MODULE 8 EXAMPLES

TAs: Nihal Reddy
Email: nireddy@ucsd.edu
OH: Thursdays 6-7pm

Slide Credits: Kira Fleischer

PROBLEM #: KEY TOPICS FROM PROBLEM

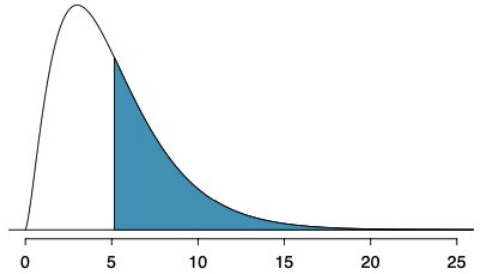
Problem setup and description.

Question

Key notes from readings/lectures needed to answer the question

Solution: written with as much detail as we expect you to give on your homework sets

PROBLEM 1: READING CHI-SQUARED TABLE



Let's get familiar with how to read p-values from χ^2 statistics. For example, using the table, we can see that $P(\chi_5^2 \geq 16) \in [0.005, 0.01]$.

Compute the following values from the table (page 417 of textbook); you do not need a specific value, but rather a range.

(a) $\boldsymbol{P(\chi_4^2 \geq 10.627)}$

Solution: 0.02 < p-value < 0.05

(b) $\boldsymbol{P(\chi_1^2 \geq 12.372)}$

Solution: p-value < 0.001

Upper tail		0.3	0.2	0.1	0.05	0.02	0.01	0.005	0.001
df	1	1.07	1.64	2.71	3.84	5.41	6.63	7.88	10.83
	2	2.41	3.22	4.61	5.99	7.82	9.21	10.60	13.82
	3	3.66	4.64	6.25	7.81	9.84	11.34	12.84	16.27
	4	4.88	5.99	7.78	9.49	11.67	13.28	14.86	18.47
	5	6.06	7.29	9.24	11.07	13.39	15.09	16.75	20.52
	6	7.23	8.56	10.64	12.59	15.03	16.81	18.55	22.46
	7	8.38	9.80	12.02	14.07	16.62	18.48	20.28	24.32
	8	9.52	11.03	13.36	15.51	18.17	20.09	21.95	26.12
	9	10.66	12.24	14.68	16.92	19.68	21.67	23.59	27.88
	10	11.78	13.44	15.99	18.31	21.16	23.21	25.19	29.59
	11	12.90	14.63	17.28	19.68	22.62	24.72	26.76	31.26
	12	14.01	15.81	18.55	21.03	24.05	26.22	28.30	32.91
	13	15.12	16.98	19.81	22.36	25.47	27.69	29.82	34.53
	14	16.22	18.15	21.06	23.68	26.87	29.14	31.32	36.12
	15	17.32	19.31	22.31	25.00	28.26	30.58	32.80	37.70
	16	18.42	20.47	23.54	26.30	29.63	32.00	34.27	39.25
	17	19.51	21.61	24.77	27.59	31.00	33.41	35.72	40.79
	18	20.60	22.76	25.99	28.87	32.35	34.81	37.16	42.31
	19	21.69	23.90	27.20	30.14	33.69	36.19	38.58	43.82
	20	22.77	25.04	28.41	31.41	35.02	37.57	40.00	45.31
	25	28.17	30.68	34.38	37.65	41.57	44.31	46.93	52.62
	30	33.53	36.25	40.26	43.77	47.96	50.89	53.67	59.70
	40	44.16	47.27	51.81	55.76	60.44	63.69	66.77	73.40
	50	54.72	58.16	63.17	67.50	72.61	76.15	79.49	86.66

PROBLEM 2: CONDUCTING CHI-SQUARED GOODNESS-OF-FIT TEST

A hospital administrator wants to find out if patient check-ins are evenly distributed across the weekdays. They randomly sample 210 records from the previous month and find patients checked in as follows: Monday–32 patients, Tuesday–40 patients, Wednesday–36 patients, Thursday–45 patients, Friday–57 patients. If check-ins were evenly distributed, the administrator would expect an equal number of patients each weekday. To test this, they conduct a chi-squared goodness-of-fit test.

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

Solution: p_1: proportion of patients on Monday; p_2: proportion of patients on Tuesday;

p_3: proportion of patients on Wednesday; p_4: proportion of patients on Thursday;

p_5: proportion of patients on Friday

H_0: $p_1 = p_2 = p_3 = p_4 = p_5 = \frac{1}{5}$

H_A: $p_1 \neq \frac{1}{5}$ or $p_2 \neq \frac{1}{5}$ or $p_3 \neq \frac{1}{5}$ or $p_4 \neq \frac{1}{5}$ or $p_5 \neq \frac{1}{5}$

PROBLEM 2: CONDUCTING CHI-SQUARED GOODNESS-OF-FIT TEST

A hospital administrator wants to find out if patient check-ins are evenly distributed across the weekdays. They randomly sample 210 records from the previous month and find patients checked in as follows: Monday–32 patients, Tuesday–40 patients, Wednesday–36 patients, Thursday–45 patients, Friday–57 patients. If check-ins were evenly distributed, the administrator would expect an equal number of patients each weekday. To test this, they conduct a chi-squared goodness-of-fit test.

(b) What are the expected values of the number of patients each day, under the null hypothesis?

Solution: Expected number of patients each day are all the same under the null hypothesis: $\frac{210}{5} = 42$ patients each day.

PROBLEM 2: CONDUCTING CHI-SQUARED GOODNESS-OF-FIT TEST

A hospital administrator wants to find out if patient check-ins are evenly distributed across the weekdays. They randomly sample 210 records from the previous month and find patients checked in as follows: Monday–32 patients, Tuesday–40 patients, Wednesday–36 patients, Thursday–45 patients, Friday–57 patients. If check-ins were evenly distributed, the administrator would expect an equal number of patients each weekday. To test this, they conduct a chi-squared goodness-of-fit test.

(c) Compute the X^2 statistic.

$X^2 = \sum_{i=1}^{k} \frac{(observed_i - expected_i)^2}{expected_i}$ for k categories

Solution: $X^2 = \frac{(32-42)^2}{42} + \frac{(40-42)^2}{42} + \frac{(36-42)^2}{42} + \frac{(45-42)^2}{42} + \frac{(57-42)^2}{42} = 8.905$

PROBLEM 2: CONDUCTING CHI-SQUARED GOODNESS-OF-FIT TEST

A hospital administrator wants to find out if patient check-ins are evenly distributed across the weekdays. They randomly sample 210 records from the previous month and find patients checked in as follows: Monday–32 patients, Tuesday–40 patients, Wednesday–36 patients, Thursday–45 patients, Friday–57 patients. If check-ins were evenly distributed, the administrator would expect an equal number of patients each weekday. To test this, they conduct a chi-squared goodness-of-fit test.

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

Two conditions for chi-squared test:

1. Independence: independence within and across groups is satisfied if the sample is random

2. Sample size: each expected count must be at least 5

Solution: Independence is satisfied within and across groups because this is a random sample; sample size is also fine since the five expected counts are all 42, which is greater than 5.

PROBLEM 2: CONDUCTING CHI-SQUARED GOODNESS-OF-FIT TEST

A hospital administrator wants to find out if patient check-ins are evenly distributed across the weekdays. They randomly sample 210 records from the previous month and find patients checked in as follows: Monday–32 patients, Tuesday–40 patients, Wednesday–36 patients, Thursday–45 patients, Friday–57 patients. If check-ins were evenly distributed, the administrator would expect an equal number of patients each weekday. To test this, they conduct a chi-squared goodness-of-fit test.

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

For k categories, there are k-1 degrees of freedom

p-value: $P(\chi^2_{df} \geq X^2)$

Solution: DOF is $k - 1 = 5 - 1 = 4$. The p-value is $P(\chi^2_4 \geq 8.905)$. Reading the table, we get that 0.05 < p-value < 0.10.

PROBLEM 2: CONDUCTING CHI-SQUARED GOODNESS-OF-FIT TEST

A hospital administrator wants to find out if patient check-ins are evenly distributed across the weekdays. They randomly sample 210 records from the previous month and find patients checked in as follows: Monday–32 patients, Tuesday–40 patients, Wednesday–36 patients, Thursday–45 patients, Friday–57 patients. If check-ins were evenly distributed, the administrator would expect an equal number of patients each weekday. To test this, they conduct a chi-squared goodness-of-fit test.

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

Solution: Since the p-value is between 0.05 and 0.10, we have p-value > $\alpha = 0.05$, so we fail to reject the null hypothesis. We do not have evidence that the number of patients who check in each day is different.