation =

-4

Contingency Tables: Solution

With 5% significance and (3 – 1) × (3 – 1) = 4 degrees of freedom and 5% significance (one-tailed), critical χ^2 statistic from a χ^2 table =

Degrees of Freedom	Probability in Right Tail					
	0.975	0.95	0.9	0.1	0.05	0.025
3	0.22	0.35	0.58	6.25	7.81	9.35
4	0.48	0.71	1.06	7.78	9.49	11.14
5	0.83	1.15	1.61	9.24	11.07	12.83

Make decision: test stat (35.74) than critical value (9.

→ I nclude th ationship between environmental rating and governance rating

-3

Solutions

Parametric Test of Correlation: Solution

State the hypotheses:

$$H_0: \rho = 0 \text{ vs. } H_a: \rho \neq 0$$

Calculate test statistic:

$$\text{t-stat} = \frac{0.3102\sqrt{36-2}}{\sqrt{1-0.3102^2}} = 1.903$$

Make decision: test stat (1.903) < critical *t*-value (2.032) → fail to reject null

Rank Correlation: Example

Time Period	Return of A	Return of B	Rank of A	Rank of B	Difference in Ranks (d_i)	d_i^2
1	12%	5%	3	10	–7	49
2	–2%	8%	36	12	24	576
…	…	…	…	…	…	…
52	0%	3%	28	18	10	100
					Sum	6,533

Perform a test of independence on the data with 5% significance:

$$r = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n(n^2 - 1)} = 1 - \frac{6(6533)}{52(52^2 - 1)} = 1 - \frac{39{,}198}{140{,}556} = 0.72$$

Rank Correlation: Solution

Perform a test of independence on above data with 5% significance:

df	One-Tailed Probabilities (p)		
	p = 0.10	p = 0.05	p = 0.025
50	1.30	1.68	2.01

$$\text{t-stat} = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} = \frac{0.72\sqrt{52-2}}{\sqrt{1-0.72^2}} = \frac{5.091}{0.694} = 7.34$$

T-stat (7.34) > critical value (2.01); hence, **reject** null hypothesis that ρ = 0

Contingency Tables: Solution

Expected frequency for cell (1,1) = [80 × 155] / 500 = 24.8

Expected frequency for cell (1,2) = [80 × 230] / 500 = 36.8

Expected frequency for cell (2,1) = [260 × 155] / 500 = 80.6

Completing the table . . .

	Governance Rating		
Environmental Rating	Progressive	Average	Poor
Progressive	24.8	36.8	18.4
Average	80.6	119.6	59.8
Poor	49.6	73.6	36.8

-3

Contingency Tables: Solution

Scaled squared deviation for cell (1,1) = $(35 - 24.8)^2 / 24.8 = 4.195$

Expected frequency for cell (1,2) = $(40 - 36.8)^2 / 36.8 = 0.278$

Expected frequency for cell (2,1) = $(80 - 80.6)^2 / 80.6 = 0.004$

Completing the table . . .

	Governance Rating		
Environmental Rating	Progressive	Average	Poor
Progressive	4.195	0.278	9.759
Average	0.004	0.904	1.606
Poor	1.858	2.513	14.626

Total scaled squared deviation = 35.74

-4

Contingency Tables: Solution

With 5% significance and (3 – 1) × (3 – 1) = 4 degrees of freedom and 5% significance (one-tailed), critical χ^2 statistic from a χ^2 table =

Degrees of Freedom	Probability in Right Tail					
	0.975	0.95	0.9	0.1	0.05	0.025
3	0.22	0.35	0.58	6.25	7.81	9.35
4	0.48	0.71	1.06	7.78	9.49	11.14
5	0.83	1.15	1.61	9.24	11.07	12.83

Make decision: test stat (35.74) greater than critical value (9.

→ Reject null, conclude there is a relationship between environmental rating and governance rating

-3