

Binary Data Analysis
Targil 2 - Solution

$$\begin{pmatrix} n_{00} \\ n_{10} \end{pmatrix} = \begin{pmatrix} 21 \\ 12 \end{pmatrix} = 293,930$$

$$\begin{array}{cc|c} 3 & 8 & 11 \\ 6 & 4 & 10 \\ \hline 9 & 12 & 21 \end{array} : \text{מספר הנושא}$$

: מספר הנושא של המידע

$$\frac{\begin{pmatrix} n_{00} \\ n_{01}-n_{11} \end{pmatrix} \begin{pmatrix} n_{10} \\ n_{11} \end{pmatrix}}{\begin{pmatrix} n_{00} \\ n_{10} \end{pmatrix}}$$

$$\begin{pmatrix} n_{00} \\ n_{01}-n_{11} \end{pmatrix} \begin{pmatrix} n_{10} \\ n_{11} \end{pmatrix} \left(\frac{1}{2}\right)^{n_{11}}$$

$$(1)(10)\left(\frac{1}{2}\right)^1 = 5.00$$

$$\frac{\binom{11}{1} \binom{10}{1}}{\binom{21}{1}} = \frac{(1)(10)}{293,930} = 0.0000$$

חסד
0 11
9 1

$$(11)(45)\left(\frac{1}{2}\right)^2 = 123.75$$

$$\frac{\binom{11}{11} \binom{10}{2}}{\binom{21}{2}} = \frac{(11)(45)}{293,930} = 0.0017$$

1 10
8 2

$$(55)(120)\left(\frac{1}{2}\right)^3 = 825.00$$

$$\frac{\binom{11}{9} \binom{10}{3}}{\binom{21}{3}} = \frac{(55)(120)}{293,930} = 0.0225$$

2 9
7 3

$$(165)(210)\left(\frac{1}{2}\right)^4 = 2165.63$$

$$\frac{\binom{11}{8} \binom{10}{4}}{\binom{21}{4}} = \frac{(165)(210)}{293,930} = 0.1179$$

3 8 ←
6 4

$$(330)(252)\left(\frac{1}{2}\right)^5 = 2598.75$$

$$\frac{\binom{11}{7} \binom{10}{5}}{\binom{21}{5}} = \frac{(330)(252)}{293,930} = 0.2829$$

4 7
5 5

$$(426)(210)\left(\frac{1}{2}\right)^6 = 1397.81$$

$$\frac{\binom{11}{6} \binom{10}{6}}{\binom{21}{6}} = \frac{(426)(210)}{293,930} = 0.3301$$

5 6
4 6

$$(462)(120)\left(\frac{1}{2}\right)^7 = 433.13$$

$$\frac{\binom{11}{5} \binom{10}{7}}{\binom{21}{7}} = \frac{(462)(120)}{293,930} = 0.1886$$

6 5
3 7

$$(330)(45)\left(\frac{1}{2}\right)^8 = 58.01$$

$$\frac{\binom{11}{4} \binom{10}{8}}{\binom{21}{8}} = \frac{(330)(45)}{293,930} = 0.0505$$

7 4
2 8

$$(165)(10)\left(\frac{1}{2}\right)^9 = 3.22$$

$$\frac{\binom{11}{3} \binom{10}{9}}{\binom{21}{9}} = \frac{(165)(10)}{293,930} = 0.0056$$

8 3
1 9

$$(55)(1)\left(\frac{1}{2}\right)^{10} = 0.05$$

$$\frac{\binom{11}{2} \binom{10}{10}}{\binom{21}{10}} = \frac{(55)(1)}{293,930} = 0.0002$$

9 2
0 10

$$7,610.35 = 2120$$

: $H_0: \omega > 1$ δ_{IN} $H_0: \omega \leq 1$ 7/28

p-value = $0.1179 + 0.2829 + 0.3301 + 0.1886 + 0.0505 + 0.0056 + 0.0002$
= 0.9758

: $H_0: \omega < 1$ δ_{IN} $H_0: \omega \geq 1$ 7/28

p-value = $0.1179 + 0.0225 + 0.0017 + 0.0000 = 0.1421$

: $H_0: \omega \neq 1$ δ_{IN} $H_0: \omega = 1$ 7/28

p-value = $0.0000 + 0.0017 + 0.0225 + 0.1179$
 $+ 0.0505 + 0.0056 + 0.0002 = 0.1984$

: $H_0: \omega > \frac{1}{2}$ δ_{IN} $H_0: \omega \leq \frac{1}{2}$ 7/28

p-value = $\frac{1}{7,610.35} [2165.63 + 2598.75 + 1397.81 + 433.13$
 $+ 58.01 + 3.22 + 0.05] = 0.8747$

: $H_0: \omega < \frac{1}{2}$ δ_{IN} $H_0: \omega \geq \frac{1}{2}$ 7/28

p-value = $\frac{1}{7,610.35} [2165.63 + 825.00 + 123.75 + 5.00] = 0.4099$

: $H_0: \omega \neq \frac{1}{2}$ δ_{IN} $H_0: \omega = \frac{1}{2}$ 7/28

p-value = $\frac{1}{7,610.35} [5.00 + 123.75 + 825.00 + 2165.63$
 $+ 1397.81 + 433.13 + 58.01 + 3.22 + 0.05]$
= 0.8163

The FREQ Procedure

Statistics for Table of a by b

Statistic	DF	Value	Prob
Chi-Square	1	2.2909	0.1301
Likelihood Ratio Chi-Square	1	2.3309	0.1268
Continuity Adj. Chi-Square	1	1.1494	0.2837
Mantel-Haenszel Chi-Square	1	2.1818	0.1396
Phi Coefficient		-0.3303	
Contingency Coefficient		0.3136	
Cramer's V		-0.3303	

WARNING: 50% of the cells have expected counts less than 5. Chi-Square may not be a valid test.

Fisher's Exact Test

Cell (1,1) Frequency (F)	3
Left-sided Pr <= F	0.1421
Right-sided Pr >= F	0.9758
Table Probability (P)	0.1179
Two-sided Pr <= P	0.1984

Estimates of the Relative Risk (Row1/Row2)

Type of Study	Value	95% Confidence Limits	
Case-Control (Odds Ratio)	0.2500	0.0400	1.5637
Cohort (Col1 Risk)	0.4545	0.1529	1.3515
Cohort (Col2 Risk)	1.8182	0.7842	4.2155

Odds Ratio (Case-Control Study)

Odds Ratio	0.2500
Asymptotic Conf Limits	
95% Lower Conf Limit	0.0400
95% Upper Conf Limit	1.5637
Exact Conf Limits	
95% Lower Conf Limit	0.0269
95% Upper Conf Limit	2.1007

Sample Size = 21

```

% Program to compute MLE of odds ratio in 2x2 table
% Assumes that the user has previously input a vector nv with the cell counts

n00=nv(1,1);
n01=nv(1,2);
n10=nv(1,3);
n11=nv(1,4);

r0 = n00 + n01;
r1 = n10 + n11;
c0 = n00 + n10;
c1 = n01 + n11;

kmin = max(c1-r0,0);
kmax = min(r1,c1);

kv = kmin:kmax;
kv2 = kv.^2;

comb1=[];
comb2=[];
for k = kmin:kmax;
    comb1 = [comb1 nchoosek(r0,c1-k)];
    comb2 = [comb2 nchoosek(r1,k)];
end;
comb = comb1.*comb2;

omega=1;
psi=log(omega);
psi_vec=psi;
diff=Inf;
eps=0.0001;

while (diff > eps);
    ppvec=exp(psi.*kv);
    ptrms=comb.*ppvec;
    prob=ptrms/sum(ptrms);
    e1=sum(kv.*prob);
    e2=sum(kv2.*prob);
    var=e2-e1^2;
    step=(n11-e1)/var;
    psi=psi+step;
    diff=abs(step);
    psi_vec=vertcat(psi_vec,psi);
end;

siz=size(psi_vec);
nitr=siz(1,1)-1;
results = [(0:nitr)' psi_vec exp(psi_vec)]
sd_phi_hat = sqrt(1/var)

%*****

>> nv=[3 8 6 4];
>> bin

results =
           0           0          1.0000
    1.0000  -1.2727    0.2801
    2.0000  -1.3147    0.2686
    3.0000  -1.3149    0.2685
    4.0000  -1.3149    0.2685

sd_phi_hat =
    0.9073

```

3. (אחרי שר הכי) וי עגור מתק (הכיוון) ויג' ונציח את העטן :

$$Pr(n_{11}=k | \text{data}) = \binom{n_{00}}{n_{11}-k} \binom{n_{01}}{k} \omega^k / \sum_l \binom{n_{00}}{n_{11}-l} \binom{n_{01}}{l} \omega^l$$

$$= \frac{n_{00}!}{(n_{11}-k)! (n_{00}-n_{11}+k)!} \frac{n_{01}!}{k! (n_{01}-k)!} \omega^k$$

$$/ \sum_l \frac{n_{00}!}{(n_{11}-l)! (n_{00}-n_{11}+l)!} \frac{n_{01}!}{l! (n_{01}-l)!} \omega^l$$

$$= \frac{1}{k! (n_{11}-k)!} \frac{1}{(n_{01}-k)! (n_{00}-n_{11}+k)!} \omega^k$$

$$/ \sum_l \frac{1}{l! (n_{11}-l)!} \frac{1}{(n_{01}-l)! (n_{00}-n_{11}+l)!} \omega^l$$

$$= \frac{n_{01}!}{k! (n_{11}-k)!} \frac{n_{00}!}{(n_{01}-k)! (n_{00}-n_{11}+k)!} \omega^k$$

$$= \binom{n_{01}}{k}$$

$$/ \sum_l \frac{n_{01}!}{l! (n_{11}-l)!} \frac{n_{00}!}{(n_{01}-l)! (n_{00}-n_{11}+l)!} \omega^l$$

$$= \binom{n_{01}}{l}$$

$$n_{00} - n_{11} = n_{00} + n_{10} - n_{10} - n_{11} \quad \text{פס}$$

$$= n_{00} - n_{11} = n_{00} + n_{01} - n_{01} - n_{11}$$

$$= n_{00} - n_{01}$$

$$(n_{00} - n_{11} + l) = (n_{00} - n_{01} + l) \quad \text{פס}$$

$$= (n_{00} - (n_{01} - l))$$

פס פס

$$\frac{n_{00}!}{(n_{01}-l)! (n_{00}-n_{11}+l)!} = \frac{n_{00}!}{(n_{01}-l)! (n_{00} - (n_{01}-l))!} = \binom{n_{00}}{n_{01}-l}$$

פס פס

$$Pr(n_{11} | \text{data}) = \frac{\binom{n_{00}}{n_{11}-k} \binom{n_{01}}{k} \omega^k}{\sum_l \binom{n_{00}}{n_{11}-l} \binom{n_{01}}{l} \omega^l}$$

☺ .d.e.n

$$\mathcal{L}(\beta) = \log \left[\binom{n_{0\cdot}}{n_{\cdot 1} - n_{11}} \binom{n_{1\cdot}}{n_{11}} \right] + n_{11} \beta - \log \sum_l \binom{n_{0\cdot}}{n_{\cdot 1} - l} \binom{n_{1\cdot}}{l} e^{l\beta}$$

$$\mathcal{L}'(\beta) = - \frac{\sum_l \binom{n_{0\cdot}}{n_{\cdot 1} - l} \binom{n_{1\cdot}}{l} l e^{l\beta}}{\sum_l \binom{n_{0\cdot}}{n_{\cdot 1} - l} \binom{n_{1\cdot}}{l} e^{l\beta}}$$

$$\left(\frac{f}{g}\right)' = \frac{f'}{g} - \frac{f}{g} \frac{g'}{g}$$

$$\mathcal{L}''(\beta) = - \left[\frac{\sum_l \binom{n_{0\cdot}}{n_{\cdot 1} - l} \binom{n_{1\cdot}}{l} l^2 e^{l\beta}}{\sum_l \binom{n_{0\cdot}}{n_{\cdot 1} - l} \binom{n_{1\cdot}}{l} e^{l\beta}} \right.$$

$$\left. - \left(\frac{\sum_l \binom{n_{0\cdot}}{n_{\cdot 1} - l} \binom{n_{1\cdot}}{l} l e^{l\beta}}{\sum_l \binom{n_{0\cdot}}{n_{\cdot 1} - l} \binom{n_{1\cdot}}{l} e^{l\beta}} \right)^2 \right]$$

$$\frac{\sum_l \binom{n_{0\cdot}}{n_{\cdot 1} - l} \binom{n_{1\cdot}}{l} l e^{l\beta}}{\sum_l \binom{n_{0\cdot}}{n_{\cdot 1} - l} \binom{n_{1\cdot}}{l} e^{l\beta}}$$

(כאן יש להוסיף נוסחה)

$$= \frac{\sum_{l'} \binom{n_{0\cdot}}{n_{\cdot 1} - l'} \binom{n_{1\cdot}}{l'} l' e^{l'\beta}}{\sum_{l'} \binom{n_{0\cdot}}{n_{\cdot 1} - l'} \binom{n_{1\cdot}}{l'} e^{l'\beta}}$$

$$= \sum_l l p_l = E[N_{11} | N_{\cdot 1} = n_{\cdot 1}]$$

$$p_l = \frac{\binom{n_{0\cdot}}{n_{\cdot 1} - l} \binom{n_{1\cdot}}{l} e^{l\beta}}{\sum_{l'} \binom{n_{0\cdot}}{n_{\cdot 1} - l'} \binom{n_{1\cdot}}{l'} e^{l'\beta}}$$

כאן

$$= \Pr(N_{11} = l | N_{\cdot 1} = n_{\cdot 1})$$

2013 / 2011-2

$$\frac{\sum_l \binom{n_0}{n_1-l} \binom{n_1}{l} l^2 e^{\beta l}}{\sum_l \binom{n_0}{n_1-l} \binom{n_1}{l} e^{\beta l}}$$

$$= \sum_l l^2 p_l = E[N_{11}^2 | N_{01} = n_{01}]$$

$$\mathcal{L}''(\beta) = - \left\{ E[N_{11}^2 | N_{01} = n_{01}] - (E[N_{11} | N_{01} = n_{01}])^2 \right\}$$

$$= - \text{Var}(N_{11} | N_{01} = n_{01})$$