Note Title

# Stat 511 Lecture 40

4/25/2008

Cases of GLMs –

1) Normal
2) Poisson
3) Binomial

a) $h(\mu) = \log \frac{p}{1-p}$ $\quad p = \frac{\exp \underset{\sim}{x}'\beta}{1+\exp \underset{\sim}{x}'\beta}$ "logit link"

("logistic regression")

b) $h(\mu) = \Phi^{-1}\left(\frac{\mu}{n}\right) = \Phi^{-1}(p)$ $\quad p = \Phi(\underset{\sim}{x}'\beta)$ "probit link"

c) $h(\mu) = \log\left(-\log\left(1-\frac{\mu}{n}\right)\right) = \log(-\log(1-p))$ "complementary log log link"

gives $p = 1 - e^{-e^{\underset{\sim}{x}'\beta}}$

"canonical" (mathematically most natural / convenient) link for a GLM is $\underset{\sim}{\theta} = \underset{\sim}{x}'\beta$ i.e. is $h = b'^{-1}$

One Last General Point (about inference in all models to date) — made in section 7.3.2 of typed outline

Big model with parameter vector $\underset{\sim}{\underset{r\times 1}{\theta}} = \begin{pmatrix} \underset{p\times 1}{\theta_1} \\ \underset{(r-p)\times 1}{\theta_2} \end{pmatrix}$

Smaller/reduced model with parameter vector $\begin{pmatrix} \underset{\sim}{0} \\ \underset{\sim}{\theta}_2 \end{pmatrix}$ this could be any fixed/known vector

To test $H_0: \underset{\sim}{\theta}_1 = \underset{\sim}{0}$ in the big model (i.e. test in the big model that the submodel that is the smaller one is adequate)
I can use a LRT and thus the statistic

$$X^2 = 2\left( \underbrace{\ell(\hat{\theta}_{MLE})}_{\text{maximized log-likelihood for big model}} - \underbrace{\ell(\underset{\sim}{0}, \hat{\underset{\sim}{\theta}}_2(\underset{\sim}{0}))}_{\text{maximized log-likelihood in small model}} \right)$$

and an approximate $\chi^2_p$ null dsn (for starting p-values)

For the LM this statistic (for $H_0: \beta_1 = \beta_2 = \cdots = \beta_p = 0$) can be shown to reduce to the Full model/Reduced Model F statistic (and I don't need to resort to the $\chi^2_p$ approximation)

In other models (nonlinear regression, MLM, GLM, etc.) this statistic and $\chi$ approximation provides a way of comparing 2 nested models – so I can use this to test all A main effects = 0 in a GLM, e.g.

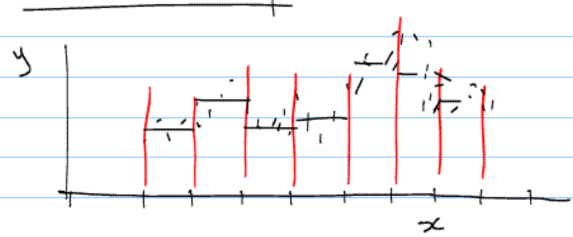
"Smoothing" / Model-Free curve fitting / Nonparametric Regression
(Chll of Faraway Book II)

$(x_i, y_i)$

polynomial regression
or nonlinear regression

There are many applications where the "obvious" ways of modeling how $y$ changes with $x$ just fail to be flexible enough — I need a flexible means of "fitting curves" — Chll discusses a number of such means

Bin "Smoothing"



partition the $x$-axis into "bins" and cook up a step function to approximate the relationship of $y$ to $x$

$\hat{y}(x) =$ mean or median or trimmed mean or ⋮ of $y_i$'s with $x_i$'s in the same bin as $x$

<u>Running Smoothers</u> use a different "bin" for each $x$

<u>k Nearest Neighbor Versions</u>

a symmetric version would be to for $\hat{y}(x)$ use

median or mean or trimmed mean or ⋮ of $\frac{k}{2}$ $y_i$'s corresponding to nearest $x_i \le x$

and $\frac{k}{2}$ $y_i$'s corresponding to nearest $x_i \ge x$

– a possibly non-symmetric version is to apply (mean or median or ⋮)

to k $y_i$'s with $x_i$'s closest to $x$ of interest (without regard to how they split in terms of being above or below $x$ –

The ? ⋮ could be the value of a fitted regression function (based on k nearest neighbors) – i.e. I could for each $x$

i) identify a set of k neighbors

ii) fit $y = b_{0,x} + b_{1,x}\, x$ to k points by OLS

iii) use $\hat{y}(x) = b_{0,x} + b_{1,x}\, x$

This "local regression" idea one version of the list

$\left(\begin{array}{c} \text{mean} \\ \text{or} \\ \text{median} \\ \text{or} \\ \vdots \end{array}\right)$ — this treats all points in the neighborhood of $x$

at which I'm trying to predict equally (and I completely ignore all points just outside the neighborhood) — because of how least squares works it is the neighbors furthest from the $x$ at which I predicting that get most influence