Universidade de Brasília
IE - Departamento de Estatística
Análise de Dados Categorizados – 1/2022

Lista Exercícios 2

*Estude o **capítulo 2** do livro –texto e as notas de aula e resolva os seguintes exercícios:*

1. Resolver o exercícios 3 a 7, páginas 55 e 56, da 2ª edição do livro-texto (2007).

2.3 According to recent UN figures, the annual gun homicide rate is 62.4 per one million residents in the United States and 1.3 per one million residents in the UK.

a. Compare the proportion of residents killed annually by guns using the (i) difference of proportions, (ii) relative risk.

b. When both proportions are very close to 0, as here, which measure is more useful for describing the strength of association? Why?

2.4 A newspaper article preceding the 1994 World Cup semifinal match between Italy and Bulgaria stated that "Italy is favored 10–11 to beat Bulgaria, which is rated at 10–3 to reach the final." Suppose this means that the odds that Italy wins are 11/10 and the odds that Bulgaria wins are 3/10. Find the probability that each team wins, and comment.

2.5 Consider the following two studies reported in the *New York Times*:

a. A British study reported (December 3, 1998) that, of smokers who get lung cancer, "women were 1.7 times more vulnerable than men to get small-cell lung cancer." Is 1.7 an odds ratio, or a relative risk?

b. A National Cancer Institute study about tamoxifen and breast cancer reported (April 7, 1998) that the women taking the drug were 45% less likely to experience invasive breast cancer compared with the women taking placebo. Find the relative risk for (i) those taking the drug compared to those taking placebo, (ii) those taking placebo compared to those taking the drug.

2.6 In the United States, the estimated annual probability that a woman over the age of 35 dies of lung cancer equals 0.001304 for current smokers and 0.000121 for nonsmokers [M. Pagano and K. Gauvreau, *Principles of Biostatistics*, Belmont, CA: Duxbury Press (1993), p. 134].

a. Calculate and interpret the difference of proportions and the relative risk. Which is more informative for these data? Why?

b. Calculate and interpret the odds ratio. Explain why the relative risk and odds ratio take similar values.

2.7 For adults who sailed on the Titanic on its fateful voyage, the odds ratio between gender (female, male) and survival (yes, no) was 11.4. (For data, see R. Dawson, *J. Statist. Educ.* **3**, no. 3, 1995.)

a. What is wrong with the interpretation, "The probability of survival for females was 11.4 times that for males"? Give the correct interpretation.

b. The odds of survival for females equaled 2.9. For each gender, find the proportion who survived.

c. Find the value of R in the interpretation, "The probability of survival for females was R times that for males."

2. Resolver o seguinte exercício (*exercício 1 da 1ª edição do livro-texto (1996)).*

2.1. A Swedish study considered the effect of low-dose aspirin on reducing the risk of stroke and heart attacks among people who have already suffered a stroke (*Lancet 338*: 1345–1349 (1991)). Of 1360 patients, 676 were randomly assigned to the aspirin treatment (one low-dose tablet a day) and 684 to a placebo treatment. During a follow-up period averaging about three years, the number of deaths due to myocardial infarction were 18 for the aspirin group and 28 for the placebo group.

a. Calculate and interpret the difference of proportions, relative risk of death, and the odds ratio.

b. Conduct an inferential analysis for these data. Interpret results.

3. Resolver o exercícios 15, 19, 27, 29, 30 e 31, páginas 58 a 63, da 2ª edição do livro-texto (2007).

2.15 A large-sample confidence interval for the log of the relative risk is

$$\log(p_1/p_2) \pm z_{\alpha/2}\sqrt{\frac{1-p_1}{n_1 p_1} + \frac{1-p_2}{n_2 p_2}}$$

Antilogs of the endpoints yield an interval for the true relative risk. Verify the 95% confidence interval of (1.43, 2.30) reported for the relative risk in Section 2.2.3 for the aspirin and heart attack study.

2.19 Table 2.14 was taken from the 2002 General Social Survey.

a. Test the null hypothesis of independence between party identification and race. Interpret.

Table 2.14. Data for Problem 2.19

	Party Identification		
Race	Democrat	Independent	Republican
White	871	444	873
Black	302	80	43

b. Use standardized residuals to describe the evidence.

c. Partition the chi-squared into two components, and use the components to describe the evidence.

2.27 A study on educational aspirations of high school students (S. Crysdale, *Int. J. Comp. Sociol.*, **16**: 19–36, 1975) measured aspirations using the scale (some high school, high school graduate, some college, college graduate). For students whose family income was low, the counts in these categories were (9, 44, 13, 10); when family income was middle, the counts were (11, 52, 23, 22); when family income was high, the counts were (9, 41, 12, 27).

a. Test independence of aspirations and family income using X^2 or G^2. Interpret, and explain the deficiency of this test for these data.

b. Find the standardized residuals. Do they suggest any association pattern?

c. Conduct a more powerful test. Interpret results.

2.29 A study (B. Kristensen et al., *J. Intern. Med.*, **232**: 237–245, 1992) considered the effect of prednisolone on severe hypercalcaemia in women with metastatic breast cancer. Of 30 patients, 15 were randomly selected to receive prednisolone, and the other 15 formed a control group. Normalization in their level of serum-ionized calcium was achieved by seven of the 15 prednisolone-treated patients and by 0 of the 15 patients in the control group. Use Fisher's exact test to find a P-value for testing whether results were significantly better for treatment than control. Interpret.

2.30 Table 2.17 contains results of a study comparing radiation therapy with surgery in treating cancer of the larynx. Use Fisher's exact test to test $H_0: \theta = 1$ against $H_a: \theta > 1$. Interpret results.

Table 2.17. Data for Problem 2.30

	Cancer Controlled	Cancer Not Controlled
Surgery	21	2
Radiation therapy	15	3

Source: W. Mendenhall et al., *Int. J. Radiat. Oncol. Biol. Phys.*, **10**: 357–363, 1984. Reprinted with permission from Elsevier Science Ltd.

2.31 Refer to the previous exercise.

a. Obtain and interpret a two-sided exact P-value.

b. Obtain and interpret the one-sided mid P-value. Give advantages of this type of P-value, compared with the ordinary one.

Bom Estudo!!!!