MATH 523: DEVIANCE COMPARISON WITH ESTIMATED DISPERSION

When ϕ is known, the Analysis of Deviance for nested models $M_0 \subset M_1$ containing $p_0 < p_1$ parameters respectively, with deviances D_0 and D_1, is based on the asymptotic theory of likelihood ratio tests: we have, for GLM with dispersion ϕ.

$$\frac{D_0 - D_1}{\phi} \dot{\sim} \chi^2_{p_1 - p_0}$$

We also have the result that if the model M (possibly, but not necessarily, model M_1) with p parameters is an adequate fitting model, then its deviance D satisfies

$$\frac{D}{\phi} \dot{\sim} \chi^2_{n-p}$$

To estimate the parameter ϕ, we can use this result to justify the moment-based estimator

$$\widehat{\phi} = \frac{D}{n-p}$$

Now, combining these results, we can establish a procedure to perform deviance tests if ϕ is unknown. Consider the statistic

$$\frac{\left[\dfrac{(D_0 - D_1)}{\phi(p_1 - p_0)}\right]}{\left[\dfrac{D}{\phi(n-p)}\right]} = \frac{(D_0 - D_1)/(p_1 - p_0)}{D/(n-p)}$$

From above, the RHS is a ratio of two approximately chi-squared random variables divided by their respective degrees of freedom; assuming (with some justification) that $D_0 - D_1$ and D are **independent**, we have by standard results that

$$\frac{(D_0 - D_1)/(p_1 - p_0)}{D/(n-p)} \dot{\sim} Fisher(p_1 - p_0, n - p)$$

that is, the statistic is approximately Fisher-F distributed. Thus for the analysis of deviance statistic with estimated dispersion

$$\frac{D_0 - D_1}{\widehat{\phi}} = (p_1 - p_0) \times \frac{(D_0 - D_1)/(p_1 - p_0)}{D/(n-p)} \dot{\sim} (p_1 - p_0)F$$

where $F \sim Fisher(p_1 - p_0, n - p)$.

Now, as $\nu \longrightarrow \infty$, if $Z_\nu \sim \chi^2_\nu$, then

$$\frac{Z_\nu}{\nu} \xrightarrow{p} 1$$

so therefore we must have that as $F \sim Fisher(\nu_1, \nu_2)$,

$$F \stackrel{d}{=} \frac{Z_{\nu_1}/\nu_1}{Z_{\nu_2}/\nu_2} \xrightarrow{d} Z_{\nu_1}/\nu_1$$

as $\nu_2 \longrightarrow \infty$. Therefore, for large n

$$\frac{D_0 - D_1}{\widehat{\phi}} \dot{\sim} \chi^2_{p_1 - p_0}.$$