9.4: Heterogeneity chi-square tests

Introduction

After finding evidence to reject the statistical null hypothesis, it may be appropriate to proceed to test a number of data sets against the same theoretical distribution. Provided these are **planned comparisons**, part of the experiment and not an exercise in **data dredging** (Norman 2014), you may proceed to test a series of nested models where groups are combined to make new groups for comparison. This approach is analogous to the **post-hoc tests** one may conduct after a statistically significant ANOVA (see Chapter 12.4) or when identifying a **best-fit** regression model (see Chapter 18.4).

Example

Imagine a scenario where a population geneticist has collected allele (gene) frequency and genotype frequency data on single nucleotide polymorphisms (SNP) for the BRCA1 locus for three USA groups: African Americans, Asian Americans, and European Americans. Data of this kind can be obtained from the SNP database at NCBI (data retrieved 14 & 15 July 2014), and I collated several SNP involving a Cytosine base for Thymine base change (Table 9.4.1). The BRCA1 locus is on chromosome 17 and some of the hundreds of mutations found for this gene are associated with high risk of breast and or ovarian cancer (Couch and Weber 1996, Tram et al 2013).

Table 9.4.1. SNP of BRCA1 locus.

SNP	Population	n	C/C	C/T	T/T	C	T
rs1060915	African American	46	0.043	0.174	0,783	0.130	0.870
	Asian American	48	0.083	0.625	0.375	0.396	0.604
	European American	48	0.167	0.458	0.375	0.396	0.604
rs3737559	African American	40	0.043	0.174	0.783	0.130	0.870
	Asian American	48	0.083	0.625	0.292	0.396	0.604
	European American	48	0.167	0.458	0.375	0.396	0.604
rs799917	African American	124	0.048	0.258	0.694	0.177	0.823
	Asian American	90	0.467	0.444	0.089	0.689	0.311
	European American	118	0.407	0.458	0.136	0.636	0.364

We may wish to test the combined data set to see if the large data set differs from Hardy Weinberg expectations before proceeding with a series of sample populations. By combining the data sets, we will be able to test whether one genotype is different from exception.

The goal then is to pool the data so that you have a more powerful test of the null hypothesis (remember our general discussion of statistical power and how increasing sample size increases your chances of correctly rejecting the null hypothesis).

For example, in our test of Hardy-Weinberg expectations (Example C), we calculated a χ^2 test of 7.8955. This test has $k-1=3-1=2$ degrees of freedom. At alpha 5%, the critical value was 5.991. We clearly would reject the null hypothesis. But do we reject because of a difference in the **aa**, **ab**, or **bb** genotypes? Simply combine categories. Let's test the homozygotes (**aa** and **bb**) versus the heterozygotes (**ab**).

Table 9.4.2. Worksheet.

	Categories		
	aa + bb	**ab**	**n**
Observed	66	34	100
Expected	52	48	100
$\chi^2=\sum_{i=1}^{k}\frac{(O_i-E_i)^2}{E_i}$	$=\frac{(66-52)^2}{52}+\frac{(34-48)^2}{48}=3.769+4.083=7.852$		

With one degree of freedom, we clearly reject the null hypothesis because the critical value at 5% and one degree of freedom is equal to 3.841. To get the p-value, use

```
pchisq(7.852,df=1,lower.tail=FALSE)
```

which will return

```
[1] 0.005076452
```

Oh, and what was the null hypothesis? That there was an equal number of homozygotes and heterozygotes.

Other tests would be possible here, but the point is that you can dissect an experiment to determine which group is causing you to reject the null original hypothesis. While this is an important tool, you should also consider issues of *a priori* and *a posterori* decisions in experiments.

Questions

1. Compare and contrast the purpose of Yates correction and heterogeneity chi-square test.

2. The SNP rs1799971 (A > G) in the μ-opioid receptor (MOR) is associated with opioid dependency. Allele frequencies for different populations are provided in the table. Global frequency for A at this SNP is 0.188 (therefore frequency of G is 0.812).

Population	n	A	G
African American	32510	0.97	0.03
Central American	2450	0.81	0.19
European	18872	0.86	0.14
Native American	1260	0.86	0.14
Native Hawaiian	4534	0.75	0.25
South Asian	856	0.59	0.41

- Combine the data sets and test whether genotype frequency is different from the reported global exception.
- Test each population separately.