Chapter 18 – Comparisons of Proportions or Odds

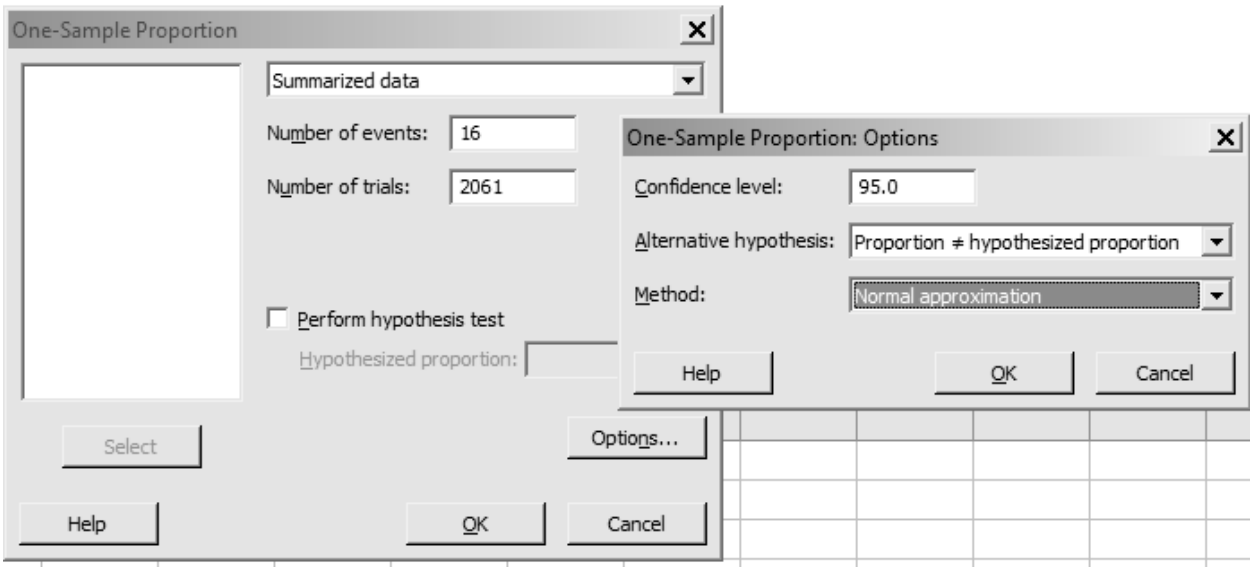
Case 18.1.1. Obesity and Heart Disease – An Observational Study. *R&S p.550.*

Step 1: Input the data directly into a worksheet. Input the data into columns directly. Name Column 1 as CVD_Yes and name Column 2 as CVD_No. Under Column 1, input the values 16 and 7 in column 1 from ***R&S Display 18.1.*** Under Column 2, input the values 2045 and 1044 in column 2 from ***R&S Display 18.1.***

Data Display

Row	CVD_Yes	CVD_No
1	16	2045
2	7	1044

Step 2: Construct a z- C.I. interval for the true proportion of CVD deaths in the population of obese Samoan women. Go to Stat → Basic Statistics → 1-Proportion; click on Summarized Data; for results, see ***R&S Display 18.4***.



Test and CI for One Proportion

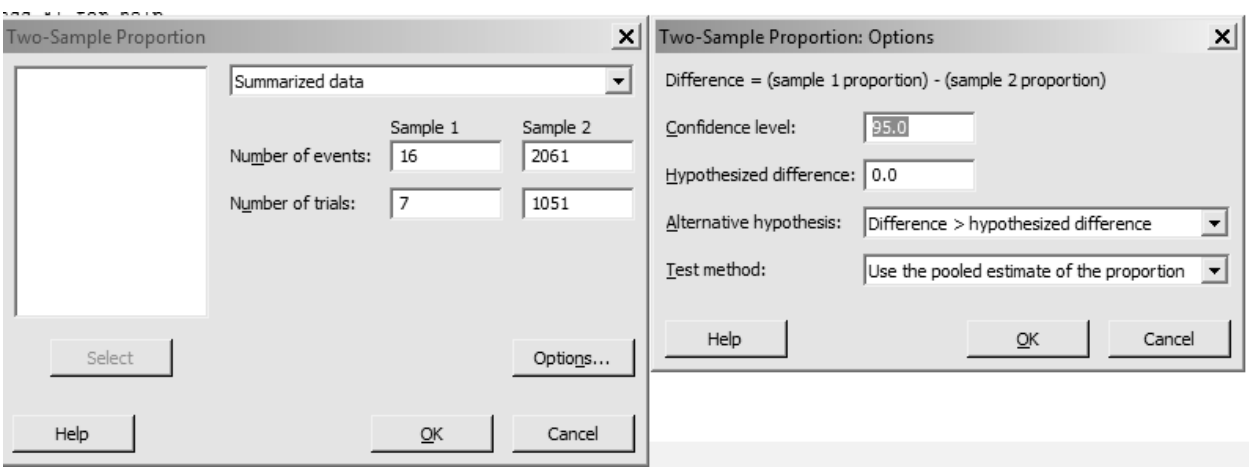
```
Sample   X     N  Sample p         95% CI
1       16  2061  0.007763  (0.003974, 0.011552)

Using the normal approximation.
```

Next, construct a z- C.I. interval for the true proportion of CVD deaths in the population of non-obese Samoan women. Go to Stat → Basic Statistics → 1-Proportion; click on Summarized Data; for results,

Test and CI for One Proportion

```
Sample  X     N  Sample p         95% CI
1       7  1051  0.006660  (0.001743, 0.011578)

Using the normal approximation.
```

Step 3: Test for equality of two true population proportions, viz., the proportion of CVD deaths in the population of obese Samoan women and the proportion of CVD deaths in the population of non-obese Samoan women. Go to Stat → Basic Statistics → 2-Proportion2; click on Summarized Data, input data for obese women into First and data for non-obese women into Second; click on options, select Alternative "greater than" and click on Use pooled estimate of p for test. For formula details, see ***R&S p. 555 -557***. For results, see ***R&S Display 18.5***.



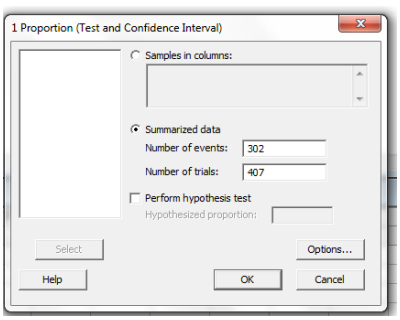
Test and CI for Two Proportions

```
Sample   X     N  Sample p
1       16  2061  0.007763
2        7  1051  0.006660


Difference = p (1) - p (2)
Estimate for difference:  0.00110290
95% lower bound for difference:  -0.00410701
Test for difference = 0 (vs > 0):  Z = 0.34  P-Value = 0.367

Fisher's exact test: P-Value = 0.462
```

Case 18.1.2. Vitamin C and the Common Cold –A Randomized Experiment. *R&S p.551-552.*

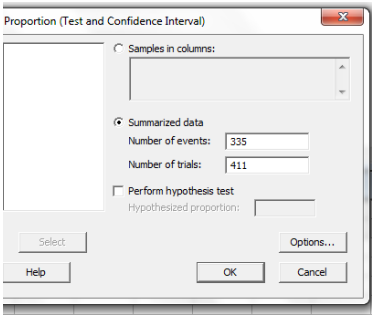
Data Display

Row	Cold	No_Cold
1	335	76
2	302	105



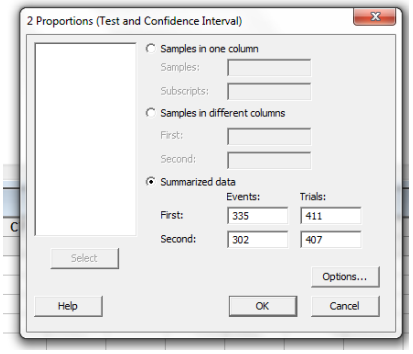
Test and CI for One Proportion

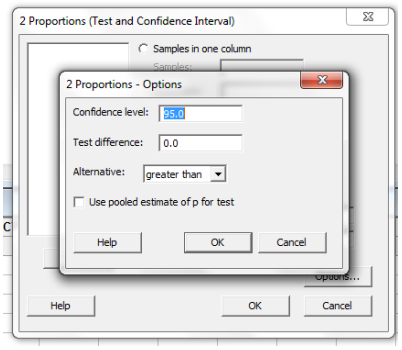
```
Sample    X    N  Sample p         95% CI
1       302  407  0.742015  (0.699508, 0.784521)

Using the normal approximation.
```



Test and CI for One Proportion

```
Sample    X    N  Sample p         95% CI
1       335  411  0.815085  (0.777552, 0.852618)

Using the normal approximation.
```





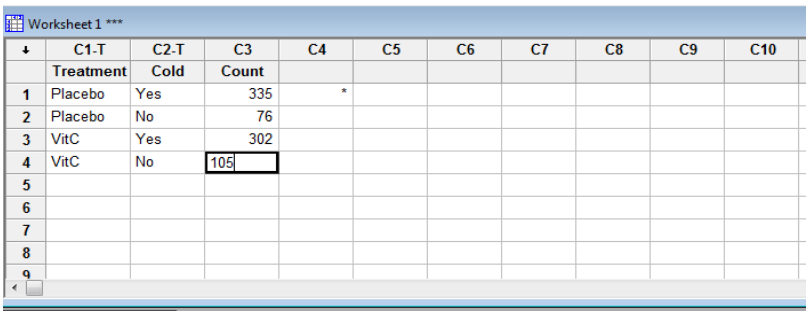
Test and CI for Two Proportions

```
Sample    X    N  Sample p
1       335  411  0.815085
2       302  407  0.742015


Difference = p (1) - p (2)
Estimate for difference:  0.0730704
95% lower bound for difference:  0.0254815
Test for difference = 0 (vs > 0):  Z = 2.53  P-Value = 0.006

Fisher's exact test: P-Value = 0.007
```

Discussion of Odds and Ratio of Odds

Enter data into Minitab as follows:

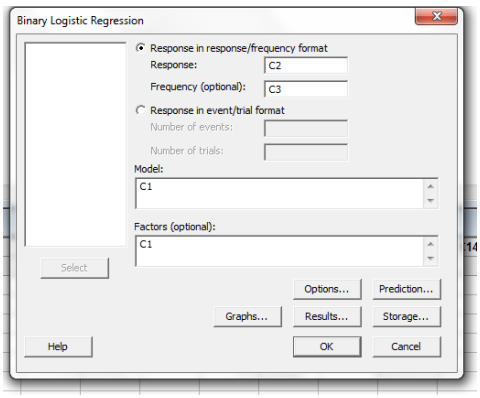
Treatment	Cold	Count
C	Yes	302
C	No	105
Placebo	Yes	335
Placebo	No	76



Worksheet 1 ***

↓	C1-T	C2-T	C3	C4	C5	C6	C7	C8	C9	C10
	Treatment	Cold	Count							
1	Placebo	Yes	335	*						
2	Placebo	No	76							
3	VitC	Yes	302							
4	VitC	No	105							
5										
6										
7										
8										
9										

Go to Stat → Regression → Binary Logistic Regression. In Response window, enter C2 and in Frequency window, enter C3. In Categorical Predictor window, center C1. Click OK.



Binary Logistic Regression: Cold versus Treatment

```
Link Function: Logit


Response Information

Variable  Value  Count
Cold      Yes      637  (Event)
          No       181
          Total    818

Frequency: Count


Logistic Regression Table

                                                Odds     95% CI
Predictor        Coef   SE Coef      Z      P  Ratio  Lower  Upper
Constant      1.48340  0.127052  11.68  0.000
Treatment
 VitC       -0.426931  0.170227  -2.51  0.012   0.65   0.47   0.91


Log-Likelihood = -429.145
Test that all slopes are zero: G = 6.357, DF = 1, P-Value = 0.012

* NOTE * No goodness of fit test performed.
* NOTE * The model uses all degrees of freedom.
```

```
Measures of Association:
(Between the Response Variable and Predicted Probabilities)

Pairs       Number  Percent  Summary Measures
Concordant   35175     30.5  Somers' D              0.11
Discordant   22952     19.9  Goodman-Kruskal Gamma  0.21
Ties         57170     49.6  Kendall's Tau-a        0.04
Total       115297    100.0
```