1 Contingency Tables

1.1

Row Marginal Totals (Total for each group):

- Without Alcohol: $120 + 60 + 20 = 200$
- With Alcohol: $60 + 100 + 40 = 200$

Interpreting the row marginal totals: These represent the total number of subjects in each group (without alcohol and with alcohol).

Column Marginal Totals (Total for each reaction time category):

- Good: $120 + 60 = 180$
- Medium: $60 + 100 = 160$
- Strong: $20 + 40 = 60$
- Delay: $20 + 40 = 60$

These represent the total number of subjects falling into each reaction time category, regardless of whether they had alcohol or not.

Overall Marginal Total (Total for the entire table):

- $180 + 160 + 60 + 60 = 460$

This represents the total number of subjects in the entire experiment, combining both groups and all reaction time categories.

1.2

The conditional relative frequency is calculated as:

$$f(\text{A — B}) = \frac{\text{Frequency of A and B}}{\text{Total frequency of B}}$$

1. f(without alcohol — good)

$$f(\text{without alcohol — good}) = \frac{120}{120 + 60} = \frac{2}{3}$$

Among subjects with a "good" reaction time, 2/3 of them did not have alcohol.

2. f(with alcohol — medium)

$$f(\text{with alcohol — medium}) = \frac{100}{60 + 100} = \frac{5}{6}$$

Among subjects with a "medium" reaction time, 5/6 of them had alcohol.

3. f(without alcohol — delay)

$$f(\text{without alcohol — delay}) = \frac{20}{20 + 40} = \frac{1}{3}$$

Among subjects with a "delay" in reaction time, 1/3 of them did not have alcohol.

1.3

Null Hypothesis: H_0 - No association between alcohol consumption and reaction time

Alternative Hypothesis: H_1 - There is an association between alcohol consumption and reaction time

Delay	Good	Medium	Strong
Without Alcohol	120	60	20
With Alcohol	60	100	40

$$\text{Chi-squared statistic: } \chi^2 = 36.67$$

$$\text{P-value: } p = 1.09 \times 10^{-8}$$

$$\text{Degrees of freedom: } df = 2$$

Expected frequencies:

$$\begin{bmatrix} 90 & 80 & 30 \\ 90 & 80 & 30 \end{bmatrix}$$

Since $\chi^2 = 36.67$ is greater than the critical value of 10.596 (for $df = 2$ and $\alpha = 0.005$), we reject the null hypothesis.

2 Feature Types

1. Determine the feature type:

Temperature (Fahrenheit):	Interval
Socio-Economic Status:	Ordinal
Temperature (Kelvin):	Ratio
Distance (cm):	Ratio
Grades in School:	Ordinal
Description:	Nominal
Date:	Ordinal

2. **Two more examples for each interval and ratio scaled:**

Interval:	Ratio:
1.IQ scores	1.Height (in centimeters)
2.pH level	2.Income (in dollars)

3 Time Series

3.1

1. L1 distance (Manhattan distance):

$$L1(A,B) = \sum_{i=1}^{n} |A_i - B_i|$$

2. L∞ distance (Chebyshev distance):

$$L\infty(A,B) = \max_{i=1}^{n} |A_i - B_i|$$

where n is the length of the time series.
Calculated Distances:

$$A = [-1.66, 0.30, -0.08, 0.10, -1.17, -0.05, 0.84, -0.66, 0.42, -0.99]$$

$$B = [0.29, 0.89, 0.82, 0.97, 0.53, 0.83, 1.06, 0.67, 0.86, 0.51]$$

1. L1 distance:

$$\begin{aligned} L1(A,B) = &|(-1.66 - 0.29)| + \\ &|(0.30 - 0.89)| + \\ &|(-0.08 - 0.82)| + \\ &|(0.10 - 0.97)| + \\ &|(-1.17 - 0.53)| + \\ &|(-0.05 - 0.83)| + \\ &|(0.84 - 1.06)| + \\ &|(-0.66 - 0.67)| + \\ &|(0.42 - 0.86)| + \\ &|(-0.99 - 0.51)|. \end{aligned}$$

$$L1(A,B) = 5.73$$

2. L∞ distance:

$$L_\infty(A,B) = \max\{\begin{matrix} |-1.66 - 0.29|, \\ |0.30 - 0.89|, \\ |-0.08 - 0.82|, \\ |0.10 - 0.97|, \\ |-1.17 - 0.53|, \\ |-0.05 - 0.83|, \\ |0.84 - 1.06|, \\ |-0.66 - 0.67|, \\ |0.42 - 0.86|, \\ |-0.99 - 0.51| \end{matrix}\}$$

$$L\infty(A,B) = 2.66$$

The L1 distance between A and B is 5.73, and the L∞ distance is 2.66.

3.2

Perform an offset translation by a constant c. The new series A' and B' would be:

$$A'_i = A_i + c$$
$$B'_i = B_i + c$$

Let $c = 2$.

$$A' = [-1.66+2, 0.30+2, -0.08+2, 0.10+2, -1.17+2, -0.05+2, 0.84+2, -0.66+2, 0.42+2, -0.99+2]$$

$$B' = [0.29+2, 0.89+2, 0.82+2, 0.97+2, 0.53+2, 0.83+2, 1.06+2, 0.67+2, 0.86+2, 0.51+2]$$

$$A' = [0.34, 2.30, 1.92, 2.10, 0.83, 1.95, 2.84, 1.34, 2.42, 1.01]$$
$$B' = [2.29, 2.89, 2.82, 2.97, 2.53, 2.83, 3.06, 2.67, 2.86, 2.51]$$

Recalculate Distances;

1. L1 distance:

$$L1(A', B') = \begin{aligned} &|(0.34 - 2.29)|+ \\ &|(2.30 - 2.89)|+ \\ &|(1.92 - 2.82)|+ \\ &|(2.10 - 2.97)|+ \\ &|(0.83 - 2.53)|+ \\ &|(1.95 - 2.83)|+ \\ &|(2.84 - 3.06)|+ \\ &|(1.34 - 2.67)|+ \\ &|(2.42 - 2.86)|+ \\ &|(1.01 - 2.51)| \end{aligned}$$

$$L1(A', B') = 20.59$$

2. L∞ distance:

$$L_\infty(A', B') = \max\{\begin{aligned} &|0.34 - 2.29|, \\ &|2.30 - 2.89|, \\ &|1.92 - 2.82|, \\ &|2.10 - 2.97|, \\ &|0.83 - 2.53|, \\ &|1.95 - 2.83|, \\ &|2.84 - 3.06|, \\ &|1.34 - 2.67|, \\ &|2.42 - 2.86|, \\ &|1.01 - 2.51| \end{aligned}\}$$

$$L\infty(A', B') = 1.83$$

After the offset translation, the L1 distance between A' and B' is 20.59, and the L∞ distance is 1.83.

1.3

Multiply each element in the series by a constant factor s. New series A'' and B'' would be:

$$A''_i = s \cdot A'_i$$
$$B''_i = s \cdot B'_i$$

Let $s = 0.5$.

$$A'' = 0.5 \times [0.34, 2.30, 1.92, 2.10, 0.83, 1.95, 2.84, 1.34, 2.42, 1.01]$$
$$B'' = 0.5 \times [2.29, 2.89, 2.82, 2.97, 2.53, 2.83, 3.06, 2.67, 2.86, 2.51]$$

$$A'' = [0.17, 1.15, 0.96, 1.05, 0.42, 0.97, 1.42, 0.67, 1.21, 0.51]$$

$$B'' = [1.15, 1.44, 1.41, 1.49, 1.27, 1.42, 1.53, 1.34, 1.43, 1.26]$$

Recalculate the distances:
1. L1 distance:

$$L1(A'', B'') = \begin{array}{l} |(0.17 - 1.15)|+ \\ |(1.15 - 1.44)|+ \\ |(0.96 - 1.41)|+ \\ |(1.05 - 1.49)|+ \\ |(0.42 - 1.27)|+ \\ |(0.97 - 1.42)|+ \\ |(1.42 - 1.53)|+ \\ |(0.67 - 1.34)|+ \\ |(1.21 - 1.43)|+ \\ |(0.51 - 1.26)|. \end{array}$$

$$L1(A'', B'') = 4.12$$

2. L∞ distance:

$$L_\infty(A'', B'') = \max\{\begin{array}{l} |0.17 - 1.15|, \\ |1.15 - 1.44|, \\ |0.96 - 1.41|, \\ |1.05 - 1.49|, \\ |0.42 - 1.27|, \\ |0.97 - 1.42|, \\ |1.42 - 1.53|, \\ |0.67 - 1.34|, \\ |1.21 - 1.43|, \\ |0.51 - 1.26| \end{array}\}$$

$$L\infty(A'', B'') = 0.92$$

After the amplitude scaling, the L1 distance between A" and B" is 4.12, and the L∞ distance is 0.92.