

5.6.1 Example: Deviance for a binomial model

If the response variables Y_1, \dots, Y_N are independent and $Y_i \sim \text{binomial}(n_i, \pi_i)$, then the log-likelihood function is

$$l(\boldsymbol{\beta}; \mathbf{y}) = \sum_{i=1}^N \left[y_i \log \pi_i - y_i \log(1 - \pi_i) + n_i \log(1 - \pi_i) + \log \binom{n_i}{y_i} \right].$$

For a saturated model, the π_i 's are all different so $\boldsymbol{\beta} = [\pi_1, \dots, \pi_N]^T$. The maximum likelihood estimates are $\hat{\pi}_i = y_i/n_i$ so the maximum value of the log-likelihood function is

$$l(\mathbf{b}_{\max}; \mathbf{y}) = \sum \left[y_i \log \left(\frac{y_i}{n_i} \right) - y_i \log \left(\frac{n_i - y_i}{n_i} \right) + n_i \log \left(\frac{n_i - y_i}{n_i} \right) + \log \binom{n_i}{y_i} \right].$$

For any other model with $p < N$ parameters, let $\hat{\pi}_i$ denote the maximum likelihood estimates for the probabilities and let $\hat{y}_i = n_i \hat{\pi}_i$ denote the fitted values. Then the log-likelihood function evaluated at these values is

$$l(\mathbf{b}; \mathbf{y}) = \sum \left[y_i \log \left(\frac{\hat{y}_i}{n_i} \right) - y_i \log \left(\frac{n_i - \hat{y}_i}{n_i} \right) + n_i \log \left(\frac{n_i - \hat{y}_i}{n_i} \right) + \log \binom{n_i}{y_i} \right].$$

Therefore the deviance is

$$\begin{aligned} D &= 2[l(\mathbf{b}_{\max}; \mathbf{y}) - l(\mathbf{b}; \mathbf{y})] \\ &= 2 \sum_{i=1}^N \left[y_i \log \left(\frac{y_i}{\hat{y}_i} \right) + (n_i - y_i) \log \left(\frac{n_i - y_i}{n_i - \hat{y}_i} \right) \right]. \end{aligned}$$

53

Table 7.1 *Frequencies for N binomial distributions.*

| | Subgroups | | | |
|-----------|-------------|-------------|-----|-------------|
| | 1 | 2 | ... | N |
| Successes | Y_1 | Y_2 | ... | Y_N |
| Failures | $n_1 - Y_1$ | $n_2 - Y_2$ | ... | $n_N - Y_N$ |
| Totals | n_1 | n_2 | ... | n_N |

The deviance, derived in Section 5.6.1, is

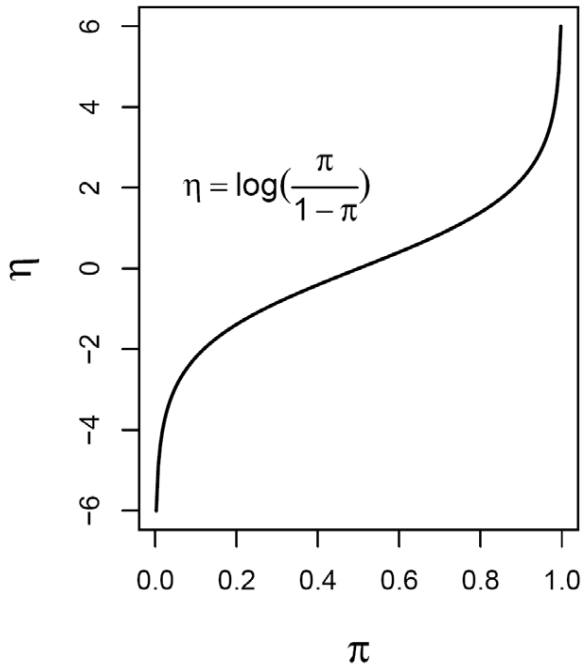
$$D = 2 \sum_{i=1}^N \left[y_i \log \left(\frac{y_i}{\hat{y}_i} \right) + (n_i - y_i) \log \left(\frac{n_i - y_i}{n_i - \hat{y}_i} \right) \right]. \quad (7.5)$$

This has the form

$$D = 2 \sum o \log \frac{o}{e}$$

54

logit function



logistic function

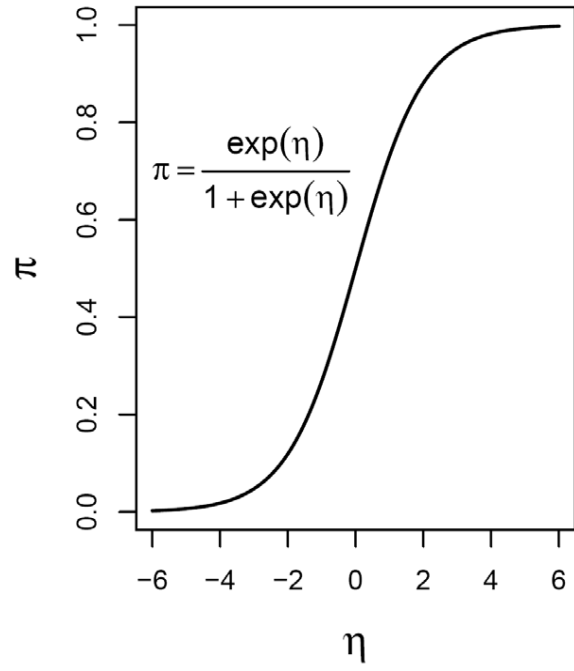
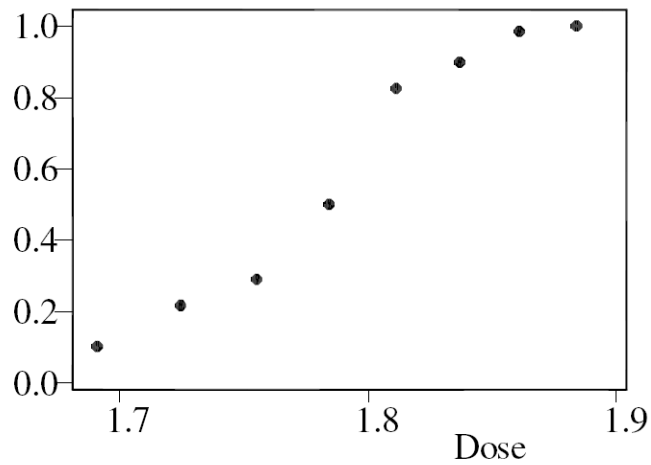


Table 7.2 *Beetle mortality data.*

| Dose, x_i ($\log_{10}\text{CS}_2\text{mg l}^{-1}$) | Number of beetles, n_i | Number killed, y_i |
|-----------------------------------------------------------|-----------------------------|-------------------------|
| 1.6907 | 59 | 6 |
| 1.7242 | 60 | 13 |
| 1.7552 | 62 | 18 |
| 1.7842 | 56 | 28 |
| 1.8113 | 63 | 52 |
| 1.8369 | 59 | 53 |
| 1.8610 | 62 | 61 |
| 1.8839 | 60 | 60 |

Proportion killed



logistic model

$$\pi_i = \frac{\exp(\beta_1 + \beta_2 x_i)}{1 + \exp(\beta_1 + \beta_2 x_i)}$$

so

$$\log\left(\frac{\pi_i}{1 - \pi_i}\right) = \beta_1 + \beta_2 x_i$$

and

$$\log(1 - \pi_i) = -\log[1 + \exp(\beta_1 + \beta_2 x_i)].$$

Therefore from equation (7.3) the log-likelihood function is

$$l = \sum_{i=1}^N \left[y_i (\beta_1 + \beta_2 x_i) - n_i \log[1 + \exp(\beta_1 + \beta_2 x_i)] + \log\left(\binom{n_i}{y_i}\right) \right]$$

and the scores with respect to β_1 and β_2 are

$$\begin{aligned} U_1 &= \frac{\partial l}{\partial \beta_1} = \sum \left\{ y_i - n_i \left[\frac{\exp(\beta_1 + \beta_2 x_i)}{1 + \exp(\beta_1 + \beta_2 x_i)} \right] \right\} = \sum (y_i - n_i \pi_i) \\ U_2 &= \frac{\partial l}{\partial \beta_2} = \sum \left\{ y_i x_i - n_i x_i \left[\frac{\exp(\beta_1 + \beta_2 x_i)}{1 + \exp(\beta_1 + \beta_2 x_i)} \right] \right\} \\ &= \sum x_i (y_i - n_i \pi_i). \end{aligned}$$

Similarly the information matrix is

$$\mathcal{J} = \begin{bmatrix} \sum n_i \pi_i (1 - \pi_i) & \sum n_i x_i \pi_i (1 - \pi_i) \\ \sum n_i x_i \pi_i (1 - \pi_i) & \sum n_i x_i^2 \pi_i (1 - \pi_i) \end{bmatrix}.$$

57

```
> dose <- c(1.6907,1.7242,1.7552,1.7842,1.8113,1.8369,1.8610,1.8839)
> ntotal <- c(59,60,62,56,63,59,62,60)
> ndead <- c(6,13,18,28,52,53,61,60)
> beetledata <- data.frame(cbind(dose,ntotal,ndead))
> beetledata
  dose ntotal ndead
1 1.6907    59     6
2 1.7242    60    13
3 1.7552    62    18
4 1.7842    56    28
5 1.8113    63    52
6 1.8369    59    53
7 1.8610    62    61
8 1.8839    60    60
> M1 <- glm(cbind(ndead,ntotal-ndead) ~ 1 + dose, data=beetledata, family=binomial(link=logit))
> summary(M1)
Call:
glm(formula = cbind(ndead, ntotal - ndead) ~ 1 + dose, family = binomial(link = logit),
    data = beetledata)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.5941  -0.3944   0.8329   1.2592   1.5940

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  -60.717     5.181  -11.72  <2e-16 ***
dose           34.270     2.912   11.77  <2e-16 ***

Null deviance: 284.202  on 7  degrees of freedom
Residual deviance:  11.232  on 6  degrees of freedom
AIC: 41.43
```

58