Analiza danych ankietowych

Sprawozdanie 3

Joanna Kusy Tomasz Srebniak

Spis treści

1 Część I oraz II

1.1 Zadanie 1

Napisano funkcję, która zwraca p–wartość w warunkowym teście symetrii w przypadku tabeli 2x2.

```
symmetry_test <- function(t) {
  n12 <- t[1, 2]
  n21 <- t[2, 1]

  if (n12 < (n12+n21)/2) {
    s <- 0
    for (i in 0:n12) {
      s <- s + choose(n12+n21, i) * 0.5^i * 0.5^(n12+n21-i)
    }
    p_value <- 2 * s
  }
  if (n12 > (n12+n21)/2) {
    s <- 0
    for (i in n12:(n12+n21)) {
      s <- s + choose(n12+n21, i) * 0.5^i * 0.5^(n12+n21-i)
    }
    p_value <- 2 * s
  }
  else {
    p_value <- 1
  }
  print(paste("p-value: ", p_value))
  }
```

1.2 Zadanie 2

W poniższej tabeli umieszczono dane dotyczące reakcji na lek po godzinie od jego przyjęcia dla dwóch różnych leków przeciwbólowych stosowanych w migrenie. Leki zostały zaaplikowane grupie pacjentów w dwóch różnych atakach bólowych. Na podstawie danych zweryfikowano hipotezę, że leki te są jednakowo skuteczne.

```
##                  Reakcja na lek B
## Reakcja na lek A Negatywna Pozytywna
##        Negatywna         1         5
##        Pozytywna         2         4
```

1.2.1 Test McNemara z poprawką na ciągłość

```
mcnemar.test(t)
```

```
##
##  McNemar's Chi-squared test with continuity correction
##
## data:  t
## McNemar's chi-squared = 0.57143, df = 1, p-value = 0.4497
```

Na poziomie istotności $\alpha = 0.05$ nie mamy podstaw do odrzucenia hipotezy zerowej o jednorodności rozkładów brzegowych. Zakładamy, że leki są jednakowo skuteczne.

1.2.2 Warunkowy test symetrii

```r
symmetry_test(t)
```

```
## [1] "p-value:  0.453125"
```

Na poziomie istotności $\alpha = 0.05$ nie mamy podstaw do odrzucenia hipotezy zerowej o jednorodności rozkładów brzegowych. Zakładamy, że leki są jednakowo skuteczne.

1.3 Zadanie 3

Przeprowadzono symulacje w celu porównania mocy testu Z i testu Z_0.

Implementacja testu Z oraz testu Z_0.

```r
z_test <- function(t) {
  n <- sum(t)
  n12 <- t[1, 2]
  n21 <- t[2, 1]
  D <- (n12 - n21) / n

  p <- prop.table(t)
  p1_plus <- rowSums(p)[1]
  p_plus1 <- colSums(p)[1]

  s <- (
    p1_plus * (1 - p1_plus) + p_plus1 * (1 - p_plus1) - 2
    * (p[1, 1] * p[2, 2] + p[1, 2] * p[2, 1])
    ) / n

  z <- D / sqrt(s)

  p_value <- 2 * (1 - pnorm(abs(z)))
  p_value |> as.numeric()
}

z0_test <- function(t) {
  z0 <- (t[1, 2] - t[2, 1]) / sqrt(t[1, 2] + t[2, 1])
```

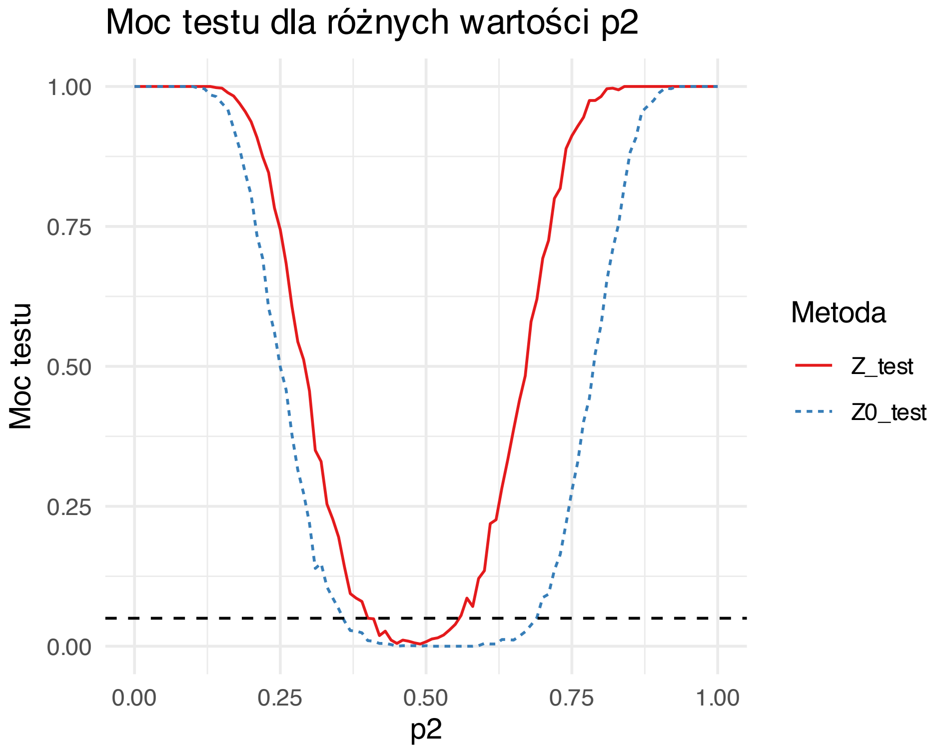
```
  p_value <- 2 * (1 - pnorm(abs(z0)))
  p_value
}
```

Funkcja szacująca moc testu Z oraz testu Z_0 w różnych punktach hipotezy alternatywnej przy pomocy metody Monte Carlo.

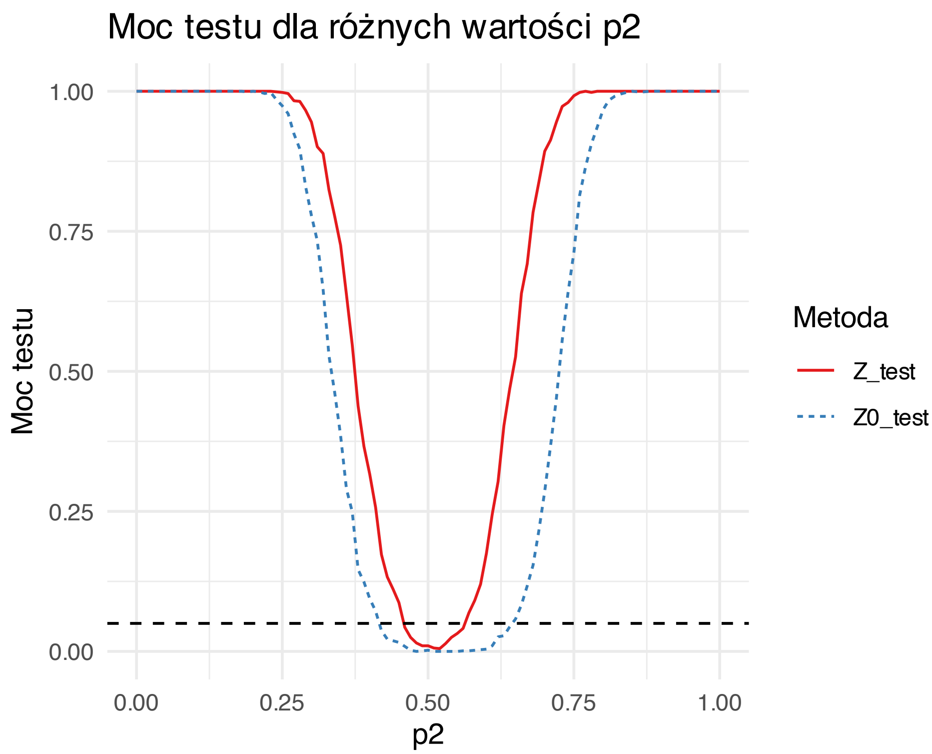
```
power <- function(n, MC = 1000) {

  p1 <- 0.5
  x <- sum(rbinom(n, 1, p1))
  X <- c(x, n-x)

  plot_data <- data.frame()
  for (p2 in seq(0, 1, 0.01)) {
    MC_z <- 0
    MC_z0 <- 0
    for (i in 1:MC) {
      y <- sum(rbinom(n, 1, p2))
      Y <- c(y, n-y)

      t <- rbind(X,Y)
      MC_z <- MC_z + (z_test(t) < 0.05)
      MC_z0 <- MC_z0 + (z0_test(t) < 0.05)
    }
    MC_z <- MC_z / MC
    MC_z0 <- MC_z0 / MC

    plot_data <- rbind(
      plot_data, cbind(data.frame(p2=p2),
                       data.frame(method = c('Z_test', 'Z0_test'),
                       pow = c(MC_z, MC_z0)))
      )
  }

  split_dfs <- split(plot_data, plot_data$method)
  z_power <- split_dfs[["Z_test"]][, "pow"]
  z0_power <- split_dfs[["Z0_test"]][, "pow"]
  ggplot(plot_data, aes(x=p2, y=pow, color=method, linetype=method)) +
    geom_line() +
    geom_hline(yintercept = 0.05, linetype="dashed") +
    scale_color_brewer(palette = "Set1", name = "Metoda") +
    scale_linetype_discrete(name = "Metoda") +
    labs(title = "Moc testu dla różnych wartości p2",
         x = "p2",
```

```
        y = "Moc testu") +
    theme_minimal()
}
```

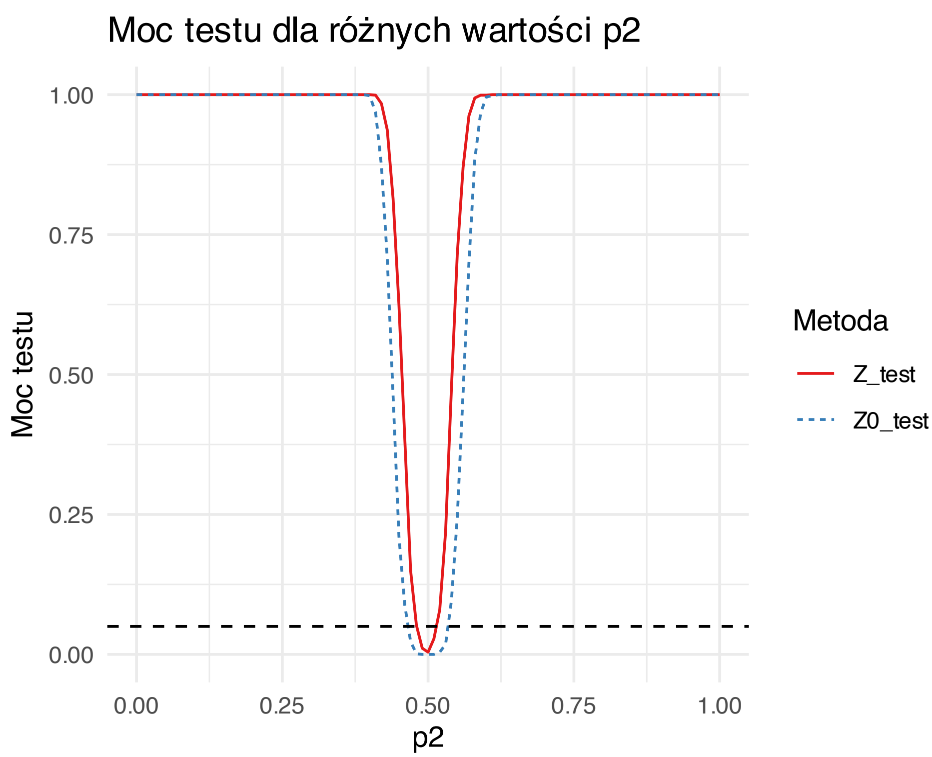
Wyniki symulacji dla $n = 50$.



Wyniki symulacji dla $n = 100$.



Wyniki symulacji dla $n = 1000$.



Na podstawie przeprowadzonych symulacji można zauważyć, moc obu testów wzrasta wraz ze wzrostem rozmiaru próby. Dodatkowo moc testu Z jest wyższa niż moc testu Z_0 dla wszystkich wartości p_2 oraz dla wszystkich rozmiarów próby.

Test Z częściej popełniał błąd I rodzaju, za to rzadziej błąd II rodzaju. W przypadku testu Z_0 było na odwrót.

1.4 Zadanie 4

Dla danych dołączonych do pierwszej listy zadań, na podstawie zmiennych **CZY_ZADW** oraz **CZY_ZADW_2**, zweryfikowano hipotezę, że zadowolenie ze szkoleń w pierwszym badanym okresie i w drugim badanym okresie pierwszego badania odpowiada modelowi symetrii. Przyjęto poziom istotności 0.05.

```
t <- table(ankieta$CZY_ZADOW, ankieta$CZY_ZADOW_2)
mcnemar.test(t)
```

```
##
##  McNemar's Chi-squared test with continuity correction
##
## data:  t
## McNemar's chi-squared = 4.3214, df = 1, p-value = 0.03764
```

Na poziomie istotności $\alpha = 0.05$ odrzucamy hipotezę zerową na rzecz hipotezy alternatywnej. Zakładamy, że zadowolenie ze szkoleń w pierwszym badanym okresie i w drugim badanym okresie nie odpowiada modelowi symetrii. Na podstawie uzyskanych wyników możemy wnioskować, że poziom zadowolenia ze szkoleń uległ zmianie.

1.5 Zadanie 5

W firmie, o której mowa w zadaniu 1 z listy 1, wdrożono pewne działania w celu poprawy komfortu pracy. Następnie badaną grupę respondentów ponownie poproszono o odpowiedź na pytanie dotyczące oceny podejścia firmy do umożliwiania wdrażania wiedzy zdobytej na szkoleniach. W poniższej tabeli przedstawiono tablicę dwudzielczą uwzględniającą odpowiedzi na pytanie w obu tych okresach. Na podstawie danych zweryfikowano hipotezę, że odpowiedzi w pierwszym badanym okresie i w drugim okresie odpowiadają modelowi symetrii.

```
##           Pytanie 2
## Pytanie 1 -2 -1  0  1  2
##        -2 10  2  1  1  0
##        -1  0 15  1  1  0
##        0   1  1 32  6  0
##        1   0  0  1 96  3
##        2   1  1  0  1 26
```

1.5.1 Test Bowkera

```
mcnemar.test(t)
```

```
##
##  McNemar's Chi-squared test
```

```
##
## data:  t
## McNemar's chi-squared = NaN, df = 10, p-value = NA
```

Test Bowkera zwrócił p–value równe NA, ponieważ w mianowniku pojawiła się wartość 0.

1.5.2 Test oparty na ilorazie wiarygodności (IW)

```
IW_test <- function(t) {
  G2 <- 0
  n <- sum(t)
  for (i in 1:nrow(t)) {
    for (j in 1:ncol(t)) {
      if (t[i, j] > 0) {
        G2 <- G2 + t[i, j] * log(t[i, j] / (t[i, j] + t[j, i]) * 2)
      }
    }
  }
  p_value <- 1 - pchisq(2 * G2, df = nrow(t) * (ncol(t) - 1) / 2)
  print(paste("p-value: ", p_value))
}
IW_test(t)
```

```
## [1] "p-value:  0.20597516357247"
```

Na poziomie istotności $\alpha = 0.05$ nie mamy podstaw do odrzucenia hipotezy zerowej o jednorodności rozkładów brzegowych. Zakładamy, że odpowiedzi w pierwszym badanym okresie i w drugim okresie odpowiadają modelowi symetrii. Na podstawie uzyskanych wyników możemy przyjąć hipotezę, że ocena podejścia firmy nie uległa zmianie.

2 Część III

2.1 Zadanie 6

W pewnym badaniu porównywano skuteczność dwóch metod leczenia – Leczenie A to nowa procedura, a Leczenie B to stara procedura. Przeanalizowano dane przedstawione w poniższych tabelach i oceniono, czy dla danych występuje paradoks Simpsona.

2.1.1 Dane dla całej grupy

```
##              Wynik leczenia
## Metoda        Poprawa Brak
##   Leczenie A      117  104
##   Leczenie B      177   44
```

P(poprawa|leczenie A) = 0.53

P(poprawa|leczenie B) = 0.8

Test chi-kwadrat dla całej grupy

```r
chisq.test(t3)
```

```
##
##  Pearson's Chi-squared test with Yates' continuity correction
##
## data:  t3
## X-squared = 35.36, df = 1, p-value = 2.74e-09
```

Na poziomie istotności $\alpha = 0.05$ odrzucamy hipotezę zerową na rzecz hipotezy alternatywnej. Zakładamy, że zmienne są zależne.

2.1.2 Dane dla pacjentów z chorobami współistniejącymi

```
##              Reakcja
## Metoda        Poprawa Brak
##   Leczenie A      17  101
##   Leczenie B       2   36
```

P(poprawa|leczenie A, choroby współ)= 0.14

P(poprawa|leczenie B, choroby współ)= 0.05

Test chi-kwadrat dla pacjentów z chorobami współistniejącymi

```r
chisq.test(t4)
```

```
##
##  Pearson's Chi-squared test with Yates' continuity correction
##
## data:  t4
## X-squared = 1.4732, df = 1, p-value = 0.2248
```

Na poziomie istotności $\alpha = 0.05$ nie mamy podstaw do odrzucenia hipotezy zerowej. Zakładamy, że zmienne są niezależne.

2.1.3 Dane dla pacjentów bez chorób współistniejących

```
##              Reakcja
## Metoda        Poprawa Brak
##   Leczenie A     100    3
##   Leczenie B     175    8
```

P(poprawa|leczenie A, brak chorób współ)= 0.97

P(poprawa|leczenie B, brak chorób współ)= 0.96

Test chi-kwadrat dla pacjentów bez chorób współistniejących

```r
chisq.test(t5)
```

```
##
##  Pearson's Chi-squared test with Yates' continuity correction
##
## data:  t5
## X-squared = 0.087398, df = 1, p-value = 0.7675
```

Na poziomie istotności $\alpha = 0.05$ nie mamy podstaw do odrzucenia hipotezy zerowej. Zakładamy, że zmienne są niezależne.

2.1.4 Wniosek

Na podstawie przeprowadzonych testów chi-kwadrat możemy zauważyć, że zależność pomiędzy metodą leczenia a wynikiem leczenia występuje dla całej grupy pacjentów. Natomiast dla pacjentów z chorobami współistniejącymi oraz bez chorób współistniejących taka zależność już nie występuje. Mamy zatem do czynienia z paradoksem Simpsona, który polega na zmianie "kierunku" związku między dwiema zmiennymi, przy uwzględnieniu trzeciej zmiennej.

2.2 Zadanie 7

Dla danych z listy 1, przyjmując za zmienną 1 zmienną **CZY_KIER**, za zmienną 2 – zmienną **PYT_2** i za zmienną 3 – zmienna **STAŻ**, podano interpretacje nastepujacych modeli log-liniowych:

2.2.1 [1 3]

Zmienne **CZY_KIER** i **STAŻ** są niezależne i **PYT_2** ma rozkład równomierny.

2.2.2 [13]

Zmienne **CZY_KIER** i **PYT_2** nie są niezależne i **STAŻ** ma rozkład równomierny.

2.2.3 [1 2 3]

Zmienne **CZY_KIER**, **PYT_2** i **STAŻ** są niezależne.

2.2.4 [12 3]

Zmienne **CZY_KIER** i **PYT_2** nie są niezależne, a zmienna **STAŻ** jest niezależna od zmiennych **CZY_KIER** i **PYT_2**.

2.2.5 [12 13]

Przy ustalonej wartości zmiennej **CZY_KIER** zmienne **PYT_2** i **STAŻ** są niezależne. Inaczej mówiąc, zmienne **PYT_2** i **STAŻ** są warunkowo niezależne.

2.2.6 [1 23]

Zmienna **CZY_KIER** jest niezależna od zmiennych **PYT_2** i **STAŻ**. Zmienne **PYT_2** i **STAŻ** nie są niezależne.

3 Część IV i V

3.1 Zadanie 8

Rozważono dwa modele log-liniowy [123] i [12 23] dla zmiennych opisanych w zadaniu 7 oszacowano prawdopobiebieństwa:

- że osoba pracująca na stanowisku kierowniczym jest zdecydowanie zadowolona ze szkoleń;
- że osoba o stażu pracy krótszym niż rok pracuje na stanowisku kierowniczym;
- że osoba o stażu pracy powyżej trzech lat nie pracuje na stanowisku kierowniczym.

Tabela z danymi

```
tab1 <- ankieta |> as_tibble()  %>%
  group_by(CZY_KIER, PYT_2, STAŻ) %>%
  summarise(count = n(), .groups = 'drop')
tab1 |> kable()
```

CZY_KIER	PYT_2	STAŻ	count
Nie	-2	1	19
Nie	-2	2	40
Nie	-2	3	5
Nie	-1	1	3
Nie	-1	2	15
Nie	2	1	18
Nie	2	2	68
Nie	2	3	5
Tak	-2	1	1
Tak	-2	2	5
Tak	-2	3	4
Tak	-1	2	2
Tak	1	3	2
Tak	2	2	10
Tak	2	3	3

3.1.1 Model [123]

Zdefiniowanie i dopasowanie modelu

```r
model <- glm(count ~ CZY_KIER * PYT_2 * STAŻ,
    data = tab1,
    family = poisson)

tab1$fitted1 <- fitted(model)
```

- osoba pracująca na stanowisku kierowniczym jest zdecydowanie zadowolona ze szkoleń

```r
subset(tab1, CZY_KIER == 'Tak') |> group_by(PYT_2) %>%
  summarise(p_dane = sum(count) / sum(subset(tab1, CZY_KIER == 'Tak')$count),
            p_model = sum(fitted1) / sum(subset(tab1, CZY_KIER == 'Tak')$fitted1),
            .groups = 'drop')
```

PYT_2	p_dane	p_model
-2	0.3703704	0.3177332
-1	0.0740741	0.1344622
1	0.0740741	0.1034585
2	0.4814815	0.4443461

Jako, że interesuje nas oszacowanie prawdopodobieństwa, że osoba pracująca na stanowisku kierowniczym jest zdecydowanie zadowolona ze szkoleń, wystarczy że spojrzymy na ostatni wiersz tabeli, gdzie **PYT_2** = 2. Liczność wyliczona z tabeli oraz wartość przewidywana przez model różnią się o około 0.0371, przy czym liczność uzyskana z tabeli była wyższa.

- osoba o stażu pracy krótszym niż rok pracuje na stanowisku kierowniczym

```r
subset(tab1, STAŻ == 1) |> group_by(CZY_KIER) %>%
  summarise(p_dane = sum(count) / sum(subset(tab1, STAŻ == 1)$count),
            p_model = sum(fitted1) / sum(subset(tab1, STAŻ == 1)$fitted1),
            .groups = 'drop')
```

CZY_KIER	p_dane	p_model
Nie	0.9756098	0.9779567
Tak	0.0243902	0.0220433

Jako, że interesuje nas oszacowanie prawdopodobieństwa, że osoba o stażu pracy krótszym niż rok pracuje na stanowisku kierowniczym, wystarczy że spojrzymy na ostatni wiersz tabeli, gdzie **CZY_KIER** = Tak. Liczność wyliczona z tabeli oraz wartość przewidywana przez model różnią się o około 0.0023, przy czym liczność uzyskana z tabeli była wyższa.

- osoba o stażu pracy powyżej trzech lat nie pracuje na stanowisku kierowniczym

```r
subset(tab1, STAŻ == 3) |> group_by(CZY_KIER) %>%
  summarise(p_dane = sum(count) / sum(subset(tab1, STAŻ == 3)$count),
```

```r
            p_model = sum(fitted1) / sum(subset(tab1, STAŻ == 3)$fitted1),
            .groups = 'drop')
```

CZY_KIER	p_dane	p_model
Nie	0.5263158	0.8064535
Tak	0.4736842	0.1935465

Jako, że interesuje nas oszacowanie prawdopodobieństwa, że osoba o stażu pracy dłuższym niż trzy lata nie pracuje na stanowisku kierowniczym, wystarczy że spojrzymy na pierwszy wiersz tabeli, gdzie **CZY_KIER** = Nie. Liczność wyliczona z tabeli oraz wartość przewidywana przez model różnią się o około 0.2801, przy czym wartość przewidywana przez model była wyższa.

3.1.2 Model [12 23]

Zdefiniowanie i dopasowanie modelu

```r
model <- glm(count ~ CZY_KIER*PYT_2 + PYT_2*STAŻ,
    data = tab1,
    family = poisson)

tab1$fitted2 <- fitted(model)
```

- osoba pracująca na stanowisku kierowniczym jest zdecydowanie zadowolona ze szkoleń

```r
subset(tab1, CZY_KIER == 'Tak') |> group_by(PYT_2) %>%
  summarise(p_dane = sum(count) / sum(subset(tab1, CZY_KIER == 'Tak')$count),
            p_model = sum(fitted2) / sum(subset(tab1, CZY_KIER == 'Tak')$fitted2),
            .groups = 'drop')
```

PYT_2	p_dane	p_model
-2	0.3703704	0.3106656
-1	0.0740741	0.1257078
1	0.0740741	0.1579920
2	0.4814815	0.4056346

Jako, że interesuje nas oszacowanie prawdopodobieństwa, że osoba pracująca na stanowisku kierowniczym jest zdecydowanie zadowolona ze szkoleń, wystarczy że spojrzymy na ostatni wiersz tabeli, gdzie **PYT_2** = 2. Liczność wyliczona z tabeli oraz wartość przewidywana przez model różnią się o około 0.0758, przy czym liczność uzyskana z tabeli była wyższa.

- osoba o stażu pracy krótszym niż rok pracuje na stanowisku kierowniczym

```r
subset(tab1, STAŻ == 1) |> group_by(CZY_KIER) %>%
  summarise(p_dane = sum(count) / sum(subset(tab1, STAŻ == 1)$count),
```

```
              p_model = sum(fitted2) / sum(subset(tab1, STAŻ == 1)$fitted2),
              .groups = 'drop')
```

CZY_KIER	p_dane	p_model
Nie	0.9756098	0.9626082
Tak	0.0243902	0.0373918

Jako, że interesuje nas oszacowanie prawdopodobieństwa, że osoba o stażu pracy krótszym niż rok pracuje na stanowisku kierowniczym, wystarczy że spojrzymy na ostatni wiersz tabeli, gdzie **CZY_KIER** = Tak. Liczność wyliczona z tabeli oraz wartość przewidywana przez model różnią się o około 0.013, przy czym wartość przewidywana przez model była wyższa.

- osoba o stażu pracy powyżej trzech lat nie pracuje na stanowisku kierowniczym

```
subset(tab1, STAŻ == 3) |> group_by(CZY_KIER) %>%
  summarise(p_dane = sum(count) / sum(subset(tab1, STAŻ == 3)$count),
              p_model = sum(fitted2) / sum(subset(tab1, STAŻ == 3)$fitted2),
              .groups = 'drop')
```

CZY_KIER	p_dane	p_model
Nie	0.5263158	0.7660311
Tak	0.4736842	0.2339689

Jako, że interesuje nas oszacowanie prawdopodobieństwa, że osoba o stażu pracy dłuższym niż trzy lata nie pracuje na stanowisku kierowniczym, wystarczy że spojrzymy na pierwszy wiersz tabeli, gdzie **CZY_KIER** = Nie. Liczność wyliczona z tabeli oraz wartość przewidywana przez model różnią się o około 0.2397, przy czym wartość przewidywana przez model była wyższa.

3.2 Zadanie 9

Dla danych wskazanych w zadaniu 7 zweryfikowano następujące hipotezy:

3.2.1 Zmienne losowe CZY_KIER, PYT_2 i STAŻ sa wzajemnie niezależne

$M_0 = [1\ 2\ 3]$

$H_0 : M = M_0$

```
M0 <- glm(count ~ CZY_KIER + PYT_2 + STAŻ,
    data = tab1,
    family = poisson)
```

3.2.1.1 Nadmodel [12 23]

$M_1 = [12\ 23]$

$H_1 : M = M_1$ i $M \neq [1\ 2\ 3]$

```
M1 <- glm(count ~ CZY_KIER * PYT_2 + PYT_2 * STAŻ,
    data = tab1,
    family = poisson)
```

```
an <- anova(M0, M1)
deviance <- an$Deviance[2]
df <- an$Df[2]
p_value <- 1 - pchisq(deviance, df)
```

P–value uzyskane w teście wynosi 0.464. Na poziomie istotności $\alpha = 0.05$ nie mamy podstaw do odrzucenia hipotezy zerowej. Zakładamy, że zmienne losowe **CZY_KIER**, **PYT_2** i **STAŻ** są wzajemnie niezależne.

3.2.1.2 Nadmodel [13 23]

$M_1 = [13\ 23]$

H_1: M_1 i $M \neq [1\ 2\ 3]$

P–value uzyskane w teście wynosi 0.5084. Na poziomie istotności $\alpha = 0.05$ nie mamy podstaw do odrzucenia hipotezy zerowej. Zakładamy, że zmienne losowe **CZY_KIER**, **PYT_2** i **STAŻ** są wzajemnie niezależne.

3.2.1.3 Nadmodel [12 13]

$M_1 = [12\ 13]$

H_1: M_1 i $M \neq [1\ 2\ 3]$

P–value uzyskane w teście wynosi 0.9151. Na poziomie istotności $\alpha = 0.05$ nie mamy podstaw do odrzucenia hipotezy zerowej. Zakładamy, że zmienne losowe **CZY_KIER**, **PYT_2** i **STAŻ** są wzajemnie niezależne.

3.2.2 Zmienna losowa PYT_2 jest niezależna od pary zmiennych CZY_KIER i STAŻ

$M_0 = [13\ 2]$

$H_0 : M = M_0$

3.2.2.1 Nadmodel [12 13]

$M_1 = [12\ 13]$

$H_1 : M = M_1$ i $M \neq [13\ 2]$

P–value uzyskane w teście wynosi 0.6862. Na poziomie istotności $\alpha = 0.05$ nie mamy podstaw do odrzucenia hipotezy zerowej. Zakładamy, że zmienna losowa **PYT_2** jest niezależna od pary zmiennych **CZY_KIER** i **STAŻ**.

3.2.2.2 Nadmodel [12 13 23]

$M_1 = [12\ 13\ 23]$

$H_1 : M = M_1$ i $M \neq [13\ 2]$

P–value uzyskane w teście wynosi 0.4279. Na poziomie istotności $\alpha = 0.05$ nie mamy podstaw do odrzucenia hipotezy zerowej. Zakładamy, że zmienna losowa **PYT_2** jest niezależna od pary zmiennych **CZY_KIER** i **STAŻ**.

3.2.2.3 Nadmodel [123]

$M_1 = [123]$

$H_1 : M = M_1$ i $M \neq [13\ 2]$

P–value uzyskane w teście wynosi 0.07. Na poziomie istotności $\alpha = 0.05$ nie mamy podstaw do odrzucenia hipotezy zerowej. Zakładamy, że zmienna losowa **PYT_2** jest niezależna od pary zmiennych **CZY_KIER** i **STAŻ**.

3.2.3 Zmienna losowa PYT_2 jest niezależna od zmiennej CZY_KIER, przy ustalonej wartości zmiennej STAŻ

$M_0 = [13\ 23]$

3.2.4 Nadmodel [12 13 23]

$M_1 = [12\ 13\ 23]$

$H_1 : M = M_1$ i $M \neq [13\ 23]$

P–value uzyskane w teście wynosi 0.5491. Na poziomie istotności $\alpha = 0.05$ nie mamy podstaw do odrzucenia hipotezy zerowej. Zakładamy, że zmienna losowa **PYT_2** jest niezależna od zmiennej **CZY_KIER**, przy ustalonej wartości zmiennej **STAŻ**.

3.2.5 Nadmodel [123]

$M_1 = [123]$

$H_1 : M = M_1$ i $M \neq [13\ 23]$

P–value uzyskane w teście wynosi 0.0572. Na poziomie istotności $\alpha = 0.05$ nie mamy podstaw do odrzucenia hipotezy zerowej. Zakładamy, że zmienna losowa **PYT_2** jest niezależna od zmiennej **CZY_KIER**, przy ustalonej wartości zmiennej **STAŻ**.