

1. The effect of age and smoking on breathing test results for workers in Texas industrial plants was investigated in a 1981 study. Workers were classified by Age, <40 and 40-59, and by their Smoking Status, whether they had never smoked, were former smokers or current smokers. The results of the breathing test classified workers as having Normal, Borderline, or Abnormal breathing patterns.

Age	Smoking Status	Breathing Test Result		
		Normal	Borderline	Abnormal
<40	Never Smoked	577	27	3
	Former Smoker	192	15	3
	Current Smoker	682	46	11
40-59	Never Smoked	164	10	4
	Former Smoker	145	15	7
	Current Smoker	245	39	27

A data frame called `breathing.df` was created and the following analysis done using S-plus:

```
> breathing.df
  Age Status      Result Cases
  1  <40  Never      Normal   577
  2  <40  Never Borderline    27
  3  <40  Never   Abnormal     3
  4  <40  Former   Normal   192
-----
18 49-50 Current   Abnormal    27

> breath.f1<-glm(Cases~.,family=poisson,data=breathing.df)
> breath.steps<-step.glm(breath.f1,scope=~.^3)
> anova(breath.steps,test="Chi")
```

```
Terms added sequentially (first to last)
      Df Deviance Resid. Df Resid. Dev    Pr(Chi)
NULL                                17    4089.237
Age  1    377.023                16    3712.214 0.000000000
Status 2    334.932               14    3377.281 0.000000000
Result 2   3245.848              12     131.433 0.000000000
Age:Status 2     55.324           10     76.109 0.000000000
Age:Result 2     50.899            8     25.210 0.000000000
Status:Result 4     21.566          4      3.644 0.000244471
```

- (a) Explain what the P-value for Age:Result indicates. Be specific about the hypothesis being tested.
- (b) Is the model selected by this analysis the maximal model? Explain.
- (c) Create an association graph based on this analysis. What can you conclude about the way these factors are related?
- (d) Can the contingency table be collapsed on any of the factors? Explain.
- (e) Why does it make sense to consider Breathing Test Result as the response for this data.
- (f) Create a table that you could use in a report to help explain the impact that Age and Smoking Status have on Breathing Test Result. Briefly comment on what you would conclude from your table.
2. A 20 year study of British male physicians noted that the proportion that died from lung cancer was 0.00140 per year for cigarette smokers and 0.00010 per year for non-smokers. The proportion who died from heart disease was 0.00669 per year for smokers and 0.00413 per year for non-smokers.
- (a) Calculate the relative risk of lung cancer for smokers compared to non-smokers. Explain what this means in terms of probabilities of dying from lung cancer.
- (b) Calculate the odds ratio of dying from heart disease for smokers relative to non-smokers (show your work). Why will the odds ratio be approximately equal to the relative risk in this situation?
- (c) Consider the reduction of deaths that would occur with an absence of smoking. Would the reduction in deaths due to heart disease be more or less than the reduction in deaths due to smoking?
3. The following data comes from a Japanese study of the effect of passive smoking on lung cancer. It summarises the results of a case-control study involving non-smoking women. The case group contains non-smoking women who developed lung cancer while the control group contains non-smoking women who did not have lung cancer. The number of women in each group who were married to smokers was recorded.

	Spouse	
	Smoked	Controls
Yes	73	188
No	21	82

- (a) Explain why the odds ratio is particularly useful for analysing data from case control studies such as this study.
- (b) Calculate the odds ratio that compares the odds that a non-smoking woman who is married to a smoker will have lung cancer with the same odds for a non-smoking woman married to a non-smoker.
- (c) Find a 95% confidence interval for odds ratio you calculated in (b). Note if  $Z \sim \text{Normal}(0, 1)$  then  $\Pr(-1.96 \leq Z \leq 1.96) = 0.95$ .
- (d) Does this data provide evidence of a link between passive smoking and lung cancer?