→ 6th Dec 10–12 SS 2128 Test

15 Nov Fields I 210 10 am

Xiao-Li 11 am

14 11 am

Finish lik. inf. (testing)

$f(y_i;\varphi) = \exp\{\varphi' t(y_i) - c(\varphi) - d(y)\}$

$f(t;\varphi) = \int_{\{\underset{\sim}{y}\,;\, t(\underset{\sim}{y}) = t\}} f(\underset{\sim}{y};\varphi)\,d\underset{\sim}{y} \qquad \underset{\sim}{y} = (y_1, \dots, y_n)$

$= \exp\{\varphi' t - nc(\varphi)\}\, h(t)$

$f(t_1 \mid t_{(-1)};\varphi_1) = \exp\{\varphi_1 t - c_{(-1)}(\varphi_1)\}\, h_{(-1)}(t_1)$

1 par. exp'l family

true for any ψ that's linear in φ

$f(t_1 \mid \underset{\sim}{t}_2 ; \varphi_1) = \exp\{\varphi_1 t - c_{t_2}(\varphi_1)\}\, h_{t_2}(t_1)$

$\underset{\sim}{t}(y) = \left(\sum_{i=1}^{n} t_1(y_i), \dots, \sum t_p(y_i)\right)$

In our asymptotics, we used

$$\sup_{\psi_2} l(\varphi_1, \psi_2) = l_p(\varphi_1)$$

found $\hat{\varphi}_1 - \varphi_1 \,\dot{\sim}\, N(0, \ldots)$

We can instead use

$$l_c(\varphi_1) = t_1\varphi_1 - c_{t_2}(\varphi_1)$$

and get w_c, $w_{c,e}$, $w_{c,u}$ etc.
same asymptotics applies

$$E l_c'(\varphi_1) = 0 \qquad \text{but } E l_p'(\varphi_1) \neq 0$$

$O(n^{-1})$

a genuine log-lik.

Example: logistic regression $Y_i \sim Ber(p_i)$

$$\log\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \beta_1 x_i$$

$$L(\beta_0, \beta_1) = \exp\left(\beta_0 \sum y_i + \beta_1 \sum x_i y_i - \ln \sum (1 + e^{\beta_0 + \beta_1 x_i})\right)$$

$\psi = \beta_1$ inference

$$f\left(s_1 = \sum x_i y_i \,\middle|\, s_0 = \sum y_i \,;\, \beta_1\right)$$

s_0 # succ.
" $s_1 = x_+$ "
(successes)

$$f(s_1 | s_0; \beta_1) = \frac{c(s_0, s_1)\, e^{\beta_1 s_1}}{\sum_u c(s_0, u)\, e^{\beta_1 u}}$$

CH Ex 5.3

$c(s_0, s_1)$ # subsets of $\{x_1, \ldots, x_n\}$ of size s_0 that add to s_1

$$x_i = \begin{cases} 1 & i = 1, \ldots, n_1 \quad \text{trt} \\ 0 & i = n_1+1, \ldots, n_1+n_2 \quad \text{control} \end{cases}$$

$\sum_{i=1}^{n} x_i y_i$ = # successes in trt group

$$\log \frac{p_i}{1-p_i} = \beta_0 + \beta_1 x_i$$

$$= \begin{cases} \beta_0 & x_i = 0 \\ \beta_0 + \beta_1 & x_i = \underline{1} \end{cases}$$

$$f(s_1 | s_0; \beta_1) = \binom{n_1}{s_1}\binom{n_2}{s_0-s_1} e^{\beta_1 s_1} \Big/ \sum_u \binom{n_1}{u}\binom{\ }{\ } e^{\beta_1 u}$$

$c(s_0, s_1)$ = # of distinct ways of choosing a set of size s_0 with s_1 1's, $s_0 - s_1$ 0's

$$f(s_1 | s_0; \beta_1 = 0) = \binom{n_1}{s_1}\binom{n_2}{s_0-s_1} \Big/ \binom{n}{s_1}$$

basis of Fisher's exact test of indep. in 2x2 table

		y_i's: 1	0	
T	1	s_1	$n_1 - s_1$	n_1
x_i's C	0	$s_0 - s_1$	$n_2 - s_0 + s_1$	n_2
		s_0	$n - s_0$	n

fix n_1 - D
n_2 - D
s_0 - C

"exact test"

3 E	1	x
4	7	
x		m

obs'd table

p-value

$$\Pr(S_1 \geq s_1 \mid s_0; \beta_1 = 0)$$

$$\sum_{\text{cells}} \frac{(O-E)^2}{E} \;\dot\sim\; \chi^2_1 \leftarrow \text{profile test}$$

$\mu_1 - \mu_2 = 0$ in cont^{s} obs$^{=}$s

$y_{11}, \dots, y_{1n_1}$ $\qquad$ $y_{21}, \dots, y_{2n_2}$

$t = (\bar{y}_1 - \bar{y}_2)/s_p/\sqrt{n}$ $\qquad$ $\Pr(T \geq t_{obs}; \mu_1 = \mu_2)$

$= 1 - pt(n-1, t^{obs}$

$Y_i = \begin{matrix} 1 \\ 0 \end{matrix}$ $\qquad$ $x_i = \begin{matrix} 1 \\ 0 \end{matrix}$ $\qquad$ $i = 1, \dots, n$

1 0 $\quad$ 1 0

Matched pairs (Y_{j1}, Y_{j2}) ind't Ber $\quad j = 1, \dots, m$

$P(Y_{j1} = 1) = \dfrac{e^{\lambda_j}}{1 + e^{\lambda_j}}$ $\qquad$ $\Pr(Y_{j2} = 1) = \dfrac{e^{\psi + \lambda_j}}{1 + e^{\psi + \lambda_j}}$

$\text{logit}(P_{j1} = 1) = \lambda_j$ $\qquad$ $\text{logit}(P_{j2} = 1) = \psi + \lambda_j$

$\uparrow$

$\underbrace{\text{"}x_j = 0\text{"}}_{\text{cases}}$ $\qquad$ $\underbrace{\text{"}x_j = 1\text{"}}_{\text{controls}}$

$\beta_0 \quad x = 0$

$\beta_0 + \beta_1 \quad 1$

$$\ell(\psi, \underset{\sim}{\lambda}) = \psi \sum_{j=1}^{m} y_{j2} + \sum_{j=1}^{m} \lambda_j (y_{j1} + y_{j2}) - c(\underset{\sim}{\lambda}, \psi)$$

profile lst: $\hat{\lambda}_{j,\psi} = \begin{cases} -\infty & 0,0 \\ \frac{1}{2}\psi & y_{j1}, y_{j2} = 0,1 \text{ or } 1,0 \\ +\infty & 1,1 \end{cases}$

$$\ell_p(\psi) = \ell(\psi, \hat{\lambda}_\psi) \qquad \ell_p'(\hat{\psi}) = 0$$

$$\cdots \quad \sum_{j=1}^{m} y_{j2} - r_0 + \frac{r_1}{1+e^{\psi/2}} = 0$$

$r_0 = \#(0,0)$

$r_1 \ \# \ (0,1) \text{ or } (1,0)$

$$1 + e^{\hat{\psi}/2} = \frac{r_1/m}{(\sum y_{j2} - r_0)/m} \xrightarrow{p} 1 + e^{\psi}$$

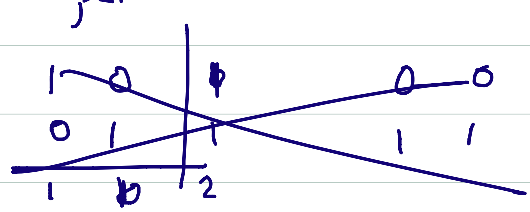
$$\hat{\psi} \xrightarrow{p} 2\psi \qquad m \to \infty \qquad \lambda_1, \ldots, \lambda_m$$

$$f\left(\sum_{j=1}^{m} y_{j2} \,\middle|\, y_{j1} + y_{j2}, j = 1, \ldots, m\right)$$

$$r_{01} \sim \text{Bin}\left(r_1 ; \frac{e^{\psi}}{1+e^{\psi}}\right)$$

unique dist^n free $\left\{ \lambda_j\text{'s} \right.$

$r_{01} = \sum_{j=1}^{m} y_{j2} \qquad r_1 = \# \ (1,0)\text{'s or } (0,1)\text{'s}$



$| \ \underset{\sim}{s}_2$ fix margins in table.

1,1 cell   margins

$$f(\underset{\sim}{t};\varphi) = \underbrace{f_c(t_1 \mid \underset{\sim}{t}_2;\varphi_1)} \boxed{f_m(\underset{\sim}{t}_2;\varphi)}$$

$\varphi_1,\dots,\varphi_p$

ignore info on φ_1 here

$\underset{\sim}{t}$ is discrete   Binomial

Bernoulli

$f(t_1 \mid \underline{t}_2)$ "more discrete" fewer pt. of +ve prob.

3	1
7	4

$Pr(T \geq t_1) = .027$ $Pr(T > t_1) = .001$

$Pr(T \geq t_1 - 1) = .053$

- Rasch model in psychometrics
- ~~xx~~ choice based sampling in econometrics

- cond'l inf. in exp'l families
  (after marginal)

$$\underset{\sim}{y} \xrightarrow{\ \ } \underset{\sim}{t} \xrightarrow{\ \ } t_1$$

$\uparrow$ integ. $\quad \uparrow$ cond'g

- marginal inference in transf. families

Example loc-scale $y_i \sim \frac{1}{\sigma} f_0\left(\frac{y_i - \mu}{\sigma}\right)$

$f_m(\hat{\sigma};\sigma).\ f(\hat{\mu} \mid \hat{\sigma}; \mu, \sigma)$

$$f(\hat{\mu}, \hat{\sigma} \mid \underset{\sim}{a}; \mu, \sigma) = \frac{L(\mu, \sigma; \hat{\mu}, \hat{\sigma}, \underset{\sim}{a})}{\int L(\mu,\sigma)\, d\mu\, d\sigma}$$

$a_i = (y_i - \hat{\mu}) / \hat{\sigma}$

$\underset{\sim}{y} \rightarrow \hat{\mu}, \hat{\sigma}$ by cond'g necessary inference

$\hat{\mu}, \hat{\sigma}$ | | $\mu \quad \sigma$ matches in dimension

$\psi = \mu$: Let $\boxed{\begin{aligned} t_1 &= (\hat{\mu} - \mu)/\hat{\sigma} \\ t_2 &= \hat{\sigma}/\sigma \end{aligned}}$ 1-1 transf.

$\mu \in \mathbb{R} \quad \sigma \in \mathbb{R}^+$

$\cancel{f(\hat{\mu}, \hat{\sigma} \mid a; \mu, \sigma) = \{f_m(t_1)\}}$

$f(t_1, t_2 \mid a; \mu, \sigma) \longleftarrow f(\hat{\mu}, \hat{\sigma} \mid \underset{\sim}{a}; \mu, \sigma)$

t_1 (t_2) is unique pivotal quantity for inference about μ (σ)

$f_m(t_1) = f_m(t_1(\mu; y))$ for fixed $\mu = \mu_0$

$f_m(t_1)$ is the only marg'l density free of σ, based on $y \mid \underline{a}$

any quantity dep. on $y \mid \underline{a}; \mu$ is dist-free of σ, is a fn of t_1.

" t_1 is a maximal invariant under the loc'n scale group is dist-free of σ "

TSH — needs more abstract argument than cond'g

$\cancel{\prod_{i=1}^{n} f_0(y_i - \theta) \rightarrow f(\hat{\theta} \mid \underset{\sim}{a}; \theta)}$

If $\quad y_i = x_i^T\beta + \sigma e_i$

$$\frac{\hat\beta_j - \beta_j}{\hat\sigma} = t_j \qquad t_{p+1} = \frac{\hat\sigma}{\sigma}$$

$f_m(t_j(\beta_j))$ free of $\beta_{(-j)}, \sigma$

$f_m\left(\frac{\hat\sigma}{\sigma}\right)$ free of β essentially unique

$$\frac{L(\beta, \sigma; \hat\beta, \hat\sigma, a)}{\int L(\beta,\sigma)\, d\beta\, d\sigma} = f(\hat\beta, \hat\sigma \mid \underline{a})$$

$f_m(\beta_1)$ requires $\displaystyle\int_{\mathbb{R}^{p-1}\times\mathbb{R}^+} d\hat\beta_{(-j)}\, d\hat\sigma$

Example $N(\mu, \sigma^2)$ $\quad \hat\mu = \bar y \quad \hat\sigma^2 = \frac{\Sigma(y_i - \bar y)^2}{n}$

$= S_y^2/n$

m.l.e. from $\ell(\mu,\sigma)$

$$f(\bar{y}, s_y^2 \mid \underset{\sim}{a}; \mu, \sigma) = f(\bar{y}, s_y^2 \; [redacted] ; \mu, \sigma)$$

$$= N\left(\mu, \frac{\sigma^2}{n}\right) \times \sigma^2 \chi^2_{n-1}$$

$$(\bar{y}, s_y^2) \longrightarrow (t_1, t_2) \qquad t_1 = \frac{\bar{y} - \mu}{s_y / [redacted]} \quad t_2 = \frac{s_y^2}{\sigma^2}$$

$$f_m(t_1) \times \boxed{\chi^2_{n-1} \text{ for } \frac{s^2}{\sigma^2}} \qquad \text{ind } t \; (N)$$

$$\Pr(T_1 \geq t_1^{obs}; \mu^0) = \cancel{\Pr(T_{n-1} \geq c\, t_1^{obs}}$$

- can be computed from T_{n-1} dist$^{=}$

$f_m\left(\frac{s^2}{\sigma^2}\right)$ is χ^2_{n-1}

$$f_m(s^2; \sigma^2) = \frac{(s^2)^{\frac{n-1}{2} - 1}}{\sigma^{2\,\frac{n-1}{2}}} \cdot e^{-s^2/2\sigma^2} \cdot \frac{1}{2^{\frac{n-1}{2}} \Gamma\left(\frac{n-1}{2}\right)}$$

$$l_m(\sigma^2) = -\frac{n-1}{2} \log \sigma^2 - \frac{s^2}{2\sigma^2}$$

$$l_m'(\sigma^2) = 0 \implies \hat{\sigma}_m^2 = \frac{s^2}{}$$

$n-1$

"mle" from marginal is better

$$\underset{\sim}{y} \to \hat{\mu}, \hat{\sigma}^2 \mid \underset{\sim}{a} \to \hat{\sigma}^2 \mid \underset{\sim}{a}$$

cond$^{\underline{n}}$ marginalize

marginal log-likelihood for s^2 gives better inference than the full log-likelihood

- extended to mixed & random effects
  (*) "REML" = marginal for variance estimates

2 model classes

1 marg. $\longrightarrow$ cond$^{\underline{n}}$

1 cond $\longrightarrow$ marg. ... examples...

Both different than profiling.

You can show that to some order of approx^n the marg'l log-lik & cond'l log. lik are approximated by

$$\ell_{ap}(\psi) = \ell_p(\psi) - \frac{1}{2}\log|\underbrace{j_{\lambda\lambda}(\psi,\hat{\lambda}_\psi)}|$$

$$= \ell(\psi,\hat{\lambda}_\psi) - \frac{1}{2}\log|j_{\lambda\lambda}(\psi,\hat{\lambda}_\psi|$$

$$j(\theta) = -\frac{\partial^2 \ell}{\partial\theta\,\partial\theta^T} = \begin{bmatrix} & \\ & j_{\lambda\lambda} \end{bmatrix}$$

(ψ,λ)

ntbc ??

example $N(\mu, \underset{\psi}{\sigma^2})$ $\quad -\frac{n}{2}\log\sigma^2 + \frac{1}{2}\log\sigma^2$

(ψ, λ) ← inference re ψ, eliminate λ

- exact calc^{n}s
- approx calc^{n}s (hint)

Summary of likelihood

$\theta = (\psi, \lambda)$

$y = (y_1, \dots, y_n)$

$f(y; \psi, \lambda)$

↑ dim q

$(\psi, \lambda) \xrightarrow{\text{change}} (\psi, \phi(\lambda, \psi))$ "interest-respecting"

und. (inv) $w(\psi) = 2\{\ell_p(\hat\psi) - \ell_p(\psi)\}$ respects 'skewness'

$w_e(\psi) = (\hat\psi - \psi)^T j_p(\hat\psi)(\hat\psi - \psi) \;\dot\sim\; \chi^2_q$ easiest to get

not inv.

$w_u(\psi) = [\ell_p'(\psi)]^T j_p^{-1}(\hat\psi)[\ell_p'(\psi)]$

$$\frac{\partial \ell(\psi, \hat\lambda_\psi)}{\partial \lambda} = 0$$

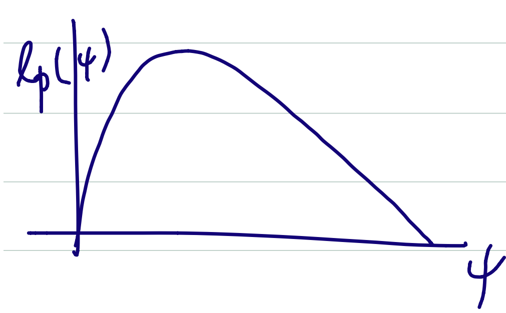
If ψ scalar $\quad \chi^2_1$

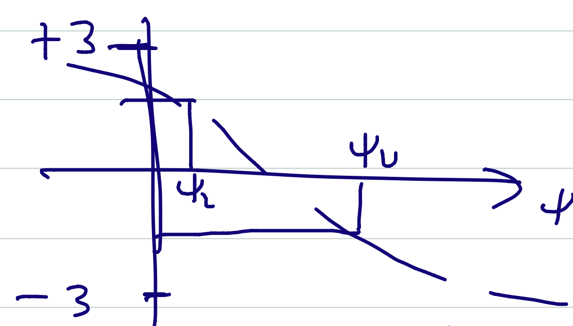
$r = \text{sign}(\hat\psi - \psi)\sqrt{w(\psi)} \quad \dot\sim N(0,1)$

$r_e = (\hat\psi - \psi) j_p^{1/2}(\hat\psi) \quad \dot\sim N(0,1)$

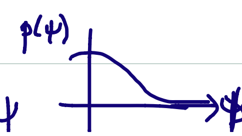
$r_u \quad \ell_p'(\psi) j_p^{-1/2}(\hat\psi) \quad \dot\sim N(0,1)$

plot $r(\psi; y^0)$ vs ψ





plot $\Phi(r)$ vs ψ



"p value fn of ψ"
confidence distn
fiducial distn ← F

$$W_e = \underline{(\hat{\psi} - \psi)^T j_p(\hat{\psi})(\hat{\psi} - \psi)}$$

$j_p(\psi)$
expected info

$$W_u \quad \ell_p'(\psi)^T j_p^{-1}(\psi) \ell_p'(\psi)$$

$j_p(\psi) = (j_{\psi\psi} - j_{\psi\lambda} j_{\lambda\lambda}^{-1} j_{\lambda\psi})^{-1}$

use $(i_{\psi\psi} - i_{\psi\lambda} i_{\lambda\lambda}^{-1} i_{\lambda\psi})^{-1}$

$\psi = 0$ of interest

$\hat{\lambda}_0$, $-\partial^2 \ell / \partial \theta^2$ @ $(0, \hat{\lambda}_0)$

score test, using 'null' pt. for information
require just one model fit
(last resort)

Example $\underset{\sim}{Y} \sim \text{Mult}(n; \underset{\sim}{\pi})$

$\underset{\sim}{\pi} = (\pi_1, \dots, \pi_m)$ $\quad \underset{\sim}{Y} = (Y_1, \dots, Y_m)$ $\quad \Sigma Y_i = n$

$\Sigma \pi_i = 1$

$$f(y; \pi) = \frac{n!}{y_1! \cdots y_m!} \pi_1^{y_1} \cdots \pi_m^{y_m}$$

$$\pi_m = 1 - \pi_1 - \dots - \pi_{m-1}$$

$$y_m = n - y_1 - \dots - y_{m-1}$$

W_1, W_e, W_u as fn of $\underset{\sim}{\pi}$

$$\left.\frac{\partial \ell(\underset{\sim}{\pi}; y)}{\partial \pi_j}\right|_{\hat{\underset{\sim}{\pi}}} = 0 \quad \text{etc}$$

$m-1$ eq's in $m-1$ unk.

$$\left.-\frac{\partial^2 \ell}{\partial \pi_j \partial \pi_k}\right|_{\hat{\pi}} = \underset{(m-1)\times(m-1)}{j(\hat{\pi})}$$

$$\hat{\pi}_j = \frac{y_j}{n} \qquad j(\pi) = i(\pi) \text{ bec exp'l family}$$

SM §4.5, CH Ex 2.17

$$i_{jk}(\pi) = \mathbb{1}\{j=k\} \cdot \frac{1}{\pi_j} - \frac{1}{\pi_m}$$

$$n \begin{bmatrix} \frac{1}{\pi_1} - \frac{1}{\pi_m} & & -\frac{1}{\pi_m} \\ & \ddots & \\ -\frac{1}{\pi_m} & & \frac{1}{\pi_{m-1}} - \frac{1}{\pi_m} \end{bmatrix}$$

$$\left[i^{-1}(\pi)\right]_{jk} = \frac{1}{n}\begin{cases} \pi_j(1-\pi_j) & j=k \\ -\pi_j\pi_k & j \neq k \end{cases}$$

$$\hat{\pi}_j = \frac{y_j}{n} \sim \cancel{Bin} \; \frac{1}{n} Bin(n, \pi_j)$$

$$w(\underline{\pi}) = 2\{\ell(\hat{\pi}) - \ell(\pi)\}$$

$$= 2 \sum_{j=1}^{m} (y_j \cdot \log \hat{\pi}_j - y_j \cdot \log \pi_j)$$

$$= 2 \sum_{j=1}^{m} y_j \cdot \log\left(\frac{y_j}{n\pi_j}\right) \qquad \sum O \log\left(\frac{O}{E}\right)$$

$$\frac{O}{E} = \frac{O-E}{E} + 1$$

$$w_e(\pi) = (\underline{\hat{\pi}} - \underline{\pi})^T \hat{j} (\hat{\pi} - \pi)$$

$$= \sum^{m} (y_j - n\pi_j)^2 \qquad \sum (O-E)^2$$

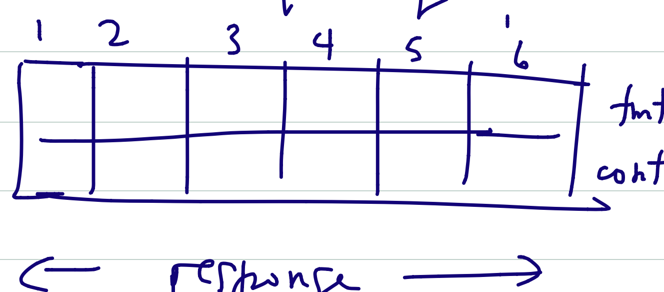
$$\sum_{j=1} \frac{\cdots}{y_j} \qquad \sum \frac{(O-E)}{O}$$

$$W_u(\pi) = \left(U(\pi) = \frac{\partial \ell}{\partial \pi} \right)^T j^{-1}(\pi) \left(U(\pi) \right)$$

$$= \cdots = \sum_{j=1}^{m} \frac{(y_j - n\pi_j)^2}{n\pi_j} \qquad \sum \frac{(O-E)^2}{E}$$

Typically $\pi = \pi(\beta)$ $\dim \beta = q < m-1$

e.g. if we had



$$\pi_{ij} = \pi_{i+} \cdot \pi_{+j}$$

$1 + 5 = q$

$$W(\beta) = 2 \sum_{i=1}^{m} y_i \log\left\{ \frac{y_i}{n\pi_i(\hat\beta)} \right\}$$

etc.

$$W_e(\beta) = \sum_{i=1}^{m} \frac{\{y_i - n\pi_i(\hat\beta)\}^2}{n\pi_i(\hat\beta)}$$