Activity 3

Quantitative Reasoning

2021-08-30

Definition of age-adjusted incidence

What would be the incidence rate in country A if, in all age groups, it had the same fraction of the population as country B? Let's introduce some notation.

- $c_{A,\text{age}}$: the number of cancer cases in A in a given age group.
- $p_{B,\text{age}}$: the population in B in the same age group

The number of cases in this age group that would occur in country A if it had the population of country B is

$$\tilde{c}_{A,\text{age}} = \frac{p_{B,\text{age}}}{p_{A,\text{age}}} \cdot c_{A,\text{age}} . \qquad \text{(I)}$$

The "age-adjusted" incidence rate in country A is

$$\frac{\sum_{\text{age}} \tilde{c}_{A,\text{age}}}{p_{B,\text{age}}} \cdot 100\,000 . \qquad \text{(II)}$$

The unadjusted incidence rate would have

- $c_{A,\text{age}}$ in the numerator instead of $\tilde{c}_{A,\text{age}}$,
- $p_{A,\text{age}}$ in the denominator instead of $p_{B,\text{age}}$.

Activity 3: Age adjusted lung cancer incidence

- Append a column `Population_UK` to the `lung_cancer` data frame with

```
uk_population <- lung_cancer$Population[lung_cancer$Country == "UK"]
lung_cancer$Population_UK <- uk_population
```

This column contains the value of $p_{B,\text{age}}$ in Equation (I).

In the second line, we're taking advantage of R's vectorization: the shorter vector `uk_population` is repeated as often as necessary to fill the entire length of the longer vector `lung_cancer$Population_UK`.

- Append another column `Cases_if_UK` that contains the hypothetical number of cases if the country in the corresponding row had the population of the UK in this age group.

This column contains the values of $\tilde{c}_{A,\text{age}}$ in Equation (I).

- Use the `aggregate()` function to calculate the total number of cases that the country would have if it had the population and age structure of the UK.

  This column contains the numerator of Equation (II).

- Calculate the age-adjusted incidence rate defined by Equation (II).

- Make a bar chart of the age-adjusted incidence rate (one bar for each country).

- Compare this plot with the bar chart for the overall (i.e. unadjusted) incidence rate that you made during the prep work for toady. Are there any noteworthy changes?