Notes for week 5

Ben Bolker

October 17, 2013

Version: 2013-10-17 23:33:25

Interpreting parameters

- continuous: units: depends whether scaled or not
- categorical: differences between groups
- depends on presence of interactions
- scale of measurement: *link scale*
  - log: proportional, if small. The argument here is that if $\mu_0 = \exp \beta_0$ and $\mu_1 = \exp \beta_0 + \beta_1 x$,

    $$\begin{aligned} \mu_1 &= \exp(\beta_0 + \beta_1 x) \\ &= \mu_0 \exp(\beta_1 x) \\ &\approx \mu_0 (1 + \beta_1 x) \qquad \text{if } \beta_1 x \ll 1 \end{aligned}$$

    so for continuous predictors β_1 is the proportional change in the mean per unit change in x (for categorical predictors it's the proportional change between categories).

    Predicted values are the expected *geometric* mean of the category.
  - logit: log-odds change.
    * for $\beta \Delta x$ small, as for log (proportional)
    * for intermediate values, linear change in probability with slope $\approx \beta/4$
    * for large values, as for $\log(1 - x)$

Inference

Single vs multi-parameter

Single-parameter *Wald* vs. *likelihood ratio* test (LRT): the former is easier (it's what you get from `summary()`), because Wald standard errors of the estimates ($\sigma_{\hat{b}eta}$) are automatically available from the Hessian of the fitted model, so we can get p-values via a Z test on $\hat{\beta}/\sigma_{\hat{b}eta}$ (this is what `summary` gives) and confidence intervals via Normal confidence intervals on $\hat{\beta}$.

The *Hauck-Donner effect* occurs in cases of extreme parameter estimates (e.g. in the case of complete or near-complete separation), when the quadratic approximation is extremely poor: the hallmark is large parameter estimates (e.g. $|\hat{\beta}| > 10$) and very large confidence intervals (leading to small Z statistics and large p values).

You can get LRTs via

- `drop1(.,test="Chisq")` (only on parameters that can be dropped while respecting marginality)
- ```
  reduced_model <- update(full_model,.~.-foo)
  anova(full_model,reduced_model,test="Chisq")
  ```

  where `foo` is the parameter you want to test.
- or by hand (having fitted these models)

  ```
  pchisq
  ```

You can get *profile confidence intervals* via `MASS::confint.glm`.

M ulti-parameter

If you want to test a hypothesis that multiple $\hat{\beta}$ values are simultaneously zero (i.e. you want to see if the overall effect of a factor is significant), you *could* do a Wald test: e.g. to test $\hat{\beta}_1 = \hat{\beta}_2 = 0$, you would calculate the sums of squares ($\hat{\beta}_1^2 + \hat{\beta}_2^2 = 0$) and the variance; the scaled result should be χ^2 distributed.

```
contr <- c(1,1)
t(contr) %*% vcov(model) %*% contr
pchisq()
```

It generally makes more sense to do this with `anova()` or `drop1()` (`anova(model)` gives *sequential* (forward/"type I") tests: `anova(model1,model2,model3)` compares a specific sequence of models); these use LRTs (if `test="Chisq"`) or F tests (if `test="F"`, which you should use when the dispersion parameter is estimated (Gaussian, Gamma, or quasi-likelihood models).

Interactions/marginality issues

You have to be very careful when testing main effects in the presence of interactions. `drop1()` generally respects marginality, although you can do `drop1(.~.)` to get `drop1` to test *all* the effects (i.e not respecting marginality). ([1] is a standard reference from one of the proponents of respecting marginality: see Section 5.)

Your options with respect to marginality are:

[1] Venables, W. N. (1998). Exegeses on linear models. 1998 International S-PLUS User Conference, Washington, DC
```

- don't test main effects at all in the presence of interactions
- test main effects, but be very careful/aware that the meaning of the main effects depends on the parameterization/contrasts used
- split the data set and run separate analyses for each category involved in the interaction

Finite-size issues

In general LRTs are better than Wald tests, but even they make a (weaker) asymptotic assumption (not that the log-likelihood surface is quadratic, but that the deviance is χ^2 distributed). People generally ignore this problem since the number of observations is usually sufficiently large that this is a reasonable approximation, but *Bartlett corrections* [2, 1] are one approach to this.

If the dispersion parameter is estimated (rather than fixed, as it is for Poisson and binomial models), then we should use F tests ("quasi-LRT" for want of a better term) rather than χ^2, because the deviance differences are now scaled by the (χ^2-distributed) $\hat{\phi}$ (note that this does *not* address the issue of whether the deviance itself is really χ^2 distributed).

You can use bootstrap or parametric bootstrap samples to get more reliable p-values/confidence intervals, if it's important: for example

```
Ldat <- read.csv("lizards.csv")
model1 <- glm(gfrac~height+diameter+light+time,
              family=binomial,weights=N,data=Ldat)
## a function to sample and refit the model
bootFun <- function() {
    bootdat <- Ldat[sample(nrow(Ldat),replace=TRUE),]
    coef(update(model1,data=bootdat))
}
library(plyr)
## do this 250 times, store the results as a matrix
bootParms <- raply(250,bootFun())
```

Get 2.5% and 97.5% quantiles of each column (`MARGIN=2` specifies columns rather than rows), and transpose the results:

```
t(apply(bootParms,MARGIN=2,quantile,c(0.025,0.975)))

##                  2.5%   97.5%
## (Intercept)   0.6709  1.9157
## height>=5ft   0.7482  1.8815
## diameter>2in -1.2676 -0.4047
```

```
## lightsunny    0.2977  1.4942
## timelate     -1.8443 -0.3236
## timemidday   -0.7874  0.6025
```

Two-sided p-values:

```
apply(bootParms,MARGIN=2,
        function(x) 2*min(mean(x<0),mean(x>0)))

##  (Intercept)  height>=5ft diameter>2in   lightsunny     timelate
##        0.000        0.000        0.024        0.000        0.000
##   timemidday
##        0.448
```

References

[1] Cordeiro, G. M. and S. L. P. Ferrari (1998, August). A note on bartlett-type correction for the first few moments of test statistics. *Journal of Statistical Planning and Inference 71*(1-2), 261–269.

[2] McCullagh, P. and J. A. Nelder (1989). *Generalized Linear Models*. London: Chapman and Hall.

[3] Venables, W. N. (1998). Exegeses on linear models. 1998 International S-PLUS User Conference, Washington, DC.