Homework 4 - Solutions

STAT 212 (Fall 2022)

1

We have one population for which one categorical variable with 8 categories is being observed. We want to use a goodness of fit test where the null probabilities of categories are pre-specified. This is the case since it is hypothesized that pigeons' direction of flight will be uniformly distributed over [0, 360] and this density has no parameters to be estimated. We have bucketed the directions into 45 degree increments to get a cateorical distribution. Our hypotheses to be tested are:

$$H_0 : p_1 = p_2 = \ldots = p_8 = \frac{1}{8} \quad \text{vs.}$$
$$H_a : p_i \neq p_j \text{ for at least one } (i, j).$$

```
freq <- c(12, 16, 17, 15, 13, 20, 17, 10)
exp_freq <- sum(freq) / 8
chi_sqr_stat <- sum((freq - exp_freq)^2) / exp_freq
critical <- qchisq(p = .9, df = 7)
pval <- 1 - pchisq(chi_sqr_stat, df = 7)
```

Test.statistic	Critival.value	P.value
4.8	12.01704	0.6843549

Since the test statistic is less than the critical value (and p-value is 0.68), we fail to reject the null at $\alpha = 0.1$.

2

We are testing the fit of a Poisson distribution by looking at the proportion of observations that fall into bins. The Poisson has a parameter θ that we must estimate from the data. We use the goodness of fit test where the probability of falling into a bin is given by the estimated Poisson distribution. The hypotheses are:

$$H_0 : p_1 = \pi_1(\theta), \ldots, p_9 = \pi_9(\theta) \quad \text{vs.}$$
$$H_a : \text{At least one of these does not hold.}$$

```
vals <- 0:9
counts <- c(6, 24, 42, 59, 62, 44, 41, 14, 6, 2)
mle_est <- sum(vals * counts / sum(counts))
counts_trunc <- c(6, 24, 42, 59, 62, 44, 41, 14, 8)
probs <- c(dpois(x = 0:7, lambda = mle_est), 1 - ppois(7, lambda = mle_est))
exp_counts <- sum(counts_trunc) * probs
chi_sqr_stat <- sum((counts_trunc - exp_counts) ^ 2 / exp_counts)
critical <- qchisq(p = .9, df = 7)
pval <- 1 - pchisq(chi_sqr_stat, df = 7)
```

Test.statistic	Critival.value	P.value
7.810415	12.01704	0.349608

The p-value is greater than 0.1, so we fail to reject the null hypothesis at the level of significance $\alpha = 0.1$, and conclude that the Poisson distribution seems to be a good fit to the observed data.

3 Problem 10, pg 360

We have three populations we are interested in that are determined the materials A, B, and C used in children's pajamas. Each population can be divided into the same two categories: ignited and did not ignite. We want to know if the probability of ignition is the same for the different materials. Hence, we are precisely in a contigency table setting characterized by $k = 3$ populations, and categorical variable with $r = 2$ categories is being observed for each population. Hence, the χ^2 test for testing homogeniety of proportions is appropriate here. Let $p_i,\ i \in \{A, B, C\}$ be the probability of ignition for the ith material. Our hypotheses are:

$$H_0 : p_A = p_B = p_C \quad \text{vs.}$$
$$H_a : \text{This hypothesis is false.}$$

We observed:

Material	Ignited	Did not ignite	Total
A	37	74	111
B	28	57	85
C	21	79	100
—	—	—	—
Total	86	210	296

If the null is true, the expected cell counts are:

	Ignite	Not ignite
A	32.25	78.75
B	24.69595	60.30405
C	29.05405	70.94595

Our test statistic is $\chi^2 = 4.756$. The critical value is $\chi^2_{2,0.05} = 5.991$. Since $\chi^2 < 5.991$, we fail to reject H_0. There is not enough evidence to conclude that the 3 materials differ with respect to probability of ignition.