Directional tests for vector parameters

Nancy Reid

September 29, 2012

with Don Fraser, Nicola Sartori, Anthony Davison

Parametric models and likelihood

- model $f(y;\theta)$, $\theta \in \mathbb{R}^p$
- data $y = (y_1, \ldots, y_n)$ independent observations
- log-likelihood function $\ell(\theta; y) = \log f(y;\theta)$
- parameter of interest $\theta = (\psi, \lambda)$, $\psi \in \mathbb{R}^d$

- likelihood inference $w(\psi) = 2\{\ell(\hat{\psi}, \hat{\lambda}) - \ell(\psi, \hat{\lambda}_\psi)\}$

Example: 2×3 contingency table

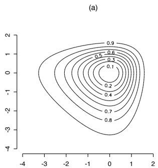
- contingency table on activity amongst psychiatric patients (Everitt, 1992 CH)

	Affective disorders	Schizophrenics	Neurotics
Retarded	12	13	5
Not retarded	18	17	25

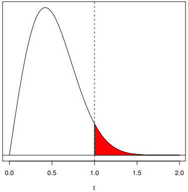
- model: log-linear $y \sim$ Poisson, $\log\{E(y)\} = X\theta, \quad \theta \in \mathbb{R}^6$
- log-likelihood $\ell(\theta; y) = \theta' X' y - 1' e^{X\theta} = \theta' s - c(\theta)$

- $\theta = (\psi, \lambda) \qquad \psi \in \mathbb{R}^2, \lambda \in \mathbb{R}^4$
- (ψ_1, ψ_2) interaction parameters
- $H_0 : \psi = \psi_0 = (0, 0)$ independence
- log-likelihood $\ell(\psi, \lambda; y) = \psi' \underline{s}_1 + \lambda' \underline{s}_2 - c(\psi, \lambda)$

Testing $\psi = 0$



$w(\psi_0) \dot{\sim} \chi^2_2$ *p*-value 0.047



directional *p*-value 0.050

'exact' conditional *p*-value 0.051

Directional tests

- log-likelihood

$$\ell(\psi, \lambda) = \psi' \underline{s}_1 + \lambda' \underline{s}_2 - c(\psi, \lambda)$$

- Step 1: get rid of nuisance parameter

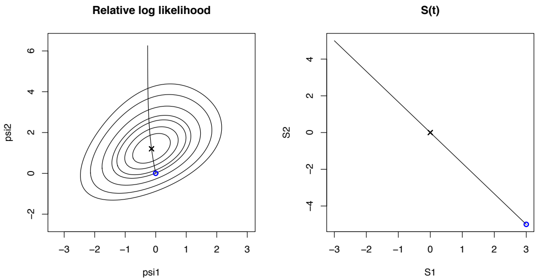
  use conditional density of $\underline{s}_1$, given $\underline{s}_2$
- Step 2: measure the directed departure from H_0

  s_ψ: expected value of $\underline{s}_1$, under H_0

  s^0: observed value of $\underline{s}_1$

  line through these two points
- $\mathcal{L}^* = ts^0 + (1-t)(s_\psi - s^0), \quad t \in \mathbb{R}$

... directional tests $\mathcal{L}^* = ts^0 + (1-t)s_\psi$



○ null hypothesis of independence $t = 0$

X observed value of s $t = 1$

$$p\text{-value} = \frac{\int_1^\infty g(t)dt}{\int_0^\infty g(t)dt}$$

like a 2-sided p-value

Pr (response $>$ observed $|$ response > 0)

... directional p-value

- p-value $= \frac{\int_1^\infty g(t)dt}{\int_0^\infty g(t)dt} = \frac{\int_1^\infty t^{d-1} f\{s(t);\psi\}dt}{\int_0^\infty t^{d-1} f\{s(t);\psi\}dt}$
- sample space, dimension $p\,(=6)$, $\quad f(s;\theta)$
- reduced space $\mathcal{L}_\psi$, dimension $d\,(=2)$, $\quad f(\underline{s}_1;\psi)$
- line $\mathcal{L}^*$, dimension 1, $\quad f\{s_1(t);\psi\}$
- directed departure – magnitude, given direction
  polar coordinates $\longrightarrow t^{d-1}$

Simplifications

- 1: $\mathcal{L}^* \subset \mathcal{L}_\psi \subset \mathbb{R}^p$
- 2: ratio of two integrals – drop any terms that don't depend on t
- 3: saddle point approximation to $f(s;\theta)$ gives a very accurate starting point

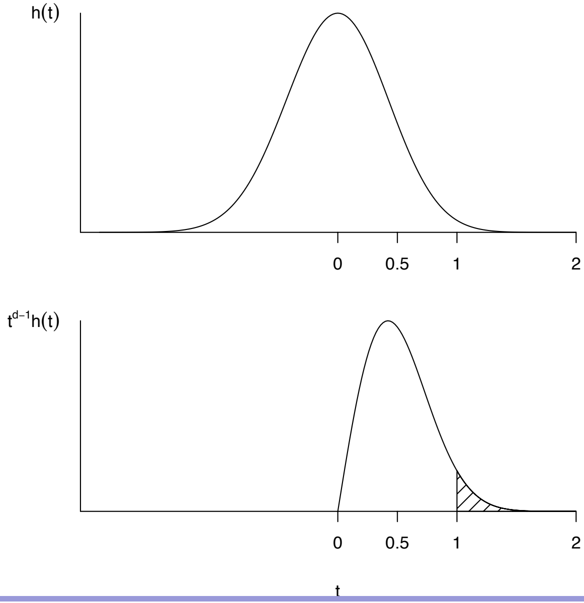
Aside: From full likelihood to density along the line

- suppose log-likelihood function $\ell(\varphi; s) = \varphi' s + \ell(\varphi; s^0)$
- such as in exponential families
- $\varphi = \varphi(\theta); \quad \theta = (\psi, \lambda) = (\psi(\varphi), \lambda(\varphi))$
- double saddle point approximation gives

$$\hat{f}(s; \psi) \doteq c \exp\left[\ell\{\varphi(\hat{\theta}^0_\psi); s\} - \ell\{\hat{\varphi}(s); s\}\right] \frac{|j_{\varphi\varphi}\{\hat{\varphi}(s); s\}|^{-1/2}}{|j_{(\lambda\lambda)}\{\varphi(\hat{\theta}^0_\psi; s)\}|^{-1/2}}$$

- $\hat{\theta}^0_\psi = (\psi, \hat{\lambda}^0_\psi)$
- $\hat{\varphi}(s)$ from $\ell_\varphi(\varphi; s) = 0$
- uses only likelihood for the full model, plus constrained maximum likelihood estimator
- density along the line: use $f\{s(t); \psi\}, t \in \mathbb{R}$
- $\hat{f}(s; \psi)$ approximates the marginal likelihood (Kalbfleisch & Sprott, 1970, 1974)

2 × 3 table



$t = 0$

10	10	10
20	20	20

$t = 0.5$

11.0	11.5	7.5
19.0	18.5	22.5

$t = 1$

12	13	5
18	17	25

$t = 2$

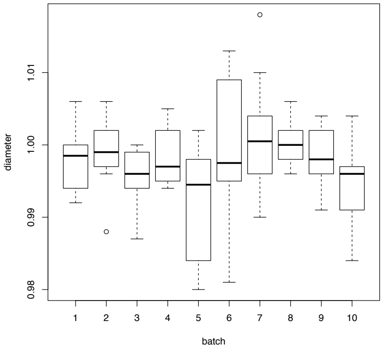
14	16	0
16	14	30

Simulations, 2 × 3 table

Nominal	0.010	0.025	0.050	0.100	0.250	0.500
LRT	0.011	0.028	0.055	0.107	0.260	0.510
Directional	0.010	0.024	0.050	0.100	0.250	0.501
Skovgaard, 2001	0.010	0.025	0.050	0.101	0.251	0.501

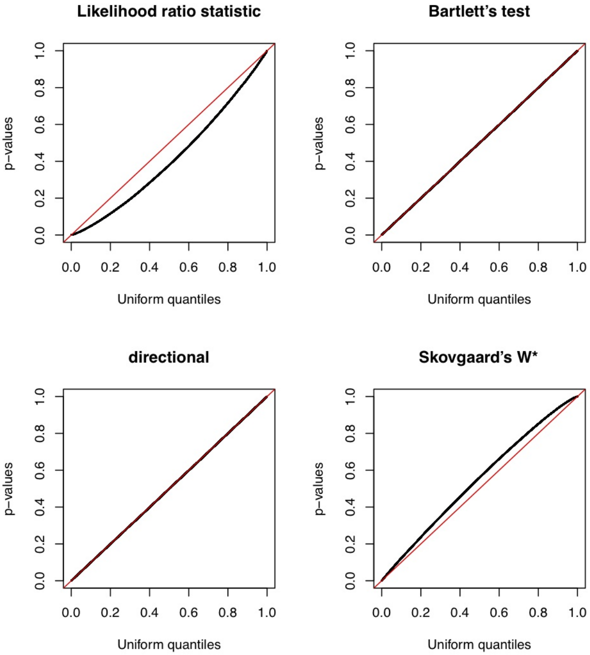
Nominal	0.750	0.900	0.950	0.975	0.990
LRT	0.757	0.905	0.952	0.974	0.992
Directional	0.752	0.902	0.950	0.973	0.992
Skovgaard, 2001	0.752	0.900	0.950	0.973	0.991

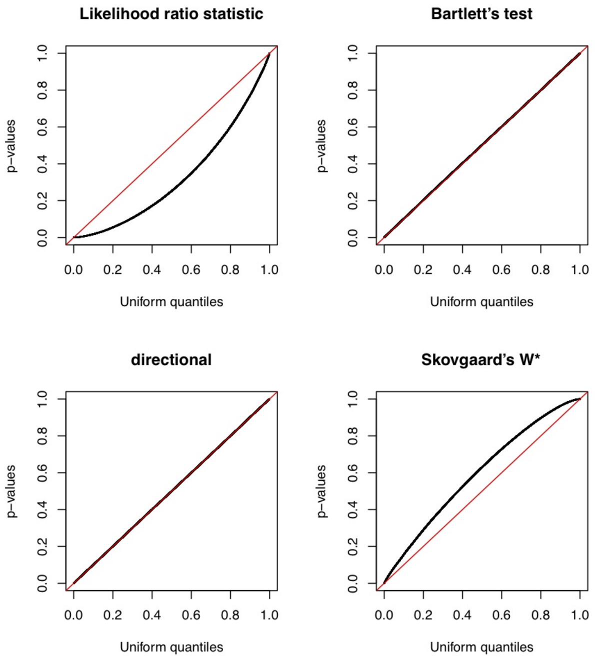
Example: comparison of normal variances



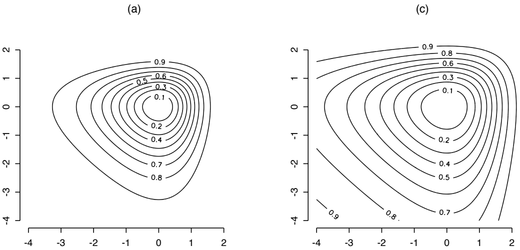
Likelihood ratio statistic	0.0042
Directional	0.0389
Skovgaard, 2001	0.0622
Bartlett's test	0.0136

F-test


Likelihood ratio statistic
Bartlett's test
directional
Skovgaard's W*
p−values
Uniform quantiles
0.0
0.2
0.4
0.6
0.8
1.0

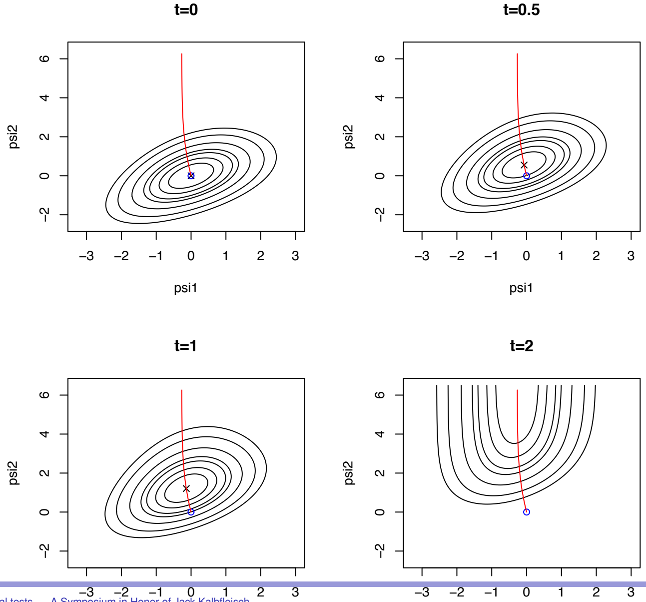

Likelihood ratio statistic
Bartlett's test
directional
Skovgaard's W*
p−values
Uniform quantiles
0.0
0.2
0.4
0.6
0.8
1.0

Aside: Bartlett correction of likelihood ratio test



- $w(\psi) = 2\{\ell(\hat{\theta}) - \ell(\hat{\theta}_\psi)\} \dot{\sim} \chi^2_d$
- $E\{w(\psi)\} = d\{1 + \dfrac{B(\psi)}{n} + O(n^{-2})$
- $\tilde{w}(\psi) = \dfrac{w(\psi)}{1 + B(\psi)/n} \dot{\sim} \chi^2_d \qquad O(n^{-3})$

Aside: compare directional test



Other examples

- binary regression
- equality of exponential rates
- covariance selection models

- comparison of normal variances
- symmetry restrictions in contingency tables

- ...

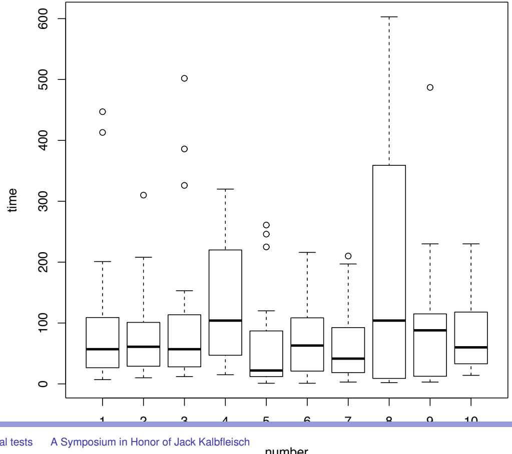
Binary regression

- binary response
  – presence/absence of bacteria   MASS, §10.4
- covariate of interest
  – week: factor variable with levels 0, 2, 4, 6, 11
- nuisance parameter
  – subject effects: factor with 50 levels
- 108 observations
  – 24 nuisance parameters, 4 parameters of interest

Likelihood ratio statistic	0.0005
Directional	0.0053
Skovgaard, 2001	0.0043

Equality of exponential rates

- $y_{ij} \sim \exp(\theta_i), \quad j = 1, \ldots, n_i; i = 1, \ldots, g$
- $H_0 : \theta_1 = \cdots = \theta_g$
- Example: air-conditioning data, Cox & Snell (1981)



... exponential rates

- p-value = $\dfrac{\int_1^\infty t^{d-1} h(t) dt}{\int_0^\infty t^{d-1} h(t) dt}$
- $h(t) \propto \prod_{i=1}^{g} \left\{ 1 - \dfrac{t(\bar{y} - \bar{y}_i)}{\bar{y}} \right\}^{n_i - 1}$

Nominal	0.010	0.025	0.050	0.100	0.250	0.500
LRT	0.013	0.031	0.059	0.115	0.275	0.529
Directional	0.010	0.025	0.049	0.099	0.248	0.500
Skovgaard, 2001	0.008	0.020	0.039	0.080	0.210	0.441

Nominal	0.750	0.900	0.950	0.975	0.990
LRT	0.770	0.910	0.955	0.977	0.990
Directional	0.750	0.899	0.949	0.975	0.989
Skovgaard, 2001	0.688	0.855	0.918	0.953	0.977

Conclusion

- new way to assess vector parameters
- incorporates information in the direction of departure
- easy to compute: two model fits, plus 1-d numerical integration
- accurate conditionally, by construction, and unconditionally (simulations)
- can be used in models of practical interest – discrete and continuous data
- exponential family model not necessary – easily generalized using approximate exponential family model

Thank you!

References

Fraser, D.A.S. and Massam, H. (1985). Conical tests. Stat. Hefte

Skovgaard, I.M. (1988). Saddlepoint expansions for directional test probabilities. JRSS B

Skovgaard, I.M. (2001). Likelihood asymptotics. Scand. J. Statist. **28**, 3–32.

Cheah, P.K., Fraser, D.A.S., Reid, N. (1994). Multiparameter testing in exponential models. Biometrika

Davison, A.C., Fraser, D.A.S., Reid, N. (2006). Improved likelihood inference for discrete data. JRSS B

Davison, A.C., Fraser, D.A.S., Reid, N., Sartori, N. (2012). Directional inference for vector parameters.