Disease Metrics

1. Measure of disease occurrence

$$\text{Prevalence} = \frac{\text{Number of existing cases at a given time}}{\text{Total population at that time}} \times K$$

(where K is a constant, e.g., 1,000 or 100,000, for ease of interpretation)

Problem 1: Prevalence

Scenario: In a town of 20,000 people, a health survey conducted on January 1st, 2023, found that 400 people had diabetes. Over the course of the year, 100 new cases of diabetes were diagnosed, and 50 people with diabetes died.

Questions:

- Calculate the point prevalence of diabetes on January 1st, 2023.
- Calculate the period prevalence of diabetes for the year 2023.

Solution:

1. **Point Prevalence:**

$$\text{Point Prevalence} = \frac{\text{Number of existing cases}}{\text{Total population}} \times K$$

$$\text{Point Prevalence} = \frac{400}{20,000} \times 1000 = 20 \text{ cases per } 1,000 \text{ population}$$

2. **Period Prevalence:**
   - Total cases during the year = Existing cases + New cases = 400 + 100 = 500.

$$\text{Period Prevalence} = \frac{\text{Total cases during the period}}{\text{Total population}} \times K$$

$$\text{Period Prevalence} = \frac{500}{20,000} \times 1000 = 25 \text{ cases per } 1,000 \text{ population}$$

Ho to perform that in RStudio?

Step 1: Create input prompt to input values using **readline()** function to input Num_of_Exist_cases, New_cases, Total_pop, and K values.

Note: readline() has argument called "prompt =" used to enter text to inform user how to enter needed value:

```
readline(prompt = "Enter No. of cases: ")
```

But the problem with this function is that it just returns string value, so we need to convert the returned value to a suitable data type that we need:

```
as.numeric(readline(prompt = "Enter No. of cases: "))
```

Step 2: calculate the Total cases during the year (Total_pop), Point prevalence (Point_prev), and Period prevalence (Period_prev).

Step 3: Print the results.

Performing the example in RStudio:

We have to write the code in code editor window (upper-left window)

```
# create input prompt to input vales
Num_of_Exist_cases <- as.numeric(readline(prompt = "Enter No. of cases: "))
New_cases <- as.numeric(readline(prompt = "Enter No. of new cases: "))
Total_cases <- Num_of_Exist_cases + New_cases
Total_pop <- as.numeric(readline(prompt = "Enter total population: "))
K <- as.numeric(readline(prompt = "Enter the constant value: "))
# Calculate prevalence
Point_prev <- (Num_of_Exist_cases/Total_pop)* K
Period_prev <- (Total_cases/Total_pop)* K
# Printing results
print(paste("Point prevalence = ", Point_prev, "per",K ))
print("------------------------------------------------------")
print(paste("Period prevalence = ", Period_prev, "per",K))
```

For optimal performance, run the code using **Source** or **Ctrl+Shift+S** instead of **Run (Ctrl+Enter)** , which is typically used for executing single lines. This ensures proper execution as follows:

```
> source("E:/Epideiology/prevalence.R")
Enter No. of cases: 400
Enter No. of new cases: 100
Enter total population: 20000
Enter the constant value: 1000
[1] "Point prevalence =  20 per 1000"
[1] "------------------------------------------------------"
```

```
[1] "Period prevalence =  25 per 1000"
```

Now you have to use R to solve the following problems:

Problem 2: Incidence Risk (Cumulative Incidence)

Scenario: A study follows 1,000 healthy individuals for 5 years to determine the incidence of hypertension. By the end of the study, 150 individuals developed hypertension.

Question:

Calculate the incidence risk of hypertension over the 5-year period.

$$Incidence\ risk = \frac{Number\ of\ new\ cases}{Population\ at\ risk\ at\ the\ start} \times K = 150$$

Problem 3: Incidence Rate (Incidence Density)

Scenario: A study follows 500 individuals for 2 years to track the development of asthma. During the study, 20 new cases of asthma are diagnosed. The total person-time observed is 900 person-years.

Question:

Calculate the incidence rate of asthma.

$$Incidence\ rate = \frac{Number\ of\ new\ cases}{Total\ person - time\ at\ risk} \times K = 22.22$$

Problem 4: Mortality Rate

Scenario: In a population of 50,000 people, there were 200 deaths from all causes in a year. Of these, 50 deaths were due to cardiovascular disease (CVD).

Questions:

Calculate the crude mortality rate.

Calculate the cause-specific mortality rate for CVD.

$$Crude\ mortality\ rate = \frac{Total\ deaths}{Total\ population} \times K = 4$$

$$Cause - spcific\ mortality\ rate = \frac{Deathsdue\ to\ specific\ cause}{Total\ population} \times K = 1$$

Problem 5: Risk Ratio (Relative Risk, RR)

Scenario: A cohort study is conducted to investigate the association between smoking and the development of lung cancer. The study follows two groups of individuals over 10 years:

Group 1 (Smokers): 1,000 smokers.

Group 2 (Non-Smokers): 2,000 non-smokers.

By the end of the study, the following outcomes are observed:

Smokers: 100 cases of lung cancer.

Non-Smokers: 20 cases of lung cancer.

Questions:

- Calculate the risk of lung cancer in smokers and non-smokers.
- Calculate the Risk Ratio (Relative Risk, RR) for lung cancer among smokers compared to non-smokers.
- Interpret the Risk Ratio in terms of the association between smoking and lung cancer.

$$Risk\ in\ exposed\ group = \frac{Number\ of\ cases\ in\ exposed\ group}{Total\ number\ of\ exposed\ group} = 0.10$$

$$Risk\ in\ non\ exposed\ group = \frac{Number\ cases\ in\ non\ exposed\ group}{Total\ number\ of\ non\ exposed\ group} = 0.01$$

$$RR = \frac{Risk\ in\ exposed\ group}{Risk\ in\ non\ exposed\ group} = 10$$

Problem 6: Odds Ratio (OR)

Scenario: A case-control study is conducted to investigate the association between alcohol consumption and the risk of liver cirrhosis. The study includes:

Cases: 200 individuals diagnosed with liver cirrhosis.

Controls: 400 individuals without liver cirrhosis.

The exposure (alcohol consumption) is assessed in both groups, and the following data is obtained:

Cases (Liver Cirrhosis):

Exposed to alcohol: 120.

Not exposed to alcohol: 80.

Controls (No Liver Cirrhosis):

Exposed to alcohol: 100.

Not exposed to alcohol: 300.

Questions:

1. Construct a **2x2 table** to summarize the data.
2. Calculate the **Odds Ratio (OR)** for the association between alcohol consumption and liver cirrhosis.
3. Interpret the Odds Ratio in terms of the strength and direction of the association.

$$Odds\ of\ exposure\ in\ cases = \frac{Edxposed\ cases}{non\ exposed\ cases} = 1.5$$

$$Odds\ of\ exposure\ in\ controls = \frac{Exposed\ controls}{non\ exposed\ controls} = 0.333$$

$$OR = \frac{Odd\ of\ exposure\ in\ cases}{Odds\ of\ exposure\ in\ controls} = 4.5$$

Problem 7: Attributable Risk (AR)

Scenario: A cohort study is conducted to investigate the association between smoking and the risk of developing cardiovascular disease (CVD). The study follows two groups of individuals over 10 years:

• **Group 1 (Smokers)**: 2,000 smokers.

• **Group 2 (Non-Smokers)**: 3,000 non-smokers.

By the end of the study, the following outcomes are observed:

• **Smokers: 200 cases of CVD.**

• **Non-Smokers: 90 cases of CVD.**

Questions:

1. Calculate the **risk of CVD** in smokers and non-smokers.
2. Calculate the **Attributable Risk (AR)** for CVD among smokers compared to non-smokers.

3. Interpret the Attributable Risk in terms of the impact of smoking on CVD.

$$Risk\ in\ exposed = \frac{Number\ of\ casees\ in\ exposed}{Total\ number\ of\ exposed} = 0.10$$

$$Risk\ in\ non\ exposed = \frac{Number\ of\ cases\ in\ non\ exposed}{Total\ number\ of\ non\ exposed} = 0.03$$

$$AR = Risk\ in\ exposed - Risk\ in\ non\ exposed = 0.07$$
