Module 8 Solutions

Xihan Qian

June, 2022

1. Let's get familiar with how to read p-values from χ^2 statistics using the tables in the book. For example, using these tables, we can see that $P\left(\chi^2_5 \geq 16\right) \in [0.005, 0.01]$. Find the following values from the tables; you do not need a specific value, but rather a range.

(a)$P(\chi^2_{10} \geq 16.536)$
(b)$P(\chi^2_3 \geq 8.20)$

Solutions: (a) $0.05 < p$-value < 0.10.
(b) $0.02 < p$-value < 0.05.

2. Towards the end of the Spring quarter, students at UCSD have made plans for their summer. Your friend thinks that 40% of students will spend their summer outside of the U.S., 35% will spend their summer in California; and 25% will spend their summer outside of California but inside of the U.S.. You then surveyed a random sample of students and found that 30 students will spend their summer outside of the U.S., 30 students will spend their summer in California, and 30 students will spend their summer outside of California but inside of the U.S. You would like to conduct a hypothesis test to check whether your friend is right.

(a) Define the parameter of interest, and write down the null and alternative hypotheses.
(b) What are the expected values of the number of students in each category, under the null hypothesis?
(c) Compute the χ^2 statistics.
(d) Before proceeding with the test, quickly check whether the conditions are met.
(e) How many degrees of freedom are associated with this setup? Find the p-value.
(f) Write down an appropriate conclusion with significance level $\alpha = 0.05$.

Solutions: (a) p_1 : proportion of students that will spend their summer outside of the U.S.;
p_2 : proportion of students that will spend their summer in California;
p_3 : proportion of students that will spend their summer outside of California but inside of the U.S.;

$H_0 : p_1 = 0.4; p_2 = 0.35; p_3 = 0.25.$
$H_A : p_1 \neq 0.4$ or $p_2 \neq 0.35$ or $p_3 \neq 0.25$.

(b) Expected number outside of U.S.: $90 \times 0.4 = 36$;
Expected number in California: $90 \times 0.35 = 31.5$;

Expected number outside of California but inside of the U.S.: $90 \times 0.25 = 22.5$

(c)

$$\chi^2 = \frac{(30-36)^2}{36} + \frac{(30-31.5)^2}{31.5} + \frac{(30-22.5)^2}{22.5} = 3.571$$

(d) Independence is satisfied within and across groups because this is a random sample; sample size is also fine since each number has at least 5 expected cases.

(e) The DOF is $k-1 = 3-1 = 2$. The p-value is computed as: $P(\chi^2 \geq X^2) = P(\chi_2^2 \geq 3.571)$. After reading from the table, we can get that p-value > 0.10.

(f) Since p-value $> 0.10.$, in particular, p-value $> 0.10 > 0.05$, we fail to reject the null hypothesis and don't have enough evidence against the proposed values.

3. Suppose that in a university, 40% are arts majors and 60% are science majors. In a random sample, it was observed that out of the 200 students, 95 are arts majors while 105 are science majors. Is this representative of the population? To conclude, we use a Chi-squared Goodness-of-fit test.

(a) Define the parameter of interest, and write down the null and alternative hypotheses.

(b) What are the expected values of the number of students in each category, under the null hypothesis?

(c) Compute the χ^2 statistics.

(d) Before proceeding with the test, quickly check whether the conditions are met.

(e) How many degrees of freedom are associated with this setup? Find the p-value.

(f) Write down an appropriate conclusion with significance level $\alpha = 0.05$.

Solutions: (a) p_1 : proportion of students that are arts majors.;
p_2 : proportion of students that are science majors;

$H_0 : p_1 = 0.4; p_2 = 0.6;$
$H_A : p_1 \neq 0.4$ or $p_2 \neq 0.6$

(b) Expected number of art students.: $200 \times 0.4 = 80;$
Expected number of science students: $200 \times 0.6 = 120;$

(c)

$$\chi^2 = \frac{(95-80)^2}{80} + \frac{(105-120)^2}{120} = 4.688$$

(d) Independence is satisfied within and across groups because this is a random sample; sample size is also fine since each number has at least 5 expected cases.

(e) The DOF is $k-1 = 2-1 = 1$. The p-value is computed as: $P(\chi^2 \geq X^2) = P(\chi_1^2 \geq 4.688)$. After reading from the table, we can get that $.02 < p$-value < 0.05.

(f) Since p-value $< 0.05.$, in particular, p-value $< \alpha$, we can reject the null hypothesis and can conclude the proportions of 40% of art majors and 60% of science majors at this college are not representative of majors selected by the students.