2-Way Hypothesis Testing Practice

1. A student noticed that her English teacher, Mrs. Jones, has one of the highest rates of dress coding people throughout the year. The student decided to look into whether the distribution of dress coding (leveled by severity A or B) is (equal) between genders. [Notes: Severity A dress code is receiving clothes to cover up, Severity B dress code is being sent home to change.] Determine whether there is an equal distribution between genders in terms of dress codes, when the level of significance is 10%.

	Severity A Dress Code	Severity B Dress Code	
Male	34	4	⌐38
Female	55	63	⌐118

89   67   [n=156]

   a. What kind of 2-Way test is this?

   Test of Homogeneity

   b. State the Hypotheses.

   H_0: The distributions are Homogeneous. H_1: The dist. are not Homogeneous.

   c. Calculate the expected counts.

   $\mathcal{E}_{M,A} = \frac{(89)(38)}{156} = 21.68$ $\mathcal{E}_{M,B} = \frac{(67)(38)}{156} = 16.32$

   $\mathcal{E}_{F,A} = \frac{(89)(118)}{156} = 67.32$ $\mathcal{E}_{F,B} = \frac{(67)(118)}{156} = 50.68$

   d. What is the degrees of freedom?

   $df = (2-1)(2-1) = 1$

   e. What is the test statistic?

   $TS = \frac{(34-21.68)^2}{21.68} + \frac{(4-16.32)^2}{16.32} + \frac{(55-67.32)^2}{67.32} + \frac{(63-50.68)^2}{50.68} = 21.55$

   f. Calculate the p-value. What is the decision? $\alpha = 0.1$

   $p = \chi^2 cdf(21.55, E99, 1) = 3.45 \times 10^{-6} \approx 0.0000 < 0.1$

   Reject H_0

   g. Choose the correct interpretation.
   ☑ At 10% level of significance, there is sufficient evidence to support that the variables are not homogeneous.

- [ ] At 10% level of significance, there is insufficient evidence to support that the variables are not homogeneous.

2. Jackson, at random, wondered whether there is a relationship between high school level (freshman, sophomore, junior, and senior) and pet species preference (cat, dog, other, or none). Complete the questions using the observed count table below.

	Cat	Dog	Other	None	Row Totals
Freshman	53	68	34	35	190
Sophomore	72	60	85	66	283
Junior	66	67	21	44	198
Senior	89	45	63	38	235
Column Totals	280	240	203	183	Overall Total
					906

a. What kind of 2-Way test is this?

Test of Independence

b. State the Hypotheses.

H_0: The variables are Independent. H_1: The variables are dependent.

c. Calculate the Expected Counts in the table below. (found via Matrix)

	Cat	Dog	Other	None
Freshman	58.72	50.331	42.572	38.377
Sophomore	87.461	74.967	63.409	57.162
Junior	61.192	52.45	44.364	39.943
Senior	72.627	62.252	52.655	47.467

Via χ^2-Test

d. What is the Test Statistic?

$$TS = \chi^2 = 52.73$$

e. What is the degrees of freedom?

$$df = 9$$

f. Calculate the p-value. What is the decision? $\alpha = 0.05$

$$p = 3.28 \times 10^{-8} \approx 0.0000 < 0.05$$

Reject H₀

g. Choose the correct interpretation.

☒ At 5% level of significance, there is **sufficient** evidence to support that there is an association between high school level and preferred pet.

☐ At 5% level of significance, there is insufficient evidence to support that there is an association between high school level and preferred pet.

3. Doctor Smith noticed that a certain gender tends to require a Tonsillectomy before the age of 10, while others require it later in life. Thus, using the data pulled from 58 previous Tonsillectomies, determine if there is a relationship between gender and when a Tonsillectomy, on average, is required. [*Note: the level of significance is 15%]

		10 & Under	Over 10	Row Totals
Male	**Observed**	21	A?	35
	Expected	B?	16.90	
Female	**Observed**	C?	14	D?
	Expected	11.90	E?	
Column Totals		30	F?	**Overall Total**
				58

$A = 35 - 21 = 14$ 　　　　 $B = \dfrac{(35)(30)}{58} = 18.16$ 　　　　 $C = 30 - 21 = 9$

$D = 58 - 35 = 23$ 　　　　 $E = \dfrac{(23)(28)}{58} = 11.10$ 　　　　 $F = 58 - 30 = 28$

a. What kind of 2-Way test is this?

Test of Independence

b. State the Hypotheses.

H_0: Variables are Independent. H_1: Variable are Dependent.

c. What is the Test Statistic?

$$TS = \frac{(21-18.10)^2}{18.10} + \frac{(14-16.90)^2}{16.90} + \frac{(9-11.90)^2}{11.90} + \frac{(14-11.10)^2}{11.10} = 2.43$$

d. What is the degrees of freedom?

$$df = (2-1)(2-1) = 1$$

e. Calculate the p-value. What is the decision? $\alpha = 0.15$

$$p = \chi^2 cdf(2.43, E99, 1) = 0.1190 < 0.15,$$

Reject H_0

f. Choose the correct interpretation.

☒ At 15% level of significance, there is sufficient evidence to support that there is an association between gender and when a Tonsillectomy is required.

❑ At 15% level of significance, there is insufficient evidence to support that there is an association between gender and when a Tonsillectomy is required.

4. John, the director of a dog shelter, noticed that out of all the dogs present, certain sizes (large, medium, small) are usually preferred more than others. Similarly, he also noticed that the status of the prospective owner (single, married, divorced, widowed) seems to play a part. Determine whether the various dog breeds are equally distributed between the owner's status when the level of significance is 1%.

		Single	Married	Divorced	Widowed	Row Totals
Large	Observed	9	7	A?	8	34
Dogs	Expected	B?	8.5	8.19	8.5	
Medium	Observed	14	11	9	11	C?
Dogs	Expected	11.67	D?	10.83	11.25	

Small	Observed	E?	9	7	8	29
Dogs	Expected	7.52	F?	6.98	7.25	
Column Totals		28	27	G?	27	**Overall Total**
						H?

$A = 34 - (9+7+8)$
$= 10$

$B = \dfrac{(34)(28)}{108}$
$= 8.81$

$C = 14+11+9+11$
$= 45$

$D = \dfrac{(45)(27)}{108}$
$= 11.25$

$E = 29 - (9+7+8)$
$= 5$

$F = \dfrac{(29)(27)}{108}$
$= 7.25$

$G = 10+9+7$
$= 26$

$H = 34+45+29$
$= 108$

a. What kind of 2-Way test is this?

Test of Homogeneity

b. State the Hypotheses.

H_0: The distributions are Homogeneous. H_1: The dist. are not Homo.

(left margin, rotated): $(A) = $ Observed $(B) = $ Expected   Using Matrix & χ^2-Test

c. What is the Test Statistic?

$TS = \chi^2 = 2.83$

d. What is the degrees of freedom?

$df = 6$

e. Calculate the p-value. What is the decision? $\alpha = 0.01$

$p = 0.8296 > 0.01 \longrightarrow$ **fail to Reject H_0**

f. Choose the correct interpretation.
- ❑ At 1% level of significance, there is sufficient evidence to support that the variables are not homogeneous.
- ☒ At 1% level of significance, there is insufficient evidence to support that the variables are not homogeneous.