

Raport nr 3

Natalia Iwańska 262270, Klaudia Janicka 262268

2023-06-16

Zadanie 1

Table 1: Tablica dwudzielcza dla zmiennych A1 i A2.

	-2	-2	0	1	2	Sum
	10	2	1	1	0	14
	0	15	1	1	0	17
	1	1	32	6	0	40
	0	0	1	96	3	100
	1	1	0	1	26	29
Sum	12	19	35	105	29	200

Test McNemary

Nie możemy skorzystać z testu McNemary, ponieważ w tablicy na odpowiadających sobie miejscach (Y_{ij} i Y_{ji}) występują zera, co “psuje” nam statystykę testową (wynika to wprost z jej definicji).

Test bazujący na ilorazie wiarygodności

Korzystając z testu bazującego na ilorazie wiarygodności na poziomie istotności $\alpha = 0.05$ otrzymana p-wartość wyniosła 0.2059752. Zatem weryfikowaną hipotezę o symetrii, która jest równoważna hipotezie o brzegowej jednorodności należy odrzucić.

Zadanie 2

Table 2: Tablica dwudzielcza dla zmiennych W1 i W2.

	-2	-1	1	2	Sum
	74	0	0	0	74
	0	19	1	0	20
	0	0	1	1	2
	0	0	0	104	104
Sum	74	19	2	105	200

Test McNemary

Podobnie jak w poprzednim zadaniu nie możemy skorzystać z testu McNemary, ponieważ w tablicy na odpowiadających sobie miejscach (Y_{ij} i Y_{ji}) występują zera, co “psuje” nam statystykę testową (wynika to wprost z jej definicji).

Test bazujący na ilorazie wiarygodności

Korzystając z testu bazującego na ilorazie wiarygodności na poziomie istotności $\alpha = 0.05$ otrzymana p-wartość wyniosła 0.8368001. Zatem nie mamy podstaw do odrzucenia hipotezy zerowej o symetrii, która jest równoważna hipotezie o brzegowej jednorodności.

Zadanie 3

```
#TEST Z
test_z <- function(tabela){
  n <- sum(rowSums(tabela))
  P <- tabela/n
  r <- rowSums(P)
  c <- colSums(P)
  D <- r[1] - c[1]
  sigma2_D <- (r[1]*(1-r[1])+c[1]*(1-c[1])-2*(P[1,1]*P[2,2]-P[1,2]*P[2,1]))/n
  Z <- D/sqrt(sigma2_D)
  p <- 2*(1 - pnorm(abs(Z)))
  return(p)
}

#TEST Z0
test_z0 <- function(tabela){
  n <- sum(rowSums(tabela))
  P <- tabela/n
  r <- rowSums(P)
  c <- colSums(P)
  D <- r[1] - c[1]
  sigma2_D0 <- (tabela[1,2]+tabela[2,1])/n^2
  Z_0 <- D/sqrt(sigma2_D0)
  p <- 2*(1 - pnorm(abs(Z_0)))
  return(p)
}
```

test	p-value
Test Z	NaN
Test Z0	NaN
McNemar test z poprawką na ciągłość	NaN
McNemar test bez poprawki	NaN

Na podstawie otrzymanych p-wartości testów przeprowadzonych na poziomie istotności $\alpha = 0.05$ stwierdzamy, że nie mamy podstaw do odrzucenia hipotezy zerowej o symetrii, która jest równoważna hipotezie o brzegowej jednorodności.

```
moc <- function(n){
  MC <- 1000
  p2 <- seq(0.01,0.99,0.01)
  p1 <- 0.5
  m <- length(p2)
  res <- rep(NA, m)
  for (i in 1:m){
    counter <- 0
    for (j in 1:MC){
```

```

X <- factor(sample(c("1","0"), n, replace=TRUE, prob = c(p1,1-p1)), levels = 0:1)
Y <- factor(sample(c("1","0"), n, replace=TRUE, prob = c(p2[i],1-p2[i])), levels=0:1)
tab <- ftable(X,Y)
if (test_z(tab) < 0.05){
  counter <- counter+1
}
}
res[i] <- counter/MC
}
return(data.frame( 'prob' = p2, 'results' = res))
}

```

```

moc(20)

```

```

##      prob results
## 1  0.01   0.993
## 2  0.02   0.995
## 3  0.03   0.984
## 4  0.04   0.977
## 5  0.05   0.953
## 6  0.06   0.943
## 7  0.07   0.917
## 8  0.08   0.907
## 9  0.09   0.875
## 10 0.10   0.853
## 11 0.11   0.820
## 12 0.12   0.816
## 13 0.13   0.778
## 14 0.14   0.741
## 15 0.15   0.719
## 16 0.16   0.708
## 17 0.17   0.679
## 18 0.18   0.621
## 19 0.19   0.627
## 20 0.20   0.564
## 21 0.21   0.545
## 22 0.22   0.526
## 23 0.23   0.476
## 24 0.24   0.452
## 25 0.25   0.413
## 26 0.26   0.407
## 27 0.27   0.355
## 28 0.28   0.339
## 29 0.29   0.310
## 30 0.30   0.315
## 31 0.31   0.268
## 32 0.32   0.261
## 33 0.33   0.223
## 34 0.34   0.213
## 35 0.35   0.194
## 36 0.36   0.188
## 37 0.37   0.173
## 38 0.38   0.169
## 39 0.39   0.152

```

##	40	0.40	0.118
##	41	0.41	0.123
##	42	0.42	0.124
##	43	0.43	0.094
##	44	0.44	0.092
##	45	0.45	0.102
##	46	0.46	0.064
##	47	0.47	0.068
##	48	0.48	0.073
##	49	0.49	0.094
##	50	0.50	0.071
##	51	0.51	0.073
##	52	0.52	0.068
##	53	0.53	0.067
##	54	0.54	0.076
##	55	0.55	0.075
##	56	0.56	0.081
##	57	0.57	0.084
##	58	0.58	0.114
##	59	0.59	0.111
##	60	0.60	0.133
##	61	0.61	0.131
##	62	0.62	0.153
##	63	0.63	0.161
##	64	0.64	0.178
##	65	0.65	0.189
##	66	0.66	0.196
##	67	0.67	0.214
##	68	0.68	0.248
##	69	0.69	0.263
##	70	0.70	0.297
##	71	0.71	0.290
##	72	0.72	0.354
##	73	0.73	0.378
##	74	0.74	0.375
##	75	0.75	0.427
##	76	0.76	0.465
##	77	0.77	0.464
##	78	0.78	0.516
##	79	0.79	0.535
##	80	0.80	0.541
##	81	0.81	0.594
##	82	0.82	0.651
##	83	0.83	0.653
##	84	0.84	0.700
##	85	0.85	0.732
##	86	0.86	0.753
##	87	0.87	0.784
##	88	0.88	0.787
##	89	0.89	0.826
##	90	0.90	0.862
##	91	0.91	0.863
##	92	0.92	0.914
##	93	0.93	0.916

```
## 94 0.94 0.945
## 95 0.95 0.959
## 96 0.96 0.976
## 97 0.97 0.974
## 98 0.98 0.991
## 99 0.99 0.995
```

Zadanie 5

- [1 3] zmienne „S” i „W1” mają dowolne rozkłady oraz zmienne te są niezależne, a zmienna „Wyk” ma rozkład równomierny

```
model_a <- glm(Freq ~ S + Wyk,
               data = df, family = poisson)
p_a <- 1-pchisq(deviance(model_a), df = df.residual(model_a))
```

P-wartość 0 jest mniejsza niż założony poziom istotności α , więc odrzucamy hipotezę zerową. Nasze dane nie pochodzą z modelu [1 3].

- [13] zmienne „S” i „W1” mają dowolne rozkłady oraz zmienne te nie są niezależne, a zmienna „Wyk” ma rozkład równomierny

```
model_b <- glm(Freq ~ S + Wyk + S*Wyk,
               data = df, family = poisson)
p_b <- 1-pchisq(deviance(model_b), df = df.residual(model_b))
cbind(model_a$data, fitted(model_b))
```

```
##      S W1 Wyk Freq fitted(model_b)
## 1  0 -2  1  19      10.00
## 2  1 -2  1   1       0.25
## 3  0 -1  1   3      10.00
## 4  1 -1  1   0       0.25
## 5  0  1  1   0      10.00
## 6  1  1  1   0       0.25
## 7  0  2  1  18      10.00
## 8  1  2  1   0       0.25
## 9  0 -2  2  40     30.75
## 10 1 -2  2   5       4.25
## 11 0 -1  2  15     30.75
## 12 1 -1  2   2       4.25
## 13 0  1  2   0     30.75
## 14 1  1  2   0       4.25
## 15 0  2  2  68     30.75
## 16 1  2  2  10       4.25
## 17 0 -2  3   5       2.50
## 18 1 -2  3   4       2.25
## 19 0 -1  3   0       2.50
## 20 1 -1  3   0       2.25
## 21 0  1  3   0       2.50
## 22 1  1  3   2       2.25
## 23 0  2  3   5       2.50
## 24 1  2  3   3       2.25
```

P-wartość rp_b jest mniejsza niż założony poziom istotności α , więc odrzucamy hipotezę zerową. Nasze dane nie pochodzą z modelu [13].

- [1 2 3] zmienne „S”, „W1”, „Wyk” są wzajemnie niezależne

```
model_c <- glm(Freq ~ S + Wyk + W1,
               data = df, family = poisson)
p_c <- 1-pchisq(deviance(model_c), df = df.residual(model_c))
cbind(model_a$data, fitted(model_b))
```

##	S	W1	Wyk	Freq	fitted(model_b)
## 1	0	-2	1	19	10.00
## 2	1	-2	1	1	0.25
## 3	0	-1	1	3	10.00
## 4	1	-1	1	0	0.25
## 5	0	1	1	0	10.00
## 6	1	1	1	0	0.25
## 7	0	2	1	18	10.00
## 8	1	2	1	0	0.25
## 9	0	-2	2	40	30.75
## 10	1	-2	2	5	4.25
## 11	0	-1	2	15	30.75
## 12	1	-1	2	2	4.25
## 13	0	1	2	0	30.75
## 14	1	1	2	0	4.25
## 15	0	2	2	68	30.75
## 16	1	2	2	10	4.25
## 17	0	-2	3	5	2.50
## 18	1	-2	3	4	2.25
## 19	0	-1	3	0	2.50
## 20	1	-1	3	0	2.25
## 21	0	1	3	0	2.50
## 22	1	1	3	2	2.25
## 23	0	2	3	5	2.50
## 24	1	2	3	3	2.25

P-wartość 6.18728×10^{-4} jest mniejsza niż założony poziom istotności α , więc odrzucamy hipotezę zerową. Nasze dane nie pochodzą z modelu [1 2 3].