

Lista 1

Alicja Wiązkowska

2023-04-14

Zadanie 1

parametry testu:

czułość - 92%, swoistość - 94%
10% populacji choruje na chorobę

Oznaczenia i wartości:

A - zdarzenie, że test wykaże chorobę (u losowo wybranej osoby)

Z - osoba jest zdrowa, C - osoba jest chora

$\mathbb{P}(A|Z)=6\%$ $\mathbb{P}(A|C)=92\%$ $\mathbb{P}(Z)=90\%$ $\mathbb{P}(C)=10\%$

a) Szukamy $\mathbb{P}(A)$

Ze wzoru na prawdopodobieństwo całkowite obliczamy $\mathbb{P}(A)$

$$\mathbb{P}(A) = \mathbb{P}(A|Z) \cdot \mathbb{P}(Z) + \mathbb{P}(A|C) \cdot \mathbb{P}(C) = 0.06 \cdot 0.9 + 0.92 \cdot 0.1$$

otrzymujemy wynik:

[1] 0.146

b) Szukamy $\mathbb{P}(C|A)$

korzystamy ze wzoru Bayesa

$$\mathbb{P}(C|A) = \frac{\mathbb{P}(A|C) \cdot \mathbb{P}(C)}{\mathbb{P}(A|C) \cdot \mathbb{P}(C) + \mathbb{P}(A|Z) \cdot \mathbb{P}(Z)} = \frac{0.92 \cdot 0.1}{0.92 \cdot 0.1 + 0.06 \cdot 0.9}$$

otrzymujemy wynik:

[1] 0.630137

c) Symulacja doświadczenia

doświadczalne sprawdzenie dla a)

Zestawienie wyników w zależności od wielkości próby dla doświadczalnie wznaczonego prawdopodobieństwa przedstawiono poniżej.

wielkosc proby n	P(A)
10	0.2000000
20	0.1500000
30	0.1333333
40	0.1250000
50	0.1000000
60	0.0666667
70	0.1000000
80	0.1500000
90	0.1333333
100	0.1200000

Wyniki te są dość bliskie wartości teoretycznej 0.146, choć nie są jej równe.

Z powodu zbyt małej liczby osób chorych w symulowanej populacji nie zrobiono tabelki w zależności od wielkości próby - za próbę przyjęto wszystkie osoby z pozytywnym wynikiem.

doświadczalnie wyznaczona wartość $P(B)$:

```
## [1] 0.75
```

Jest ona również zbliżona do teoretycznie wyznaczonej wartości 0.630137.

Zadanie 2

a) Zgadywanie wyniku

1. Powiedziano ci, że co najmniej jedno z dzieci jest dziewczyną. Jakie jest prawdopodobieństwo, że mają dwie dziewczynki?
33,(3)%
2. Teraz założmy, że powiedziano ci, iż starsze dziecko jest dziewczyną. Jakie jest prawdopodobieństwo, że mają dwie dziewczynki?
50%
3. Tym razem powiedziano ci, że co najmniej jedno z dzieci jest dziewczyną urodzoną we wtorek. Jakie jest prawdopodobieństwo, że mają dwie dziewczynki?
33,(3)%
4. Jakie jest prawdopodobieństwo, że mają dwie dziewczynki gdy wiesz, że co najmniej jedno z dzieci jest dziewczyną urodzoną we wtorek lub w piątek?
33,(3)%

b) Symulacje

1. Powiedziano ci, że co najmniej jedno z dzieci jest dziewczyną. Jakie jest prawdopodobieństwo, że mają dwie dziewczynki?

Odpowiedź symulacyjna:

```
## [1] 0.3170732
```

2. Teraz załóżmy, że powiedziano ci, iż starsze dziecko jest dziewczyną. Jakie jest prawdopodobieństwo, że mają dwie dziewczynki?

Odpowiedź symulacyjna:

[1] 0.474645

3. Tym razem powiedziano ci, że co najmniej jedno z dzieci jest dziewczyną urodzoną we wtorek. Jakie jest prawdopodobieństwo, że mają dwie dziewczynki?

Odpowiedź symulacyjna:

[1] 0.4295775

4. Jakie jest prawdopodobieństwo, że mają dwie dziewczynki gdy wiesz, że co najmniej jedno z dzieci jest dziewczyną urodzoną we wtorek lub w piątek?

Odpowiedź symulacyjna:

[1] 0.4269663

c) Wnioski

Prawdopodobieństwo obliczone w symulacji w podpunkcie 4. (0.43) wydaje się odbiegać od spodziewanego. Założyłam, że dzień urodzenia dziecka nie ma znaczenia, więc informacja ta została pominięta, co okazało się błędem. Z tego powodu podpunkty 3 i 4 należy obliczyć teoretycznie, gdyż ich wyniki mogą być sprzeczne z intuicją.

Oznaczenia:

- D_m - Dzień tygodnia urodzenia młodszego dziecka
- D_s - Dzień tygodnia urodzenia starszego dziecka
- P_m - Płeć urodzenia młodszego dziecka
- P_s - Płeć urodzenia starszego dziecka

Dane:

- $\mathbb{P}(P_x = K) = \frac{1}{2}$ dla $x \in \{s, m\}$
- $\mathbb{P}(P_x = d) = \frac{1}{7}$ dla $x \in \{s, m\}$ \wedge $d \in \{pon, wt, \acute{s}r, czw, pt, sb, nd\}$

Poprawne obliczenie prawdopodobieństwa:

3. Co najmniej jedno z dzieci jest dziewczyną urodzoną we wtorek. Jakie jest prawdopodobieństwo, że mają dwie dziewczynki?

$$\begin{aligned} \mathbb{P}[(P_m = K \wedge P_s = K) \mid ((P_m = K \wedge D_m = wt) \vee (P_s = K \wedge D_s = wt))] = \\ = \frac{\mathbb{P}[(P_m = K \wedge P_s = K) \wedge (D_m = wt \vee D_s = wt)]}{\mathbb{P}[(P_m = K \wedge D_m = wt) \vee (P_s = K \wedge D_s = wt)]} = \end{aligned}$$

$$\frac{\mathbb{P}(P_m = K \wedge P_s = K) \cdot [\mathbb{P}(D_m = wt \wedge D_s = wt) + \mathbb{P}(D_m = wt \wedge D_s \neq wt) + \mathbb{P}(D_m \neq wt \wedge D_s = wt)]}{\mathbb{P}(P_m = K) \cdot \mathbb{P}(D_m = wt) + \mathbb{P}(P_s = K) \cdot \mathbb{P}(D_s = wt) - \mathbb{P}(P_m = K) \cdot \mathbb{P}(P_s = K) \cdot \mathbb{P}(D_m = wt) \cdot \mathbb{P}(D_s = wt)} =$$

$$= 0.5 \cdot 0.5 \cdot \frac{\frac{1}{7} \cdot \frac{1}{7} + \frac{1}{7} \cdot \frac{6}{7} + \frac{6}{7} \cdot \frac{1}{7}}{0.5 \cdot \frac{1}{7} + 0.5 \cdot \frac{1}{7} - 0.5 \cdot 0.5 \cdot \frac{1}{7} \cdot \frac{1}{7}} = \frac{1}{4} \cdot \frac{\frac{13}{49}}{\frac{2}{14} - \frac{1}{4} \cdot \frac{1}{49}} = \frac{13}{\frac{4 \cdot 49}{7} - 1} = \frac{13}{28 - 1} = \frac{13}{27} \approx 0.48$$

Wyliczona powyżej wartość teoretyczna jest bliższa wartości z symulacji (0.47) niż założone 33,(3)%. Z tego powodu podpunkt 3.

4. Jakie jest prawdopodobieństwo, że mają dwie dziewczynki gdy wiesz, że conajmniej jedno z dzieci jest dziewczyną urodzoną we wtorek lub w piątek?

Przyjmijmy oznaczenie:

$$\begin{aligned} & \mathbb{P}[(P_m = K \wedge P_s = K) \mid ((P_m = K \wedge D_m \in \{wt, pt\}) \vee (P_s = K \wedge D_s \in \{wt, pt\}))] = \\ &= \frac{\mathbb{P}[(P_m = K \wedge P_s = K) \wedge (D_m \in \{wt, pt\} \vee D_s \in \{wt, pt\})]}{\mathbb{P}[(P_m = K \wedge D_m \in \{wt, pt\}) \vee (P_s = K \wedge D_s \in \{wt, pt\})]} = \end{aligned}$$

$$\frac{\mathbb{P}(P_m = K \wedge P_s = K) \cdot [\mathbb{P}(D_m, D_s \in \{wt, pt\}) + \mathbb{P}(D_m \in \{wt, pt\} \wedge D_s \notin \{wt, pt\}) + \mathbb{P}(D_m \notin \{wt, pt\} \wedge D_s \in \{wt, pt\})]}{\mathbb{P}(P_m = K) \cdot \mathbb{P}(D_m = wt) + \mathbb{P}(P_s = K) \cdot \mathbb{P}(D_s = wt) - \mathbb{P}(P_m = K) \cdot \mathbb{P}(P_s = K) \cdot \mathbb{P}(D_m = wt) \cdot \mathbb{P}(D_m = wt)}$$

$$= \frac{\frac{1}{2} \cdot \frac{1}{2} \cdot (\frac{2}{7} \cdot \frac{2}{7} + \frac{2}{7} \cdot \frac{5}{7} + \frac{5}{7} \cdot \frac{2}{7})}{\frac{1}{2} \cdot \frac{2}{7} + \frac{1}{2} \cdot \frac{2}{7} - \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{2}{7} \cdot \frac{2}{7}} = \frac{\frac{1}{4} \cdot \frac{24}{49}}{\frac{2}{7} - \frac{1}{49}} = \frac{\frac{6}{49}}{\frac{14-1}{49}} = \frac{6}{13} \approx 0.46$$

Wartość wyliczona teoretycznie jest bliższa wartości symulacyjnej 0.43 niż założone 33,(3)%.