Stat 135
7/22/2021

Lab 10 Solution

1. Goodness of Fit
With a perfectly balanced roulette wheel, in the long run, red numbers should show up 18 times in 38. To test its wheel, one casino records the results of 3800 plays, finding 1890 red numbers. Is that too many reds? Or chance variation? Formulate the test, starting your hypotheses, significance level, and the p-value. Recall that a roulette wheel has 38 numbers: 18 red, 18 black and 2 green.

$H_0: P_{red} = \frac{18}{38}, \; P_{not\ red} = \frac{20}{38}$ versus H_1: H_0 is not true.

We want to test if the wheel we are observing follows what we would expect from a roulette wheel in the long run. Our hypothesis is that the probability of observing red is 18/38, our alternative is that the distribution of reds is not as specified. Using the generalized likelyhood ratio test for multinomials with 2 labels (RED, and NOT RED), we use pearson's chisquare statistic with 1 degrees of freedom and aim to reject the null at the $\alpha = .05$ level. We observed 1890 red numbers and 1910 non reds thus using the chisquare statistic with 1 degrees of freedom (1 estimated parameter under alternative - 0 estimated parameters under null) and the expected observations of 1800 and 2000 respectively. The chisquare statistic is 8.55. The p-value of observing a result as extreme as this found using the code (pchisq(8.55, df=1, lower.tail = FALSE)); The p-value of this estimate under the null is 0.003, so we can reject the null hypothesis that the data was observed from a multinomial with parameters (18/38, 20/38).

2. Goodness of Fit (R)
Rice 9.38: Yip et al. (2000) studied seasonal variations in suicide rates in England and Wales during 1982-1996, collecting counts shown in the following table:

Month	Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec
Male	3755	3251	3777	3706	3717	3660	3669	3626	3481	3590	3605	3392
Female	1362	1244	1496	1452	1448	1376	1370	1301	1337	1351	1416	1226

Do either the male or female data show seasonality? (Hint: the null hypothesis should assume the same death probability for each DAY)

Solution see the Shiny app.

3. Test of Homogeneity

Rice 13.2: Phillips and Smith (1990) conducted a study to investigate whether people could briefly postpone their deaths until after the occurrence of a significant occasion. The senior-woman of the household plays a central ceremonial role in the Chinese Harvest Moon Festival. Phillips and Smith compared the mortality patterns of old Jewish women and old Chinese women who died of natural causes for the weeks immediately preceding and following the festival, using records from California for the years 19601984. Compare the mortality patterns shown in the table. (Week 1 is the week preceding the festival, week 1 is the week following, etc.)

Week	Chinese	Jewish	**Total**
-2	55	141	196
-1	33	145	178
1	70	139	209
2	49	161	210
Total	207	586	793

This problem lends itself to a Chisquare test for Homogenity; that is, our null hypothesis is that the death rates of Jewish and Chinese Women in 1 and 2 weeks before and after the Lunar Festival should be drawn from the same distribution. I.e. the null is that we expect should no difference between the mortality rates in this, otherwise, random interval. The alternative would suggest that the presence of the Lunar Festival had some effect on mortality rates for the Chinese women in comparison to the Jewish women. Setting an $\alpha = 0.05$.

Here is a modified table with the row totals added in:

Week	Jewish	Chinese	
-2	55	141	196
-1	33	145	178
1	70	139	209
2	49	161	210
	207	586	793

The Chisquare statistic of 3 degrees of freedom (4 - 1 labels) x (2 - 1 groups) is as follows:

$$X^2 = (55 - 51.16)^2/(51.16) + ... + (161 - 155.18)^2/(155.18) = 12.421$$

This yields a p-value of 0.006072317, which is far below our significance level. It is then resonable to say that Chinese Women have different mortality rates in the the interval of weeks around the Lunar Festival than Jewish women.

4. Test of Independence

It is conventional wisdom in military squadrons that pilots tend to father more girls than boys. Snyder (1961) gathered data for military fighter pilots. The sex of the pilots offspring were tabulated for three kinds of flight duty during the month of conception, as shown in the following table. Is there any significant difference between the three groups? In the United States in 1950, 105.37 males were born for every 100 females. Are the data consistent with this sex ratio?

Father's Activity	Female Offspring	Male Offspring	**Total**
Flying Fighters	51	38	89
Flying Transports	14	16	30
Not Flying	38	46	84
Total	103	100	203

First we want to do another Homogenity test among the three groups, and compare if their distribution of female and male offspring is the same. The statistic is as follows:

Observed

	Females	Males	
flying fighters	51	38	89
flying transports	14	16	30
not flying	38	46	84
	103	100	203

Expected

	Female	Males
flying fighters	45.15764	43.84236
flying transports	15.22167	14.77833
not flying	42.62069	41.37931

This yields a statistic of:

$$X^2 = (51 - 45.15764)^2/(45.15764) + ... + (46 - 41.37931)^2/(41.37931) = 2.75$$

The p-value – $P(X^2_{df=2} > 2.75) = 0.2528$, below our preset alpha – so it is resonable to hold to the belief that the offspring distributions of these three military groups are the same.

Second, we get the goodness of fit given the null hypothesis that the probabilities of female offspring and male offspring are 100/205.37 and 105.37/205.37. Find E_ij:

	Expected_Flying_fighters	Expected_Flying_transport	Expected_Not_flying
Female Offspring	43.33642	14.60778	40.90179
Male Offspring	45.66358	15.39222	43.09821

Let $\alpha = 0.05$. Using Pearson's chi-square test, the null distribution X^2 is approximately chi-square with 3 degrees of freedom (3 parameters under alternative hypothesis and 0 free parameter under null hypothesis).

$$X^2 = \sum_{i=1}^{2}\sum_{j=1}^{3}\frac{(O_{ij} - E_{ij})^2}{E_{ij}} = 3.09$$

Since 3.09 is less than the 95% percentile of chi square distribution with 3 degrees of freedom, which is 7.81, we accept the null hypothesis. Therefore, the data of pilots are consistent with the birth ratio.