

Odds ratio can be estimated from *both*

- | | |
|--|--|
| Prospective sampling | Retrospective sampling |
| <ul style="list-style-type: none"> Select exposed group and comparable unexposed group. Observe outcomes over time (e.g. Disease). Exposure totals fixed; incidence of outcome random. | <ul style="list-style-type: none"> Select outcome and non-outcome group and investigate whether exposure occurred. Frequency of exposure is random. <i>Cannot estimate proportions of outcomes at each level of exposure for retrospective studies.</i> |

Deviance as analog to RSS

- Deviance residuals are based on the loglikelihood of binary response variables. (see p. 578)
- Deviance component for each observation i is
$$2[l_i(y_i) - l_i(\pi_i)]$$
- Deviance residual is square root of this quantity, and is negative if $y_i < \pi_i$, positive if $y_i > \pi_i$
- Model deviance* is the sum of squared deviance residuals, on $n-p$ degrees of freedom.

Drop in deviance Chi-Squared Test

- Compare a model with K parameters (full) to model with $K-H$ parameters (reduced)
- Test: $H_o: \beta_1 = \beta_2 = \dots = \beta_H = 0$ vs. $H_a: \beta_1, \dots, \beta_H$ free
- Test statistic is:
 $D = (\text{deviance from reduced}) - (\text{deviance from full})$ on H d.f.
- Under the null (reduced is adequate), D behaves like a χ_H^2 .
- p-value: $P(\chi_H^2 > D \mid H_o \text{ is true})$

From last class

Comparison of null model and univariate model:

- Null model (deviance = 83.18)
- owl~pctring1 (deviance = 71.10)

Drop in Deviance Test

- $D = 12.08$ on 1 d.f.
 - p-value: $P(\chi_1^2 > 12.08 \mid H_o) \approx 0.0005$ (from Tables)
- Result: Reject H_o and conclude slope term is significantly different from zero. There appears to be a linear association between the log odds that a site is a nest site and density of mature forest in ring 1.

Sleuth then considers selecting a model with successively larger rings to evaluate significance of each additional increment of mature forest. (not covered here).

Pesticide Exposure

- Pesticide residues in human milk, Western Australia 1979-80
- 45 donors
 - age in years
 - new suburb (yes/no; coded as yes=1, no=0)
 - house treated for termites in past 3 years (yes/no)
 - milk contained >0.009 ppm dieldrin (yes/no)

Write out fitted model for logistic regression of dieldrin on age, new suburb and termites.

- What predictors are significant at $\alpha=0.05$?
- Interpret each coefficient in a sentence.
- Compare presence of above-average levels of dieldrin for new and old suburbs using an odds ratio.
- What *probability* of high dieldrin is predicted for a woman of average age, in old-suburb house not treated for termites? (avg. age=30)
- Test the joint significance of age and new suburb in the model.

Dieldrin ~ Age + New + Treated

```
*** Generalized Linear Model ***
Call: glm(formula = Dieldrin ~ Age + New + Treated, family =
binomial(link = logit), data = dieldrin, na.action =
na.exclude, control = list(epsilon= 0.0001, maxit = 50, trace =
F))

Coefficients:
(Intercept) -8.9727497  3.8532397 -2.328625
Age          0.2083864  0.1172264  1.777640
New          2.1406872  0.9308404  2.299736
Treated      2.5969714  0.9683143  2.681951

(Dispersion Parameter for Binomial family taken to be 1 )

Null Deviance: 56.76518 on 42 degrees of freedom
Residual Deviance: 41.08711 on 39 degrees of freedom
```

Dieldrin ~ Age + New + Treated

Analysis of Deviance Table

Binomial model

Response: Dieldrin

Terms added sequentially (first to last)

	Df	Deviance	Resid.	Df	Resid. Dev
NULL				42	56.76518
Treated	1	7.086807		41	49.67837
New	1	4.851417		40	44.82696
Age	1	3.739850		39	41.08711

Dieldrin ~ Treated

```
Call: glm(formula = Dieldrin ~ Treated, family = binomial(link
= logit), data =
  dieldrin, na.action = na.exclude, control = list(epsilon =
  0.0001,
    maxit = 50, trace = F))
Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.249127 -0.9177273 -0.5863281  1.107343  1.921256

Coefficients:
            Value Std. Error  t value
(Intercept) -1.673723   0.6233116  -2.685211
  Treated    1.840777   0.7458884   2.467899

(Dispersion Parameter for Binomial family taken to be 1 )

Null Deviance: 56.76518 on 42 degrees of freedom

Residual Deviance: 49.67837 on 41 degrees of freedom

Number of Fisher Scoring Iterations: 3
```