

Lab 1: Revision

12.03.2015

1. The table presents the distribution of age in a population living in an area (in thousands)

Age	< 15	15-24	25-34	35-44	45-54	55-64	65-74	> 74
Population	475	304	182	190	208	170	111	72

- (a) Calculate the proportion of population in each age group.

Age	< 15	15-24	25-34	35-44	45-54	55-64	65-74	> 74
Proportion								

- (b) Suppose that you are to compare the distribution above with another distribution, for example distribution of age in a town in this area and examine the goodness-of-fit of those distributions. What would be the null hypothesis in statistical significance test?

H_0 :

- (c) Suppose that in that town there live 12500 people. What are the expected numbers of people in are groups?

Age	< 15	15-24	25-34	35-44	45-54	55-64	65-74	> 74
Expected number								

- (d) Test if the distribution of age in a town in the area, given in table below, fits the distribution in the whole area.

Age	< 15	15-24	25-34	35-44	45-54	55-64	65-74	> 74
Population	3016	2438	2037	2031	1253	977	585	163

2. Statistical data about Titanic sinking says that:

- 124 women died, 323 survived,
- 694 men died, 175 survived.

Is there a statistical dependence between gender and survival?

.....			
.....			
.....			

3. χ^2 tasks in lab01.xls

- (a) Task 1
- (b) Task 2

4. Entropy

Compare results of 100 tosses of three different coins, one perfectly fair, second unbalanced, third fake with 2 sides with heads.

	#heads	#tails	p(coin toss = heads)	p(coin toss = tails)	entropy of coin toss
perfectly fair coin	50	50			
unbalanced coin	25	75			
fake coin	100	0			

What is the maximal possible entropy value? What is the minimal value? What do they mean?