

# Laboratoria 9

Krystian Baran 145000

27 kwietnia 2021

## Spis treści

<b>1</b>	<b>Zadanie 3</b>	<b>3</b>
<b>2</b>	<b>Zadanie 7</b>	<b>3</b>
<b>3</b>	<b>Zadanie 10</b>	<b>4</b>
3.1	a) . . . . .	6
3.2	b) . . . . .	7
<b>4</b>	<b>Zadanie 13</b>	<b>8</b>
4.1	a) . . . . .	8
4.2	b) . . . . .	9
<b>5</b>	<b>Zadanie 16</b>	<b>9</b>
5.1	a) . . . . .	10
5.2	b) . . . . .	11
5.3	c) . . . . .	11
5.4	d) . . . . .	12

## 1 Zadanie 3

Wytwórnia cukierków paczkuje w torebki po około 200 sztuk mieszanek złożoną z dwóch rodzajów cukierków, przy czym paczkowane są dwa typy mieszanek. Mieszanka typu  $A$  zawiera 40% cukierków pierwszego rodzaju i 60% drugiego rodzaju, natomiast mieszanka typu  $B$  zawiera jednakowe liczby cukierków obydwu rodzajów.

Do weryfikacji hipotezy  $H_0 : p = 40\%$ , że mieszanka jest typu  $A$ , wobec hipotezy alternatywnej  $H_1 : p = 50\%$ , zaproponowano następującą procedurę:

jeśli wśród 5 cukierków wylosowanych z torebki znajdą się więcej niż 3 cukierki pierwszego rodzaju, to odrzuca się hipotezę zerową na rzecz hipotezy alternatywnej. W przeciwnym przypadku przyjmuje się hipotezę zerową.

Przy tak określonej procedurze testowej, znaleźć prawdopodobieństwa błędów obydwu rodzajów oraz moc testu.

Ponieważ znamy ilość cukierków w torebce i procentowość każdego rodzaju cukierków w obu typach paczek, możemy zastosować rozkład hypergeometryczny z parametrami  $n = 200 * 40/100 = 80$ ,  $m = 200 * 60/100 = 120$  dla typu  $A$  i  $n = 100$ ,  $m = 100$  dla typu  $B$ . Nazwijmy je odpowiednio  $X$  i  $Y$ .

Aby sprawdzić błąd pierwszego rodzaju oznaczony jako  $\alpha$  potrzebujemy przedział krytyczny dla typu  $A$ . Skoro podany został typ testu, jeżeli wylosujemy więcej niż 3 cukierki pierwszego rodzaju jest to typ  $B$ , zatem przedział krytyczny dla  $H_0$  jest:

$$R = \{4, 5\}$$

Wtedy, zakładając że hipoteza zerowa jest prawdziwa, czyli korzystamy z rozkładu  $X$ , można obliczyć  $\alpha$  jako:

$$\begin{aligned}\alpha &= P(U_n \in R | H_0) = P(X > 3) = 1 - P(X \leq 3) \\ &\stackrel{R}{=} 1 - \text{phyper}(3, 80, 120, 5) \approx 0.08432931\end{aligned}$$

Aby obliczyć błąd drugiego rodzaju  $\beta$  zakładamy że hipoteza alternatywna jest prawdziwa, zatem korzystamy z rozkładu  $Y$ . Wtedy można obliczyć szukane  $\beta$  z następującego wzoru:

$$\begin{aligned}\beta &= 1 - P(U_n \in R | H_1) = 1 - P(Y > 3) = P(Y \leq 3) \\ &\stackrel{R}{=} \text{phyper}(3, 100, 100, 5) \approx 0.8156646\end{aligned}$$

Moc testu oznacza się jako  $1 - \beta = 1 - 0.8156646 = 0.1843354$ . Zatem moc testu wynosi około 0.184.

## 2 Zadanie 7

Zbadano czułość 80 telewizorów i uzyskano  $\bar{x} = 348[mV]$  i  $s = 107[mV]$ . Na poziomie istotności  $\alpha = 0,05$  zweryfikować hipotezę, że odchylenie standardowe

czułości jest większe od nominalnej wartości wynoszącej 100[mV].

Wyznamy najpierw hipotezę zerową. Ponieważ szukamy aby odchylenie standardowe było większe od pewnej wartości przyjmujemy że to hipoteza alternatywna. Zatem szukane hipotezy będą wyglądać następująco:

$$H_0 : \sigma \leq \sigma_0 = 100[mV]$$

$$H_1 : \sigma > \sigma_0 = 100[mV]$$

Nie jest znany rozkład czułości telewizora i nieznane są także parametry tego rozkładu. Zatem zastosujemy statystykę:

$$Z = \frac{S_n^2 - \sigma_0^2}{\sigma_0^2} \sqrt{\frac{n}{2}}$$

Dla wartości  $\sigma_0 = 100[mV]$ ,  $S_{80} = 107[mV]$  i  $n = 80$ . Zakładając że  $n$  jest wystarczająco duże rozkład ten jest zbliżony do rozkładu  $N(0, 1)$ .

Obliczymy teraz wartość  $Z_0$  która pozwoli nam obliczyć wartość  $p$  value aby sprawdzić prawdziwość hipotezy alternatywnej.

$$Z_0 = \frac{107^2 - 100^2}{100^2} \sqrt{\frac{80}{2}} \approx 0.916428$$

Wtedy można obliczyć  $p$  value następująco:

$$p \text{ value} = 1 - \Phi(Z_0) \stackrel{R}{=} 1 - pnorm(0.916428, 0, 1) \approx 0.1797212$$

Ponieważ wartość  $p$  jest większa od  $\alpha$  nie mamy podstaw żeby odrzucić hipotezę zerową. Zatem odchylenie standardowe może być mniejsze od wartości nominalnej.

### 3 Zadanie 10

Wzrost losowo wybranej osoby z pewnej populacji ma rozkład normalny o nieznanym parametrach. Pobrano próbę losową o liczności  $n = 26$  i po obliczeniu przedziału ufności na poziomie 0,9 otrzymano następujący wynik: (162;178)(cm). Wygenerować próbę złożoną z 26 pomiarów według rozkładu  $N(\bar{x}, s_{26})$  i na poziomie istotności 0,05 zweryfikować hipotezy

- a) średni wzrost ludzi z badanej populacji jest większy od 178 cm.
- b) odchylenie standardowe wzrostu ludzi z badanej populacji jest mniejsze od 24 cm.

Wyznamy najpierw parametr  $\alpha$ . Ponieważ ufność wynosi 0.9 wtedy  $\alpha = 0.1$ .

Następnie za pomocą tabeli na przedział ufności wartości oczekiwanej wyznaczmy średnią i wariancję z próby.

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

$$t_{0.95;25} \stackrel{R}{=} qt(0.95, 25) \approx 1.708141$$

$$\begin{cases} \bar{X}_{26} - t_{0.95;25} \frac{S_{26}}{\sqrt{26}} = 162 \\ \bar{X}_{26} + t_{0.95;25} \frac{S_{26}}{\sqrt{26}} = 178 \end{cases}$$

$$\begin{cases} \bar{X}_{26} = 0.334994 \cdot S_{26} + 162 \\ \bar{X}_{26} = -0.334994 \cdot S_{26} + 178 \end{cases}$$

$$\begin{cases} \bar{X}_{26} = 0.334994 \cdot S_{26} + 162 \\ 0.334994 \cdot S_{26} + 162 = -0.334994 \cdot S_{26} + 178 \end{cases}$$

$$\begin{cases} \bar{X}_{26} = 0.334994 \cdot S_{26} + 162 \\ 0.669988 \cdot S_{26} = 16 \end{cases}$$

$$\begin{cases} \bar{X}_{26} = 0.334994 \cdot S_{26} + 162 \\ S_{26} = 23.881025 \end{cases}$$

$$\begin{cases} \bar{X}_{26} = 170 \\ S_{26} = 23.881025 \end{cases}$$

Próbę losową według rozkładu  $N(\bar{x}, s_{26})$  wygenerowano w R i przedstawiona poniżej; wartości zaokrąglone do 5 liczb bo przecinku.

Lp	Vart
1	206.75631
2	150.65441
3	209.90324
4	135.2722
5	164.51461
6	208.97946
7	205.83026
8	133.27125
9	160.10169
10	151.44174
11	141.06375
12	159.56375
13	185.89217
14	173.9433
15	169.70846
16	168.77696
17	180.32007
18	145.65124
19	232.2268
20	163.71309
21	167.86651
22	191.62654
23	180.19727
24	162.64296
25	158.25598
26	186.41663

$\bar{X}_{26}$  i  $S_{26}^2$  obliczono zgodnie z odpowiednimi wzorami będą wykorzystane do dalszych obliczeń i wynoszą:

$$\bar{X}_{26} = 172.8688712 \approx 173$$

$$S_{26}^2 = 631.9613666$$

$$S_{26} = 25.13884179$$

### 3.1 a)

Zakładamy że średni wzrost populacji jest wartością  $m$ . Wtedy hipoteza że  $m > 178$  jest hipotezą alternatywną i, zgodnie z normami statystyki można wyznaczyć hipotezę zerową.

$H_0$	$m \leq 178$
$H_1$	$m > 178$

Ponieważ znamy rozkład ale nie znamy jego parametrów wyznaczmy sta-

tystykę zgodnie z tabelami.

$$t = \frac{\bar{X}_n - m_0}{\frac{S_n}{\sqrt{n}}}$$

Obliczymy teraz  $t_0$  podstawiając  $m_0$  z hipotezy i wartości obliczone w wygenerowanej próbie.

$$t_0 = \frac{173 - 178}{\frac{25.13884179}{\sqrt{26}}} \approx -1.014172$$

Statystyka ta ma rozkład statystyczny zbliżony do rozkładu t-Studenta z  $n-1$  stopniami swobody. Można teraz obliczyć przedział krytyczny dla  $\alpha = 0.05$ .

$$t_{1-0.05;25} \stackrel{R}{=} qt(0.95, 25) \approx 1.708141$$

$$(1.708141, \infty)$$

Wartość  $t_0$  nie należy do przedziału krytycznego, zatem nie możemy odrzucić hipotezę zerową; zatem nie wiem czy hipoteza alternatywna jest prawdziwa lub nie.

### 3.2 b)

Hipoteza że odchylenie standardowe jest mniejsze od 24 jest hipoteza alternatywna. Zatem jak poprzednio wyznaczmy hipotezę zerową.

$H_0$	$\sigma \geq 24$
$H_1$	$\sigma < 24$

Przeprowadzimy natomiast test dla wariancji i z tego testu wywnioskujemy hipotezę dla odchylenia standardowego

Ponieważ znamy rozkład ale nie znamy jego parametrów, zgodnie z tabelami skorzystamy z następującej statystyki:

$$\chi^2 = \frac{(n-1)S_n^2}{\sigma_0^2}$$

Statystyka ta ma w przybliżeniu rozkład statystyki chi kwadrat z  $n-1$  stopniami swobody.

Obliczymy teraz  $\chi_0^2$  podstawiając odpowiednie wartości.

$$\chi_0^2 = \frac{25 \cdot 631.9613666}{24^2} \approx 27.428879$$

Zgodnie z tabelami wyznaczmy przedział krytyczny.

$$\chi_{\alpha;n-1}^2 \stackrel{R}{=} qchisq(0.05, 25) \approx 14.61141$$

$$(0, 14.61141)$$

Ponownie wartość  $\chi_0^2$  nie należy do przedziału krytycznego zatem nie możemy odrzucić hipotezę zerową. Nie wiemy po za tym czy jest ona prawdziwa czy nie.

## 4 Zadanie 13

Czuły przyrząd pomiarowy powinien mieć niewielką wariancję błędów pomiaru. W próbie 41 błędów pomiaru stwierdzono wariancję  $102 \text{ [j.m.]}^2$ . Na poziomie istotności  $\alpha_1 = 0,05$  i  $\alpha_2 = 0,01$  zweryfikować hipotezy:

- a) wariancja błędów pomiaru wynosi  $120 \text{ [j.m.]}^2$ ;
- b) wariancja błędów pomiaru wynosi poniżej  $120 \text{ [j.m.]}^2$ .

### 4.1 a)

Hipoteza że wariancja błędów pomiaru wynosi  $120 \text{ [j.m.]}^2$  jest hipotezą zerową, zatem można wyznaczyć hipotezę alternatywną jako jej przeciwieństwo.

$H_0$	$\sigma^2 = 120 \text{ [j.m.]}^2$
$H_1$	$\sigma^2 \neq 120 \text{ [j.m.]}^2$

Ponieważ nie znamy rozkład błędów pomiaru ani ich parametrów skorzystamy z następującej statystyki:

$$Z = \frac{S_n^2 - \sigma_0^2}{\sigma_0^2} \sqrt{\frac{n}{2}}$$

Która ma w przybliżeniu rozkład statystyki  $N(0, 1)$ .

Obliczymy teraz  $Z_0$  potrzebne do dalszych rozważań postawiając znaną wariancję z próby i  $\sigma_0^2$ .

$$Z_0 = \frac{102 - 120}{120} \sqrt{\frac{41}{2}} \approx -4.482416$$

Wyznamy teraz obszary krytyczne zgodnie z tabelami.

Dla  $\alpha_1$ :

$$\begin{aligned} z_{\frac{0.05}{2}} &\stackrel{R}{=} qnorm(0.05/2, 0, 1) \approx -1.959964 \\ z_{1-\frac{0.05}{2}} &\stackrel{R}{=} qnorm(1 - 0.05/2, 0, 1) \approx 1.959964 \\ &(-\infty, -1.959964) \cup (1.959964, \infty) \end{aligned}$$

Dla  $\alpha_2$ :

$$\begin{aligned} z_{\frac{0.01}{2}} &\stackrel{R}{=} qnorm(0.01/2, 0, 1) \approx -2.575829 \\ z_{1-\frac{0.01}{2}} &\stackrel{R}{=} qnorm(1 - 0.01/2, 0, 1) \approx 2.575829 \\ &(-\infty, -2.575829) \cup (2.575829, \infty) \end{aligned}$$

Wartość  $Z_0$  należy do obszary krytycznego dla oby  $\alpha$ , zatem odrzucamy hipotezę zerową i przyjmujemy hipotezę alternatywną czyli  $\sigma^2 = 120 \text{ [j.m.]}^2$ .



## 4.2 b)

Hipotezą że  $\sigma^2 > 120[\text{j.m.}]^2$  wariancja błędów jest hipotezą alternatywną, zatem, zgodnie z normami statystyki wyznaczamy hipotezę zerową.

$H_0$	$\sigma^2 \leq 120[\text{j.m.}]^2$
$H_1$	$\sigma^2 > 120[\text{j.m.}]^2$

Statystyka dla tego podpunktu jest taka sama jak w poprzednim podpunkcie, natomiast zmienia się obszar krytyczny.

Zgodnie z tabelami, dla wartości  $\alpha_1$ :

$$z_{1-0.05} \stackrel{R}{=} qnorm(0.95, 0, 1) \approx 1.644854$$

$$(1.281552, \infty)$$

Dla  $\alpha_2$ :

$$z_{1-0.01} \stackrel{R}{=} qnorm(0.99, 0, 1) \approx 2.326348$$

$$(1.281552, \infty)$$

Wartość  $Z_0$  nie należy do obszaru krytycznego, zatem nie możemy odrzucić hipotezy zerowej. Obliczymy zatem wartość  $p$  value jako:

$$p \text{ value} = 1 - \Phi(Z_0) = 1 - \Phi(-4.482416) \stackrel{R}{=} 1 - pnorm(-4.482416) \approx 0.9999963$$

Ponieważ jest to wartość większa od obu  $\alpha$  nie możemy odrzucić hipotezę zerową.

## 5 Zadanie 16

Dla wylosowanej próby studentów otrzymano następujący rozkład tygodniowego czasu nauki (w godz.):

Czas nauki	[0, 2)	[2, 4)	[4, 6)	[6, 8)	[8, 10)	[10, 12)
Liczba studentów	10	28	42	30	15	7

Na poziomach istotności  $\alpha_1 = 0,1$  i  $\alpha_2 = 0,01$  sprawdzić hipotezy:

- średni czas poświęcony tygodniowo na naukę dla badanej populacji studentów wynosi 6 godz.
- średni czas poświęcony tygodniowo na naukę dla badanej populacji studentów wynosi poniżej 6 godz.;
- wariancja tego czasu wynosi 4godz.<sup>2</sup>;
- wariancja tego czasu wynosi ponad 4godz.<sup>2</sup>.

Jako pierwsze obliczono średnią z podanej próby i wariancję korzystając z następujących wzorów i zakładając środek przedziały jako przedstawiciel przedziału:

$$\begin{aligned}\bar{X} &= \frac{\sum x_i \cdot n_i}{N} = \frac{726}{132} = 5.5 \\ S_n^2 &= \frac{\sum (x_i - \bar{X})^2 \cdot n_i}{N} = \frac{851}{132} \approx 6.496183 \\ S_n &= \sqrt{S_n^2} \approx 2.548761\end{aligned}$$

### 5.1 a)

Ponieważ nie znany jest rozkład zmiennej losowej opisującej czas poświęcony tygodniowo na naukę przez studenta, ani nie są znane jego parametry zastosujemy statystykę następującą:

$$Z = \frac{\bar{X}_n - m_0}{\frac{S_n}{\sqrt{n}}}$$

Która ma rozkład statystyki zbliżony do rozkładu normalnego  $N(0, 1)$ . Podane hipotezy są następujące:

	$H_0$	$H_1$
$\alpha_1$	$m = 6$	$m \neq 6$
$\alpha_2$	$m = 6$	$m \neq 6$

Obliczymy teraz  $Z_0$  podstawiając wartości hipotezy do statystyki:

$$Z_0 = \frac{5.5 - 6}{\frac{2.548761}{\sqrt{132}}} \approx -2.253866$$

Wyznamy obszar krytyczny zgodnie z tabelami. Dla  $\alpha_1$ .

$$z_{\frac{0.1}{2}} \stackrel{R}{=} qnorm(0.1/2, 0, 1) \approx -1.644854$$

$$\begin{aligned}z_{1-\frac{0.1}{2}} &\stackrel{R}{=} qnorm(1 - 0.1/2, 0, 1) \approx 1.644854 \\ &(-\infty, -1.644854) \cup (1.644854, \infty)\end{aligned}$$

Dla  $\alpha_2$ .

$$z_{\frac{0.01}{2}} \stackrel{R}{=} qnorm(0.01/2, 0, 1) \approx -2.575829$$

$$\begin{aligned}z_{1-\frac{0.01}{2}} &\stackrel{R}{=} qnorm(1 - 0.01/2, 0, 1) \approx 2.575829 \\ &(-\infty, -2.575829) \cup (2.575829, \infty)\end{aligned}$$

Ponieważ wartość  $Z_0$  należy do obszaru krytycznego dla  $\alpha_1$ , odrzucamy hipotezę zerową; natomiast dla  $\alpha_2$  wartość  $Z_0$  nie wpada pod obszar krytyczny, zatem nie mamy mocy aby odrzucić hipotezę zerową.

## 5.2 b)

Obliczenia przechodzą jak w poprzednim podpunkcie ale zmieniają się hipotezy. Hipoteza że  $m < 6$  godzin jest hipotezą alternatywną, zatem:

	$H_0$	$H_1$
$\alpha_1$	$m \geq 6$	$m < 6$
$\alpha_2$	$m \geq 6$	$m < 6$

Wartość  $Z_0$  pozostaje nie zmieniona, natomiast zmienia się obszar krytyczny który zgodnie z tabelami obliczymy.

Dla  $\alpha_1$ :

$$z_{0.1} \stackrel{R}{=} qnorm(0.1, 0, 1) \approx -1.281552$$

$$(-\infty, -1.644854)$$

Dla  $\alpha_2$ :

$$z_{0.01} \stackrel{R}{=} qnorm(0.01, 0, 1) \approx -2.326348$$

$$(-\infty, -2.326348)$$

Jak poprzednio, dla  $\alpha_1$  odrzucamy hipotezę zerową, natomiast dla  $\alpha_2$  nie mamy mocy aby to zrobić.

## 5.3 c)

Hipoteza że  $\sigma^2 = 4$  jest hipotezą zerową, zatem można wyznaczyć hipotezę alternatywną:

$H_0$	$\sigma^2 = 4$
$H_1$	$\sigma^2 \neq 4$

Dla sprawdzenia wariancji, ponieważ nie znamy rozkładu ani jego parametrów a mamy wystarczająco dużą próbę skorzystamy ze statystyki:

$$Z = \frac{S_n^2 - \sigma_0^2}{\sigma_0^2} \sqrt{\frac{n}{2}}$$

Która ma rozkład statystyki zbliżony do rozkładu  $N(0, 1)$ .

Obliczymy teraz  $Z_0$  podstawiając  $\sigma_0^2 = 4$  i wartości wcześniej obliczone.

$$Z_0 = \frac{6.496183 - 4}{4} \sqrt{\frac{132}{2}} \approx 5.069772$$

Ponieważ obszary krytyczne są takie same jak w poprzednich podpunktach wnioskujemy że, skoro wartość  $Z_0$  dla obu  $\alpha$  należy do obszaru krytycznego, odrzucamy hipotezę zerową i przyjmujemy hipotezę alternatywną  $\sigma^2 \neq 4$ .

## 5.4 d)

Hipoteza że  $\sigma^2 > 4$  jest hipotezą alternatywną, zatem można wyznaczyć hipotezę zerową korzystając z praw statystyki:

$H_0$	$\sigma^2 \leq 4$
$H_1$	$\sigma^2 > 4$

Statystyka i wartość  $Z_0$  są takie same jak poprzednim podpunkcie, natomiast zmienia się obszar krytyczny który obliczymy zgodnie z tabelami.

Dla  $\alpha_1$ :

$$z_{1-0.1} \stackrel{R}{=} qnorm(0.9, 0, 1) \approx 1.281552$$
$$(1.281552, \infty)$$

Dla  $\alpha_2$ :

$$z_{1-0.01} \stackrel{R}{=} qnorm(0.99, 0, 1) \approx 2.326348$$
$$(2.326348, \infty)$$

Możemy zatem powiedzieć, skoro wartość  $Z_0$  należy do obszarów krytycznych, że odrzucamy hipotezę zerową i przyjmujemy hipotezę alternatywną że  $\sigma^2 > 4$ .