

# Chi-Square Tests

## Homework 10

### Problems

1. In 1920, Rutherford and Geiger counted the number of  $\alpha$  particles emitted by a mass of polonium during 2608 disjoint 7.5 second time periods. Here are their data

0<-c(57,203,383,525,532,408,273,139,45,27,10,4,0,1,1)

for  $c = 0, 1, \dots, 14$  or more counts in each of the time periods.

- Find the mean for these data.
  - Find expected number of counts for a goodness of fit test hypothesis against a Poisson distribution.
  - Find the values for  $\chi^2$  statistic. What are the degrees of freedom for this test?
  - What is the  $p$ -value for this test?
  - What do you conclude from the  $p$ -value?
2. A researcher wants to see if different drug dependence therapies work differently depending on gender. Here, “postive” means drug free for at least 6 month. Thus, we have a three way table with categorical variables - therapy type, success of therapy, and gender. Here is the table that summarizes the data.

	Therapy	1	2	3
Positive	Male	59	55	107
	Female	32	24	80
Negarive	Male	9	12	17
	Female	16	33	56

- Write the hypothesis for the situation for the independence of the three categorical variables.
- Give the dimension of the parameter space  $\Theta$  for this situation.
- Give the dimension of the null hypothesis space  $\Theta_0$  for this situation.
- Find the marginal values for each of the three parameter values.
- Find the three way table of expected counts.
- Find the values for  $\chi^2$  statistic. What are the degrees of freedom for this test?
- What is the  $p$ -value for this test?
- What do you conclude from the  $p$ -value?

3. Hemoglobin E is a variant of hemoglobin with a mutation in the  $\beta$  globin gene causing substitution of glutamic acid for lysine at position 26 of the  $\beta$  globin chain. HbE (E is the one letter abbreviation for glutamic acid.) is the second most common abnormal hemoglobin after sickle cell hemoglobin (HbS). HbE is common from India to Southeast Asia. The  $\beta$  chain of HbE is synthesized at a reduced rate compare to normal hemoglobin (HbA) as the HbE produces an alternate splicing site within an exon. It has been suggested that Hemoglobin E provides some protection against malaria virulence when heterozygous, but is causes anemia when homozygous. The circumstance in which the heterozygotes for the alleles under consideration have a higher adaptive value than the homozygote is called **balancing selection**.

The table below gives the counts of differing hemoglobin genotypes on two Indonesian islands.

genotype	AA	AE	EE
Flores	128	6	0
Sumba	119	78	4

Because the heterozygotes are rare on Flores, it appears malaria is less prevalent there since the heterozygote does not provide an adaptive advantage.

- Is the  $\chi$  square test appropriate to this situation/
- Carry out a Fisher exact test for these data.
- What is the  $p$ -value for this test?
- What do you conclude from the  $p$ -value?