NMFM 402, Practical 5, GLM - variable reduction

Ex 1. Recall that for normal distr. with $w_m = 1$ we have

$$f_{Y_m}(y) = \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left\{ -\frac{(y-\mu_m)^2}{2\sigma^2} \right\} =$$

$$= \exp\left\{ \frac{y\cdot\mu_m - \frac{\mu_m^2}{2}}{\sigma^2} - \frac{y^2}{2\sigma^2} - \log(\sqrt{2\pi}\,\sigma) \right\}$$

and so: $\theta_m = \mu_m \qquad b(\theta) = \frac{\theta^2}{2} \qquad b'(\theta) = \theta \qquad b''(\theta) = 1$

$\varphi = \sigma^2 \qquad\qquad h(\mu) = (b')^{-1}(\mu) = \mu$

(a) Deviance statistics:

$$\underline{D(Y, \hat{\mu})} = \varphi \cdot D^*(Y, \hat{\mu}) = 2 \cdot \sum_{m=1}^{M} w_m \Big[Y_m \cdot h(Y_m) - b(h(Y_m)) -$$

$$- Y_m\, h(\hat{\mu}_m) + b(h(\hat{\mu}_m)) \Big] = 2 \cdot \sum_m \Big[Y_m^2 - \frac{Y_m^2}{2} - Y_m \cdot \hat{\mu}_m +$$

$$+ \frac{(\hat{\mu}_m)^2}{2} \Big] = \cancel{2} \cdot \sum_m Y_m^2 - 2Y_m\,\hat{\mu}_m + (\hat{\mu}_m)^2 = \underline{\sum_{m=1}^{M} (Y_m - \hat{\mu}_m)^2}$$

$\Rightarrow$ in our setting, D coincides with residual sum of squares.

(b) F-statistics for sub-model testing:

$$F = \frac{D(Y, \hat{\mu}_{H_0}) - D(Y, \hat{\mu}_{full})}{D(Y, \hat{\mu}_{full})} \cdot \frac{M-k-1}{p} =$$

$$= \frac{\dfrac{\sum_{m=1}^{M} (Y_m - \hat{\mu}_{m,H_0})^2 - \sum_{m=1}^{M} (Y_m - \hat{\mu}_{m,full})^2}{p}}{\dfrac{\sum_{m=1}^{M} (Y_m - \hat{\mu}_{m,full})^2}{M-k-1}}$$

$\Rightarrow$ coincides with F-statistics based on the increment of sum of squares of residuals, used for sub-model testing in classical linear regression

NOTE: Similarly to linear regression, the F-statistics in GLM enables us to perform hierarchical testing of series of submodels. In variable reduction analysis, this would be done via backward stepwise selection.

(c) Deviance residuals:

recall $D(Y, \hat{\mu}) = \sum_{m=1}^{M} d(Y_m, \hat{\mu}_m)$

$\rightarrow d(Y_m, \hat{\mu}_m) = (Y_m - \hat{\mu}_m)^2$ ... distance function for normal distribution

compare: distance function for Poisson distr.:

$$d(Y_m, \hat{\mu}_m) = 2 w_m \cdot \left[Y_m \log \frac{Y_m}{\hat{\mu}_m} + \hat{\mu}_m - Y_m \right]$$

$$r_m^D = \text{sgn}(Y_m - \hat{\mu}_m) \cdot \sqrt{d(Y_m, \hat{\mu}_m)} = \text{sgn}(Y_m - \hat{\mu}_m) \cdot |Y_m - \hat{\mu}_m| =$$

$$= Y_m - \hat{\mu}_m$$

Pearson's residuals:

$$r_m^P = \frac{Y_m - b'(\hat{\theta}_m)}{\sqrt{\frac{b''(\hat{\theta}_m)}{w_m}}} = Y_m - \hat{\mu}_m$$

$\Rightarrow$ in our setting, both concepts of residuals coincide and they correspond to residuals in classical linear regression

(d) $\hat{\varphi}_P = \frac{1}{M-K-1} \sum_{m=1}^{M} (r_m^P)^2 = \frac{1}{M-K-1} \sum_{m=1}^{M} (Y_m - \hat{\mu}_m)^2$

$$\hat{\varphi}_D = \frac{D(Y, \hat{\mu})}{M-K-1} = \frac{1}{M-K-1} \sum_{m=1}^{M} (r_m^D)^2 = \frac{1}{M-K-1} \sum_{m=1}^{M} (Y_m - \hat{\mu}_m)^2$$

$\Rightarrow$ in our setting, the two ~~exp~~ estimators of φ coincide and they correspond to the residual variance s^2, which is an unbiased estimator of σ^2 in classical lin. regr.

Ex. 2. Recall that for gamma distr. with $w_m = 1$ we have

$$b(\theta) = -\log(-\theta) \quad b'(\theta) = -\frac{1}{\theta} \quad b''(\theta) = \frac{1}{\theta^2}$$

$$\theta = h(\mu) = (b')^{-1}(\mu) = -\frac{1}{\mu}$$

(a) $D(Y, \hat{\mu}) = 2 \cdot \sum_{m=1}^{M} w_m \left[Y_m h(Y_m) - b(h(Y_m)) - Y_m h(\hat{\mu}_m) + \right.$
$\left. + b(h(\hat{\mu}_m)) \right] = 2 \cdot \sum_{m=1}^{M} Y_m \cdot \left(-\frac{1}{Y_m}\right) + \log\left(\frac{1}{Y_m}\right) - Y_m \cdot \left(-\frac{1}{\hat{\mu}_m}\right) -$
$- \log\left(\frac{1}{\hat{\mu}_m}\right) = 2 \cdot \sum_{m=1}^{M} \frac{Y_m - \hat{\mu}_m}{\hat{\mu}_m} - \log\left(\frac{Y_m - \hat{\mu}_m}{\hat{\mu}_m} + 1\right).$

(b) Deviance residuals

$$D(Y, \hat{\mu}) = \sum_{m=1}^{M} d(Y_m, \hat{\mu}_m) \Rightarrow$$

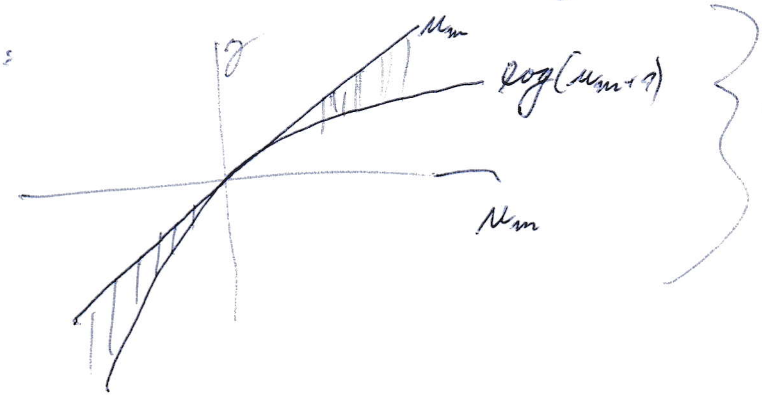
$$\Rightarrow d(Y_m, \hat{\mu}_m) = 2 \cdot \left[\frac{Y_m - \hat{\mu}_m}{\hat{\mu}_m} - \log\left(\frac{Y_m - \hat{\mu}_m}{\hat{\mu}_m} + 1\right) \right]$$

NOTE: verify that $d(Y_m, \hat{\mu}_m) \geq 0$:

set $u_m := \frac{Y_m - \hat{\mu}_m}{\hat{\mu}_m}$ ... relative difference.

~~then~~ $\Rightarrow d(Y_m, \hat{\mu}_m) = 2 \cdot [u_m - \log(u_m + 1)] \geq 0$

by picture:



$$r_m^D = \operatorname{sgn}(Y_m - \hat{\mu}_m) \cdot \sqrt{2 \cdot \left[\frac{Y_m - \hat{\mu}_m}{\hat{\mu}_m} - \log\left(\frac{Y_m - \hat{\mu}_m}{\hat{\mu}_m} + 1 \right) \right]}$$

Pearson's residuals:

recall: $\hat{\mu}_m = b'(\hat{\theta}_m) = -\frac{1}{\hat{\theta}_m} \quad \Rightarrow \hat{\theta}_m = -\frac{1}{\hat{\mu}_m}$

$$b''(\hat{\theta}_m) = \frac{1}{\left(-\frac{1}{\hat{\mu}_m}\right)^2} = (\hat{\mu}_m)^2$$

$$r_m^P = \frac{Y_m - b'(\hat{\theta}_m)}{\sqrt{\frac{b''(\hat{\theta}_m)}{w_m}}} = \frac{Y_m - \hat{\mu}_m}{|\hat{\mu}_m|}$$