1 Election Votes Distribution

A city records election votes for three candidates: A, B, and C. The expected votes were equal, but the observed data shows differences. Does the vote distribution match expectations?

```
import numpy as np
import scipy.stats as stats

observed = np.array([500, 600, 550])
expected = np.array([550, 550, 550])


#  H₀: The observed vote distribution follows the expected uniform
distribution.
#  H₁: The observed vote distribution does not match the expected
distribution.

chi2,p_value = stats.chisquare(observed,expected)

print(f'p_value {p_value}\n')

alpha = 0.05

if p_value < alpha:
  print('Reject the null hypothesis: The observed vote distribution do
not follows the expected uniform distribution')
else:
  print('Faild to Reject the null hypothesis: The observed vote
distribution follows the expected uniform distribution')
```

2 Factory Defective Product Rates

A factory expects 5% defective products, but observed values differ. Does the defect rate follow the expected distribution?

```
observed = np.array([50, 950])  # Defective, Non-defective
expected = np.array([1000 * 0.05, 1000 * 0.95])  # 5% defective


# H₀: The defect rate follows the expected 5% rate.
# H₁: The defect rate does not match expectations.

chi2_stats , p_value = stats.chisquare(observed,expected)

print(f'chi2_stats = {chi2_stats} \np_value = {p_value} \n{50*"="}')

alpha = 0.05

if p_value < alpha:
```

```
  print('Reject The Null Hypothesis :The defect rate do not follows
the expected 5% rate')
else:
  print('Fail to reject the Null Hypothesis: defect rate follows the
expected 5% rate')
```

3 Car Color Preference

A survey recorded customer preferences for car colors (Red, Blue, Black, White). Are they equally preferred?

```
observed = np.array([250, 300, 280, 270])
expected = np.array([275, 275, 275, 275])  # Equal expected values
print((250+300+280+270)/4)

# H₀: Customers equally prefer all colors.
# H₁: Preferences vary.

chi2_stats , p_value = stats.chisquare(observed,expected)

print(f'chi2_stats = {chi2_stats} \np_value = {p_value} \n{50*"="}')

alpha = 0.05

if p_value < alpha:
  print('Reject The Null Hypothesis : Customers are not equally prefer
all colors.')
else:
  print('Fail to reject the Null Hypothesis: Customers equally prefer
all colors')
```

4 Lottery Number Frequency

A lottery company expects each number (1–5) to appear equally.

Observed counts suggest a deviation. Is the draw fair?

```
observed = np.array([220, 180, 210, 190, 200])
expected = np.array([200, 200, 200, 200, 200])

# H₀: The numbers occur with equal probability.
# H₁: Some numbers are favored.

chi2_stats , p_value = stats.chisquare(observed,expected)

print(f'chi2_stats = {chi2_stats} \np_value = {p_value} \n{50*"="}')
```

```
alpha = 0.05

if p_value < alpha:
  print('Reject The Null Hypothesis : The numbers are not occur with
equal probability.')
else:
  print('Fail to reject the Null Hypothesis: The numbers occur with
equal probability')
```

5 Online Shopping Preferences

A retailer predicts 30% customers prefer electronics, 25% prefer clothing, 20% prefer books, and 25% prefer groceries. Is this correct?

```
observed = np.array([300, 250, 200, 250])
expected = np.array([1000 * 0.30, 1000 * 0.25, 1000 * 0.20, 1000 *
0.25])


# H₀: Customer preferences match the given percentages.
# H₁: The preferences differ.

chi2_stats , p_value = stats.chisquare(observed,expected)

print(f'chi2_stats = {chi2_stats} \np_value = {p_value} \n{50*"="}')

alpha = 0.05

if p_value < alpha:
  print('Reject The Null Hypothesis : Customer The preferences
differ.')
else:
  print('Fail to reject the Null Hypothesis: Customer preferences
match the given percentages')
```

6 Coin Toss Fairness

A coin is flipped 1000 times, and we get 520 heads and 480 tails. Is the coin fair?

```
observed = np.array([520, 480])
expected = np.array([500, 500])

# H₀: The coin is fair (50% heads, 50% tails).
# H₁: The coin is biased.

chi2_stats , p_value = stats.chisquare(observed,expected)

print(f'chi2_stats = {chi2_stats} \np_value = {p_value} \n{50*"="}')
```

```
alpha = 0.05

if p_value < alpha:
  print('Reject The Null Hypothesis : The coin is biased ')
else:
  print('Fail to reject the Null Hypothesis: The coin is fair (50%
heads, 50% tails)')
```

7 Dice Fairness

A six-sided die is rolled 600 times. Expected frequency is 100 per face, but the observed data differs. Is the die fair?

```
observed = np.array([90, 110, 95, 105, 100, 100])
expected = np.array([100, 100, 100, 100, 100, 100])

# H₀: The die is fair (each side appears equally)
# H₁: The die is biased.

chi2_stats , p_value = stats.chisquare(observed,expected)

print(f'chi2_stats = {chi2_stats} \np_value = {p_value} \n{50*"="}')

alpha = 0.05

if p_value < alpha:
  print('Reject The Null Hypothesis :The die is not fair')
else:
  print('Fail to reject the Null Hypothesis: The die is fair')
```

8 Customer Purchase Preferences

A supermarket expects 30% of purchases in dairy, 20% in bakery, 25% in vegetables, and 25% in snacks. Do actual sales match?

```
observed = np.array([300, 200, 250, 250])
expected = np.array([1000 * 0.30, 1000 * 0.20, 1000 * 0.25, 1000 *
0.25])


# H₀: The purchase distribution follows expectations
# H₁: The purchase distribution differs.

chi2_stats , p_value = stats.chisquare(observed,expected)

print(f'chi2_stats = {chi2_stats} \np_value = {p_value} \n{50*"="}')

alpha = 0.05
```

```
if p_value < alpha:
  print('Reject The Null Hypothesis : The purchase distribution do not
follows expectations')
else:
  print('Fail to reject the Null Hypothesis: The purchase distribution
follows expectations')
```

9 Student Grades Distribution

A teacher believes that A (20%), B (30%), C (30%), and D (20%) should be distributed among students. Observed data suggests otherwise. Are grades assigned fairly?

```
observed = np.array([40, 60, 70, 30])
expected = np.array([200 * 0.20, 200 * 0.30, 200 * 0.30, 200 * 0.20])


# H₀: Grade distribution follows expectations
# H₁: Grade distribution differs

chi2_stats , p_value = stats.chisquare(observed,expected)

print(f'chi2_stats = {chi2_stats} \np_value = {p_value} \n{50*"="}')

alpha = 0.05


if p_value < alpha:
  print('Reject The Null Hypothesis : Grade distribution do not
follows expectations')
else:
  print('Fail to reject the Null Hypothesis: Grade distribution
follows expectations')
```

Traffic Light Timings

A city claims that traffic lights turn red (40%), yellow (10%), and green (50%) in a controlled way. Observed data suggests a deviation. Is the traffic control system fair?

```
observed = np.array([400, 150, 450])
expected = np.array([1000 * 0.40, 1000 * 0.10, 1000 * 0.50])


# H₀: The traffic light timings match expectations.
# H₁: The timings differ

chi2_stats , p_value = stats.chisquare(observed,expected)

print(f'chi2_stats = {chi2_stats} \np_value = {p_value} \n{50*"="}')
```

```
alpha = 0.05


if p_value < alpha:
  print('Reject The Null Hypothesis : The traffic light timings do not
match expectations')
else:
  print('Fail to reject the Null Hypothesis: The traffic light timings
match expectations')
```

11 Social Media Usage

A company claims that 40% use Facebook, 30% use Instagram, and 30% use Twitter. A survey gives different results. Is the claim correct?

```
observed = np.array([420, 280, 300])
expected = np.array([1000 * 0.40, 1000 * 0.30, 1000 * 0.30])


# H₀: The usage follows the claimed distribution.
# H₁: The usage differs


chi2_stats , p_value = stats.chisquare(observed,expected)

print(f'chi2_stats = {chi2_stats} \np_value = {p_value} \n{50*"="}')

alpha = 0.05


if p_value < alpha:
  print('Reject The Null Hypothesis : The usage do not follows the
claimed distribution')
else:
  print('Fail to reject the Null Hypothesis: The usage follows the
claimed distribution')
```

12 Website Traffic Sources

A website predicts its traffic comes 50% from search engines, 30% from social media, and 20% from direct visits. Actual data suggests otherwise. Is the prediction accurate?

```
observed = np.array([500, 250, 250])
expected = np.array([1000 * 0.50, 1000 * 0.30, 1000 * 0.20])


# H₀: The traffic sources follow predictions.
# H₁: The traffic sources are different.
```

```
chi2_stats , p_value = stats.chisquare(observed,expected)

print(f'chi2_stats = {chi2_stats} \np_value = {p_value} \n{50*"="}')

alpha = 0.05


if p_value < alpha:
  print('Reject The Null Hypothesis : The traffic sources do not
follow predictions')
else:
  print('Fail to reject the Null Hypothesis: The traffic sources
follow predictions')
```

13 Sports Preference Survey

A survey asked people to choose Cricket, Football, or Basketball. The expected preference is 40%, 35%, and 25%. Is this correct?

```
observed = np.array([400, 370, 230])
expected = np.array([1000 * 0.40, 1000 * 0.35, 1000 * 0.25])


# H₀: The distribution follows expectations.
# H₁: The preference distribution is different.

chi2_stats , p_value = stats.chisquare(observed,expected)

print(f'chi2_stats = {chi2_stats} \np_value = {p_value} \n{50*"="}')

alpha = 0.05


if p_value < alpha:
  print('Reject The Null Hypothesis :  The distribution do not follows
expectations')
else:
  print('Fail to reject the Null Hypothesis:  The distribution follows
expectations')
```

21 Does Gender Affect Shopping Preference?

A survey records the shopping preferences (Online/In-store) of men and women. Are they independent?

```
import pandas as pd
from scipy.stats import chi2_contingency
```

```
data = np.array([[300, 200], [250, 250]])
df = pd.DataFrame(data, columns=["Online", "In-store"], index=["Male",
"Female"])


# H₀: Gender and shopping preference are independent.
# H₁: Gender and shopping preference are related.

chi2_stats, p_value, dof,expected = chi2_contingency(df)

print(f"Chi-Square: {chi2_stats:.2f},\np-value: {p_value:.4f}\
n{80*'='}")

alpha = 0.05


if p_value < alpha:
  print('Reject The Null Hypothesis : Gender and shopping preference
are not independent')
else:
  print('Fail to reject the Null Hypothesis:  Gender and shopping
preference are independent')
```

22 Does Smoking Affect Lung Disease?

A study checks if smoking is related to lung disease.

```
data = np.array([[200, 150], [180, 470]])
df = pd.DataFrame(data, columns=["Lung Disease", "No Disease"],
index=["Smoker", "Non-Smoker"])


# H₀: Smoking and lung disease are independent.
# H₁: Smoking is related to lung disease.

chi2_stats, p_value, dof,expected = chi2_contingency(df)

print(f"Chi-Square: {chi2_stats:.2f},\np-value: {p_value:.4f}\
n{80*'='}")

alpha = 0.05


if p_value < alpha:
  print('Reject The Null Hypothesis : Smoking and lung disease are not
independent')
else:
  print('Fail to reject the Null Hypothesis: Smoking and lung disease
are independent')
```

23 Does Education Level Affect Voting Preference?

A survey checks if education level influences voting choice.

```
data = np.array([[250, 300], [200, 250], [150, 100]])
df = pd.DataFrame(data, columns=["Candidate A", "Candidate B"],
index=["High School", "College", "Postgraduate"])


# H₀: Education level and voting preference are independent.
# H₁: Education level affects voting preference.

chi2_stats, p_value, dof,expected = chi2_contingency(df)

print(f"Chi-Square: {chi2_stats:.2f},\np-value: {p_value:.4f}\
n{80*'='}")

alpha = 0.05


if p_value < alpha:
  print('Reject The Null Hypothesis : Education level and voting
preference are not independent')
else:
  print('Fail to reject the Null Hypothesis: Education level and
voting preference are independent')
```

24 Does Diet Affect Heart Disease?

A study examines whether a vegetarian or non-vegetarian diet is linked to heart disease.

```
data = np.array([[180, 220], [250, 350]])
df = pd.DataFrame(data, columns=["Heart Disease", "No Heart Disease"],
index=["Vegetarian", "Non-Vegetarian"])


# H₀: Diet and heart disease are independent.
# H₁: Diet and heart disease are related.

chi2_stats, p_value, dof,expected = chi2_contingency(df)

print(f"Chi-Square: {chi2_stats:.2f},\np-value: {p_value:.4f}\
n{80*'='}")

alpha = 0.05


if p_value < alpha:
  print('Reject The Null Hypothesis : Education level and voting
```

```
preference are not independent')
else:
  print('Fail to reject the Null Hypothesis: Education level and
voting preference are independent')
```

25 Does Work Experience Affect Promotion?

A company wants to check if promotion depends on years of experience.

```
data = np.array([[40, 60], [70, 30], [90, 10]])
df = pd.DataFrame(data, columns=["Promoted", "Not Promoted"],
index=["<5 years", "5-10 years", ">10 years"])


# H₀: Work experience and promotion are independent.
# H₁: Work experience affects promotion.

chi2_stats, p_value, dof,expected = chi2_contingency(df)

print(f"Chi-Square: {chi2_stats:.2f},\np-value: {p_value:.4f}\
n{80*'='}")

alpha = 0.05


if p_value < alpha:
  print('Reject The Null Hypothesis : Work experience and promotion
are not independent')
else:
  print('Fail to reject the Null Hypothesis : Work experience and
promotion are independent')
```

26 Does Smartphone Brand Preference Vary by Age Group?

A survey asks different age groups about their preferred smartphone brand.

```
data = np.array([[100, 150, 200], [130, 170, 100], [90, 120, 140]])
df = pd.DataFrame(data, columns=["Apple", "Samsung", "OnePlus"],
index=["18-25", "26-40", "41+"])


# H₀: Age group and brand preference are independent.
# H₁: Age group affects brand preference.

chi2_stats, p_value, dof,expected = chi2_contingency(df)
```

```
print(f"Chi-Square: {chi2_stats:.2f},\np-value: {p_value:.4f}\
n{80*'='}")

alpha = 0.05


if p_value < alpha:
  print('Reject The Null Hypothesis : Work experience and promotion
are not independent')
else:
  print('Fail to reject the Null Hypothesis : Work experience and
promotion are independent')
```

27 Does City Influence Internet Usage?

A survey checks if people from different cities use the internet differently.

```
data = np.array([[300, 200], [250, 250], [200, 300]])
df = pd.DataFrame(data, columns=["Heavy User", "Light User"],
index=["City A", "City B", "City C"])


# H₀: City and internet usage are independent.
# H₁: City affects internet usage.

chi2_stats, p_value, dof,expected = chi2_contingency(df)

print(f"Chi-Square: {chi2_stats:.2f},\np-value: {p_value:.4f}\
n{80*'='}")

alpha = 0.05


if p_value < alpha:
  print('Reject The Null Hypothesis : City and internet usage are not
independent')
else:
  print('Fail to reject the Null Hypothesis : City and internet usage
are independent')
```

28 Does Gender Influence Car Type Preference?

A survey records car type preferences among men and women.

```
data = np.array([[150, 200, 250], [180, 220, 200]])
df = pd.DataFrame(data, columns=["Sedan", "SUV", "Truck"],
index=["Male", "Female"])
```

```
# H₀: Gender and car type preference are independent.
# H₁: Gender and car type preference are related.

chi2_stats, p_value, dof,expected = chi2_contingency(df)

print(f"Chi-Square: {chi2_stats:.2f},\np-value: {p_value:.4f}\
n{80*'='}")

alpha = 0.05


if p_value < alpha:
  print('Reject The Null Hypothesis : Gender and car type preference
are not independent')
else:
  print('Fail to reject the Null Hypothesis : Gender and car type
preference are independent')
```

29 Does Education Level Affect Job Sector Choice?

A study explores if people with different education levels prefer government or private jobs.

```
data = np.array([[100, 200], [150, 250], [200, 150]])
df = pd.DataFrame(data, columns=["Government", "Private"],
index=["High School", "Bachelor's", "Master's"])


# H₀: Education level and job sector choice are independent.
# H₁: Education level influences job sector choice.

chi2_stats, p_value, dof,expected = chi2_contingency(df)

print(f"Chi-Square: {chi2_stats:.2f},\np-value: {p_value:.4f}\
n{80*'='}")

alpha = 0.05


if p_value < alpha:
  print('Reject The Null Hypothesis : Education level and job sector
choice are not independent')
else:
  print('Fail to reject the Null Hypothesis : Education level and job
sector choice are independent')
```

30 Is Political Affiliation Linked to News Source?

A survey asks participants about their political party and preferred news source.

```
data = np.array([[130, 170, 200], [150, 180, 170]])
df = pd.DataFrame(data, columns=["TV", "Online", "Newspaper"],
index=["Party A", "Party B"])


# H₀: Political affiliation and news source are independent.
# H₁: Political affiliation influences news source choice.

chi2_stats, p_value, dof,expected = chi2_contingency(df)

print(f"Chi-Square: {chi2_stats:.2f},\np-value: {p_value:.4f}\
n{80*'='}")

alpha = 0.05


if p_value < alpha:
  print('Reject The Null Hypothesis : Political affiliation and news
source are not independent')
else:
  print('Fail to reject the Null Hypothesis : Political affiliation
and news source are independent ')
```

31 Does Exercise Frequency Affect Obesity?

Researchers analyze whether exercise frequency influences obesity rates.

```
data = np.array([[200, 300], [150, 250], [100, 400]])
df = pd.DataFrame(data, columns=["Obese", "Not Obese"], index=["No
Exercise", "Occasional", "Regular"])


# H₀: Exercise frequency and obesity are independent.
# H₁: Exercise frequency influences obesity.

chi2_stats, p_value, dof,expected = chi2_contingency(df)

print(f"Chi-Square: {chi2_stats:.2f},\np-value: {p_value:.4f}\
n{80*'='}")

alpha = 0.05


if p_value < alpha:
  print('Reject The Null Hypothesis : Exercise frequency and obesity
are not independent')
```

```
else:
  print('Fail to reject the Null Hypothesis : Exercise frequency and
obesity are independent.')
```

32 Does Smoking Habit Depend on Age Group?

A survey studies the relationship between smoking habits and age groups.

```
data = np.array([[120, 180], [150, 250], [200, 300]])
df = pd.DataFrame(data, columns=["Smoker", "Non-Smoker"], index=["18-
30", "31-50", "51+"])


# H₀: Smoking habit and age group are independent.
# H₁: Smoking habit is related to age group.

chi2_stats, p_value, dof,expected = chi2_contingency(df)

print(f"Chi-Square: {chi2_stats:.2f},\np-value: {p_value:.4f}\
n{80*'='}")

alpha = 0.05


if p_value < alpha:
  print('Reject The Null Hypothesis : Smoking habit and age group are
not independent.')
else:
  print('Fail to reject the Null Hypothesis : Smoking habit and age
group are independent.')
```

33 Does Social Media Usage Depend on Profession?

Researchers investigate if social media usage is linked to different professions.

```
data = np.array([[300, 200], [250, 250], [200, 300]])
df = pd.DataFrame(data, columns=["Heavy User", "Light User"],
index=["IT", "Education", "Healthcare"])


# H₀: Profession and social media usage are independent.
# H₁: Profession influences social media usage.

chi2_stats, p_value, dof,expected = chi2_contingency(df)

print(f"Chi-Square: {chi2_stats:.2f},\np-value: {p_value:.4f}\
n{80*'='}")
```

```
alpha = 0.05


if p_value < alpha:
  print('Reject The Null Hypothesis : Profession and social media
usage are not independent.')
else:
  print('Fail to reject the Null Hypothesis : Profession and social
media usage are independent.')
```

34 Does Internet Usage Depend on Gender?

A study examines whether men and women use the internet differently.

```
data = np.array([[250, 300], [200, 250]])
df = pd.DataFrame(data, columns=["Frequent User", "Occasional User"],
index=["Male", "Female"])


# H₀: Gender and internet usage are independent.
# H₁: Gender affects internet usage.

chi2_stats, p_value, dof,expected = chi2_contingency(df)

print(f"Chi-Square: {chi2_stats:.2f},\np-value: {p_value:.4f}\
n{80*'='}")

alpha = 0.05


if p_value < alpha:
  print('Reject The Null Hypothesis : Gender and internet usage are
not independent.')
else:
  print('Fail to reject the Null Hypothesis : Gender and internet
usage are independent.')
```

35 Does Marital Status Affect Financial Stability?

A financial study checks if being married impacts financial stability.

```
data = np.array([[180, 220], [240, 160]])
df = pd.DataFrame(data, columns=["Stable", "Unstable"],
index=["Married", "Single"])
```

```
# H₀: Marital status and financial stability are independent.
# H₁: Marital status affects financial stability.

chi2_stats, p_value, dof,expected = chi2_contingency(df)

print(f"Chi-Square: {chi2_stats:.2f},\np-value: {p_value:.4f}\
n{80*'='}")

alpha = 0.05


if p_value < alpha:
  print('Reject The Null Hypothesis :Marital status and financial
stability are not independent.')
else:
  print('Fail to reject the Null Hypothesis : Marital status and
financial stability are independent')
```

36 Does Region Affect Voting Preference?

A political survey examines if voting patterns differ by region.

```
data = np.array([[100, 150, 250], [200, 180, 120], [170, 140, 190]])
df = pd.DataFrame(data, columns=["Party A", "Party B", "Party C"],
index=["North", "South", "West"])


# H₀: Region and voting preference are independent.
# H₁: Region influences voting preference.

chi2_stats, p_value, dof,expected = chi2_contingency(df)

print(f"Chi-Square: {chi2_stats:.2f},\np-value: {p_value:.4f}\
n{80*'='}")

alpha = 0.05


if p_value < alpha:
  print('Reject The Null Hypothesis :  Region and voting preference
are not independent.')
else:
  print('Fail to reject the Null Hypothesis :  Region and voting
preference are independent')
```

37 Does Food Preference Depend on Ethnicity?

A study looks at food preferences among different ethnic groups.

```
data = np.array([[130, 170, 200], [150, 180, 170], [140, 160, 180]])
df = pd.DataFrame(data, columns=["Vegetarian", "Non-Vegetarian",
"Vegan"], index=["Group A", "Group B", "Group C"])


# H₀: Ethnicity and food preference are independent.
# H₁: Ethnicity affects food preference.

chi2_stats, p_value, dof,expected = chi2_contingency(df)

print(f"Chi-Square: {chi2_stats:.2f},\np-value: {p_value:.4f}\
n{80*'='}")

alpha = 0.05


if p_value < alpha:
  print('Reject The Null Hypothesis : Ethnicity and food preference
are not independent')
else:
  print('Fail to reject the Null Hypothesis : Ethnicity and food
preference are independent')
```

38 Does Location Affect Transportation Choice?

A survey examines if people from urban, suburban, and rural areas prefer different transportation modes.

```
data = np.array([[200, 250, 150], [180, 220, 200], [140, 160, 250]])
df = pd.DataFrame(data, columns=["Car", "Public Transport", "Bike"],
index=["Urban", "Suburban", "Rural"])


# H₀: Location and transportation choice are independent.
# H₁: Location affects transportation preference.

chi2_stats, p_value, dof,expected = chi2_contingency(df)

print(f"Chi-Square: {chi2_stats:.2f},\np-value: {p_value:.4f}\
n{80*'='}")

alpha = 0.05


if p_value < alpha:
  print('Reject The Null Hypothesis : Location and transportation
choice are not independent.')
else:
  print('Fail to reject the Null Hypothesis :Location and
```

```
transportation choice are independent.')
```

39 Does Age Affect Online Shopping Frequency?

A study examines if age groups influence online shopping habits.

```
data = np.array([[250, 150, 100], [200, 180, 150], [140, 200, 160]])
df = pd.DataFrame(data, columns=["Frequent", "Occasional", "Rare"],
index=["18-30", "31-50", "51+"])


# H₀: Age group and online shopping frequency are independent.
# H₁: Age group affects online shopping habits.


chi2_stats, p_value, dof,expected = chi2_contingency(df)

print(f"Chi-Square: {chi2_stats:.2f},\np-value: {p_value:.4f}\
n{80*'='}")

alpha = 0.05


if p_value < alpha:
  print('Reject The Null Hypothesis : Age group and online shopping
frequency are not independent.')
else:
  print('Fail to reject the Null Hypothesis : Age group and online
shopping frequency are independent.')
```

40 Does Region Affect Internet Speed Satisfaction?

A telecom survey investigates if internet speed satisfaction varies by region.

```
data = np.array([[100, 200, 250], [150, 220, 180], [200, 170, 160]])
df = pd.DataFrame(data, columns=["Satisfied", "Neutral",
"Dissatisfied"], index=["City A", "City B", "City C"])


# H₀: Region and internet speed satisfaction are independent.
# H₁: Region influences internet speed satisfaction.

chi2_stats, p_value, dof,expected = chi2_contingency(df)

print(f"Chi-Square: {chi2_stats:.2f},\np-value: {p_value:.4f}\
n{80*'='}")
```

```
alpha = 0.05


if p_value < alpha:
  print('Reject The Null Hypothesis : Region and internet speed
satisfaction are not independent.')
else:
  print('Fail to reject the Null Hypothesis : Region and internet
speed satisfaction are independent.')
```

Is a six-sided die fair?

A die was rolled 60 times, and the results were:

1 → 8 times

2 → 10 times

3 → 12 times

4 → 14 times

5 → 9 times

6 → 7 times

```
observed_roll = [8,10,12,14,9,7]
expected_rolls = [10, 10, 10, 10, 10, 10]

import scipy.stats  as stats

chi2_stat, p_value = stats.chisquare(observed_roll,expected_rolls)
print(f"Chi-Square Statistic: {chi2_stat}")
print(f"P-value: {p_value}\n")

if p_value < 0.05:
  print('Reject the null hypothesis: The die is not fair')
else:
  print('Fail to reject the null hypothesis The die appears fair')
```

Are customer visits to a store evenly distributed across the week?

Observed visits:

Monday → 120

Tuesday → 150

Wednesday → 180

Thursday → 160

Friday → 140

Saturday → 130

Sunday → 110

```
# Observed frequencies
observed_visits = [120, 150, 180, 160, 140, 130, 110]
expected_visits = [sum(observed_visits) / 7] *7

chi_stats, p_value = stats.chisquare(observed_visits,expected_visits)

if p_value < 0.05:
    print("Reject Null Hypothesis: Customer visits are NOT evenly
distributed.")
else:
    print("Fail to Reject Null Hypothesis: Customer visits appear
evenly distributed.")
```