Practice questions

1. For the 23 space shuttle flights that occurred before the Challenger mission disaster in 1986, the data below shows the temperature in fahrenheit at the time of the flight and whether at least one primary O-ring suffered thermal distress (Yes=1 and 0=No)

```
Flight Temperature (x) ThermalDistress (y)
 1          66              0
 2          70              1
 3          69              0
 4          68              0
 5          67              0
 6          72              0
 7          73              0
 8          70              0
 9          57              1
 10         63              1
 11         70              1
 12         78              0
 13         67              0
 14         53              1
 15         67              0
 16         75              0
 17         70              0
 18         81              0
 19         76              0
 20         79              0
 21         75              1
 22         76              0
 23         58              1
```

A logistic regression model

$$\text{logit}(\pi(x)) = \beta_0 + \beta_1 x$$

where $\pi(x) = P(y = 1|x)$ was used to analyze these data and the the result is

```
> fit<-glm(y~x, family=binomial)
> fit

Call:  glm(formula = y ~ x, family = binomial)

Coefficients:
(Intercept)            x
    15.0429      -0.2322

Degrees of Freedom: 22 Total (i.e. Null);  21 Residual
```

```
Null Deviance:     28.27
Residual Deviance: 20.32  AIC: 24.32
```

(a) Give the estimated logistic regression equation and interpret b_1 the estimate of the slope β_1
Answer:

$$logit(\hat{\pi}(x)) = 15.0429 - 0.2322x$$

Interpretation of b_1: if the temperature increases by on degree, the odds of a fatal accident change by a multiplicative factor of about 0.7920 (or they decrease by about 20.80%)

(b) Estimate the $\pi(65)$ using a 95% confidence interval if the estimated covariance of b_0 and b_1 is given by

```
> vcov(fit)
            (Intercept)           x
(Intercept)  54.4441826 -0.79638547
x            -0.7963855  0.01171512
```

Answer: Recall that

$$\pi(65) = \frac{e^{\beta_0+\beta_1(65)}}{1+e^{\beta_0+\beta_1(65)}}$$

We need to find first a 95% confidence interval for $\beta_0 + \beta_1 65$. We have

$$\begin{aligned} var(b_0 + b_1(65)) &= var(b_0) + 2(65)cov(b_0, b_1) + (65)^2 var(b_1) \\ &= var(b_0) + 130cov(b_0, b_1) + 4225var(b_1) \end{aligned}$$

An estimate of this variance is then given by

$$\widehat{var}(b_0 + b_1(65)) = 54.4441826 + 130(-0.79638547) + 4225(0.01171512) = 0.4104535.$$

Therefore a 95% confidence interval for

$$b_0 + b_1 65 \pm 1.96\sqrt{0.4104535} = [-1.305806, 1.205606]$$

and a 95% for $\pi(65)$ is then given by

$$\left[\frac{e^{-1.305806}}{1+e^{-1.305806}}, \frac{e^{1.205606}}{1+e^{1.205606}}\right] = [0.2131895, 0.7695206]$$

2. In a logistic regression with five predictors it is found that the residual deviances of the following models

$$\begin{aligned} \text{logit}(\pi(x_1, x_2, x_3, x_4, x_5)) &= \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 \\ \text{logit}(\pi(x_1, x_2, x_3)) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 \end{aligned}$$

are 14.60 and 21.70, respectively. Test

$$\begin{aligned} H_0 &: \text{logit}(\pi(x)) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 \quad \text{against} \\ H_a &: \text{logit}(\pi(x)) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 \end{aligned}$$

using $\alpha = 0.05$.
Answer:
The value of the test statistic is this case is 21.70-14.60=7.10. The degrees of freedom equals 3. The cut off point is $\chi^2_{0.05}(3) = 7.814728$. Since $7.10 < 7.814728$, we fail to reject H_0.