ICEF Probability theory and Statistics

Home assignment 8
23.05.2025

Problem 1. A large retail company wants to understand whether income level affects quality of groceries which customers choose. In order to do that records of 150 clients were analysed to determine which category of groceries they purchase more often: regular or ''ecological" (with special label "Eco"). The following results were obtained:

	Regular	Eco
Poor	41	24
Middle Class	22	18
Rich	20	25

At 1% significance level test whether income level is independent of groceries preferences.

Problem 2. An advertisement agency randomly sampled people in a town to find out which newspapers are preferred by different social groups. Each respondent has been classified by his/her income level and favorite newspaper.

Social Class	Newspaper A	Newspaper B	Newspaper C
Poor	28	20	19
Middle Class	47	48	40
Rich	20	36	42

(a) Is it true that these newspapers have significantly different numbers of readers in the town?
(b) Is it true that the proportions of town residents belonging to each of the three income groups are not equal?
(c) Is it true that the preferences do not depend on the income?

Use 10% significance level.

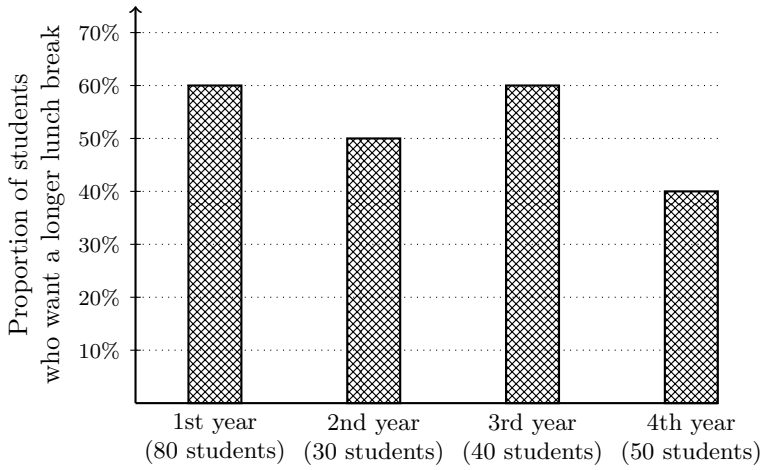
Problem 3. In order to study the dynamics of telemedicine development (remote consultation of specialist) the data for 300 patients was collected: those you attended East and West hospitals and those who used telemedicine:

	Therapist	Surgeon	Psychologist
East Hospital	39	35	18
West Hospital	34	50	23
Telemedicine	35	27	39

At 5% significance level test whether choice of specialist depends on medicine type (remote/in-person visit).

Problem 4. In some university, a group of students asked the administration to increase the duration of the lunch break. In order to determine the proportion of students who would like to have a longer lunch break, the student union conducted a survey in a sample of randomly selected students. The graph below shows the proportion of students who want the lunch break to be increased, and the number of students of each year of study who participated in the survey.

Is there statistical evidence that the opinions among the students of different years of study differ? Use the p-value approach and assume that all the conditions for applying the appropriate statistical test are met.



Problem 5. A large company gathered a random sample of 600 of its employees to determine whether they like the design of their new web page. The table below summarises the responses of the employees.

Gender	Sample size	Positive view on the new web page
Males	240	130
Females	360	211

Do the employees' responses indicate a difference between males and females in whether they like the design of the latest web page? Conduct a suitable hypothesis test at 1% and 5% significance levels.

(a) Use test for proportions.

(b) Use χ^2 test for homogeneity.

Compare results. Did you get the same conclusions?

Additional problems

Problem 6. In some survey, 100 young people were asked to select the most preferred form of sports activity from 5 choices. Test the hypothesis that the choice is independent of the gender of the respondent. Use the 5% significance level. Assume that the conditions for applying the χ^2 test are met (despite that in this problem some expected counts will be smaller than 10).

	Soccer	Basketball	Swimming	Running	Tennis
Male	21	5	9	12	13
Female	9	3	2	14	12

Answer: $\chi^2_{st} = 6.2$, do not reject H_0 at 5% significance level.

Problem 7. 500 individual investors were asked to specify their risk profile (aggressive, moderate or conservative) and the main asset type they invest into (stocks, bonds or derivatives). Results are aggregated in a table

	Stocks	Bonds	Derivatives
Aggressive	54	40	89
Moderate	84	46	31
Conservative	53	98	5

Use 5% significance level to test a hypothesis that joint distribution of risk profile and main asset type is

	Stocks	Bonds	Derivatives
Aggressive	0.1	0.07	0.16
Moderate	0.16	0.1	0.05
Conservative	0.12	0.22	0.02

In solution you should specify test used and justify your choise.

Answer: $\chi^2_{st} = 8.633$, do not reject H_0 at 5% significance level.