

Case study | Palenie a nowotwór płuc

- a) Na podstawie danych można wnioskować, iż zmienna zależną jest historia palenia papierosów pacjenta, natomiast zmienną niezależną jest obecność raka płuc u danego pacjenta.
- b) Biorąc pod uwagę charakter zaprezentowania danych można wnioskować, że były to badania retrospektywne. Grupa badawcza biorąca udział w badaniu była grupą pacjentów chorujących na raka płuc. Na tej podstawie badano zależność palenia wobec tej grupy a grupy kontrolnej, co jest przykładem badań retrospektywnych. Badania te sięgają po zachowania danej grupy z przeszłości, szukając pewnych wzorców.
- c) Na podstawie danych można wnioskować, iż powyższe dane nie mogą być dobrą formą porównania osób palących a niepalących względem proporcji pacjentów cierpiących na raka płuc. Jest to spowodowane sposobem w jakim dany eksperyment został przeprowadzony. Dystrybucje marginalne dla zmiennej obecności raka płuc są ściśle ustalone względem prowadzonego badania (liczba pacjentów z próby kontrolnej odpowiada liczbie pacjentów z próby badawczej). Charakter eksperymentu sprawia, iż nie jesteśmy w stanie określić prawdopodobieństwa obecności raka płuc, biorąc pod uwagę ich historie palenia papierosów. Możemy natomiast wyznaczyć zależność logicznie odwrotną (prawdopodobieństwo bycia palaczem przez pacjenta, wiedząc iż cierpi na nowotwór płuc) lub współczynnik szans.
- d) Wyniki analizy za pomocą programu SAS poniżej; Test chi-squared na niezależność zmiennych wykazał, iż na 5% stopniu istotności, możemy odrzucić hipotezę zerową (Nawyki palenia papierosów pacjenta są niezależne wobec rozwoju u danego pacjenta nowotworu płuc) i tym samym przyjąć hipotezę alternatywną (Nawyki palenia papierosów pacjenta są zależne wobec rozwoju u danego pacjenta nowotworu płuc). Ważniejszą statystyką w tym badaniu będzie natomiast iloraz szans, ponieważ do każdego pacjenta z grupy badawczej dobrano pacjenta do grupy kontrolnej. Iloraz szans (odds ratio) przybiera wartości większe od 1 w 95% przedziale ufności. Dolna wartość przedziału ufności tego współczynnika informuje nas, iż szansa że pacjent jest palaczem, biorąc pod uwagę obecność u niego nowotworu jest minimum około 1,8 razy większa niż szansa pacjenta nie będącego palaczem, biorąc pod uwagę obecność u niego nowotworu. Ponieważ wartość współczynnika szans nie zmienia się względem zmian orientacji pierwotnej tabeli, możemy wnioskować że szansa rozwoju nowotworu u osób palących jest minimum 1,8 razy większa niż szansa rozwoju nowotworu u osób niepalących, natomiast nie możemy wyznaczyć bezpośrednich wartości prawdopodobieństwa takiego zdarzenia.

Table of Have Smoked by Lung Cancer			
Have Smoked	Lung Cancer		
Frequency Expected	Cases	Control	Total
Yes	688 669	650 669	1338
No	21 40	59 40	80
Total	709	709	1418

Statistics for Table of Have Smoked by Lung Cancer

Statistic	DF	Value	Prob
Chi-Square	1	19.1292	<.0001
Likelihood Ratio Chi-Square	1	19.8780	<.0001
Continuity Adj. Chi-Square	1	18.1357	<.0001
Mantel-Haenszel Chi-Square	1	19.1157	<.0001
Phi Coefficient		0.1161	
Contingency Coefficient		0.1154	
Cramer's V		0.1161	

Fisher's Exact Test	
Cell (1,1) Frequency (F)	688
Left-sided Pr <= F	1.0000
Right-sided Pr >= F	<.0001
Table Probability (P)	<.0001
Two-sided Pr <= P	<.0001

Statistic	Value	ASE
Gamma	0.4967	0.0979
Kendall's Tau-b	0.1161	0.0243
Stuart's Tau-c	0.0536	0.0122
Somers' D C R	0.2517	0.0511
Somers' D R C	0.0536	0.0122
Pearson Correlation	0.1161	0.0243
Spearman Correlation	0.1161	0.0243
Lambda Asymmetric C R	0.0536	0.0123
Lambda Asymmetric R C	0.0000	0.0000
Lambda Symmetric	0.0482	0.0108
Uncertainty Coefficient C R	0.0101	0.0044
Uncertainty Coefficient R C	0.0323	0.0137
Uncertainty Coefficient Symmetric	0.0154	0.0066

Odds Ratio and Relative Risks			
Statistic	Value	95% Confidence Limits	
Odds Ratio	2.9738	1.7867	4.9494
Relative Risk (Column 1)	1.9589	1.3517	2.8387
Relative Risk (Column 2)	0.6587	0.5716	0.7591

Odds Ratio	
Odds Ratio	2.9738
Asymptotic Conf Limits	
95% Lower Conf Limit	1.7867
95% Upper Conf Limit	4.9494
Exact Conf Limits	
95% Lower Conf Limit	1.7556
95% Upper Conf Limit	5.2107

Sample Size = 1418