Measure\_of\_associations\_HW

Elena Marochkina

2024-10-04

Table of Contents

# Task 1

Doctors decided to study how body mass index (BMI) is associated with the risk of developing type 2 diabetes. The file diabetes.csv contains data on a random sample of 200 residents of settlement N. For each respondent, the BMI (high or normal) and diabetes status (presence/absence of type 2 diabetes) are known. Determine how high BMI is associated with the development of type 2 diabetes, indicate the relative risk and the absolute risk difference. How do you interpret the results?

# Load the diabetes data  
data\_diabetes <- read.csv("data/raw/diabetes.csv")  
  
# Create a contingency table for BMI and Diabetes status  
contingency\_table <- table(data\_diabetes$ИМТ, data\_diabetes$Диабет)  
  
# The structure of the table is:  
# Диабет (Есть) Диабет (Нет)  
# ИМТ Высокий a b  
# ИМТ Нормальный c d  
  
# Extract values from the contingency table  
High\_BMI\_Diabetes <- contingency\_table["Высокий", "Есть"]  
High\_BMI\_Healthy <- contingency\_table["Высокий", "Нет"]  
Normal\_BMI\_Diabetes <- contingency\_table["Нормальный", "Есть"]  
Normal\_BMI\_Healthy <- contingency\_table["Нормальный", "Нет"]  
  
# Calculate Incidence Rates (IR)  
IR\_High\_BMI\_Diabetes <- High\_BMI\_Diabetes / (High\_BMI\_Diabetes + High\_BMI\_Healthy)  
IR\_Normal\_BMI\_Diabetes <- Normal\_BMI\_Diabetes / (Normal\_BMI\_Diabetes + Normal\_BMI\_Healthy)  
  
# Calculate Relative Risk (RR)  
RR <- IR\_High\_BMI\_Diabetes / IR\_Normal\_BMI\_Diabetes  
  
# Calculate Risk Difference (RD)  
RD <- IR\_High\_BMI\_Diabetes - IR\_Normal\_BMI\_Diabetes

**Incidence Rate (IR) for High BMI = 0.59** : This means that 59% of individuals with a high BMI developed type 2 diabetes over the study period.

**Incidence Rate (IR) for Normal BMI = 0.26** : This indicates that 26% of individuals with a normal BMI developed type 2 diabetes during the same period.

**Relative Risk (RR) = 2.24**: Individuals with a high BMI are 2.24 times more likely to develop type 2 diabetes compared to individuals with a normal BMI. This shows a strong association between high BMI and the likelihood of developing diabetes.

**Risk Difference (RD) = 0.33**: The absolute risk of developing type 2 diabetes is 33% higher in individuals with a high BMI compared to those with a normal BMI. This highlights a substantial increase in risk associated with high BMI.

The results demonstrate that high BMI significantly increases both the relative and absolute risk of developing type 2 diabetes. With a relative risk of 2.24 and a risk difference of 33%, these findings suggest that managing BMI is crucial for reducing the risk of type 2 diabetes in the population.

# Task 2

Task 2. An outbreak of pneumonia has been recorded in city N. 250 people living in different houses have been affected. All of them have visited different places over the past two weeks: shopping centers, restaurants, and social events. 750 people who did not get pneumonia were taken as a control. A survey was conducted about which places each person visited (pneumonia.csv). Using an appropriate measure of association, determine which place of visit is most likely associated with the occurrence of pneumonia.

data\_pneumonia <- read.csv("data/raw/pneumonia.csv")  
  
# Create contingency tables for each place  
restaurant\_table <- table(data\_pneumonia$`Ресторан`, data\_pneumonia$`Группа`)  
mall\_table <- table(data\_pneumonia$`Торговый.центр`, data\_pneumonia$`Группа`)  
soc\_events\_table <- table(data\_pneumonia$`Общественные.мероприятия`, data\_pneumonia$`Группа`)  
  
# Calculate odds ratio manually for each place  
calculate\_odds\_ratio <- function(contingency\_table) {  
 # Values in the table: (row, column)  
 # (1,1) -> exposed controls (visited place and did not get pneumonia)  
 # (1,2) -> exposed cases (visited place and got pneumonia)  
 # (2,1) -> non-exposed controls (did not visit place and did not get pneumonia)  
 # (2,2) -> non-exposed cases (did not visit place and got pneumonia)  
   
 exposed\_cases <- contingency\_table[1, 2]  
 exposed\_controls <- contingency\_table[1, 1]  
 non\_exposed\_cases <- contingency\_table[2, 2]  
 non\_exposed\_controls <- contingency\_table[2, 1]  
   
 # Calculate odds for exposed and non-exposed  
 odds\_exposed <- exposed\_cases / exposed\_controls  
 odds\_non\_exposed <- non\_exposed\_cases / non\_exposed\_controls  
   
 # Odds ratio  
 odds\_ratio <- odds\_exposed / odds\_non\_exposed  
 return(odds\_ratio)  
}  
  
# Calculate odds ratio for each place  
restaurant\_or <- calculate\_odds\_ratio(restaurant\_table)  
mall\_or <- calculate\_odds\_ratio(mall\_table)  
soc\_events\_or <- calculate\_odds\_ratio(soc\_events\_table)

**Restaurant OR = 1.11**

The odds of getting pneumonia for people who visited a restaurant are 11% higher than for those who did not visit a restaurant. Since the odds ratio is greater than 1 (1.11), it suggests a weak positive association between visiting a restaurant and getting pneumonia. However, this association is not strong, meaning that visiting a restaurant might slightly increase the risk of pneumonia, but the effect is minimal.

**Mall OR = 1.55**

The odds of getting pneumonia for people who visited a mall are 55% higher than for those who did not visit a mall. An OR of 1.55 suggests a moderate positive association between visiting a mall and getting pneumonia. This indicates that people who visited the mall were somewhat more likely to develop pneumonia compared to those who did not visit the mall.

**Social Events OR = 0.98**

The odds of getting pneumonia for people who attended social events are almost the same as for those who did not attend social events. An OR of 0.98 indicates virtually no association between attending social events and the risk of getting pneumonia. This means that attending social events neither increases nor decreases the likelihood of developing pneumonia.