① 9 Oct. 2001

Outliers, data values which cause surprise in relation to the majority of the sample

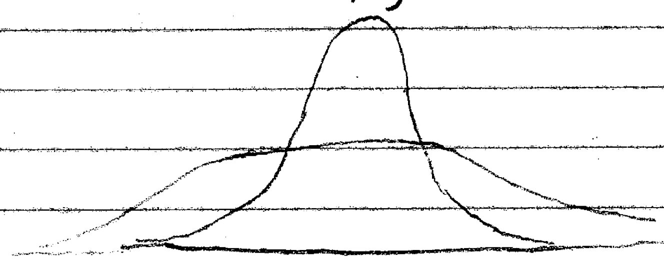
V and R

* Amazon example, midterm example

Long-tails. Students-t

Probability element

$$f(t)\,dt = \frac{1}{\sqrt{\nu}\,\beta\left(\frac{1}{2}, \frac{\nu}{2}\right)} \frac{1}{\left(1+\frac{t^2}{\nu}\right)^{(\nu+1)/2}}\,dt$$

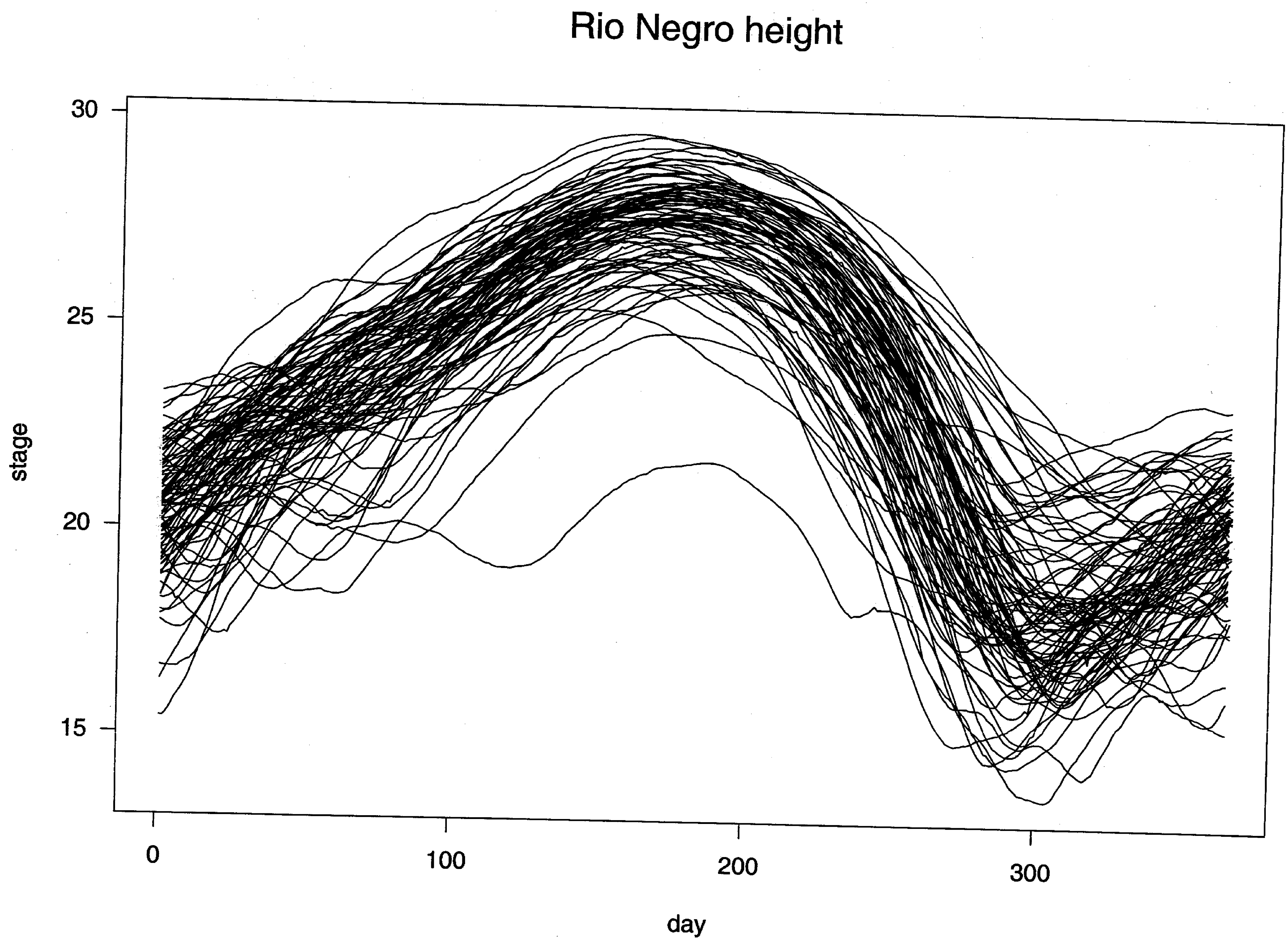


Dies off like $t^{-(\nu+1)}$ Cauchy: $\nu=1$

Normal dies off like $e^{-x^2/2}$

Add location and scale parameters

$$f\left(\frac{x-\mu}{\sigma}\right)\frac{1}{\sigma}$$



Rio Negro height
stage
30
25
20
15
day
0
100
200
300

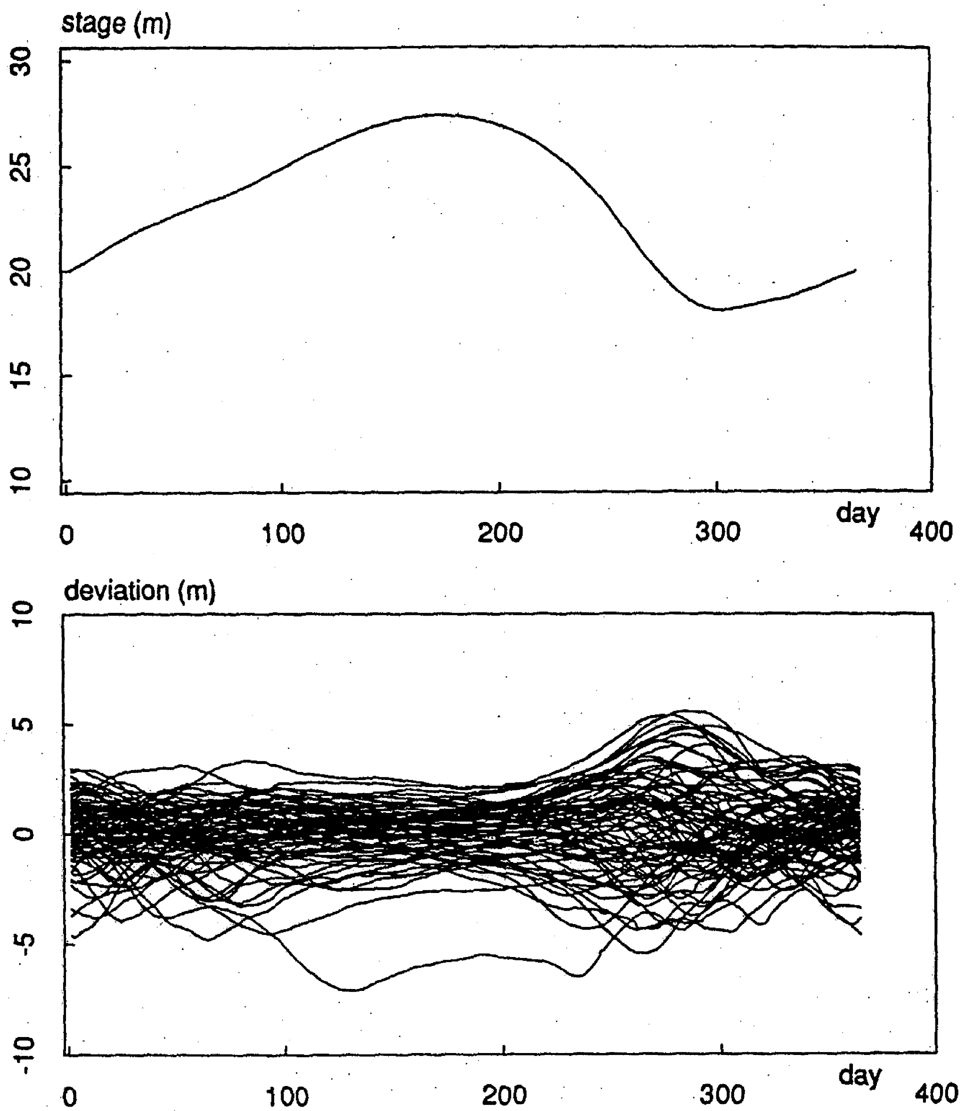


Figure 3

The top graph is the 50% trimmed mean of the annual curves, as defined by (2.1). The bottom graph provides the deviations of each year's curve from the 50% trimmed mean curve. The very low curve corresponds to a year of a serious forest fire.

19 Oct 2001

Breakdown point

Sample $\underset{\sim}{y} = (y_1, \dots, y_n)$ and ~~estimator~~ $\hat{\theta}(\underset{\sim}{y})$

The breakdown point is m/n where m is the smallest number such that if we are allowed to change m data values in any way we can force the absolute value of $\hat{\theta}(.)$ for the "perturbed" sample to $\pm\infty$.

A measure of the robustness of an estimator

Mean: $1/n$ upset by a single outlier

Median: $\frac{n+1}{2}$ if n is odd $n=3$, $\frac{n+1}{2} = \frac{2}{3}$

tolerates 50% gross errors $.5 \times 3 = 1.5 \equiv 1$

Problem of scale.

9 Oct. 2001

MM-estimate (Yohai)

γ function

1. $\gamma(0)=0$, $\gamma(-u)=\gamma(u)$
2. $0 \le u \le v \Rightarrow \gamma(u) \le \gamma(v)$
3. γ is continuous, $0 < \sup \gamma(u) < \infty$
4. $\gamma(u) < \sup \gamma(\,)$ and $0 \le u < v \Rightarrow \gamma(u) < \gamma(v)$

Given robust $\hat{\epsilon}_1, \ldots, \hat{\epsilon}_n$ estimate scale via

$$\frac{1}{n} \sum_{i=1}^{n} \gamma\left(\hat{\epsilon}_i / s\right) = B$$

eg. $B = \int \gamma(z)\varphi(z)\,dz$

Rousseeuw & Leroy (1987). Robust Regression and Outlier Detection. Wiley

9 Oct. 2005

robust & resistant Splus functions.

in library (mass), or statlib

l1fit()	"min abs residual (L1) regression"
ltsreg()	"least trimmed squares robust regression"
rreg()	"M-estimates of regression"
glm(, family = robust())	
rlm()	"fit robust linear regression model"
lmsreg()	"least median of squares regression"
mean(, trim=)	"mean value"
location.m()	"robust M-estimate of location"
mad()	"median absolute deviation"
cov.mve()	"minimum volume ellipsoid covariance estimation"
loess(, "symmetric")	"fit a local regression model"
median()	
minreg()	

10 Oct 2001

Missing values

Plague most kinds of data to some degree

What to do varies widely
forgetting is rarely desirable
time series - equal spacing
anova - symmetries in design

One approach:
insert rather bad values and use a robust technique

Questions:
How frequently missing?
Individually or in blocks of what length?
. . .

Sometimes need to model the missingness mechanism. (probability model)
* eg. China quakes

Trouble if missingness depends on Y
censoring (cp. Winsorizing)

10 Oct 2001

Approach

Suppose Y is a missing value and $\underset{\sim}{X} = (X_1, \dots, X_m)$ are the available data.

~~Estimate~~ by

$$\hat{E}(Y \mid \underset{\sim}{X})$$

Old anova technique:
"make" residual $\hat{\epsilon} = Y - \hat{Y}$ zero

Idea.
Make the influence 0

10 Oct 2001

Splus. Denoted by NA

Various functions contain argument

na.action

= na.fail creates error when missing found

= na.omit deletes observations containing one or more missing values

= na.gam.replace

is.na( )

sort(, na.last)

11 Oct 2001

Back to confidence and testing

Likelihood ratio test an extension of anova

Prob function $f(y|\theta)$ $\theta \in \Theta$

Data $y_1, \ldots, y_n$

Null hypothesis $H_0: \theta \in \Theta_0 \subseteq \Theta$ nested

Likelihood ratio statistic

$$\lambda_n = \frac{\max_{\theta \in \Theta_0} \prod_1^n f(y_i|\theta)}{\max_{\theta \in \Theta} \prod_1^n f(y_i|\theta)}$$

$$-2 \log \lambda_n = 2[\ell(\hat{\theta}) - \ell(\hat{\underline{\theta}}_0)]$$ difference of deviances

as $n \to \infty$

$\sim \chi^2_{k-s}$ or χ^2_r, r: number of restrictions, s: number of free parameters

if $\Theta_0 = \{(\theta_1, \ldots, \theta_r, \theta_{r+1}, \ldots, \theta_k); \theta_j = \theta_{0j}, j = 1, \ldots, r\}$

and model in family, i.e. model correct

There may be a sequence of nested hypotheses

$$\Theta_1 \subset \Theta_2 \subset \Theta$$

$$\ell(\theta) = \sum_i \log f(y_i|\theta)$$

11 Oct. 2001

Proof of χ^2,

a) full mle, k unknowns

Set $\overset{k\times 1}{\underset{\sim}{D}} = \sqrt{n}\,(\hat{\theta}-\theta_0) \sim N_k(\underset{\sim}{0}, \underset{\sim}{I}_0^{-1})$

$$\overset{k\times 1}{\underset{\sim}{V}} = \frac{1}{\sqrt{n}} \sum_i \left.\frac{\partial \log f(y_i|\theta)}{\partial \theta}\right|_{\theta_0} \qquad (*)$$

Then $\underset{\sim}{V} \sim N(\underset{\sim}{0}, \underset{\sim}{I}_0)$

$$\text{where } \underset{\sim}{I}_0 \sim \text{var}\left\{ \left.\frac{\partial \log f(Y|\theta)}{\partial \theta}\right|_{\theta_0} \right\}$$
$$= -E\left\{ \frac{\partial^2}{\partial\theta\,\partial\theta^\tau} \right\}$$

By Taylor expansion

$$\overset{k\times 1}{\underset{\sim}{0}} = \sum_i \left.\frac{\partial \log f(y_i|\theta)}{\partial \theta}\right|_{\hat{\theta}} \simeq \sum_i \overset{k\times 1}{\left.\frac{\partial \log f(y_i|\theta)}{\partial \theta}\right|_{\theta_0}}$$
$$+ \sum_i \overset{k\times k}{\frac{\partial^2 \log f(y_i|\theta)}{\partial\theta\,\partial\theta^\tau}} \overset{k\times 1}{(\hat{\theta}-\theta_0)}$$

So $\underset{\sim}{V} \sim \underset{\sim}{I}_0\, \underset{\sim}{D}$

$\underset{\sim}{D} \sim \underset{\sim}{I}_0^{-1}\, \underset{\sim}{V}$

11 Oct 2001

$$l(\hat{\theta}) = \sum_i \log f(y_i | \hat{\theta})$$

$$l(\theta_0) \simeq \sum_i \log f(y_i | \hat{\theta}) + \left(\sum_i \frac{\partial \log f(y_i|\theta)}{\partial \theta} \bigg|_{\hat{\theta}} \right)^T (\theta_0 - \hat{\theta})$$

$$+ \frac{1}{2} (\theta_0 - \hat{\theta})^T \left(\sum_i \frac{\partial^2 \log f(y_i|\theta)}{\partial \theta \, \partial \theta^T} \right)_0 (\theta_0 - \hat{\theta})$$

So $\quad l(\theta_0) \simeq l(\hat{\theta}) - \frac{1}{2} (\theta_0 - \hat{\theta})^T n \underset{\sim}{I}_0 (\theta_0 - \hat{\theta})$

or $\quad 2 \left[l(\hat{\theta}) - l(\theta_0) \right] \sim \underset{\sim}{D}^T \underset{\sim}{I}_0 \underset{\sim}{D}$

$$\sim \chi^2_k$$

$$\sim \underset{\sim}{V}^T \underset{\sim}{I}_0^{-1} \underset{\sim}{V} \quad \text{also}$$

This result can be used to:

i) examine the simple hypothesis $\theta = \theta_0$

ii) construct a confidence region for θ_0 (using $\hat{\underset{\sim}{I}}_0$)

iii) corresponds to $r = k$ restrictions

11 Oct. 2001

Now lets suppose that there are $k-r$ free parameters only.

Have restrictions such that the original $\theta_1, \ldots, \theta_k$ are functions of $s = k-r$ new ~~parameters~~

eg. $\theta_j = g_j(\beta_1, \ldots, \beta_s)$ $\quad j = 1, \ldots, k$

$r = k-s$ restrictions

Suppose, for convenience

$$\begin{bmatrix} \beta_1 \\ \vdots \\ \beta_s \end{bmatrix} = \begin{bmatrix} \theta_{r+1} \\ \vdots \\ \theta_k \end{bmatrix}$$

We will need

$$\underset{\sim}{\overset{k \times s}{M}} = \left[\frac{\partial g_i}{\partial \beta_j} \right] = \begin{bmatrix} \underset{\sim}{O} \\ \underset{\sim}{I}_s \end{bmatrix}$$

Write underscore when referring to last s coordinates of θ, eg $\underline{\theta}$, $\underline{\theta}_0$

Will repeat previous development only using last s coordinates.

18 Oct. 2001

Let $\underset{\sim}{U}^{s\times 1} = \sum_i \frac{1}{\sqrt{n}} \left.\frac{\partial \log f(y_i|\theta)}{\partial \underline{\theta}}\right|_{\underline{\theta}_0}$ (**)

Let $\underset{\sim}{F} = \sqrt{n}(\hat{\underline{\theta}} - \underline{\theta}_0)$

$\underset{\sim}{\mathcal{J}}_0$: information matrix for $\underline{\theta}$

As before

$$\underset{\sim}{U} \sim \underset{\sim}{\mathcal{J}}_0 \underset{\sim}{F}, \quad \underset{\sim}{F} \sim \underset{\sim}{\mathcal{J}}_0^{-1} \underset{\sim}{U}$$

$$2[\ell(\hat{\underline{\theta}}_0) - \ell(\underline{\theta}_0)] \sim \underset{\sim}{F}^T \underset{\sim}{\mathcal{J}}_0 \underset{\sim}{F}$$

Lets connect $\underset{\sim}{U}$ and $\underset{\sim}{V}$

From (*) and (**) $\quad \underset{\sim}{U} = \underset{\sim}{M}^T \underset{\sim}{V}$ and so

$$\underset{\sim}{\mathcal{J}}_0 = \underset{\sim}{M}^T \underset{\approx}{I}_0 \underset{\sim}{M} \qquad (***\,*)$$

Continuing

$$2[\ell(\hat{\underline{\theta}}_0) - \ell(\underline{\theta}_0)] \sim \underset{\sim}{F}^T \underset{\sim}{\mathcal{J}}_0 \underset{\sim}{F}$$

$$\sim \underset{\sim}{U}^T \underset{\sim}{\mathcal{J}}_0^{-1} \underset{\sim}{U}$$

$$\sim \underset{\sim}{V}^T \underset{\sim}{M} \underset{\sim}{\mathcal{J}}_0^{-1} \underset{\sim}{M}^T \underset{\sim}{V}$$

11 Oct. 2001

Subtracting

$$2\left[l(\hat{\theta}) - l(\hat{\mathbf{\theta}})\right] \sim \mathbf{V}^T\left(\mathbf{I}_0^{-1} - \mathbf{M}\,\mathbf{D}_0^{-1}\mathbf{M}^T\right)\mathbf{V} \qquad (***)$$

But $\mathbf{V} \sim N_k(\mathbf{0}, \mathbf{I}_0)$

write $\mathbf{V} = \mathbf{I}_0^{1/2}\mathbf{Z}$ with $\mathbf{Z} \sim N(\mathbf{0}, \mathbf{I})$

So

$$(***) = \mathbf{Z}^T\left(\mathbf{I} - \mathbf{I}_0^{1/2}\mathbf{M}\,\mathbf{D}_0^{-1}\mathbf{M}^T\mathbf{I}_0^{1/2}\right)\mathbf{Z}$$

Note that $\mathbf{D}_0$ from $(****)$

$$\mathbf{I}_0^{1/2}\mathbf{M}\,\mathbf{D}_0^{-1}\underbrace{\mathbf{M}^T\mathbf{I}_0^{1/2}\,\mathbf{I}_0^{1/2}\mathbf{M}}_{\mathbf{D}_0}\mathbf{D}_0^{-1}\mathbf{M}^T\mathbf{I}_0^{1/2}$$

$$= \mathbf{I}_0^{1/2}\mathbf{M}\,\mathbf{D}_0^{-1}\mathbf{M}^T\mathbf{I}_0^{1/2} \quad \text{i.e. idempotent}$$

So $(***)$ is χ^2_ν

$$\nu = k - \operatorname{tr}\left(\mathbf{I}_0^{1/2}\mathbf{M}\,\mathbf{D}_0^{-1}\mathbf{M}^T\mathbf{I}_0^{1/2}\right)$$

$$= k - \operatorname{tr}\left(\mathbf{D}_0^{-1}\mathbf{M}^T\mathbf{I}_0\mathbf{M}\right)$$

$$= k - \operatorname{tr}\left(\mathbf{D}_0^{-1}\mathbf{D}_0\right) \quad \text{from } (****)$$

$$= k - s = r$$

The nested sequence result follows similarly.

11 Oct 2001

The final deviance of the fitted model $f(y|\hat{\theta})$ is

$$2[\ell(\tilde{\theta}) - \ell(\hat{\theta})] \geq 0$$

where $\tilde{\theta}$ is the mle for the saturated model, viz. parameter for each observation

By a big stretch

$$2[\ell(\tilde{\theta}) - \ell(\hat{\theta})] \sim \chi^2_{n-k}$$

People often compare final deviance to

$E\chi^2_{n-k} = n-k$, but $\exists$ problems with approx.

We have seen that

$$2[\ell(\hat{\theta}) - \ell(\hat{\underline{\theta}})] \sim \chi^2_{k-s} = \chi^2_p \text{ under null}$$

i.e. the difference in deviances between two nested models has a chi-squared distribution under the null hypothesis.

i.e.

$$2[\ell(\tilde{\theta}) - \ell(\hat{\underline{\theta}})] - 2[\ell(\tilde{\theta}) - \ell(\hat{\theta})]$$

H_0 $\qquad\qquad\qquad$ H

18 Oct 2001

Might have $\Theta_1 \subset \Theta_0 \subset \Theta$

free params	s_1 <	s_0 <	k
constraints/restrictions	r_1	r_0	0

$$2\left[\ell(\tilde{\theta}) - \ell(\hat{\underline{\theta}})\right] - 2\left[\ell(\tilde{\theta}) - \ell(\hat{\theta}_-)\right] = D_{\Theta_1} - D_{\Theta_0}$$

$$\sim \chi^2_{s_0 - s_1}$$

Can set up an ANODEV table

Hypothesis	Deviance	df	Deviance difference	Δ df
Θ_1	D_{Θ_1}	$n - s_1$	$D_{\Theta_1} - D_{\Theta_0}$	$s_0 - s_1$
Θ_0	D_{Θ_0}	$n - s_0$	$D_{\Theta_0} - D_{\Theta}$	$k - s_0$
Θ	D_{Θ}	$n - k$		

Deviance gets smaller as bring in more parameters

11 Oct. 2001

Advantages of deviance.

1. Invariant under 1-1 parametrizations of the model

2. Unifies a variety of model selection problems: linear model (anova), glm, graphical models . . .

3. Has interpretation in terms of K-L divergence

4. Lot known about large sample distribution

18 Oct. 2001

Distribution of the deviance / likelihood ratio ~~statistic~~ when the model is not true.

We saw that

$$\text{mle } \hat{\theta} \xrightarrow{p} \theta_* \quad \text{where } \theta_* \text{ maximizes}$$

$$E_p\{\log f(Y|\theta)\}$$

$$\frac{1}{n} \ell(\theta) = \frac{1}{n} \sum \log f(Y_i|\theta)$$

$$\xrightarrow{p} E_p\{\log f(Y|\theta)\}$$

so can anticipate

$$\frac{1}{n} \ell(\hat{\theta}) \xrightarrow{p} E_p\{\log f(Y|\theta_*)\}$$

Likewise

$$\frac{1}{n} \ell(\hat{\underline{\theta}}) \longrightarrow E_p\{\log f(Y|\underline{\theta}_*)\}$$

so $$-\frac{2}{n} \log \lambda_n \longrightarrow 2 E_p\{\log f(Y|\theta_*)\} - 2 E_p\{\log f(Y|\underline{\theta}_*)\}$$

Going to ∞ when $\theta_* \notin \Theta_0$

$\underline{\theta}_*$ gives $\max_{\theta \in \Theta_0} E_p\{\log f(Y|\theta)\}$

11 Oct. 2001

When $\theta_* \in \Theta_0$

$-2 \log \lambda_n \sim c_1 z_1^2 + \dots + c_r z_r^2$

Foutz & Srivastava (1978) Canadian J. Stat 6, 273-9.

14 Oct. 2001

Example. Gauss-Markov + normal model

$$Y_i = \alpha + (\underset{\sim}{x}_i - \bar{\underset{\sim}{x}})' \underset{\sim}{\beta} + \epsilon_i \qquad \epsilon_i : IN(0, \sigma^2)$$

H_0: $\underset{\sim}{\beta} = \underset{\sim}{\beta}_0$

$$\theta = \begin{bmatrix} \alpha \\ \underset{\sim}{\beta} \\ \sigma \end{bmatrix}$$

$$L(\theta) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left\{ -\frac{1}{2\sigma^2} \| y_i - \alpha - (\underset{\sim}{x}_i - \bar{\underset{\sim}{x}})' \underset{\sim}{\beta} \|^2 \right\}$$

$$\ell(\theta) = -\frac{n}{2} \log \sigma^2 - \frac{1}{2\sigma^2} \sum_{i=1}^{n} \| y_i - \alpha - (\underset{\sim}{x}_i - \bar{\underset{\sim}{x}})' \underset{\sim}{\beta} \|^2 + C$$

"Full" mle: $\hat{\alpha} = \bar{y}$

$$\tilde{\underset{\sim}{X}}' \tilde{\underset{\sim}{X}} \hat{\underset{\sim}{\beta}} = \tilde{\underset{\sim}{X}}' \underset{\sim}{Y} \qquad \tilde{\underset{\sim}{X}} = \underset{\sim}{X} - \bar{\underset{\sim}{X}}$$

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} \| y_i - \bar{y} - (\underset{\sim}{x}_i - \bar{\underset{\sim}{x}})' \hat{\underset{\sim}{\beta}} \|^2$$

$$\ell(\hat{\theta}) = -\frac{n}{2} \log \hat{\sigma}^2 + C'$$

14 Oct. 01

Mle under H_0: $\hat{\alpha} = \bar{y}$

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n} \|y_i - \bar{y}\|^2$$

$$l(\hat{\theta}_0) = -\frac{n}{2}\log \hat{\sigma}^2$$

$$2\left[l(\hat{\theta}) - l(\hat{\theta}_0)\right] = n \log \frac{\sum_i \|y_i - \bar{y}\|^2}{\sum_i \|y_i - \bar{y} - (x_i - \bar{x})'\hat{\beta}\|^2}$$

Now

$$\sum_{i=1}^{n} \|y_i - \bar{y}\|^2 = \sum_{i=1}^{n} \|(x_i - \bar{x})'\hat{\beta}\|^2 + \sum_{i=1}^{n} \|y_i - \bar{y} - (x_i - \bar{x})'\hat{\beta}\|^2$$

$$SST = SSR + SSE$$

Test statistic is based on SSR/SSE

We saw its exact distribution under H_0

$$\frac{SSR/p}{SSE/(n-p-1)} \sim F_{p,\,n-p-1}$$

14 Oct. 01

Relation to previous approximation:

It was based on $n \longrightarrow \infty$

$$\hat{\sigma}^2 \xrightarrow{\text{Prob}} \sigma^2$$

$$n \log \frac{SSR + SSE}{SSE} = n \log\left(1 + \frac{SSR}{SSE}\right)$$

$$\sim \frac{SSR}{SSE/n}$$

$$\sim \frac{SSR}{\sigma^2}$$

$$\sim \chi^2_p$$

Table 2

List of earthquakes in 'Central China'

NO	YEAR	MO	DY	HR:MN: SEC	LAT	LONG	DEPTH	MAG	REGION
1	-1177				34.5 N	107.8 E		4.5	SHENSI
2	-780				34.5 N	107.8 E ?		6.5	SHENSI
3	-231				36.5 N	111.5 E		6.5	SHANSI
4	-193	FEB			35.4 N	103.9 E		6.5	KANSU
5	-186	FEB	22		33.4 N	104.8 E		6.5	KANSU
6	-159	JUN			32.2 N	110.4 E		5.	HUPEH
7	-154				32.2 N	110.4 E		5.	HUPEH
8	-142?				32.2 N	110.4 E		5.	HUPEH
9	- 70	JUN	1		36.3 N	119.0 E		7.0	SHANTUNG
10	- 47	APR	17		35.1 N	104.6 E		6.75	KANSU
11	- 35	JUL			34.4 N	109.0 E		5.	SHENSI
12	46	OCT	23		33.0 N	112.5 E		6.5	HONAN
13	128	FEB	23		34.7 N	105.4 E		6.5	KANSU
14	138	MAR	1		35.5 N	104.0 E		6.75	KANSU
15	143	OCT			34.7 N	105.3 E ?		7.0	KANSU
16	294	JUL			32.6 N	116.8 E		5.5	ANHWEI
17	294	SEP			40.3 N	116.0 E		5.5	HOPEH
18	319	JUN	17		34.0 N	105.2 E		4.5	KANSU
19	344				36.3 N	114.5 E		5.5	HOPEH
20	373	AUG			36.6 N	101.8 E		4.75	TSINGHAI
21	406	JUN			36.3 N	104.5 E		5.5	KANSU
22	408				36.8 N	118.3 E		5.	SHANTUNG
23	408				39.0 N	100.5 E		4.75	KANSU
24	416				34.3 N	105.5 E		5.	KANSU
25	421				41.6 N	120.4 E		5.	LIAONING
26	462	AUG	16		35.6 N	116.8 E		5.5	SHANTUNG
27	495	MAR	31		37.5 N	121.2 E		5.5	SHANTUNG
28	506	AUG	30		37.9 N	102.6 E		4.75	KANSU
29	512	MAY	21		39.0 N	113.0 E		7.5	SHANSI
30	575	JAN	14		37.9 N	102.6 E		5.5	KANSU
31	600	DEC	13		34.3 N	108.9 E		5.5	SHENSI
32	638	FEB	11		32.6 N	103.6 E		5.	SZECHWAN
33	649	SEP	12		36.1 N	111.5 E		5.5	SHANSI
34	692	MAR?			37.5 N	117.5 E		5.	SHANTUNG
35	734	MAR	19		34.7 N	106.3 E		7.0	KANSU
36	756	NOV	27		39.0 N	100.5 E		6.0	KANSU
37	777				37.8 N	115.2 E		6.0	HOPEH
38	788	MAR	8		32.5 N	109.2 E		6.5	SHENSI
39	793	MAY	27		34.5 N	109.7 E ?		6.0	SHENSI
40	835	APR	11		34.3 N	108.9 E		4.75	SHENSI
41	916	FEB	25		34.3 N	108.9 E		4.75	SHENSI
42	839				34.4 N	104.0 E		6.5	KANSU
43	865	DEC			35.9 N	111.4 E		5.5	SHANSI
44	867	FEB	14		35.9 N	111.4 E		5.5	SHANSI
45	876	JUL	14		37.8 N	105.9 E		6.5	NINGSIA

Table 2 (*continued*)

NO	YEAR	MO	DY	HR:MN: SEC	LAT	LONG	DEPTH	MAG	REGION
46	879	MAR			34.2 N	109.3 E		4.75	SHENSI
47	880	FEB			34.5 N	107.8 E		4.75	SHENSI
48	953	NOV			36.3 N	115.1 E		4.75	HOPEH
49	999	OCT			31.8 N	119.9 E		5.5	KIANGSU
50	1010	NOV			38.1 N	106.4 E		5.5	NINGSIA
51	1011	AUG			38.2 N	114.6 E		4.75	HOPEH
52	1022	APR			39.7 N	113.1 E		6.5	SHANSI
53	1038	JAN	9		38.4 N	112.9 E		7.25	SHANSI
54	1046	APR	18		37.8 N	120.7 E		5.	SHANTUNG
55	1057				39.5 N	116.3 E ?		6.75	HOPEH
56	1068	AUG	14		38.5 N	116.1 E		6.0	HOPEH
57	1069	JAN	18		38.3 N	116.8 E		4.75	HOPEH
58	1076	DEC			39.9 N	116.4 E		5.	HOPEH
59	1092	DEC?			37.9 N	102.6 E		4.75	KANSU
60	1125	AUG	30		36.0 N	103.9 E		7.0	KANSU
61	1143	APR			38.5 N	106.3 E		6.5	NINGSIA
62	1169	JAN	24		31.9 N	104.4 E		4.75	SZECHWAN
63	1209	DEC	04		36.0 N	111.8 E		6.5	SHANSI
64	1219	MAY	21		36.0 N	106.2 E		6.5	NINGSIA
65	1290	SEP	27		41.5 N	119.3 E		6.75	LIAONING
66	1291	AUG	25		36.1 N	111.5 E		6.5	SHANSI
67	1303	SEP	17		36.3 N	111.7 E		8.0	SHANSI
68	1304	FEB			36.1 N	111.5 E		5.5	SHANSI
69	1304	SEP	3		37.5 N	112.6 E		4.75	SHANSI
70	1305	MAY	3		39.8 N	113.1 E		6.5	SHANSI
71	1306	SEP	12		35.9 N	106.1 E		6.5	NINGSIA
72	1314	OCT	5		36.5 N	113.8 E		6.0	HOPEH
73	1316				36.4 N	111.1 E		5.5	SHANSI
74	1322				40.6 N	115.0 E		4.5	HOPEH
75	1336	MAR?			30.1 N	115.9 E		4.7	HUPEH
76	1337	SEP	8		40.4 N	115.7 E ?		6.5	HOPEH
77	1338	AUG	2		40.4 N	115.2 E		5.	HOPEH
78	1342	MAY	5		37.9 N	112.6 E		5.5	SHANSI
79	1346	APR			37.1 N	118.0 E		4.75	SHANTUNG
80	1351	MAY	14		37.3 N	113.0 E ?		5.5	SHANSI
81	1352	APR	18		35.6 N	105.3 E ?		7.0	KANSU
82	1368	JUL	8		37.6 N	112.5 E		6.0	SHANSI
83	1372	AUG	16		32.0 N	118.8 E		4.75	KIANGSU
84	1378	APR	30		38.5 N	106.3 E		5.75	NINGSIA
85	1399	APR	29		32.0 N	118.8 E		4.75	KIANGSU
86	1407	NOV			31.2 N	112.6 E		5.5	HUPEH
87	1425	MAR	7		31.7 N	116.5 E		5.75	ANHWEI
88	1433				30.5 N	115.2 E		4.75	HUPEH
89	1440	OCT	26		36.2 N	103.4 E		6.25	KANSU
90	1448	SEP	30		38.3 N	109.7 E		5.	SHENSI