Contigency

Marc Los Huertos

2025-10-04

Introduction

Freshwater contamination is a major public health concern, particularly in regions where people rely on untreated water sources. Previous studies have shown that exposure to contaminated water may increase the risk of gastrointestinal disease.

This analysis examines whether there is a **statistically significant association** between freshwater contamination and disease occurrence using hypothetical survey data.

Data

The following table summarizes household survey results:

Water Contamination	Disease Reported	No Disease Reported	Total
Contaminated	35	15	50
Clean	10	40	50
Total	**45**	**55**	**100**

We will use a **Chi-square test of independence** to determine whether there is a significant association between water contamination and disease occurrence.

Create contingency table

```
data <- matrix(c(35, 15, 10, 40),
               nrow = 2,
               byrow = TRUE)

colnames(data) <- c("Disease", "No_Disease")
rownames(data) <- c("Contaminated", "Clean")

# Display the table
data
```

```
##              Disease No_Disease
## Contaminated      35         15
## Clean             10         40
```

```
# Perform chi-square test
chisq_result <- chisq.test(data)
```

```r
# Display results
chisq_result
```

```
##
##  Pearson's Chi-squared test with Yates' continuity correction
##
## data:  data
## X-squared = 23.273, df = 1, p-value = 1.406e-06
```