# STAC51 TUT02

Mar 18, 2021

(Azen and Walker) In the table below $x$ represents a continuous predictor and $\pi$ represents a probability. (for example $x$ may be assumed to be a measure of pollution levels and $\pi$ may be the probability of getting a disease which is believed to due to pollutants).

| $x$ | $\pi$ |
|---|---|
| 3 | 0.03 |
| 4 | 0.06 |
| 5 | 0.12 |
| 6 | 0.23 |
| 7 | 0.40 |
| 8 | 0.60 |
| 9 | 0.77 |
| 10 | 0.88 |
| 11 | 0.94 |
| 12 | 0.97 |

```
## eg1

x <- 3:12
pi <- c(0.03,0.06,0.12,0.23,0.40,0.60,0.77,0.88,0.94,0.97)
```

(a) (3 points) Create a scatter plot of the data (with values of $x$ on the horizontal axis and values of $\pi$ on the vertical axis). Describe the relationship between $x$ and $\pi$. Is it linear throughout?

(b) (3 points) Plot $\log(\pi)$ versus $x$. Describe the relationship between $x$ and $\log(\pi)$. Is it linear throughout?

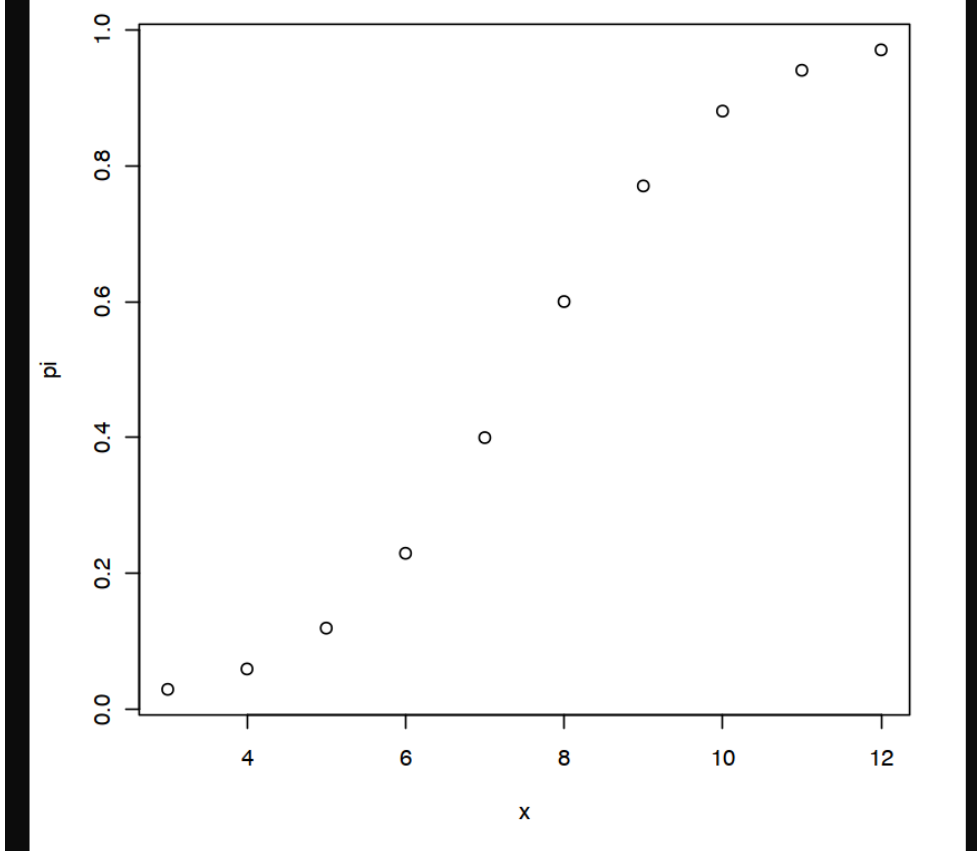
(c) (3 points) Plot $\text{logit}(\pi)$ versus $x$. Describe the relationship between $x$ and $\text{logit}(\pi)$. Is it linear throughout? Calculate the correlation between $x$ and $\text{logit}(\pi)$

(d) (3 points) Plot $\Phi^{-1}(\pi)$ versus $x$ where $\Phi$ is the standard Normal c.d.f. Describe the relationship between $x$ and $\Phi^{-1}(\pi)$. Is it linear throughout? Calculate the correlation between $x$ and $\Phi^{-1}(\pi)$.

(e) (2 points) Which of the above GLMs best fits this data? Given reasons.

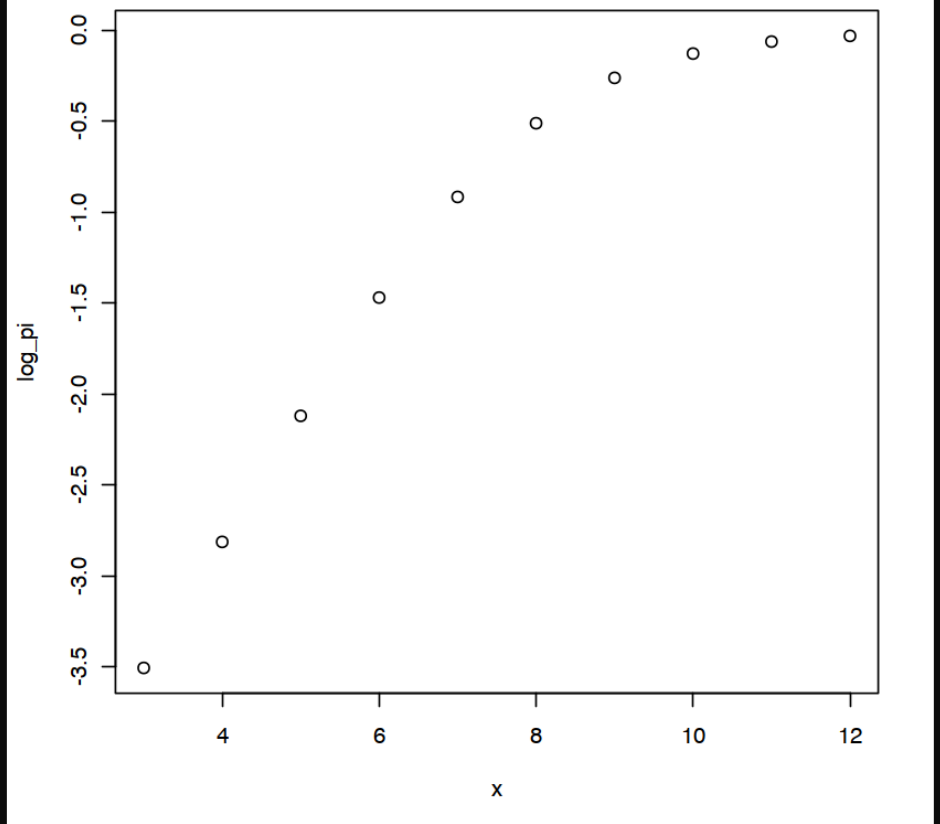
(3 points) Create a scatter plot of the data (with values of $x$ on the horizontal axis and values of $\pi$ on the vertical axis). Describe the relationship between $x$ and $\pi$. Is it linear throughout?

```r
#(a)
plot(x,pi)

# A positive relationship but nonlinear.
```



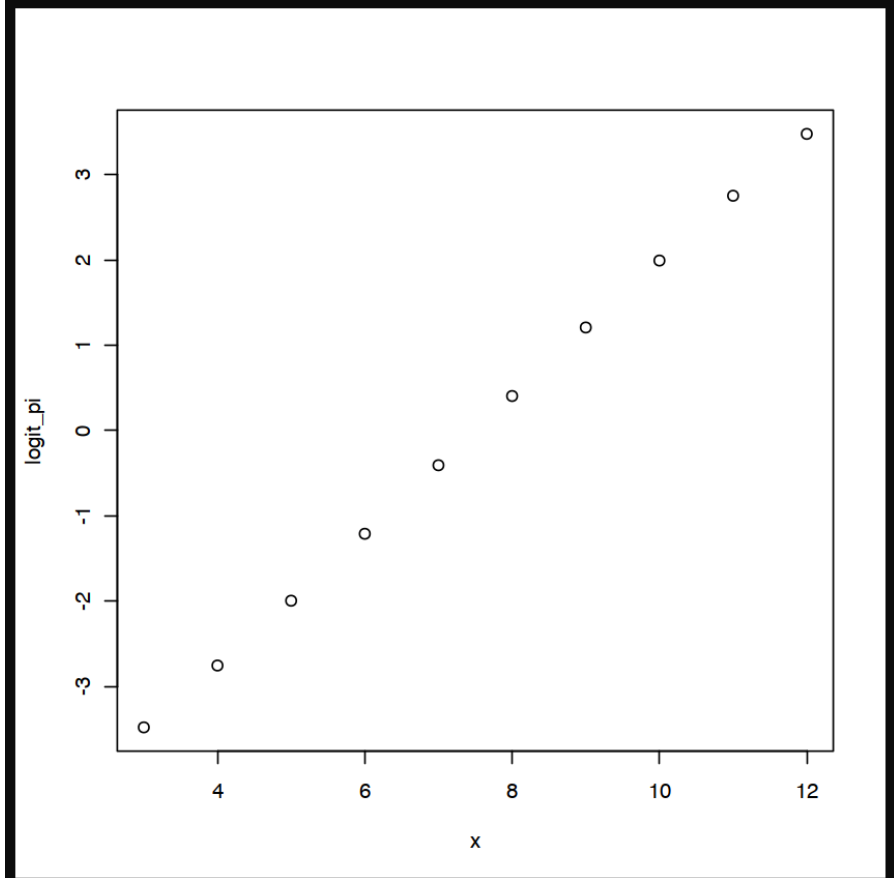
(3 points) Plot $\log(\pi)$ versus $x$. Describe the relationship between $x$ and $\log(\pi)$. Is it linear throughout?

```
# (b)
log_pi <- log(pi)
plot(x, log_pi)

#A positive relationship but nonlinear
```



(3 points) Plot logit($\pi$) versus $x$. Describe the relationship between $x$ and logit($\pi$). Is it linear throughout? Calculate the correlation between $x$ and logit($\pi$)

```
# (c)
logit_pi <- log(pi/(1-pi))
plot(x, logit_pi)

# A positive relationship. Very close to linear
```

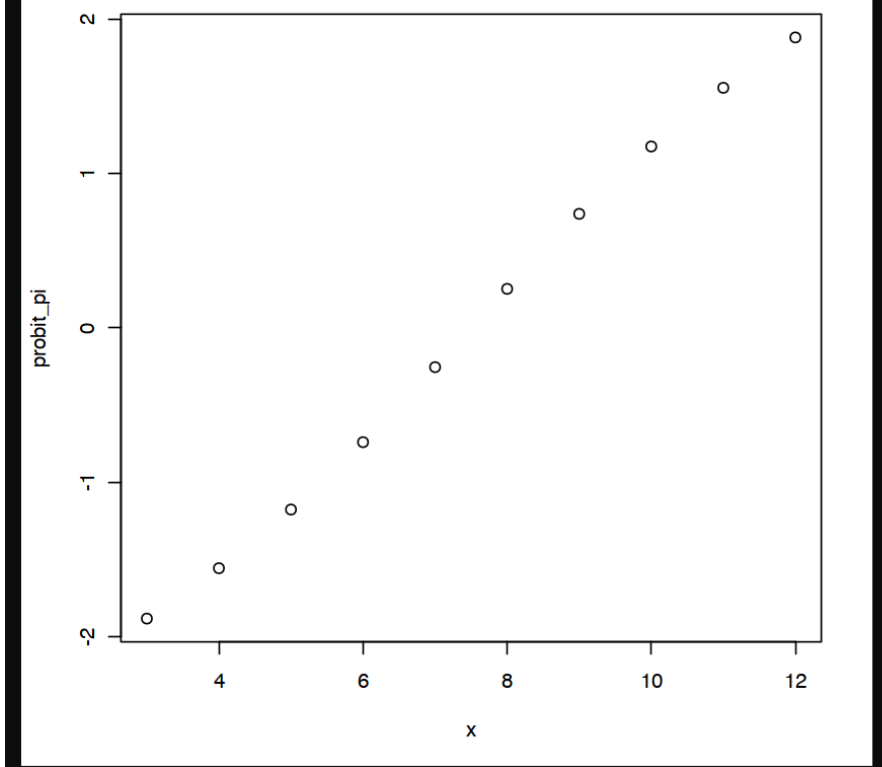



```
cor(x, logit_pi)
```

```
0.999902847008527
```

(d) (3 points) Plot $\Phi^{-1}(\pi)$ versus $x$ where $\Phi$ is the standard Normal c.d.f. Describe the relationship between $x$ and $\Phi^{-1}(\pi)$. Is it linear throughout? Calculate the correlation between $x$ and $\Phi^{-1}(\pi)$.

```r
# (d)
probit_pi <- qnorm(pi)
plot(x, probit_pi)
# A positive relationship. Very close to linear.
```




```r
cor(x, probit_pi)
```

```
0.998558798395567
```

Which of the above GLMs best fits this data? Given reasons.

The table below shows the data (source: Agresti, Collectt ) from a study about $y =$ whether a patient having surgery experienced a sore throat on waking ($1 =$ yes, $0 =$ no) as a function of $d =$ duration of the surgery (in minutes) and $t =$ type of device used to secure the airway ($1 =$ tracheal tube, $0 =$ laryngeal mask airway).

| Patient | d | t | y |
|---|---|---|---|
| 1 | 45 | 0 | 0 |
| 2 | 15 | 0 | 0 |
| 3 | 40 | 0 | 1 |
| 4 | 83 | 1 | 1 |
| 5 | 90 | 1 | 1 |
| 6 | 25 | 1 | 1 |
| 7 | 35 | 0 | 1 |
| 8 | 65 | 0 | 1 |
| 9 | 95 | 0 | 1 |
| 10 | 35 | 0 | 1 |
| 11 | 75 | 0 | 1 |
| 12 | 45 | 1 | 1 |
| 13 | 50 | 1 | 0 |
| 14 | 75 | 1 | 1 |
| 15 | 30 | 0 | 0 |
| 16 | 25 | 0 | 1 |
| 17 | 20 | 1 | 0 |
| 18 | 60 | 1 | 1 |
| 19 | 70 | 1 | 1 |
| 20 | 30 | 0 | 1 |
| 21 | 60 | 0 | 1 |
| 22 | 61 | 0 | 0 |
| 23 | 65 | 0 | 1 |
| 24 | 15 | 1 | 0 |
| 25 | 20 | 1 | 0 |
| 26 | 45 | 0 | 1 |
| 27 | 15 | 1 | 0 |
| 28 | 25 | 0 | 1 |
| 29 | 15 | 1 | 0 |
| 30 | 30 | 0 | 1 |
| 31 | 40 | 0 | 1 |
| 32 | 15 | 1 | 0 |
| 33 | 135 | 1 | 1 |
| 34 | 20 | 1 | 0 |
| 35 | 40 | 1 | 0 |

1.Fit a main effects model using these predictors. Interpret parameter estimates.

2.Conduct inference about the D effect in (1).

3.Fit a model permitting interaction. Report the prediction equation for the effect of D when (i) T = 1, (ii) T = 0. Interpret.

4.Conduct inference about whether you need the interaction term in (c).

5. Overlay the fitted logistic curves corresponding to the main effects model in the left-hand plot and the inter-action model in the right-hand plot, and using different line types for T= 0 versusT= 1. Include legends on your plots

**1.Fit a main effects model using these predictors. Interpret parameter estimates. Conduct inference about the D effect**

```
# Q1
summary(model1)

Call:
glm(formula = Y ~ D + T, family = binomial, data = SoreThroat)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.3802  -0.5358   0.3047   0.7308   1.7821

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept) -1.41734    1.09457  -1.295  0.19536
D            0.06868    0.02641   2.600  0.00931 **
T           -1.65895    0.92285  -1.798  0.07224 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 46.180  on 34  degrees of freedom
Residual deviance: 30.138  on 32  degrees of freedom
AIC: 36.138

Number of Fisher Scoring iterations: 5
```

**Fit a model permitting interaction. Report the prediction equation for the effect of D when (i) T = 1, (ii) T = 0. Interpret.**

```
# (3)
model2 <- glm(Y~D+T+D:T, data=SoreThroat, family=binomial)
summary(model2)
```

```
Call:
glm(formula = Y ~ D + T + D:T, family = binomial, data = SoreThroat)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.9707  -0.3779   0.3448   0.7292   1.9961

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  0.04979    1.46940   0.034   0.9730
D            0.02848    0.03429   0.831   0.4062
T           -4.47224    2.46707  -1.813   0.0699 .
D:T          0.07460    0.05777   1.291   0.1966
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 46.180  on 34  degrees of freedom
Residual deviance: 28.321  on 31  degrees of freedom
AIC: 36.321

Number of Fisher Scoring iterations: 6
```

**Conduct inference about whether you need the interaction term in (c).**

```
anova(model1, model2, test="Chisq")
```

A anova: 2 × 5

| | Resid. Df | Resid. Dev | Df | Deviance | Pr(>Chi) |
|---|---|---|---|---|---|
| | <dbl> | <dbl> | <dbl> | <dbl> | <dbl> |
| 1 | 32 | 30.13794 | NA | NA | NA |
| 2 | 31 | 28.32105 | 1 | 1.816886 | 0.1776844 |

```
drop1(model2,test="Chisq")
```

A anova: 2 × 5

| | Df | Deviance | AIC | LRT | Pr(>Chi) |
|---|---|---|---|---|---|
| | <dbl> | <dbl> | <dbl> | <dbl> | <dbl> |
| <none> | NA | 28.32105 | 36.32105 | NA | NA |
| D:T | 1 | 30.13794 | 36.13794 | 1.816886 | 0.1776844 |

```r
# (5)
x <- range(SoreThroat$D)
x <- seq(x[1], x[2])
par(mfrow=c(1,2)); set.seed(111);
plot(jitter(Y,.2) ~ D, pch=2-T, data=SoreThroat, ylab="P(SoreThroat)",xlab="Duration", main="Main effects model")
curve(predict(model1, data.frame(D=x,T=1), type="response"), lty=1, add=T)
curve(predict(model1, data.frame(D=x,T=0), type="response"), lty=2, add=T)
legend("bottomright", pch=1:2, lty=1:2,legend=c("Tracheal tube", "Laryngeal mask"),cex = 0.6)
plot(jitter(Y,.2) ~ D, pch=2-T, data=SoreThroat, ylab="P(SoreThroat)",xlab="Duration", main="Interaction model")
curve(predict(model2, data.frame(D=x,T=1), type="response"), lty=1, add=T)
curve(predict(model2, data.frame(D=x,T=0), type="response"), lty=2, add=T)
legend("bottomright", pch=1:2, lty=1:2,legend=c("Tracheal tube", "Laryngeal mask"),cex = 0.6)
```

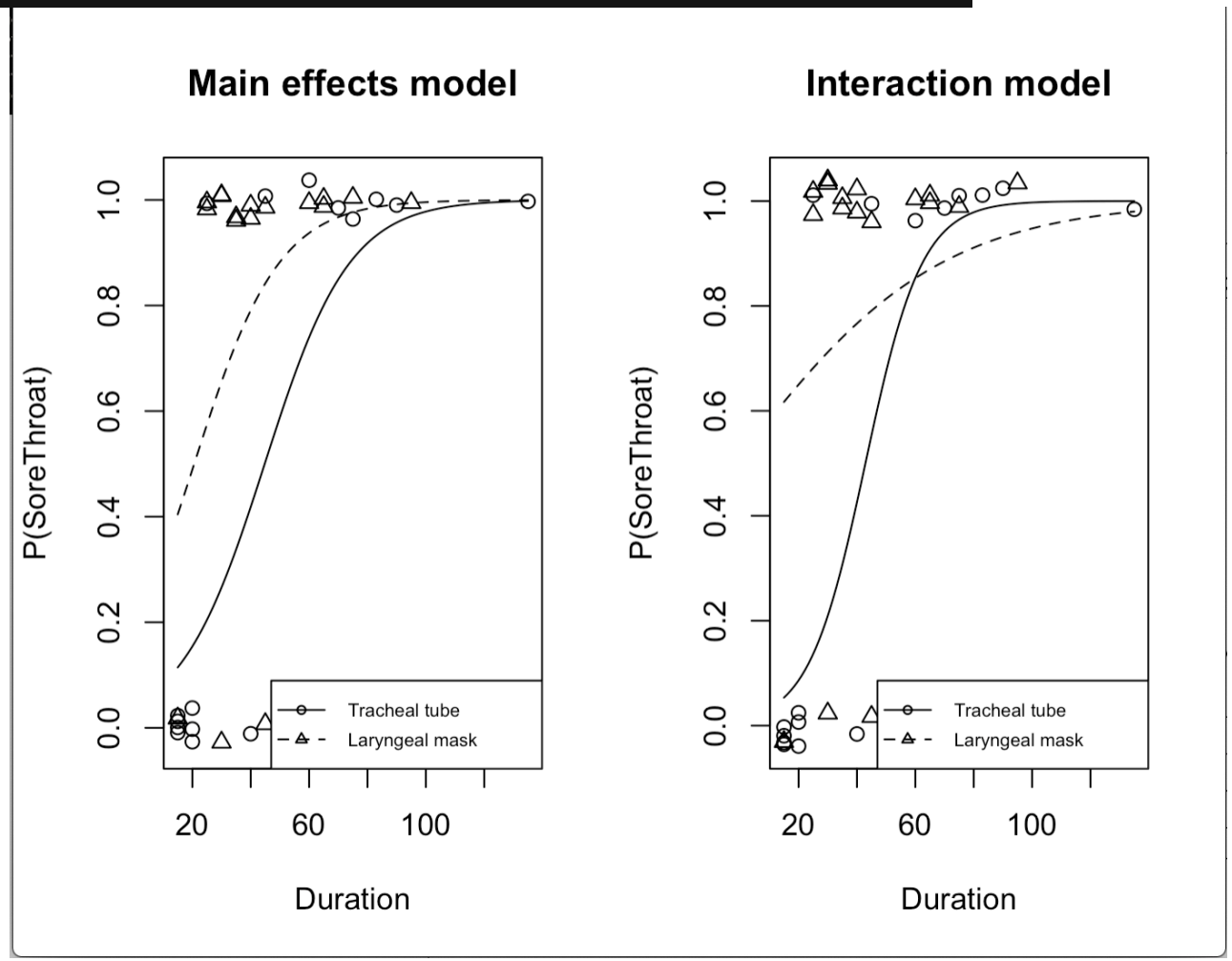


Table 4.18 shows estimated effects for a fitted logistic regression model with squamous cell esophageal cancer (1 = yes, 0 = no) as the response variable $Y$. Smoking status ($S$) equals 1 for at least one pack per day and 0 otherwise, alcohol consumption ($A$) equals the average number of alcoholic drinks consumed per day, and race ($R$) equals 1 for blacks and 0 for whites.

**a.** To describe the race-by-smoking interaction, construct the prediction equation when $R = 1$ and again when $R = 0$. Find the fitted $YS$ conditional odds ratio for each case. Similarly, construct the prediction equation when $S = 1$ and again when $S = 0$. Find the fitted $YR$ conditional odds ratio for each case. Note that, for each association, the coefficient of the cross-product term is the difference between the log odds ratios at the two fixed levels for the other variable.

**Table 4.18. Table for Problem 4.23 on Effects on Esophageal Cancer**

| Variable | Effect | $P$-value |
|---|---|---|
| Intercept | −7.00 | <0.01 |
| Alcohol use | 0.10 | 0.03 |
| Smoking | 1.20 | <0.01 |
| Race | 0.30 | 0.02 |
| Race × smoking | 0.20 | 0.04 |

**b.** In Table 4.18, explain what the coefficients of $R$ and $S$ represent, for the coding as given above. What hypotheses do the $P$-values refer to for these variables?

**Table 2.7. Infant Malformation and Mother's Alcohol Consumption**

| Alcohol Consumption | Malformation | | Total | Percentage Present | Standardized Residual |
|---|---|---|---|---|---|
| | Absent | Present | | | |
| 0 | 17,066 | 48 | 17,114 | 0.28 | −0.18 |
| <1 | 14,464 | 38 | 14,502 | 0.26 | −0.71 |
| 1–2 | 788 | 5 | 793 | 0.63 | 1.84 |
| 3–5 | 126 | 1 | 127 | 0.79 | 1.06 |
| ≥6 | 37 | 1 | 38 | 2.63 | 2.71 |

*Source*: B. I. Graubard and E. L. Korn, *Biometrics*, **43**: 471–476, 1987. Reprinted with permission from the Biometric Society.

Refer to Table 2.7 on mother's drinking and infant malformations.

**a.** Fit the logistic regression model using scores {0, 0.5, 1.5, 4, 7} for alcohol consumption. Check goodness of fit.

**b.** Test independence using the likelihood-ratio test for the model in (**a**). (The trend test of Section 2.5.1 is the score test for this model.)

**c.** The sample proportion of malformations is much higher in the highest alcohol category because, although it has only one malformation, its sample size is only 38. Are the results sensitive to this single observation? Re-fit the model without it, entering 0 malformations for 37 observations, and compare the results of the likelihood-ratio test. (Because results are sensitive to a single observation, it is hazardous to make conclusions, even though $n$ was extremely large.)

**d.** Fit the model and conduct the test of independence for all the data using scores {1, 2, 3, 4, 5}. Compare the results with (**b**). (Results for highly unbalanced data can be sensitive to the choice of scores.)

```
#
mydata = data.frame(drinks = c(0,0.5,1.5,4,7),
                    absent = c(17066, 14464, 788, 126, 37),
                    present = c(48, 38, 5, 1, 1) )
mydata$total = with(mydata, absent + present)
mydata$proportion = with(mydata, present/total)

glm.logit = glm(proportion~drinks, family=binomial, weight=total, data=mydata)
summary(glm.logit)
```

```
Call:
glm(formula = proportion ~ drinks, family = binomial, data = mydata,
    weights = total)

Deviance Residuals:
      1       2       3       4       5
 0.5921  -0.8801   0.8865  -0.1449   0.1291

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  -5.9605     0.1154 -51.637   <2e-16 ***
drinks        0.3166     0.1254   2.523   0.0116 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 6.2020  on 4  degrees of freedom
Residual deviance: 1.9487  on 3  degrees of freedom
AIC: 24.576

Number of Fisher Scoring iterations: 4
```

```
drop1(glm.logit,test = "Chisq")
```

A anova: 2 × 5

| | Df | Deviance | AIC | LRT | Pr(>Chi) |
|---|---|---|---|---|---|
| | <dbl> | <dbl> | <dbl> | <dbl> | <dbl> |
| **<none>** | NA | 1.948721 | 24.57552 | NA | NA |
| **drinks** | 1 | 6.201998 | 26.82880 | 4.253277 | 0.03917467 |

```r
mydata_v1 = data.frame(drinks = c(0,0.5,1.5,4,7),
                       absent = c(17066, 14464, 788, 126, 37),
                       present = c(48, 38, 5, 1, 0) )
mydata_v1$total = with(mydata_v1, absent + present)
mydata_v1$proportion = with(mydata_v1, present/total)

glm.logit_v1 = glm(proportion~drinks, family=binomial, weight=total, data=mydata_v1)
summary(glm.logit_v1)
```

```
Call:
glm(formula = proportion ~ drinks, family = binomial, data = mydata_v1,
    weights = total)

Deviance Residuals:
      1        2        3        4        5
 0.3232  -0.6893   1.2065   0.3509  -0.8279

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  -5.9207     0.1188  -49.85   <2e-16 ***
drinks        0.1776     0.1709    1.04    0.299
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 3.7225  on 4  degrees of freedom
Residual deviance: 2.8438  on 3  degrees of freedom
AIC: 23.497

Number of Fisher Scoring iterations: 5
```

```r
drop1(glm.logit_v1,test = "Chisq")
```

A anova: 2 × 5

| | Df | Deviance | AIC | LRT | Pr(>Chi) |
|---|---|---|---|---|---|
| | <dbl> | <dbl> | <dbl> | <dbl> | <dbl> |
| **<none>** | NA | 2.843807 | 23.49716 | NA | NA |
| **drinks** | 1 | 3.722519 | 22.37587 | 0.8787127 | 0.3485545 |

```r
mydata_v2 = data.frame(drinks = 1:5,
                       absent = c(17066, 14464, 788, 126, 37),
                       present = c(48, 38, 5, 1, 1) )
mydata_v2$total = with(mydata_v2, absent + present)
mydata_v2$proportion = with(mydata_v2, present/total)

glm.logit_v2 = glm(proportion~drinks, family=binomial, weight=total, data=mydata_v2)
summary(glm.logit_v2)
```

```
Call:
glm(formula = proportion ~ drinks, family = binomial, data = mydata_v2,
    weights = total)

Deviance Residuals:
      1        2        3        4        5
 0.7302  -1.1983   0.9636   0.4272   1.1692

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  -6.2089     0.2873 -21.612   <2e-16 ***
drinks        0.2278     0.1683   1.353    0.176
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 6.2020  on 4  degrees of freedom
Residual deviance: 4.4473  on 3  degrees of freedom
AIC: 27.074

Number of Fisher Scoring iterations: 5
```

```r
drop1(glm.logit_v2,test = "Chisq")
```

A anova: 2 × 5

| | Df | Deviance | AIC | LRT | Pr(>Chi) |
|---|---|---|---|---|---|
| | <dbl> | <dbl> | <dbl> | <dbl> | <dbl> |
| **<none>** | NA | 4.447316 | 27.07412 | NA | NA |
| **drinks** | 1 | 6.201998 | 26.82880 | 1.754682 | 0.1852892 |

The following are true–false questions.

**a.** A model for a binary response has a continuous predictor. If the model truly holds, the deviance statistic for the model has an asymptotic chi-squared distribution as the sample size increases. It can be used to test model goodness of fit.

**b.** For the horseshoe crab data, when width or weight is the sole predictor for the probability of a satellite, the likelihood-ratio test of the predictor effect has $P$-value $<0.0001$. When both weight and width are in the model, it is possible that the likelihood-ratio tests for the partial effects of width and weight could both have $P$-values larger than 0.05.

**c.** For the model, $\text{logit}[\pi(x)] = \alpha + \beta x$, suppose $y = 1$ for all $x \leq 50$ and $y = 0$ for all $x > 50$. Then, the ML estimate $\hat{\beta} = -\infty$.

# Textbook: 3.4.3, 5.2.3

$i$ of snoring.

Because the saturated model has additional parameters, its maximized log likelihood $L_S$ is at least as large as the maximized log likelihood $L_M$ for a simpler model $M$. The *deviance* of a GLM is defined as

$$\text{Deviance} = -2[L_M - L_S]$$

The deviance is the likelihood-ratio statistic for comparing model $M$ to the saturated model. It is a test statistic for the hypothesis that all parameters that are in the saturated model but not in model $M$ equal zero. GLM software provides the deviance, so it is not necessary to calculate $L_M$ or $L_S$.

For some GLMs, the deviance has approximately a chi-squared distribution. For example, in Section 5.2.2 we will see this happens for binary GLMs with a fixed number of explanatory levels in which each observation is a binomial variate having relatively large counts of successes and failures. For such cases, the deviance



provides a goodness-of-fit test of the model, because it tests the hypothesis that all possible parameters not included in the model equal 0. The residual $df$ equals the number of observations minus the number of model parameters. The $P$-value is the right-tail probability above the observed test statistic value, from the chi-squared distribution. Large test statistics and small $P$-values provide strong evidence of model lack of fit.

When calculated for logistic regression models fitted with continuous or nearly continuous predictors, the $X^2$ and $G^2$ statistics do *not* have approximate chi-squared distributions. How can we check the adequacy of a model for such data? One way creates categories for each predictor (e.g., four categories according to where a value falls relative to the quartiles) and then applies $X^2$ or $G^2$ to observed and fitted counts for the grouped data. As the number of explanatory variables increases, however, simultaneous grouping of values for each variable produces a contingency table with a very large number of cells. Most cells then have fitted values that are too small for the chi-squared approximation to be good.

An alternative way of grouping the data forms observed and fitted values based on a partitioning of the estimated probabilities. With 10 groups of equal size, the first pair of observed counts and corresponding fitted counts refers to the $n/10$ observations having the highest estimated probabilities, the next pair refers to the $n/10$ observations having the second decile of estimated probabilities, and so forth. Each group has an observed count of subjects with each outcome and a fitted value for each outcome. The fitted value for an outcome is the sum of the estimated probabilities for that outcome for all observations in that group.

The *Hosmer–Lemeshow test* uses a Pearson test statistic to compare the observed and fitted counts for this partition. The test statistic does not have exactly a limiting chi-squared distribution. However, Hosmer and Lemeshow (2000, pp. 147–156) noted that, when the number of distinct patterns of covariate values (for the original data) is close to the sample size, the null distribution is approximated by chi-squared with $df =$ number of groups $-2$.

```
Call:
glm(formula = y ~ width, family = binomial(link = "logit"), data = crab)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.0281  -1.0458   0.5480   0.9066   1.6942

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept) -12.3508     2.6287  -4.698 2.62e-06 ***
width         0.4972     0.1017   4.887 1.02e-06 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 225.76  on 172  degrees of freedom
Residual deviance: 194.45  on 171  degrees of freedom
AIC: 198.45

Number of Fisher Scoring iterations: 4

Call:
glm(formula = y ~ weight, family = binomial(link = "logit"),
    data = crab)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.1108  -1.0749   0.5426   0.9122   1.6285

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  -3.6947     0.8802  -4.198 2.70e-05 ***
weight        1.8151     0.3767   4.819 1.45e-06 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 225.76  on 172  degrees of freedom
Residual deviance: 195.74  on 171  degrees of freedom
AIC: 199.74

Number of Fisher Scoring iterations: 4

Call:
glm(formula = y ~ width + weight, family = binomial(link = "logit"),
    data = crab)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.1127  -1.0344   0.5304   0.9006   1.7207

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  -9.3547     3.5280  -2.652  0.00801 **
width         0.3068     0.1819   1.686  0.09177 .
weight        0.8338     0.6716   1.241  0.21445
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 225.76  on 172  degrees of freedom
Residual deviance: 192.89  on 170  degrees of freedom
AIC: 198.89

Number of Fisher Scoring iterations: 4
```

```
> set.seed(123)
> dat_c <- data.frame(y = c(rep(1,50),rep(0,50)),x = c(runif(50,0,50),runif(50,50,100)))
> summary(glm(y~x, family = binomial,data = dat_c))

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

Deviance Residuals:
       Min          1Q      Median          3Q         Max
-6.889e-04  -2.000e-08   0.000e+00   2.000e-08   7.206e-04

Coefficients:
             Estimate Std. Error z value Pr(>|z|)
(Intercept)   4774.18  270318.12   0.018    0.986
x              -95.73    5420.95  -0.018    0.986

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 1.3863e+02  on 99  degrees of freedom
Residual deviance: 9.9393e-07  on 98  degrees of freedom
AIC: 4

Number of Fisher Scoring iterations: 25

Warning messages:
1: glm.fit: algorithm did not converge
2: glm.fit: fitted probabilities numerically 0 or 1 occurred
>
```

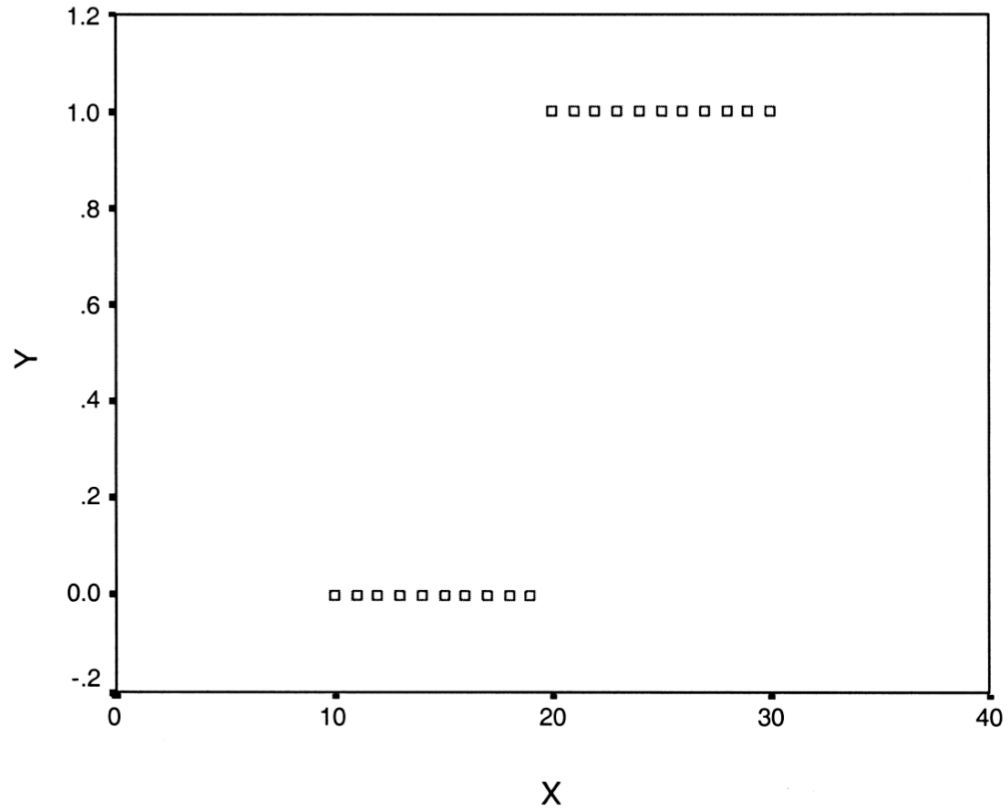



FIGURE 2. *Plot of artificial data from Ryan (1996).*

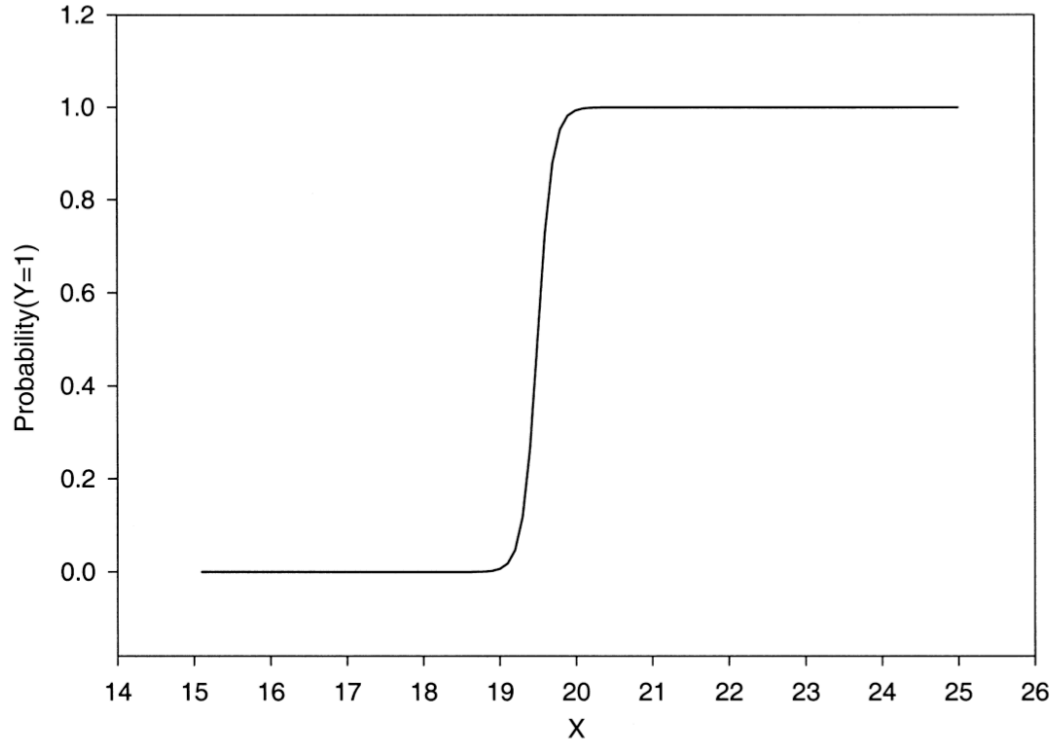



FIGURE 3. *Plot of logistic curve with slope equal to 10.*
*Note: This curve provides a nearly perfect fit to the Ryan (1996) data.*

1. Rindskopf D. Infinite Parameter Estimates in Logistic Regression: Opportunities, Not Problems. Journal of Educational and Behavioral Statistics. 2002;27(2):147-161. doi:10.3102/10769986027002147