1.

(a) Fit the logistic regression of Failure (1 for failure) on Temperature. Report the estimated coefficients and their standard errors.

Log odds of coefficient “Temperature” is -0.1713, with a p-value of 0.04, indicating that it is statistically significant.

code input:

```
> model<-glm(Failure~Temperature,family="binomial",data=s)

> summary(model)


Call:

glm(formula = Failure ~ Temperature, family = "binomial", data = s)


Deviance Residuals:

    Min       1Q   Median       3Q      Max

-1.2125  -0.8253  -0.4706   0.5907   2.0512


Coefficients:

            Estimate Std. Error z value Pr(>|z|)

(Intercept) 10.87535    5.70291   1.907   0.0565 .

Temperature -0.17132    0.08344  -2.053   0.0400 *

---

Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1


(Dispersion parameter for binomial family taken to be 1)


    Null deviance: 28.975  on 23  degrees of freedom

Residual deviance: 23.030  on 22  degrees of freedom

AIC: 27.03
```

Number of Fisher Scoring iterations: 4

(b) Test whether the coefficient of Temperature is 0, using both Wald's and drop-in-deviance test.

<Wald's Test>

The chi-square test statistics value is 4.2, with a p-value of 0.04 and 1 degree of freedom, which shows that "Temperature" is statistical significant. Thus, the coefficient of temperature is not 0.

> wald.test(b=coef(model),Sigma=vcov(model),Terms=2)

Wald test:

Chi-squared test:

X2 = 4.2, df = 1, P(> X2) = 0.04

<Drop-in-deviance>

Null deviance = 28.975

Residual deviance = 23.030

28.975-23.030= 5.945 (difference = 5.945) = deviance is explained by the model, residual deviance: 5.945.

Finding the cut-off points at the .05 and .001 level, we have a p-value 0.014759 —> $p<0.05$

This model is significant at the 0.05 level, which we could also see from the p-value for the coefficient because it was a single variable regression.

Code input:

> 1-pchisq(5.945, df=1)

0.01475909

(c) What is the estimated logit failure probability and the estimated failure probability at 31°F?

The predicted logit: -1.1028 an the predicted probability : 0.9962

Code input:

> n.data<-data.frame(Temperature=c(31))

```
> predict(model,n.data,type="terms")
  Temperature
1    6.667223
attr(,"constant")
[1] -1.102809
> predict(model,n.data,type="response")
        1
0.9961828
```

(d) Why must the answer to part (c) be treated cautiously?

It is a prediction outside the range of the available explanatory variable values

(e) Fit the log-linear regression of Number.Incidents on Temperature. Does this model offer evidence that the number of incidents increases with decreasing temperature?

$$\log(\mu) = 6.388 + (-0.10894)Temperature$$

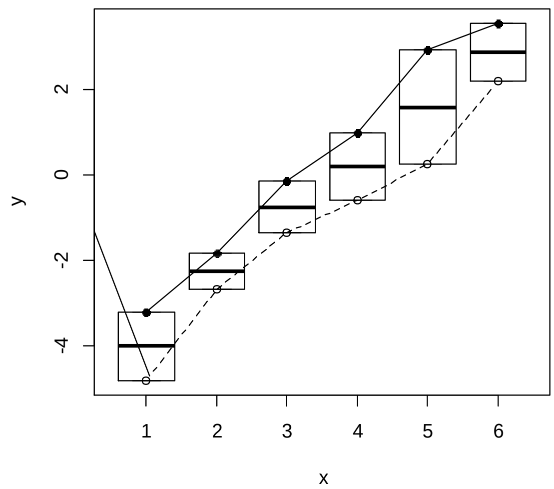
The model does serve as an evidence that the numbers of the incidents increases as the temperature decreases.

```
> model2=glm(Number.Incidents ~ Temperature, data=s, family=poisson)
> summary(model2)


Call:
glm(formula = Number.Incidents ~ Temperature, family = poisson,
    data = s)


Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.1155  -0.8158  -0.5495  -0.2731   2.4972


Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  6.38844    2.50849   2.547  0.01087 *
```

```
Temperature -0.10894    0.03937  -2.767  0.00566 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for poisson family taken to be 1)

    Null deviance: 26.874  on 23  degrees of freedom
Residual deviance: 19.290  on 22  degrees of freedom
AIC: 38.896

Number of Fisher Scoring iterations: 6

> cbind(exp(coef(model2)), exp(confint(model2)))
Waiting for profiling to be done...
                          2.5 %       97.5 %
(Intercept) 594.9306860 3.665752 82710.191332
Temperature   0.8967882 0.827404     0.968477
```

2.

(a) Plot the logits of the observed proportions versus the level of aggravation. The logit, however, is undefined for the rows where the proportion is 0 or 1, so compute the empirical logit = log[(y + 0.5) / (m – y + 0.5)] and plot this versus aggravation level, using different plotting symbols to distinguish proportions based on white and black victims.



Solid line: White ppl; dashed line: black ppl

legends:x=aggravation level, y=empirical logit of the proportion of death penalty sentences

> pportion<-log((0.5+Death)/(0.5+Nodeath))

> plot(Aggravation,pportion,type="n")

> points(Aggravation, pportion, pch=ifelse(Victim=="Black", 1, 16))

> lines(Aggravation[Victim=="Black"], pportion[Victim=="Black"], lty=2)

> lines(Aggravation[Victim=="White"], pportion[Victim=="White"], lty=1)

(b) Fit the logistic regression of death sentence proportions on aggravation level and an indicator variable for race of victim.

$$logit(\hat{\pi}) = \beta_0 + \beta_0 Aggravation + \beta_2 Victim$$

$logit(\hat{\pi}) = -3.4207 - 1.7409 Black$(for Baseline class: Aggravation 1)

Code input & output

> model1<-glm(formula=cbind(Death,Nodeath)~as.factor(Aggravation)+Victim,family=binomial,data=case1902)

> summary(model1)

Call:

glm(formula = cbind(Death, Nodeath) ~ as.factor(Aggravation) +

```
    Victim, family = binomial, data = case1902)

Deviance Residuals:
       1          2          3          4          5          6          7          8
 0.02705  -0.03705  -0.27695   0.46062  -0.22255   0.33222   0.02846  -0.03695
       9         10         11         12
 1.21437  -0.55797   0.00006   0.00007

Coefficients:
                        Estimate Std. Error z value Pr(>|z|)
(Intercept)              -3.4207     0.6144  -5.567 2.59e-08 ***
as.factor(Aggravation)2   1.6090     0.8506   1.892  0.05855 .
as.factor(Aggravation)3   3.3902     0.7474   4.536 5.74e-06 ***
as.factor(Aggravation)4   4.5004     0.7858   5.727 1.02e-08 ***
as.factor(Aggravation)5   5.8814     0.9128   6.443 1.17e-10 ***
as.factor(Aggravation)6  26.2636  8772.8073   0.003  0.99761
VictimBlack              -1.7409     0.5426  -3.208  0.00134 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 212.2838  on 11  degrees of freedom
Residual deviance:   2.2391  on  5  degrees of freedom
AIC: 38.105

Number of Fisher Scoring iterations: 19
```

(c) Report the p-value from the deviance goodness-of-fit test for this fit.

p-value=0.8151731=> there is no evidence that the model is inadequate.

```
> deviance(s)
[1] 2.239068
> s$df.residual
[1] 5
> 1-pchisq(deviance(s), s$df.residual)
[1] 0.8151731
>
```

(d) Test whether the coefficient of the indicator variable for race is equal to 0, using Wald's test.

From Wald's test of model fitted in (b), p-value=0.001, with z-statistics=-3.2. Therefore, the coffeficient of the indicator variable for race is not equal to 0.

```
> confint(mod2)
                2.5 %      97.5 %
(Intercept)    -4.8662156   -2.3805658
Aggravation2   -0.1363627    3.3534370
Aggravation3    2.0024122    5.0184128
Aggravation4    3.0636702    6.2102157
Aggravation5    4.2518697    7.8875449
Aggravation6 -277.1886527 3308.1137064
VictimBlack    -2.8840469   -0.7251278
```

(e) Construct a confidence interval for the same coefficient, and interpret it in a sentence about the odds of death sentence for white-victim murderers relative to black-victim murderers, accounting for aggravation level of crime.

The odds of death penalty for white victim murderers are estimated to be 6 times the odds of death penalty for black victim murderers with similar aggravation level(95% confidence interval).

```
> cbind(exp(coef(model1)), exp(confint(model1)))
Waiting for profiling to be done...
                                     2.5 %       97.5 %
(Intercept)             3.268960e-02  7.702460e-03 9.249823e-02
as.factor(Aggravation)2 4.997577e+00  8.725261e-01 2.860087e+01
```

```
as.factor(Aggravation)3 2.967202e+01  7.406902e+00 1.511712e+02
as.factor(Aggravation)4 9.005042e+01  2.140598e+01 4.978086e+02
as.factor(Aggravation)5 3.583025e+02  7.023661e+01 2.663896e+03
as.factor(Aggravation)6 2.547689e+11 4.154299e-121          Inf
VictimBlack             1.753671e-01  5.590805e-02 4.842627e-01
```

(f) Refit the model by treating the aggravation level as a factor. How would you interpret the results of this model?

By including the products of the race variable with the aggravation level indicators to model interaction, and determining the drop in deviance for including them. This constitutes a test of equal odds ratio assumption, which could be used to check for the Mantel-Haenszel test.

Code input & output

```
> myTable <- array(rbind(Death, Nodeath), dim=c(2,2,6),dimnames=list(Penalty=c("Death","No Death"),
Victim=c("White","Black"),Aggravation=c("1","2","3","4","5","6")))

> mantelhaen.test(myTable, alternative="greater", correct=FALSE) # 1-sided p-value


        Mantel-Haenszel chi-squared test without continuity correction


data:  myTable
Mantel-Haenszel X-squared = 11.26, df = 1, p-value = 0.000396
alternative hypothesis: true common odds ratio is greater than 1
95 percent confidence interval:
 2.264218      Inf
sample estimates:
common odds ratio
          5.49258


> mantelhaen.test(myTable, alternative="greater") # with continuity correction


        Mantel-Haenszel chi-squared test with continuity correction
```

data: myTable

Mantel-Haenszel X-squared = 9.6983, df = 1, p-value = 0.0009222

alternative hypothesis: true common odds ratio is greater than 1

95 percent confidence interval:

 2.264218   Inf

sample estimates:

common odds ratio

        5.49258

> mantelhaen.test(myTable) # two.sided (default) for confidence interval

Mantel-Haenszel chi-squared test with continuity correction

data: myTable

Mantel-Haenszel X-squared = 9.6983, df = 1, p-value = 0.001844

alternative hypothesis: true common odds ratio is not equal to 1

95 percent confidence interval:

  1.910687 15.789312

sample estimates:

common odds ratio

        5.49258

> mod1<-
glm(formula=cbind(Death,Nodeath)~Aggravation+Victim+Victim:Aggravation,family=binomial,data=case1902)

> summary(mod1)

Call:

```
glm(formula = cbind(Death, Nodeath) ~ Aggravation + Victim +
    Victim:Aggravation, family = binomial, data = case1902)

Deviance Residuals:
 [1]  0  0  0  0  0  0  0  0  0  0  0  0

Coefficients:
                         Estimate Std. Error z value Pr(>|z|)
(Intercept)            -3.401e+00  7.188e-01  -4.732 2.23e-06 ***
Aggravation2            1.386e+00  1.041e+00   1.332 0.182890
Aggravation3            3.247e+00  9.090e-01   3.572 0.000354 ***
Aggravation4            4.500e+00  9.804e-01   4.590 4.43e-06 ***
Aggravation5            2.943e+01  9.061e+04   0.000 0.999741
Aggravation6            3.000e+01  8.785e+04   0.000 0.999728
VictimBlack            -1.797e+00  1.234e+00  -1.457 0.145185
Aggravation2:VictimBlack  7.677e-01  1.771e+00   0.433 0.664673
Aggravation3:VictimBlack  4.474e-01  1.563e+00   0.286 0.774698
Aggravation4:VictimBlack  5.540e-03  1.648e+00   0.003 0.997318
Aggravation5:VictimBlack -2.394e+01  9.061e+04   0.000 0.999789
Aggravation6:VictimBlack  5.626e-01  1.314e+05   0.000 0.999997
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 2.1228e+02  on 11  degrees of freedom
Residual deviance: 2.6206e-10  on  0  degrees of freedom
AIC: 45.866
```

Number of Fisher Scoring iterations: 23