**Kaczmarek Wiktoria, Aktywność 07.**

**5.** Linia lotnicza chce oszacować frakcję Polaków, którzy będą korzystać z nowo otwartego połączenia między Poznaniem a Londynem. Wybrano losową próbę 347 pasażerów korzystających z tego połączenia, z których 201 okazało się Polakami.

a) Wyznaczyć 90% przedział ufności dla frakcji Polaków wśród pasażerów korzystających z nowo otwartego połączenia. **Odp.:** (0,536; 0,623).

b) Wygenerować 347 elementową próbę według rozkładu $B(0{,}58)$ identyfikującą polskich pasażerów i na tej podstawie wyznaczyć 90% przedział ufności.

Odpowiedzi:

Zmienna losowa opisuje czy wybrany pasażer jest Polakiem, czy nie. Zmienne losowe są niezależne od siebie. Możemy zatem stwierdzić, że zmienna losowa ma rozkład dwumianowy, oceniamy, że sukcesem będzie wybranie pasażera, który jest Polakiem, jednak nie znamy prawdopodobieństwa wylosowania takiego pasażera.

a) Dla przedziału ufności 90% mamy:

$$1-\alpha = 0{,}9 \Rightarrow \alpha = 0{,}1$$

A więc korzystamy z:

| | | |
|---|---|---|
| $X \sim B(p)$, $p$ nieznane, $0 < \bar{p}_n \mp 3 \cdot \sqrt{\frac{\bar{p}_n(1-\bar{p}_n)}{n}} < 1$ | $p$ | $\bar{P}_n \mp z_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{\bar{P}_n(1-\bar{P}_n)}{n}}$ |

Otrzymujemy:

$$\overline{p_n} \mp 3\sqrt{\frac{\overline{p_n}(1-\overline{p_n})}{n}} = \frac{201}{347} \mp 3\sqrt{\frac{\frac{201}{347}\left(1-\frac{201}{347}\right)}{347}} \approx 0{,}5792507 \mp 0{,}07950629$$

```
> 201/347
[1] 0.5792507
> 3*sqrt((201/347)*(1-201/347)/347)
[1] 0.07950629
```

Zatem spełniony jest warunek, możemy przejść do wyznaczenia poziomu ufności:

$$\overline{P_n} \mp z_{1-\frac{\alpha}{2}}\sqrt{\frac{\overline{p_n}(1-\overline{p_n})}{n}} = 0{,}5792507 \mp 0{,}04359207$$

```
> p<-201/347
> z<-qnorm(0.95,0,1)*sqrt((201/347)*(1-201/347)/347)
> p+z
[1] 0.6228428
> p-z
[1] 0.5356586
```

Stąd wnioskujemy, że przedział dla 90% ufności to (0,5356586; 0,6228428).

b) Tutaj mamy podany rozkład B(0.58). Przy pomocy wspomagania komputerowego – programu R generujemy 347 elementową próbę

p-próba

Najpierw sprawdzamy warunek:

$$\overline{p_n} \mp 3\sqrt{\frac{\overline{p_n}(1-\overline{p_n})}{n}} = \frac{194}{347} \mp 3\sqrt{\frac{\frac{194}{347}\left(1-\frac{194}{347}\right)}{347}} \approx\ 0{,}5590778 \mp 0{,}07996015$$

Warunek jest spełniony więc obliczamy przedziały ufności:

$$\overline{P_n} \mp z_{1-\frac{\alpha}{2}}\sqrt{\frac{\overline{p_n}(1-\overline{p_n})}{n}} = 0{,}5590778 \mp 0{,}043884092$$

Otrzymujemy, że przedział ufności wynosi (0,5152369;0,6029187).

```
> p<-rbinom(1,347,0.58)
> p
[1] 194
> 3*sqrt((p/347)*(1-p/347)/347)
[1] 0.07996015
> p/347
[1] 0.5590778
> z<-qnorm(0.95,0,1)*sqrt((p/347)*(1-p/347)/347)
> (p/347)-z
[1] 0.5152369
> (p/347)+z
[1] 0.6029187
```

**15.** (**Studium przypadku**). Z partii kondensatorów wybrano losowo 12 kondensatorów i zmierzono ich pojemności, otrzymując wyniki (w *pF*):

4,45; 4,40; 4,42; 4,38; 4,44; 4,36; 4,40; 4,39; 4,45; 4,35; 4,40; 4,35.

a) Znaleźć ocenę wartości oczekiwanej $\bar{x}_{12}$ i wariancji $s_{12}^2$ pojemności kondensatora pochodzącego z danej partii.

b) Wygenerować 100 elementową próbę według rozkładu $\mathcal{N}(\bar{x}_{12}, s_{12})$.

c) Znaleźć ocenę wskaźnika kondensatorów, które nie spełniają wymagań technicznych, przyjmując, że kondensator nie spełnia tych wymagań, gdy jego pojemność jest mniejsza od 4,39 *pF*.

d) Znaleźć ocenę wariancji pojemności kondensatorów.

e) Wyznaczyć 90-procentową ocenę przedziału ufności dla wartości oczekiwanej pojemności kondensatora pochodzącego z danej partii.

f) Wyznaczyć 90-procentową realizację przedziału ufności dla wskaźnika kondensatorów, które nie spełniają wymagań technicznych w badanej partii.

Odpowiedzi:

a) $\overline{x_{12}} = \frac{1}{12}\sum_{i=1}^{12} x_i = \frac{52{,}79}{12} = 4{,}3991667 \approx 4{,}4$

$$s_{12}^2 = \frac{1}{11}\sum_{i=1}^{12}(x_i - \overline{x_{12}})^2 = \frac{0{,}14091667}{11} = 0{,}001281061 \approx 0{,}0013$$

b)

```
> x<-rnorm(100,4.4,0.0013)
> x<-sort(x)
> x
  [1] 4.396133 4.396462 4.396872 4.396981 4.397007 4.397185 4.397248 4.397445
  [9] 4.397560 4.397623 4.397751 4.397778 4.397823 4.398100 4.398448 4.398537
 [17] 4.398591 4.398656 4.398809 4.399019 4.399036 4.399057 4.399091 4.399129
 [25] 4.399132 4.399134 4.399139 4.399156 4.399162 4.399351 4.399356 4.399388
 [33] 4.399407 4.399447 4.399474 4.399483 4.399528 4.399536 4.399538 4.399554
 [41] 4.399562 4.399587 4.399595 4.399611 4.399620 4.399631 4.399714 4.399734
 [49] 4.399759 4.399767 4.399768 4.399835 4.399835 4.399843 4.399862 4.399951
 [57] 4.399965 4.400001 4.400061 4.400072 4.400082 4.400126 4.400173 4.400195
 [65] 4.400197 4.400238 4.400250 4.400305 4.400332 4.400350 4.400360 4.400474
 [73] 4.400522 4.400524 4.400547 4.400590 4.400694 4.400716 4.400726 4.400743
 [81] 4.400771 4.400899 4.400931 4.400987 4.401074 4.401165 4.401186 4.401238
 [89] 4.401358 4.401363 4.401420 4.401559 4.401588 4.401721 4.401791 4.401901
 [97] 4.402066 4.402094 4.402481 4.402482
```

c) Ilość wadliwych kondensatorów – 4, n=12

$$P_n = \frac{4}{12} = \frac{1}{3} \approx 0,33$$

d) $s_{12}^2 = \frac{1}{11}\sum_{i=1}^{12}(x_i - \overline{x_{12}})^2 = \frac{0,14091667}{11} = 0,001281061 \approx 0,0013$

e) $1 - \alpha = 0,9 => \ \alpha = 0,1$

Następnie korzystamy z:

| | | |
|---|---|---|
| $X \sim \mathcal{N}(m, \sigma)$, $m$ nieznane, $n$ dowolne | $\sigma^2$ | $\left(\frac{(n-1)S_n^2}{\chi^2_{1-\frac{\alpha}{2};n-1}}; \frac{(n-1)S_n^2}{\chi^2_{\frac{\alpha}{2};n-1}}\right)$ |

Korzystając z R obliczamy przedział:

```
> x<-11*0.0013
> y1<-qchisq(0.95,11)
> y2<-qchisq(0.05,11)
> x/y1
[1] 0.0007268056
> x/y2
[1] 0.003125811
```

Ostatecznie mamy (0,0007268056; 0,003125811).

Aby obliczyć dla m korzystamy z:

| | | |
|---|---|---|
| $X \sim \mathcal{N}(m, \sigma)$, $\sigma$ nieznane, $n$ dowolne | $m$ | $\bar{X}_n \mp t_{1-\frac{\alpha}{2},n-1} \cdot \frac{S_n}{\sqrt{n}}$ |

Obliczamy oczywiście z wykorzystaniem R:

```
> x<-qt(0.95,11)
> 4.4-x*sqrt(0.0013/12)
[1] 4.381308
> 4.4+x*sqrt(0.0013/12)
[1] 4.418692
```

i otrzymujemy przedział (4.381308; 4.418692).

f) $1-\alpha=0{,}9 \Rightarrow \alpha=0{,}1$

Korzystamy tym razem ze wzorów:

| | | |
|---|---|---|
| $X\sim B(p)$, $p$ nieznane, $0<\bar{p}_n \mp 3\cdot\sqrt{\frac{\bar{p}_n(1-\bar{p}_n)}{n}}<1$ | $p$ | $\bar{P}_n \mp z_{1-\frac{\alpha}{2}}\cdot\sqrt{\frac{\bar{P}_n(1-\bar{P}_n)}{n}}$ |

Zaczynam od sprawdzenia warunku:

$$\overline{p_n} \mp 3\sqrt{\frac{(1-\overline{p_n})\overline{p_n}}{n}} = \frac{4}{12} \mp \sqrt{\frac{\frac{4}{12}\cdot\frac{8}{12}}{12}} = 0.333333 \mp 0.408248$$

Widzimy, że warunek jest spełniony, ale tylko dla wartości dodatniej, więc możemy przejść do obliczenia przedziału ufności, jednak tylko prawostronnego:

```
> pn<-4/12
> z<-qnorm(0.95,0,1)
> pn+z*sqrt((4/12*8/12)/12)
[1] 0.5571696
```

Ostatecznie otrzymujemy przedział (0; 0.5571696).

Zadanie 11 z książki K.Andrzejczak strona 176.

Biolog wyłowił 20 nowo narodzonych cierników, więc n=20. Zakładamy, że zmienna losowa T opisująca długość hodowlanych ryb ma rozkład t-Studenta z v=n-1=19 stopniach swobody.

Będziemy korzystać z:

| | | |
|---|---|---|
| $X\sim\mathcal{N}(m,\sigma)$, $\sigma$ nieznane, $n$ dowolne | $m$ | $\bar{X}_n \mp t_{1-\frac{\alpha}{2},n-1}\cdot\frac{S_n}{\sqrt{n}}$ |

Chcąc zbudować 95% realizację przedziału ufności mamy $95\%=(1-\alpha)100\%$ więc wnioskujemy, że $\alpha=0.05$. Następnie obliczamy wartości dla kwantyli za pomocą R:

```
> qt(0.025,19)
[1] -2.093024
> qt(0.975,19)
[1] 2.093024
```

$t_{0.025}=-2.093024,\ t_{0.975}=\ 2.093024$

$AVE(x)=27.5 \Rightarrow \bar{X}_n=27.5$

$S(x)=2.6$

Ostatecznie wyznaczamy przedział ufności:

$$27.5+(-2.093)\frac{2.6}{\sqrt{20}}=26.28318$$

$$27.5+(2.093)\frac{2.6}{\sqrt{20}}=28.71682$$

Obliczone wartości za pomocą R:

```
> 27.5-2.093*(2.6/sqrt(20))
[1] 26.28318
> 27.5+2.093*(2.6/sqrt(20))
[1] 28.71682
```

Więc ostatecznie otrzymujemy przedział (26.28318; 28.71682).

Odp. 95% realizację przedziału ufności dla nieznanej wartości oczekiwanej długości ciernika hodowlanego 10 tygodni w wodzie zawierającej chlor jest przedział od 26.28 do 28.72mm.

Zadanie

10 . Na ilu potencjalnych klientach należy przeprowadzić ankietę, aby oszacować odsetek osób mających zamiar zakupić nowy samochód w ciągu najbliższych 2 lat? Przyjąć poziom ufności 0,95 oraz maksymalny dopuszczalny błąd szacunku 6%. Odp.: 267

$$1-\alpha = 0{,}95 => \alpha = 0{,}05$$

$$d = 6\% = 0.06$$

Szukamy n=?

Musimy skorzystać z wzoru: $n = \left\lceil \frac{z^2_{1-\frac{\alpha}{2}}}{4d^2} \right\rceil$.

```
> z<-qnorm(0.95,0,1)
> ceiling(z^2/(4*0.06^2))
[1] 188
```

Ankietę trzeba przeprowadzić na potencjalnych 188 klientach.