Laboratoria 7 + 8

Krystian Baran 145000

20 kwietnia 2021

Spis treści

Część I

Laboratoria 07 - Estymacja punktowa

1 Zadanie 1

Wyznaczyć estymator parametru p w rozkładzie Bernoulliego.

Rozkład Bernoulliego ma rozkład następujący:

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, k = 0, 1, \ldots, n$$

Aby wyznaczyć estymator parametru p skorzystamy z metody momentów. Dla rozkładu Bernoulliego $\mathbb{E}X = np$ i $\mathbb{D}^2(X) = np(1-p)$. Oznaczymy momenty punktowe jako:

$$\mathbb{E}X = \overline{X}$$

$$\mathbb{D}^2(X) = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \overline{X})^2 = s^2$$

Wtedy:

$$\begin{cases} np = \overline{X} \\ np(1-p) = s^2 \end{cases}$$

$$\begin{cases} np = \overline{X} \\ \overline{X}(1-p) = s^2 \end{cases}$$

$$\begin{cases} np = \overline{X} \\ 1 - p = \frac{s^2}{\overline{X}} \end{cases}$$

$$\begin{cases} np = \overline{X} \\ p = 1 - \frac{s^2}{\overline{X}} \end{cases}$$

$$\begin{cases} n = \frac{\overline{X}^2}{\overline{X} - s^2} \\ p = 1 - \frac{s^2}{\overline{X}} \end{cases}$$

Zatem estymator parametru p jest $1 - \frac{s^2}{\overline{X}}$.

2 Zadanie 2

Wyznaczyć MM oraz MNW estymatory parametrów rozkładu normalnego.

Rozkład normalny definiowany jest w następujący sposób:

$$N(\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Parametr $\mu = \mathbb{E}X$ natomiast $\sigma = \mathbb{D}X$. Drugi moment zwykły tego rozkładu jest następujący:

$$\mathbb{E}(X^2) = \sigma^2 + \mu^2$$

Wyznaczymy estymatory parametrów metodą momentów. Niech $\mathbb{E}X = \overline{X}$ i $\mathbb{E}(X^2) = \frac{\sum_{i=1}^n X_i^2}{n}$, wtedy:

$$\begin{cases} \overline{X} = \mu \\ \frac{\sum_{i=1}^n X_i^2}{n} = \sigma^2 + \mu^2 \end{cases}$$
$$\begin{cases} \overline{X} = \mu \\ \frac{\sum_{i=1}^n X_i^2}{n} = \sigma^2 + \overline{X}^2 \end{cases}$$
$$\begin{cases} \overline{X} = \mu \\ \frac{\sum_{i=1}^n X_i^2}{n} = \sigma^2 + \mu^2 \end{cases}$$

Metodą największej wiarygodności natomiast potrzebujemy wyznaczyć funkcje wiarygodności.

$$\begin{aligned} L(x_1, x_2, \ldots, x_n|\mu,\sigma) &= \prod_{i=1}^n \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x_i-\mu)^2}{2\sigma^2}} \\ &= \frac{1}{\sigma^n (2\pi)^{n/2}} e^{-\sum_{i=1}^n \frac{(x_i-\mu)^2}{2\sigma^2}} \end{aligned}$$
$$\ln(L) = -n\ln(\sigma) - \frac{n}{2}\ln(2\pi) - \sum_{i=1}^n \frac{(x_i-\mu)^2}{2\sigma^2}$$

Obliczymy najpierw estymator parametru μ.

$$\begin{aligned} \frac{\partial(\ln(L))}{\partial\mu} &= \sum_{i=1}^n \frac{2(x_i-\mu)}{2\sigma^2} \\ &= \frac{\sum_{i=1}^n (x_i-\mu)}{\sigma^2} = 0 \end{aligned}$$
$$\sum_{i=1}^n (x_i-\mu) = 0$$
$$\sum_{i=1}^n x_i - n\mu = 0$$
$$\mu = \frac{\sum_{i=1}^n x_i}{n}$$

Dla parametru σ natomiast:

$$\begin{aligned}
\frac{\partial(\ln(L))}{\partial\sigma} &= -\frac{n}{\sigma} + \sum_{i=1}^{n} \frac{2(x_i-\mu)^2}{2\sigma^3} \\
&= -\frac{n}{\sigma} + \frac{\sum_{i=1}^{n}(x_i-\mu)^2}{\sigma^3} = 0 \\
\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{\sigma^3} &= \frac{n}{\sigma} \\
\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{n} &= \sigma^2 \\
\sigma &= \sqrt{\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{n}}
\end{aligned}$$

Widzimy zatem że parametry $\mu 4$ i σ są, odpowiednio, średnią i odchyleniem standardowym populacji szeregu punktowego.

3 Zadanie 3

Wyznaczyć MNW estymator parametru rozkładu Poissona.

Rozkład Poissona definiowany jest w następujący sposób:

$$P(X=k) = \frac{\lambda^k e^{-\lambda}}{k!}, k = 0, 1, 2, \ldots$$

Jest to rozkład jedno parometrowy. Aby wyliczyć estymator parametru λ potrzebujemy obliczyć funkcję wiarygodności.

$$\begin{aligned}
L(k_1, k_2, \ldots, k_n|\lambda) &= \prod_{i=1}^{n} P(X=k_i) = \prod_{i=1}^{n} \frac{\lambda^{k_i} e^{-\lambda}}{k_i!} \\
&= \lambda^{\sum_{i=1}^{n} k_i} e^{-n\lambda} \prod_{i=1}^{n} \frac{1}{k_i!} \\
\ln(L) &= \ln(\lambda) \sum_{i=1}^{n} k_i - n\lambda - \sum_{i=1}^{n} \ln k_i!
\end{aligned}$$

Estymator parametru jest maksimum tej funkcji po zmiennej λ, zatem przyrównamy pierwszą pochodną do zera i znajdziemy szukany estymator.

$$\begin{aligned}
\frac{\partial(\ln(L))}{\partial\lambda} &= \frac{\sum_{i=1}^{n} k_i}{\lambda} - n = 0 \\
n &= \frac{\sum_{i=1}^{n} k_i}{\lambda} \\
\lambda &= \frac{\sum_{i=1}^{n} k_i}{n}
\end{aligned}$$

Sprawdzimy teraz drugą pochodną.

$$\frac{\partial^2(\ln(L))}{\partial\lambda^2} = -\frac{\sum_{i=1}^{n} k_i}{\lambda^2} < 0, \quad \forall\lambda$$

Zatem estymator parametru λ jest średnia arytmetyczna populacji.

4 Zadanie 5

Wygenerować 50 elementową próbę prostą z populacji, w której cecha X ma rozkład o gęstości $f(x) = \frac{x}{8}\mathbf{1}_{(0;4)}(x)$

a) Sporządzić histogram.

b) Wyznaczyć wartość oczekiwaną i wariancję oraz ich oceny na podstawie wygenerowanej próby.

4.1 a)

Aby wygenerować próbę korzystając z twierdzenia obrócenia dystrybuanty musimy obliczyć dystrybuantę

$$\begin{aligned} F_X(x) &= \int_{\mathbb{R}} \frac{t}{8}\mathbf{1}_{(0;4)}(t)dt = \int_0^x \frac{t}{8}dt \\ &= \left.\frac{t^2}{16}\right|_0^x \\ &= \frac{x^2}{16} \end{aligned}$$

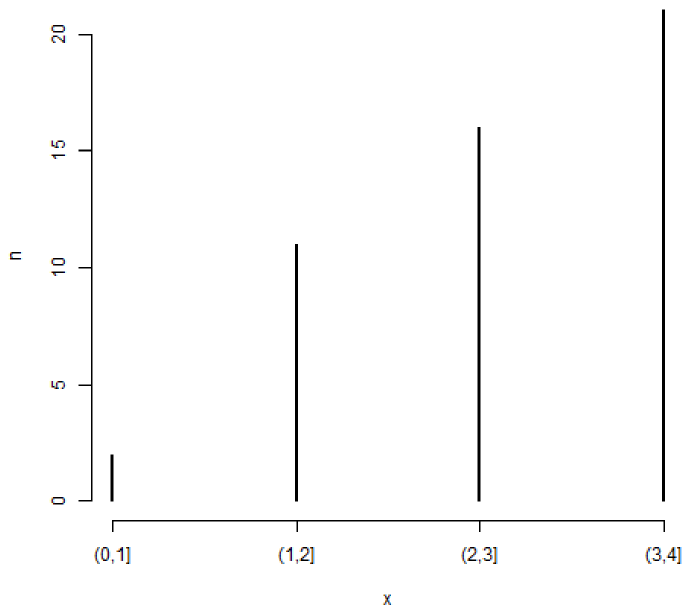
Zatem dystrybuanta jest następująca:

$$F_X(x) = \begin{cases} 0 & , \quad x \leqslant 0 \\ \frac{x^2}{16} & , \quad 0 < x < 4 \\ 1 & , \quad x \geqslant 4 \end{cases}$$

Odwrócimy dystrybuantę.

$$\begin{aligned} y &= \frac{x^2}{16} \\ 16y &= x^2 \\ x &= 4\sqrt{y} \end{aligned}$$

Poniżej przedstawiony został histogram przedziałowy wygenerowanej próby:



4.2 b)

Obliczymy teraz wartość oczekiwaną i wariancję z podanej funkcji i z wygenerowanej próby:

$$\begin{aligned}\mathbb{E}X &= \int_{\mathbb{R}} \frac{x^2}{8}\mathbb{I}_{(0,4)}(x)dx = \int_0^4 \frac{x^2}{8}dx \\ &= \frac{x^3}{24}\Big|_0^4 \\ &= \frac{64}{24} \approx 2.666666667\end{aligned}$$

$$\begin{aligned}\mathbb{E}(X^2) &= \int_{\mathbb{R}} \frac{x^3}{8}\mathbb{I}_{(0,4)}(x)dx = \int_0^4 \frac{x^3}{8}dx \\ &= \frac{x^4}{32}\Big|_0^4 \\ &= \frac{64}{32} = 8\end{aligned}$$

$$\mathbb{D}^2(X) = \mathbb{E}(X^2) - \mathbb{E}X^2 = 8 - \frac{4096}{576} \approx 0.88888888889$$

Wartość oczekiwana z próby będzie średnią z próby, natomiast wariancje oznaczymy następującym wzorem:

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{X})^2$$

Obliczymy szukane wartości w R, gdzie *prob* jest tablicą zawierającą 50-elementową próbę.

$$\overline{X} \stackrel{R}{=} mean(prob) \approx 2.751196$$

$$s^2 \stackrel{R}{=} var(prob) \approx 0.8295317$$

Widzimy że że oba wartości są do siebie blisko, zatem możemy stwierdzić że obliczyliśmy poprawnie, gdzie $\bar{(}X) = 2.7 \pm 0.1$ i $s^2 = 0.8 \pm 0.1$.

5 Zadanie 6

Korzystając z dostępnego oprogramowania wygenerować 100 elementową próbę według rozkładu

a) bin(20; 0,8),

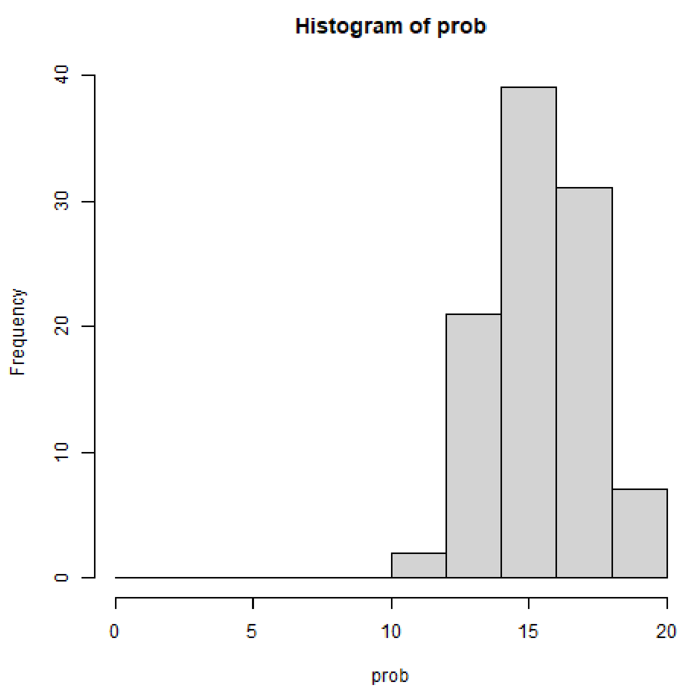
b) nbin(3;0,1),

c) Poisson(5).

Sporządzić histogram i dokonać ocenę punktową parametrów.

5.1 a)

Aby wygenerować losową próbę 100 elementową rozkładu Dwumianowego skorzystamy z funkcji R-owskiej *rbinom()*. Poniżej przestawiony został histogram dla losowej próby.



Wartość oczekiwana i wariancja teoretyczna wynoszą:

$$\mathbb{E}(X) = np = 20 \cdot 0.8 = 16$$

$$\mathbb{D}^2(X) = np(1-p) = 20 \cdot 0.8 \cdot 0.2 = 3.2$$

Natomiast, korzystając z wygenerowanej próby i z funkcji na średnią i wariancje w R otrzymujemy następujące wartości:

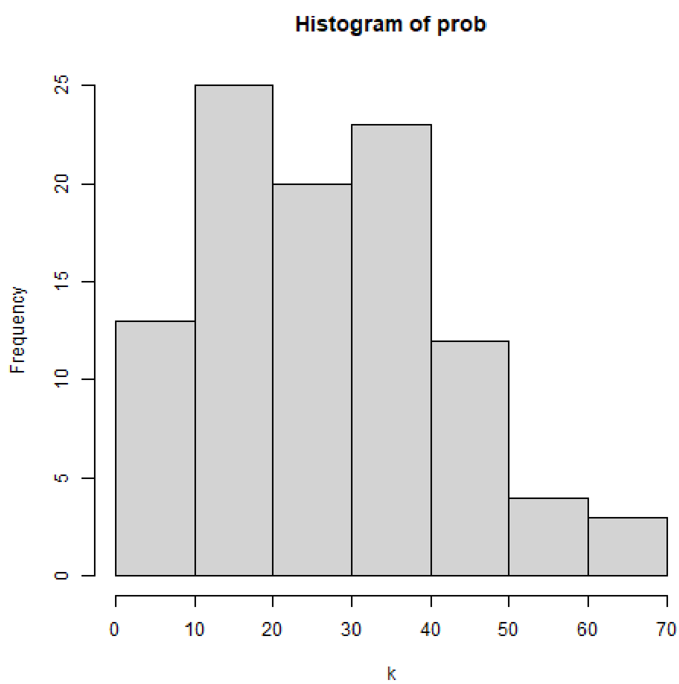
$$\overline{X} \overset{R}{=} mean(prob) \approx 15.87$$

$$s^2 \overset{R}{=} var(prob) \approx 3.104141$$

Widzimy że wartości tę są blisko wartości teoretycznej, zatem można stwierdzić że estymowana wartość oczekiwana wynosi 16 ± 0.2 a wariancja wynosi 3.1 ± 0.1.

5.2 b)

Podobnie jak w podpunkcie **a** wygenerujemy losową próbę 100-elementową w R za pomocą funkcji wbudowanej *rnbinom()*. Poniżej przedstawiono histogram.



Wartość oczekiwana i wariancja teoretyczna wynoszą:

$$\mathbb{E}(X) = \frac{(1-p)r}{p} = \frac{3 \cdot 0.9}{0.1} \approx 27$$

$$\mathbb{D}^2(X) = \frac{(1-p)r}{p^2} = \frac{3 \cdot 0.9}{0.01} \approx 270$$

Natomiast, korzystając z wygenerowanej próby i z funkcji na średnią i wariancje w R otrzymujemy następujące wartości:

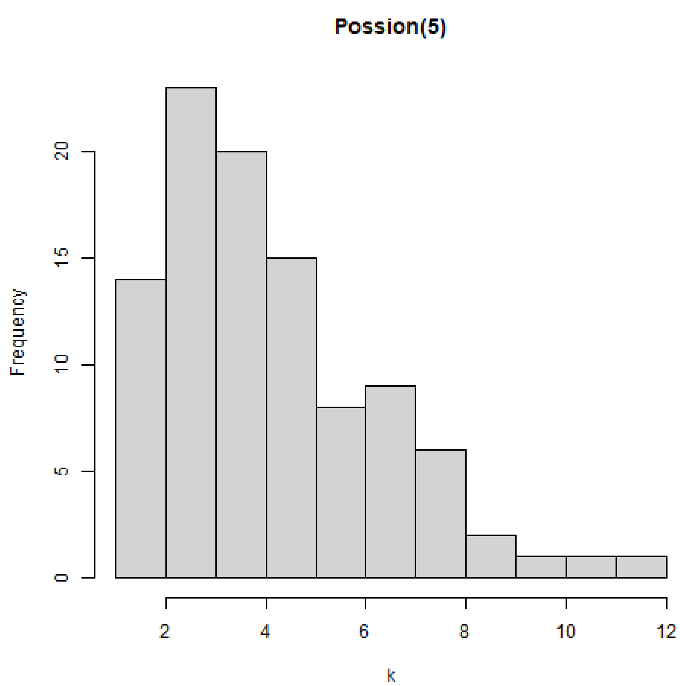
$$\overline{X} \stackrel{R}{=} mean(prob) \approx 27.75$$

$$s^2 \stackrel{R}{=} var(prob) \approx 216.8965$$

Dla wartości oczekiwanej widzimy że wartości tę są blisko wartości teoretycznej, zatem można stwierdzić że estymowana wartość oczekiwana wynosi 27 ± 0.7. Natomiast wariancja jest znacznie inna niż wartość teoretyczna; może wynikać to z tego że dla dużych wartości nie uzyskujemy znaczną dokładność, zatem na pewno pierwsza liczba została obliczona dokładnie a reszta już nie.

5.3 c)

Podobnie jak w podpunkcie **a** i **b** wygenerujemy losową próbę 100-elementową w R za pomocą funkcji wbudowanej *rpois()*. Poniżej przedstawiono histogram.



Wartość oczekiwana i wariancja teoretyczna wynoszą:

$$\mathbb{E}(X) = \lambda = 5$$

$$\mathbb{D}^2(X) = \lambda = 5$$

Natomiast, korzystając z wygenerowanej próby i z funkcji na średnią i wariancje w R otrzymujemy następujące wartości:

$$\overline{X} \stackrel{R}{=} mean(prob) \approx 4.59$$

$$s^2 \stackrel{R}{=} var(prob) \approx 4.870606$$

W tym przypadku widzimy że przy aproksymacji do liczby całkowitej uzyskamy dobrą wartość estymowanych parametrów. Zatem uznajemy że estymowana wartość oczekiwana wynosi 5 ± 1 a wariancja tak samo.

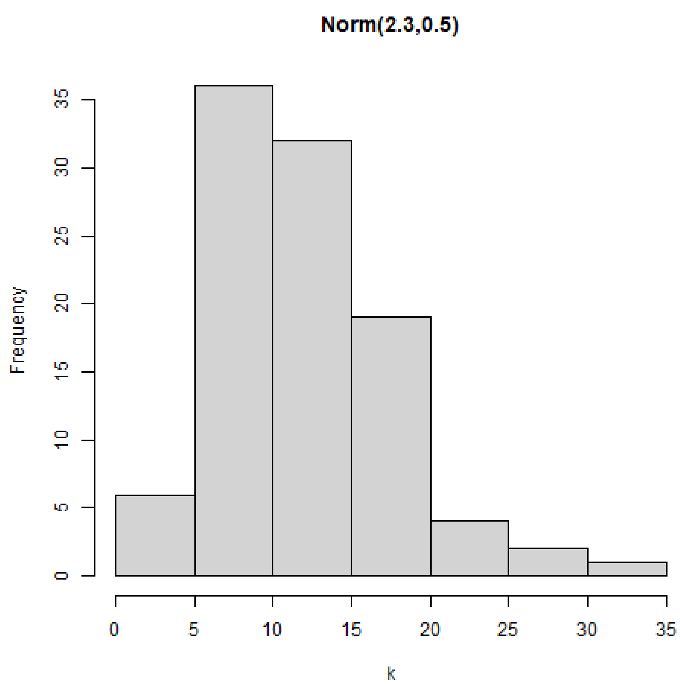
6 Zadanie 7

Wygenerować 100 elementową próbę według rozkładu logarytmiczno-normalnego z parametrami $\mu = 2.3$ i $\sigma = 0.5$.

a) Sporządzić histogram.

b) Dokonać estymacji parametrów, ocenić wartość oczekiwaną i wariancję oraz porównać te wartości z wartościami teoretycznymi.

Aby wygenerować losową próbę 100-elementową skorzystamy z dostępnej funkcji R-owskiej dla rozkładu logarytmiczno-normalnego *rlnorm()*. Poniżej przedstawiony został histogram z wygenerowanej próby:



Dla rozkładu logarytmiczno-normalnego wartość oczekiwana i wariancja są następujące:

$$\mathbb{E}(X) = e^{\mu} = e^{2}.3 \approx 9.974182455$$

$$\mathbb{D}^2(X) = (e^{\sigma^2} - 1)e^{2\mu+\sigma^2} = (e^{0.25} - 1)e^{4.6+0.25} \approx 36.28151745$$

Korzystając z wygenerowanej próby i z funkcji na średnią i wariancje w R otrzymujemy następujące wartości:

$$\overline{X} \stackrel{R}{=} mean(prob) \approx 11.72756$$

$$s^2 \overset{R}{=} var(prob) \approx 30.83267$$

Wartości te różnią się nie wiele od wartości teoretycznych.

Część II

Laboratoria 8 - Estymacja przedziałowa

7 Zadanie 3

Rozkład wyników pomiarów głębokości morza w pewnym rejonie jest normalny. Dokonano 5 niezależnych pomiarów głębokości morza w tym rejonie i otrzymano następujące wyniki (w [m]): 871, 862, 870, 876, 866. Na poziomie ufności 0,90 wyznaczyć CI dla wartości oczekiwanej oraz dla wariancji głębokości morza w badanym rejonie.

Obliczymy najpierw średnią i wariancję z podanej próby:

$$\overline{X}_5 = \frac{1}{n}\sum_{i=1}^{n} X_i = \frac{4345}{5} = 869$$

$$S_5^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2 = \frac{112}{4} = 28$$

$$S_5 = \sqrt{S_5^2} \approx 5.291502622$$

Wyznaczymy teraz α wiedząc że poziom ufności jest 0.90.

$$1 - \alpha = 0.90 \Rightarrow \alpha = 0.1$$

Ponieważ pomiary głębokości morza mają rozkład normalny gdzie nie znane są parametry m i σ skorzystamy najpierw z przedziału ufności dla σ^2 podany poniżej:

$$\Big(\frac{(n-1)S_n^2}{\chi^2_{1-\frac{\alpha}{2};n-1}}; \frac{(n-1)S_n^2}{\chi^2_{\frac{\alpha}{2};n-1}}\Big)$$

Obliczymy teraz wartości kwantyli rozkładu χ^2:

$$\chi^2_{1-0.05;4} \stackrel{R}{=} qchisq(0.95, 4) \approx 9.487729$$

$$\chi^2_{0.05;4} \stackrel{R}{=} qchisq(0.05, 4) \approx 0.710723$$

Wtedy szukany przedziały ufności dla σ^2:

$$(11.80472166; 157.5860075)$$

Teraz możemy obliczyć przedziały ufności dla wartości oczekiwanej z poniższego wzoru:

$$\overline{X}_n \mp t_{1-\frac{\alpha}{2};n-1}\frac{S_n}{\sqrt{n}}$$

$t_{1-\frac{\alpha}{2};n-1}$ to kwantyl rozkładu Studenta z $n-1$ stopniami swobody.

$$t_{0.95;4} \overset{R}{=} qt(0.95, 4) \approx 2.131847$$

Wtedy szukany przedział to:

$$(863.9551292; 874.0448708)$$

8 Zadanie 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.

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

8.1 a)

Wyznaczmy α widząc ze szukamy przedział ufności 90%.

$$1-\alpha = 0.9 \Rightarrow \alpha = 0.1$$

Niech każdy pasażer ma narodowość niezależną od innych pasażerów, wtedy każdy pasażer X_i będzie miał rozkład Bernouilliego z nieznanym parametrem p opisując czy jest polakiem czy nie. Zatem $X = \sum X_i$ będzie także rozkładem Bernoulliego i będzie opisywało liczbę polaków z pośród pasażerów.
Aby wyznaczyć przedział ufności dla parametru p sprawdzimy poniższy warunek:

$$1 < \overline{p}_n \mp 3\sqrt{\frac{\overline{p}_n(1-\overline{p}_n)}{n}} < 1$$

$$\overline{p}_n \mp 3\sqrt{\frac{\overline{p}_n(1-\overline{p}_n}{n})} = \frac{201}{347} \mp 3\sqrt{\frac{201/347 \cdot 146/347}{347}} \approx 0.5792507205 \mp 0.079506292$$

Spełniony jest warunek dla obu wartości, zatem możemy wyznaczyć szukany przedział ufności ze wzoru.

$$\overline{P}_n \mp z_{1-\frac{\alpha}{2}}\sqrt{\frac{\overline{P}_n(1-\overline{P}_n)}{n}}$$

$$z_{1-\frac{\alpha}{2}} \overset{R}{=} qnorm(0.95, 0, 1) \approx 1.644854$$

$$\overline{P}_n \mp z_{1-\frac{\alpha}{2}}\sqrt{\frac{\overline{P}_n(1-\overline{P}_n}{n}} = 0.5792507205 \mp 0.0435920808$$

$$\Big(0.5356586397; 0.6228428013\Big)$$

Podany powyżej jest szukany przedział z 90% ufności.

8.2 b)

Aby wygenerować losową próbę wykorzystamy funkcję R-owską dla rozkładu Bernoulliego *rbinom()*. Następne kroki jak w poprzednim przypadku. Oznaczmy próbę jako:

$$\text{prob} \overset{R}{=} rbinom(1, 347, 0.58) = 189$$

Dla takiej liczby sprawdzimy warunek:

$$\overline{p}_n \mp 3\sqrt{\frac{\overline{p}_n(1-\overline{p}_n}{n}} = \frac{189}{347} \mp 3\sqrt{\frac{189/347 \cdot 158/347}{347}} \approx 0.5446685879 \mp 0.0267340794$$

Warunek jest spełniony zatem obliczamy jak poprzednio:

$$\overline{P}_n \mp z_{1-\frac{\alpha}{2}}\sqrt{\frac{\overline{P}_n(1-\overline{P}_n)}{n}} = 0.5446685879 \mp 0.0439736574$$

Wtedy przedział ufności wynosi:

$$\Big(0.5006949305; 0.5886422453\Big)$$

9 Zadanie 6

Frekwencja widzów na seansie filmowym w jednym z kin ma rozkład $N(\mu = ?; \sigma = 30)$. Na podstawie rejestru liczby widzów na 25 losowo wybranych seansach filmowych oszacowano przedział liczbowy (184; 216) dla nieznanej przeciętnej frekwencji na wszystkich seansach.

a) Obliczyć średnią liczbę widzów w badanej próbie.

b) Jaki poziom ufności przyjęto przy estymacji?

9.1 a)

Dla rozkładu normalnego ze znanym parametrem σ przedział ufności dla wartości oczekiwanej jest następujący:

$$\overline{X}_n \mp z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}$$

Gdzie n jest liczebność próby i α jest parametrem ufności.
Można wtedy wyznaczyć szukany parametr $\overline{X}_n$ i α.

$$\begin{cases} 184 = \overline{X}_{25} - z_{1-\frac{\alpha}{2}} \frac{30}{\sqrt{25}} \\ 216 = \overline{X}_{25} + z_{1-\frac{\alpha}{2}} \frac{30}{\sqrt{25}} \end{cases}$$

$$\begin{cases} \overline{X}_{25} = 184 + z_{1-\frac{\alpha}{2}} 6 \\ \overline{X}_{25} = 216 - z_{1-\frac{\alpha}{2}} 6 \end{cases}$$

$$\begin{cases} \overline{X}_{25} = 184 + z_{1-\frac{\alpha}{2}} 6 \\ 184 + z_{1-\frac{\alpha}{2}} 6 = 216 - z_{1-\frac{\alpha}{2}} 6 \end{cases}$$

$$\begin{cases} \overline{X}_{25} = 184 + z_{1-\frac{\alpha}{2}} 6 \\ 12 z_{1-\frac{\alpha}{2}} = 32 \end{cases}$$

$$\begin{cases} \overline{X}_{25} = 184 + z_{1-\frac{\alpha}{2}} 6 \\ 1 - \frac{\alpha}{2} = \Phi(2.66666667) \end{cases}$$

$$\begin{cases} \overline{X}_{25} = 184 + z_{1-\frac{\alpha}{2}} 6 \\ \alpha = 2 - 2\Phi(2.66666667) \stackrel{R}{=} 2 - 2 * pnorm(2.67, 0, 1) \approx 0.007585125 \end{cases}$$

$$\begin{cases} \overline{X}_{25} \stackrel{R}{=} 184 + qnorm(1 - 0.007585125/2, 0, 1) * 6 \approx 200.02 \\ \alpha \approx 0.007585125 \end{cases}$$

Zatem szukana średnia wynosi 200.

9.2 b)

Jak wyliczono w podpunkcie **a** α wynosi około 0.008, zatem procent ufności wynosi 99.2%.

10 Zadanie 17

A random sample of 64 observation from a population produced the following summary statistics: $\sum x_i = 500$, $\sum (x_i - \overline{x})^2 = 3,566$.

a) Find 95% confidence interval for m.

b) Interpret the confidence interval you found in part (a).

11 a)

First we calculate α from the confidence level:

$$1 - \alpha = 0.95 \Rightarrow \alpha = 0.05$$

Because the standard deviation is not given we will use the formula given below to calculate our confidence interval borders:

$$\overline{X}_n \mp t_{1-\frac{\alpha}{2};n-1}\frac{S_n}{\sqrt{n}}$$

For $\overline{X}_n$ we just divide the given sum by the number of observations:

$$\overline{X}_64 = \frac{500}{64} \approx 7.8125$$

For S_n we need to take the square root of the variance which is given by the formula below:

$$S_n^2 = \frac{1}{n-1}\sum(x_i - \overline{x})^2 = \frac{3.566}{63} \approx 0.56444444$$

$$S_n = \sqrt{0.56444444} \approx 0.751295$$

Because $t_{1-\frac{\alpha}{2};n-1}$ is the quantile function of Student-t distribution with $n-1$ degrees of freedom we can calculate it using the R programming language:

$$t_{1-\frac{\alpha}{2};n-1} \overset{R}{=} qt(0.975, 63) \approx 1.998341$$

Finally we can plug in the calculated values to find the confidence interval:

$$\overline{X}_n \mp t_{1-\frac{\alpha}{2};n-1}\frac{S_n}{\sqrt{n}} = 7.8125 \mp 1.998341 \cdot \frac{0.751295}{8} \approx 7.8125 \mp 0.187668$$

$$(7.624832; 8.000168)$$

The above interval is our confidence interval.

11.1 b)

The purpose of the confidence interval is to find an interval where almost for sure we can find our searched parameter. In this case we can say that our searched m parameter is for sure between 7.7 and 8. We can then grab a value from this interval and say that our population is normally distributed with the grabbed parameter m.

12 Zadanie 19

Jak liczna powinna być próba, aby na jej podstawie można było z prawd. 0,99 oszacować średni wzrost noworodków przy maksymalnym błędzie szacunku 1cm? Zakładamy, że rozkład wzrostu noworodków jest rozkładem normalnym z odchyleniem standardowym 2,5cm.

Aby obliczyć minimalną liczebność próby skorzystamy z następującego wzoru:

$$n = \left\lceil \frac{z^2_{1-\frac{\alpha}{2}}\sigma^2}{d^2} \right\rceil$$

Gdzie $z_{1-\frac{\alpha}{2}}$ to kwantyl rozkładu $N(0,1)$, $d = 1[cm]$, $\sigma = 2.5[cm]$ i $1 - \alpha = 0.99 \Rightarrow \alpha = 0.01$.

Obliczymy najpierw kwantyl.

$$z^2_{1-\frac{\alpha}{2}} \stackrel{R}{=} qnorm(0.995, 0, 1)^2 \approx 6.634897$$

Wtedy możemy wyznaczyć n.

$$n = \left\lceil \frac{6.634897 \cdot 6.25}{1} \right\rceil = \lceil 41.468106 \rceil = 42$$

Zatem minimalna liczebność próby aby z prawdopodobieństwem 0.99 oszacować wzrost noworodków przy maksymalnym błędzie szacunku 1 [cm] jest 42.