Worksheet week 7

Q1. A genetics engineer was attempting to cross a tiger and a cheetah. She predicted a phenotypic outcome of the traits she was observing to be in the following ratio 4 stripes only: 3 spots only: 9 both stripes and spots. When the cross was performed and she counted the individuals she found 50 with stripes only, 41 with spots only and 85 with both. According to the Chi-square test, did she get the predicted outcome?

Solution:

Chi-square = $\sum (O-E)^2/E$

D.F.	Value
1	3.841
2	5.991
3	7.815

Set up a table to keep track of the calculations:

Expected ratio	Observed #	Expected #	O-E	$(O-E)^2$	$(O-E)^2/E$
4 stripes	50	44	6	36	0.82
3 spots	41	33	8	64	1.94
9 stripes/spots	85	99	-14	196	1.98
16 total	176 total	176 total	0 total		Sum = 4.74

4/16 * 176 = expected # of stripes = 44 3/16 * 176 = expected # of spots = 33 9/16 * 176 = expected # stripes/spots = 99

Degrees of Freedom = 3 - 1 = 2 (3 different characteristics - stripes, spots, or both)

Since 4.74 is less than 5.991, I can accept the null hypothesis put forward by the engineer.

Q2. Nadir is testing an octahedral die to see if it is unbiased. The results are given in the table below.

Score	1	2	3	4	5	6	7	8
Frequency	7	10	11	9	12	10	14	7

Test the hypothesis that the die is fair.

Solution:

Using χ^2, the number of degrees of freedom is $8 - 1 = 7$, so at the 5% significance level the critical value of χ^2 is 14.07. As before, a table of values is drawn up, the expected frequencies being based on a uniform distribution which gives

$$\text{frequency for each result} = \frac{1}{8}(7+10+11+9+12+10+14+7) = 10$$

O_i	E_i	$O_i - E_i$	$(O_i - E_i)^2$	$\frac{(O_i - E_i)^2}{E_i}$
7	10	−3	9	0.9
10	10	0	0	0
11	10	1	1	0.1
9	10	−1	1	0.1
12	10	2	4	0.4
10	10	0	0	0
14	10	4	16	1.6
7	10	−3	9	0.9
				4.0

The calculated value of $|^2$ is 4.0. This is well within the critical value, so Nadir could conclude that there is evidence to support the hypothesis that the die is fair.

Q3. Employers particularly want to know which days of the week employees are absent in a five day work week. Most employers would like to believe that employees are absent equally during the week. That is, the average number of times an employee is absent is the same on Monday, Tuesday, Wednesday, Thursday, or Friday. Suppose a sample of 20 absent days was taken and the days absent were distributed as follows:

Day of the Week Absent

	Monday	**Tuesday**	**Wednesday**	**Thursday**	**Friday**
Number of Absences	5	4	2	3	6

Problem

For the population of employees, do the absent days occur with equal frequencies during a five day work week? Test at a 5% significance level.

Solution

The null and alternate hypotheses are:

- H_o: The absent days occur with equal frequencies, that is, they fit a uniform distribution.
- H_a: The absent days occur with unequal frequencies, that is, they do not fit a uniform distribution.

If the absent days occur with equal frequencies, then, out of 20 absent days, there would be 4 absences on Monday, 4 on Tuesday, 4 on Wednesday, 4 on Thursday, and 4 on Friday. These numbers are the **expected** (E) values. The values in the table are the **observed** (O) values or data.

This time, calculate the χ^2 test statistic by hand. Make a chart with the following headings:

- Expected (E) values
- Observed (O) values
- $(O - E)$
- $(O-E)^2$
- $\frac{(O-E)^2}{E}$

Now add (sum) the last column. Verify that the sum is 2.5. This is the χ^2 test statistic. To find the p-value, calculate $P(\chi^2 > 2.5)$. This test is right-tailed.
The dfs are the number of cells − 1 = 4.

Use a computer or calculator to find the p-value. You should get p-value = 0.6446. The decision is to not reject the null hypothesis.

Conclusion: At a 5% level of significance, from the sample data, there is not sufficient evidence to conclude that the absent days do not occur with equal frequencies.