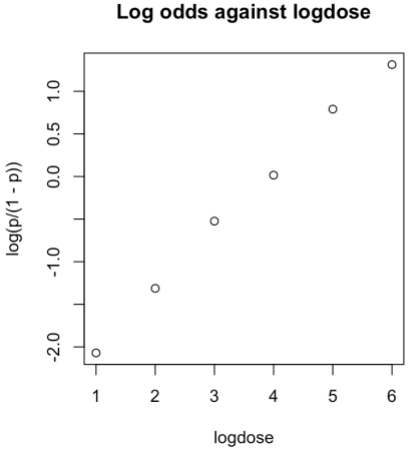
Yunlu Li
STAT 5120
Homework 9

1. (a) Logistic regression is appropriate because log odds and log dose seem to have linear relationship.



(b) $\log\left(\frac{\pi}{1-\pi}\right) = -2.64367 + 0.67399 * logdose$

```
Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept) -2.64367    0.15610  -16.93   <2e-16 ***
logdose      0.67399    0.03911   17.23   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 383.0695  on 5  degrees of freedom
Residual deviance:   1.4491  on 4  degrees of freedom
AIC: 39.358

Number of Fisher Scoring iterations: 3
```

(c) The estimated log odds of death increases by 0.67399 if log dose level increases by 1.
(d) estimated odds $= e^{-2.64367+0.67399*2} = 0.27371$
(e) $\pi = \frac{e^{-2.64367+0.67399*2}}{[1+ e^{-2.64367+0.67399*2}]} = 0.21489$
(f) $0.27371 * e^{0.67399} = 0.53703$
(g) 95% CI $= 0.67399 \pm 1.96 * 0.03911 = (0.5973344, 0.7506456)$
By exponentiating the limits of CI, we have (1.81726822823, 2.1183671931).
We are 95% confident that odds of death increases by a multiplicative factor with values between 1.81726822823 and 2.1183671931 if log dose level increases by 1.
(h) H_0: Our model is reasonably adequate; H_a: Our model does not fit data well.
Using Pearson's test, we have X^2= 1.451786. The associated p-value is 0.8351462, so we cannot reject the null. We conclude our model is reasonably adequate.
Using deviance goodness, we have G^2=1.4491. The associated p-value is 0.8356191, so we cannot reject the null. We conclude our model is reasonably adequate.

2. (a) The estimated log odds of a client ever receiving a flu shot for males is 0.43397 higher than for females, for given age and health awareness.

(b) Z = 0.43397/0.52179 = 0.83169. The p-value is 0.40540 > 0.05, so we cannot reject the null. We conclude that gender is not significant in predicting the probability of getting a flu shot, for given age and health awareness.

(c) H_0: $\beta_1 = \beta_3 = 0$; H_a: at least one of β_1 and β_3 is non zero.
Test statistics: $\Delta G^2 = 113.2 - 105.09 = 8.11$
The associated p-value is 0.01733548. which is less than 0.05. Thus, we reject the null.
We conclude that we cannot drop both age and gender from the model.