Residuals in glm

**Raw or response residuals** are the difference between the observed and the fitted value: $y_i - \hat{y}_i$. Not very informative in the glm framework.

In R: `resid(object, type='response')*object$prior.weights`

**Pearson residuals** are raw residuals divided by the standard error of the observed value $se(y_i)$

$$X_i = \frac{y_i - n_i\hat{p}_i}{\sqrt{n_i\hat{p}_i(1 - \hat{p}_i)}}$$ (1)

In R: `resid(object, type='pearson')`

**Standardized Pearson residuals** are the raw residuals divided by their standard error $se(y_i - \hat{y}_i)$

$$r_{Pi} = \frac{y_i - n_i\hat{p}_i}{\sqrt{n_i\hat{p}_i(1 - \hat{p}_i)(1 - h_i)}}$$ (2)

Standardized Pearson residuals are also called studentized Pearson residuals, standardized residuals (Dunteman and Ho, 2006), studentized residuals (GLM procedure in SPSS and SAS1), internally studentized residuals2

In R: Not possible to obtain directly(?). They can be calculated dividing the Pearson residuals by $\sqrt{1 - h_i}$. Note that the residuals should be further scaled by the dispersion parameter $\phi$ if $\phi \neq 1$

`resid(object, type='pearson')/sqrt(1 - hatvalues(object)))`

Available in the boot package: `glm.diag(object)$rp`

**Deviance residuals** The deviance residuals are the signed square roots of the ith observation to the overall deviance

$$d_i = \text{sgn}(y_i - \hat{y}_i) \left\{ 2 y_i \log \left( \frac{y_i}{\hat{y}_i} \right) + 2 (n_i - y_i) \log \left( \frac{n_i - y_i}{n_i - \hat{y}_i} \right) \right\}^{(1/2)}$$ (3)

In R: `resid(object)` (deviance residuals are default in `residuals.glm`)
**Standardized deviance residuals** are the deviance residuals divided by $\sqrt{(1 - h_i)}$

$$r_{Di} = \frac{d_i}{\sqrt{(1 - h_i)}} \quad (4)$$

The standardized deviance residuals are also called studentized deviance residuals.

In R: `rstandard(object)`

**Likelihood residuals** is the change in deviance when the $i$th observation is omitted from the data. In order to obtain exact values the model is fitted $n + 1$ times. The likelihood residuals can be approximated by

$$r_{Li} = \text{sgn}(y_i - \hat{y}_i)\sqrt{h_ir^2_i + (1 - h_i)r^2_{Di}} \quad (5)$$

The likelihood residuals are also called studentized residuals (Fox, 2002), externally studentized residuals (cf Fox, 2002), deleted studentized residuals (cf Fox, 2002), jack-knife residuals.

In R: `rstudent(object)`

**The terminology**

**In the normal case**

Several authors mention the inconsistency in the terminology regarding studentized residuals. Some of the residuals are called several names and the term *studentized* are used about different types of residuals.

Margolin (1977) defines the term Studentization as dividing a scale dependent statistic with a scale estimate so that the resulting ratio has a distribution free of the nuisance scale parameters. According to David (1981) the numerator and denominator are dependent in internal Studentization and independent in external Studentization.

According to Cook and Weisberg (1982) the internally Studentized residuals (isr) are defined by

$$r_i = \frac{e_i}{\hat{\sigma}}\sqrt{(1 - h_i)} \quad (6)$$

---

3http://data.princeton.edu/wws509/notes/c3s8.html
where $\hat{\sigma}^2 = \Sigma e_i^2 / (n-k)$ is the residual mean square. The isr are subsequently termed Studentized residuals by the same authors and others (i.e. Gray and Woodall (1994)). The isr divided by the residual degrees of freedom follows a Beta distribution. These residuals are called standardized residuals by Belsley et al (1980) and Collett (2003) (p. 132 nederst).

The externally Studentized residuals (esr) are defined by

$$t_i = \frac{e_i}{\hat{\sigma}(i) \sqrt{(1 - h_i)}}$$

where $\hat{\sigma}^2_{(i)}$ is computed without the $i$th residual leading to independency between the numerator and denominator. The esr conviently follows Student’s $t$ distribution with $n-p-1(?)$ degrees of freedom (Cook and Weisberg, 1982).

Another way to compute the externally Studentized residuals is via the mean-shift outlier model (see Williams (1987) and Fox (2002)). In the normal case it can be showed that the t-statistic equals the esr. The mean-shift outlier model corresponds to the change in scaled deviance $\phi^{-1}D$ when re-fitting a model omitting the $i$th observation$^4$ (Williams, 1987). The residuals are therefore also called (studentized) deletion residual.

In case of normality the standardized deviance residual is further divided by $s$ and, hence, the standardized deviance residuals equal the standardized Pearson residuals and are termed standardized residuals. According to Collett (2003) the standardized residuals also equal the likelihood residuals but since the likelihood residuals in R are divided by $s(i)$ instead of $s$ the two quantities do not equal each other in R.

### 0.1 Extending to generalized linear models

Cook and Weisberg, p. 199: One possible general extension of the Studentized residual $r_i^2$ (…) is $\tilde{r}_i^2 = \hat{s}_i^2 / \tilde{w}_i(1 - h_i)$.

Hardin and Hilbe (2007) addresses, like several others, the issue of terminology. They define standardization and studentization in the glm framework as follows:

- The adjective **standardized** means that the variance of the residual has been standardized to take into account the correlation between $y$ and $\hat{\mu}$. The base residual has been multiplied by the factor $(1 - h_i)^{-1/2}$.

- The adjective **studentized** means that the residuals have been scaled by an estimate of the unknown scale parameter. The base residuals have

$^4\phi$ is the dispersion parameter. For normal data $\phi$ is the variance and for binomial data $\phi = 1$
been multiplied by the factor $\phi^{-1/2}$. In many texts and papers, it is the standardized residuals that are then scaled (studentized). However, the result is often called studentized residuals instead of the more accurate standardized studentized residuals.

Referencer:


