M20732, Statistics methods in Health Research and Social Science Coursework

$\boxed{\text{np930537}}$

A.1

1) • Six of the studies, "Ang 2016", "Bang 2016", "Bapaye 2016", "Dayyeh 2017", "Mukai 2014" and "Sharaiha 2015" suggest a positive association, since RR > 1.0 [Their Risk Ratios are 1.03, 1.03, 1.28, 1.10, 1.05, 1.10 respectively]

• One study, "Lee 2014", suggests a negative association, between use of metal stents and clinical success, since RR < 1.0 [RR = 0.96]

• "Ang 2016", "Bang 2016", "Dayyeh 2017", "Mukai 2014" and "Sharaiha 2015" are not significant, since 1.0 is included in the 95% CI.

• "Bapaye 2016" and "Sharaiha 2015" are significant because 1.0 is not included in the 95% CI.

2) $$\text{I-Squared} = \frac{\chi^2 - d.f}{\chi^2} \times 100\%$$

$$\boxed{\begin{array}{l}\text{I-Squared} = 25.4\% \\ d.f = n-1 = 7-1 = 6\end{array}}$$

$$25.4 = \frac{\chi^2 - 6}{\chi^2} \times 100$$

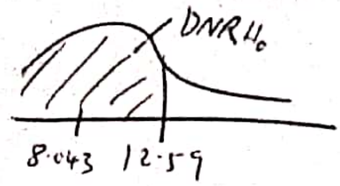
$$0.254\chi^2 = \chi^2 - 6$$

$$-0.746\chi^2 = -6 \Rightarrow \chi^2 = 8.04289...$$

$\Rightarrow \chi^2 \simeq 8.043$

$\boxed{\text{Critical Value}}$ $\chi^2_{crit} = 12.59$

$v = 7 - 1 = 6$ $\quad v = 6, 5\%$



$\boxed{\text{Conclusion}}$ $\chi^2 < 12.59$

The chi square value shows that the test of homogenity is not significant at the 5% level of significance.

this indicates no significant evidence of homogenity between the two strata (metal vs plastic stents).

3) $\boxed{\text{Overall Conclusion}}$

- The ~~ass~~ pooled RR of 1.08 indicates a positive association between the use of metal stents and clinical success.
- The RR 95% CI of (1.02, 1.14) does not include 1.0, which ~~shows that~~ confirms that there is a significant positive association between use of metal stents and clinical success at the 5% level of significance.

Scanned with CamScanner

A.2

1) Test of association between stress and heart disease:

Chi-Square Test

H_0: There is no association between stress and heart disease
H_1: There is an association between stress and heart disease

	Never	Sometimes	Often	Always	
Cases	99 [106.036]	227 [238.506]	113 [102.390]	45 [37.067]	484
Controls	250 [242.964]	558 [546.494]	224 [234.610]	77 [84.933]	1109
	349	785	337	122	1593

Expected values:

$$E = \frac{\text{column total} \times \text{row total}}{\text{grand total}}$$

$$E_1 = \frac{349 \times 484}{1593} = 106.0364\ldots \simeq 106.036$$

$$E_2 = \frac{785 \times 484}{1593} = 238.5059\ldots \simeq 238.506$$

$\vdots$

$$E_8 = \frac{122 \times 1109}{1593} = 84.9328\ldots \simeq 84.933$$

Test Statistic

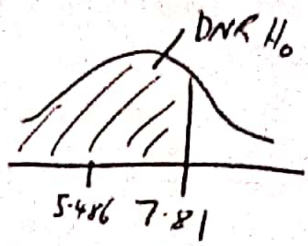
$$\chi^2 = \sum \frac{(O-E)^2}{E} = \frac{(99-106.036)^2}{106.036} + \frac{(227-238.506)^2}{238.506} + \frac{(113-102.390)^2}{102.390}$$

$$+ \frac{(45-37.067)^2}{37.067} + \frac{(250-242.964)^2}{242.964} + \frac{(558-546.494)^2}{546.494}$$

$$+ \frac{(224-234.610)^2}{234.610} + \frac{(77-84.933)^2}{84.933}$$

$\Rightarrow \chi^2 = 5.4859908... \simeq 5.486$

$\boxed{\text{Critical Value}}$ $\chi^2_{crit,\ \nu=3,\ 5\%} = 7.81$

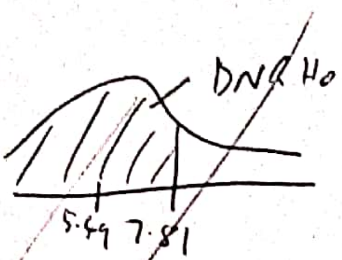


$\nu = (r-1)(c-1) = (2-1)(4-1) = 1(3) = 3$ d.f

$\boxed{\text{Conclusion}}$ $\chi^2 < 7.84$

there is insufficient evidence to reject the null hypothesis of no association between stress and heart disease at the 5% level of significance.

We conclude that there is no association between stress at work and heart disease.



$\boxed{\text{Conclusion}}$ $\chi^2 < 7.81$

There is insufficient evidence to reject the null hypothesis of no association between stress and heart disease at the 5% level of significance.
We conclude that there is no association between stress at work and heart disease.

---

2) Let "Never" = 1
"Sometimes" = 2
"Often" = 3
"Always" = 4

$$odds_1 = \frac{99}{250} = 0.396$$

$$odds_2 = \frac{227}{558} = 0.407$$

$$odds_3 = \frac{113}{224} = 0.504$$

$$odds_4 = \frac{45}{77} = 0.584$$

The odds increase as exposure to stress at work increases.
Therefore, the likelihood of heart attack increases as the amount of stress at work increases.

3) Odds Ratios relative to "Never"

$$OR_{sometimes} = \frac{Odds_2}{odds_1} = \frac{227}{558} \div \frac{99}{250} = \frac{227(250)}{558(99)} = 1.027298...$$

$$\simeq 1.027$$

"Never" = 1
"Sometimes" = 2
"Often" = 3
"Always" = 4

$\boxed{95\% \text{ CI}}$ Using Transformation method:

$$y = \ln(1.027) = 0.02664.. \simeq 0.027$$

$$V(y) = \frac{1}{99} + \frac{1}{250} + \frac{1}{227} + \frac{1}{558} = 0.0202984.. \simeq 0.020$$

OR 95% CI Transformation method:

$y \pm 1.96\sqrt{V(y)}$

$y = \ln(OR)$

$V(y) = \frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}$

$$95\% \text{ CI for } \ln(OR_{sometimes}) = 0.027 \pm 1.96\sqrt{0.020}$$

$$= (-0.25018..., 0.30418...)$$

$$\simeq (-0.250, 0.304)$$

Back Transforming

$$OR_{sometimes} \ 95\% \text{ CI} = (e^{-0.250}, e^{0.304}) = (0.7788.., 1.35526..)$$

$$\simeq (0.779, 1.355)$$

$$OR_{often} = \frac{Odds_3}{odds_1} = \frac{113}{224} \div \frac{99}{250} = \frac{113(250)}{224(99)} = 1.27389... \simeq 1.274$$

$\boxed{95\% \text{ CI}}$ Using Transformation Method:

$$y = \ln(1.274) = 0.24216.. \simeq 0.242$$

$$V(y) = \frac{1}{99} + \frac{1}{250} + \frac{1}{113} + \frac{1}{224} = 0.02741... \simeq 0.027$$

$$95\%\ CI \text{ for } \ln(OR_{often}) = 0.242 \pm 1.96\sqrt{0.027}$$
$$= (-0.08006\ldots, 0.546406\ldots)$$
$$\simeq (-0.080, 0.564)$$

Back Transforming:

$$\boxed{OR_{often}\ 95\%\ CI = \left(e^{-0.080}, e^{0.564}\right) = (0.9231\ldots, 1.7576\ldots) \simeq (0.923, 1.758)}$$

$$OR_{always} = \frac{odds_4}{odds_1} = \frac{45}{77} \div \frac{99}{250} = \frac{45(250)}{77(99)} = 1.47579\ldots \simeq 1.476$$

95% CI Using Transformation Method:

$$y = \ln(1.476) = 0.3893\ldots \simeq 0.389$$

$$V(y) = \frac{1}{99} + \frac{1}{250} + \frac{1}{45} + \frac{1}{77} = 0.04931\ldots \simeq 0.049$$

$$95\%\ CI \text{ for } \ln(OR_{always}) = 0.389 \pm 1.96\sqrt{0.049} = (-0.04486\ldots, 0.822864\ldots)$$
$$\simeq (-0.045, 0.823)$$

Back Transforming:

$$OR_{always}\ 95\%\ CI = \left(e^{-0.045}, e^{0.823}\right) = (0.95599\ldots, 2.27732\ldots)$$

$$\cancel{\simeq (0.956, 2.278)}$$

$$\simeq (0.956, 2.277)$$

4) Test for Trend:

Rows and columns rearranged:

	Cases	Controls	n_i	O_i
Never	99	250	349	99
Sometimes	227	558	785	227
Often	113	224	337	113
Always	45	77	122	45

Table:

x_i	n_i	O_i
1	349	99
2	785	227
3	337	113
4	122	45
	1593	484

$N = \sum n_i = 1593$

$\sum O_i = 484$

$\sum O_i x_i = 1(99) + 2(227) + 3(113) + 4(45)$
$= 1072$

$\sum n_i x_i = 1(349) + 2(785) + 3(337) + 4(122)$
$= 3418$

$\sum n_i x_i^2 = 1^2(349) + 2^2(785) + 3^2(337) + 4^2(122)$
$= 8474$

$\bar{p} = \dfrac{\sum O_i}{N} = \dfrac{484}{1593} = 0.3038$

$\bar{x} = \dfrac{\sum n_i x_i}{N} = \dfrac{3418}{1593} = 2.1456$

Test Statistic

$$\chi^2_{trend} = \frac{\{1072 - 2.1456(484)\}^2}{0.3038(1-0.3038)\{8474 - 1593(2.1456)^2\}}$$

$$\leftarrow \boxed{\chi^2_{trend} = \frac{\{\sum O_i x_i - \bar{x}\sum O_i\}^2}{\bar{p}(1-\bar{p})\{\sum n_i x_i^2 - N\bar{x}^2\}}}$$

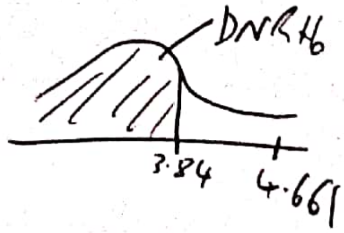
$$= \frac{1124.234076}{241.2149464}$$

$$= 4.660714\ldots$$

$$\simeq 4.661$$

Critical Value $\chi^2_{crit} = 3.84$
$v = 1$ $\quad v=1,\ 5\%$

$\Rightarrow$



Conclusion $\chi^2_{trend} > 3.84$

There is sufficient evidence of a significant ~~monotonic trend~~ linear monotonic trend between stress at work and heart attack.

There is evidence that risk of heart attack increases with increasing stress at work at the 5% level of significance.

C

1) Hazard rate for patients in control group:

$O = 19$ events

$\sum t_i = 1+2+2+3+4+4+5+5+8+8+10+11+11$
$+12+14+15+16+19+24+27+29+31+36$
$= 297$

$$\hat{\lambda} = \frac{\cancel{\Sigma} O}{\sum t_i} = \frac{19}{297} = 0.063973\ldots \simeq 0.064$$

$\hat{\lambda} \simeq 0.064$ ~~events~~ relapses per week

$\boxed{\text{95\% CI for Hazard Rate}}$

$$\frac{O \pm 1.96\sqrt{O}}{\sum t_i} = \cancel{\frac{0.064 \pm 1.96\sqrt{0.064}}{297}}$$

$\boxed{\text{Using this formula and not the logarithmic transformation for CI's.}}$

$$= \frac{19 \pm 1.96\sqrt{19}}{297}$$

$$\simeq (0.035, 0.093)$$

Median time to relapse for patients in control group:

$$\text{Median} = -\frac{1}{\lambda}\ln(0.5) = \text{mean} \times \ln(2)$$

$$= 15.625 \times \ln(2)$$

$$\simeq 10.830$$

$\boxed{\lambda = 0.064 \quad \text{mean} = \frac{1}{\lambda} = \frac{1}{0.064} \simeq 15.625}$

$\Rightarrow$ median time to relapse for patients in control group is 10.830 weeks.

$\boxed{\text{95\% CI for median}}$

$$\left(\frac{1}{\lambda_2}, \frac{1}{\lambda_1}\right) \quad \text{where} \quad \begin{array}{l}\lambda_1 = 0.035 \\ \lambda_2 = 0.093\end{array} \leftarrow \boxed{\begin{array}{l}\text{from Hazard Rate} \\ \text{95\% CI}\end{array}}$$

$$= \left(\frac{1}{0.093}, \frac{1}{0.035}\right) = (10.75268\ldots, 28.57142\ldots)$$

$$\simeq (10.753, 28.571)$$

<u>Hazard rate for patients in treated group:</u>

$O = 9$

$$\Sigma t_i = 2+6+6+6+7+9+2(10)+11+13+17+18+19+20+22$$
$$+24+25+2(32)+34+35+36$$

$$= 394$$

$$\hat{\lambda} = \frac{O}{\Sigma t_i} = \frac{9}{394} = 0.022842\ldots \simeq 0.023 \text{ relapses per week.}$$

Test for difference between treatments with regards to time to recurrence

	Test group		Control group				
Time (weeks)	At risk n_1	Recurrences d_1	At risk n_2	Recurrences d_2	n	d	E_1
1	22	0	23	1	45	1	0.489
2	22	1	22	2	44	3	1.500
3	21	0	20	1	41	1	0.512
4	21	0	19	2	40	2	1.050
5	21	0	17	2	38	2	1.105
6	21	2	15	0	36	2	1.167
7	18	1	15	0	33	1	0.545
8	17	0	15	2	32	2	1.063
10	16	1	13	1	29	2	1.103
11	14	0	12	1	26	1	0.538
12	13	0	10	1	23	1	0.565
13	13	1	9	0	22	1	0.591
14	12	0	9	1	21	1	0.571
15	12	0	8	1	20	1	0.600
18	11	1	6	0	17	1	0.647
19	10	0	6	1	16	1	0.625
22	8	1	5	0	13	1	0.615
24	7	1	5	1	12	2	1.167
27	5	0	4	1	9	1	0.556
31	5	0	2	1	7	1	0.714

Log rank test

$E_1 = 0.489 + 1.500 + 0.512 + 1.050 + 1.105 + \dots + 0.714$

$= 15.723$ ← Using table

$E_2 = E - E_1 = 28 - 15.723 = 12.277$

$O_1 = 9$
$O_2 = 19$ $\Rightarrow E = 9 + 19 = 28$

H_0: There is no significant difference in the survival functions between the two treatments.

H_1: Otherwise.

$H_0: S_1(t) = S_2(t)$
$H_1: S_1(t) \neq S_2 t$

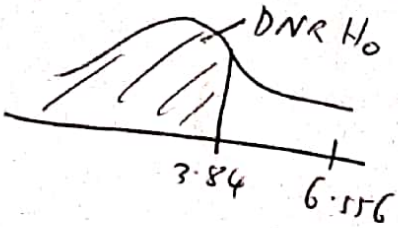
Test statistic

$$\chi^2 = \sum \frac{(O-E)^2}{E} = \frac{(O_1 - E_1)^2}{E_1} + \frac{(O_2 - E_2)^2}{E_2}$$

$$= \frac{(9 - 15.723)^2}{15.723} + \frac{(19 - 12.277)^2}{12.277} = 6.5562\ldots \simeq 6.556$$

Critical Value $\chi^2_{crit} = 3.84$

$v = n - 1 = 2 - 1 = 1$ $\quad v = 1, 5\%$



Conclusion $\chi^2 > 3.84 \Rightarrow$

There is sufficient evidence to reject the null hypothesis of no significant difference ~~between~~ in survival functions between the two treatments at the 5% level of significance.

We can conclude that there is a significant difference between the two treatments with regards to time to recurrence.

Hazard Ratio relative to Control group:

$$HR = \frac{O_1/E_1}{O_2/E_2} = \frac{9/15.723}{19/12.277} = \frac{9(12.277)}{15.723(19)} = 0.36986\ldots \simeq \cancel{0.36}\ 0.370$$

The chance of recurrence in treatment group is ~~0.332~~ 0.370 times lower compared to the control group ~~Chance is 63%~~

(Chance is 63% lower in treatment group compared to control group.)

95% CI for HR:

$$95\%\,CI = HR^{1 \pm 1.96 \div \sqrt{\chi^2}}$$

$$= 0.370^{1 \pm 1.96 \div \sqrt{6.556}}$$

$$\boxed{\begin{array}{l} HR = 0.370 \\ \chi^2 = 6.556 \end{array}}$$

$$= (0.17284\ldots, 0.79202\ldots)$$

$$\simeq (0.173, 0.792)$$

The chance of recurrence in the treatment group is between 0.173 and 0.792 times lower compared to the control group.

(Chance is between 20.8% and 82.7% lower in treatment group compared to control group.)