

AUTHOR BY:

Sana Ur Rehman Arain

LEARNING JOURNAL 3

University of the People

2024

INSTRUCTOR: ANSON XUAN

MATH 1281-01 Statistical Inference – AY2025-T3

Part 1: Barking Deer Foraging Preferences

a. Hypotheses for testing if barking deer prefer to forage in certain habitats over others:

- **Null Hypothesis (H_0):** Barking deer do not prefer any specific habitat for foraging. The observed distribution of foraging sites across habitats is consistent with the distribution of available habitats in the region.
- **Alternative Hypothesis (H_1):** Barking deer prefer certain habitats for foraging. The observed distribution of foraging sites across habitats is not consistent with the distribution of available habitats in the region.

b. Type of test:

We can use a **chi-squared goodness-of-fit** test to determine if the observed distribution of foraging sites matches the expected distribution based on the availability of habitats.

c. Assumptions and conditions:

1. **Independence:** The foraging sites are assumed to be independent of each other.
2. **Sample size:** Each expected frequency should be at least 5. We will check this after calculating the expected frequencies.

d. Hypothesis test:

1. Calculate expected frequencies:

- Total sites = 530
- Expected frequency for Woods = 4.8% of 530 = $0.048 * 530 = 25.44$



- Expected frequency for Cultivated Grassplot = 14.7% of 530 = $0.147 * 530 = 77.91$
- Expected frequency for Deciduous Forests = 39.6% of 530 = $0.396 * 530 = 209.88$
- Expected frequency for Other = $100\% - (4.8\% + 14.7\% + 39.6\%) = 40.9\%$ of 530
 $= 0.409 * 530 = 216.77$

2. Check assumptions:

- All expected frequencies are greater than 5, so the chi-squared test is appropriate.

3. Calculate the chi-squared statistic:

$$\chi^2 = \sum [(O_i - E_i)^2 / E_i]$$

$$\chi^2 = (6 - 25.44)^2 / 25.44 + (18 - 77.91)^2 / 77.91 + (71 - 209.88)^2 / 209.88 + (435 - 216.77)^2 / 216.77$$

$$\chi^2 = 14.855 + 44.0686 + 91.898 + 219.67$$

$$\chi^2 = 370.492$$

4. Degrees of freedom:

$$df = \text{Number of categories} - 1 = 4 - 1 = 3$$

5. Conclusion:

- Given that the p-value < 0.001 , we reject the null hypothesis.
- There is strong evidence to conclude that barking deer prefer certain habitats for foraging over others.

a. Calculate the actual number of site visitors in each of the six response categories:

- Total visitors = 501
- **Position 1:**
 - Download: 16.0% of 501 = $0.16 * 501 = 80.16 \approx 80$
 - No Download: 20.9% of 501 = $0.209 * 501 = 104.71 \approx 105$
- **Position 2:**
 - Download: 14.8% of 501 = $0.148 * 501 = 74.15 \approx 74$
 - No Download: 21.2% of 501 = $0.212 * 501 = 106.21 \approx 106$
- **Position 3:**
 - Download: 11.9% of 501 = $0.119 * 501 = 59.62 \approx 60$
 - No Download: 15.2% of 501 = $0.152 * 501 = 76.15 \approx 76$

b. Test for group imbalance:

1. Hypotheses:

- Null Hypothesis (H_0): The groups are balanced. The distribution of site visitors across positions is equal.
- Alternative Hypothesis (H_1): The groups are imbalanced. The distribution of site visitors across positions is not equal.

2. Conditions:

- The expected frequency for each group is $501 / 3 = 167$.
- All expected frequencies are greater than 5, so the chi-squared test is appropriate.

3. Test statistics:

$$\chi^2 = \sum [(O_i - E_i)^2 / E_i]$$

$$\chi^2 = (80 - 167)^2 / 167 + (105 - 167)^2 / 167 + (74 - 167)^2 / 167 + (106 - 167)^2 / 167 + (60 - 167)^2 / 167 + (76 - 167)^2 / 167$$

$$\chi^2 = 45.32 + 23.02 + 51.79 + 22.28 + 68.56 + 49.59$$

$$\chi^2 = 260.56$$

4. Degrees of freedom:

$$df = \text{Number of categories} - 1 = 6 - 1 = 5$$

5. Conclusion:

- Given that the p-value = 0.01215, which is less than the significance level of 0.05, we reject the null hypothesis.
 - There is evidence to suggest that the groups were imbalanced. The distribution of site visitors across positions is not equal.
-