Higher-order accuracy of multiscale double-bootstrap resampling for testing regions

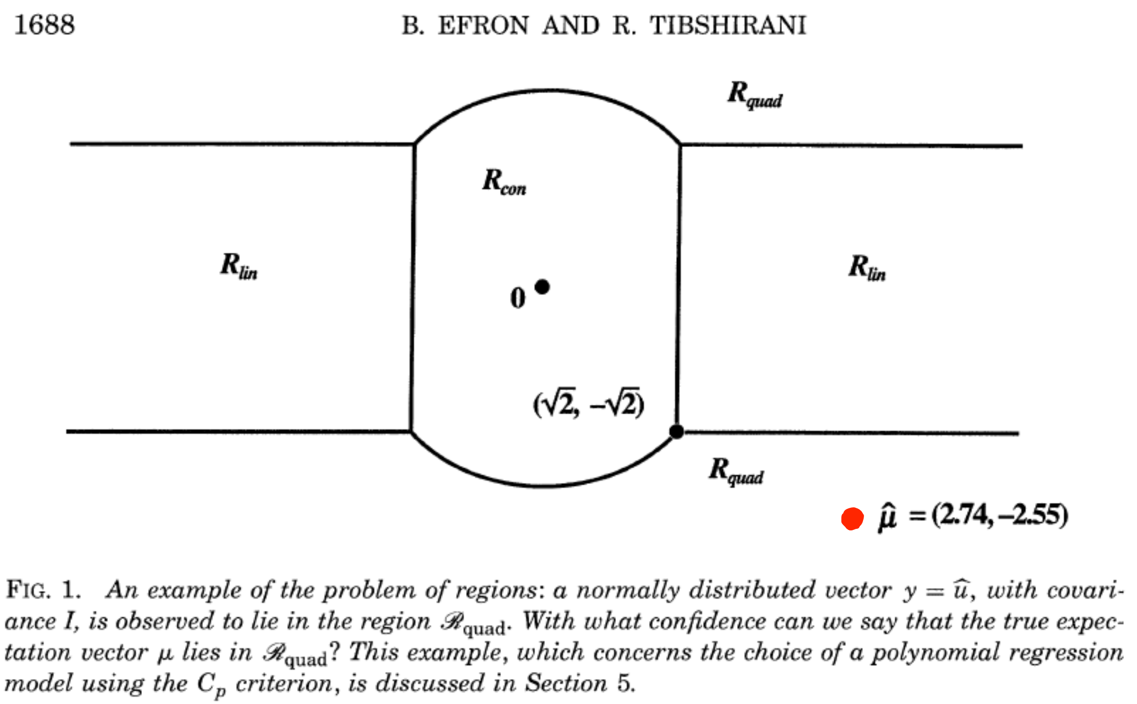
Hidetoshi Shimodaira
Osaka University

Bernoulli Society Satellite Meeting to the ISI World Statistics Congress 2013
Asymptotic Statistics and Related Topics: Theories and Methodologies
Sanjo Conference Hall, The University of Tokyo
2-4 September 2013

“The problem of regions”
Efron and Tibshirani (1998)

(example) Model Selection of Polynomial Regression

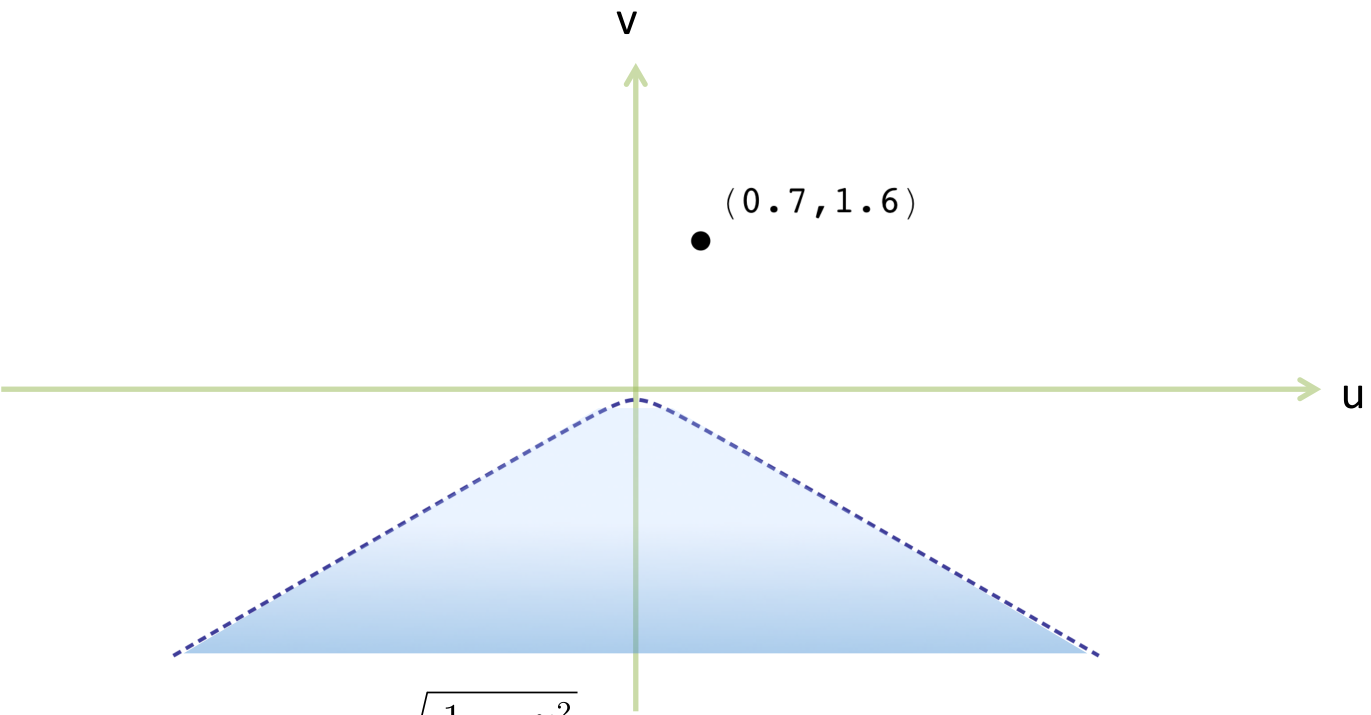
0: constant, 1: linear, 2: quadratic



1688 B. EFRON AND R. TIBSHIRANI

FIG. 1. *An example of the problem of regions: a normally distributed vector* $y = \widehat{u}$, *with covariance* I, *is observed to lie in the region* $\mathscr{R}_{\text{quad}}$. *With what confidence can we say that the true expectation vector* μ *lies in* $\mathscr{R}_{\text{quad}}$? *This example, which concerns the choice of a polynomial regression model using the* C_p *criterion, is discussed in Section* 5.

Region and a data-point



$$v \leq -\sqrt{\frac{1}{64} + \frac{u^2}{3}}$$

Multivariate Normal Model

$$y = (u, v) \in \mathbb{R}^{q+1} \qquad q = \dim u$$

$$Y \sim N_{q+1}(\mu, I_{q+1})$$

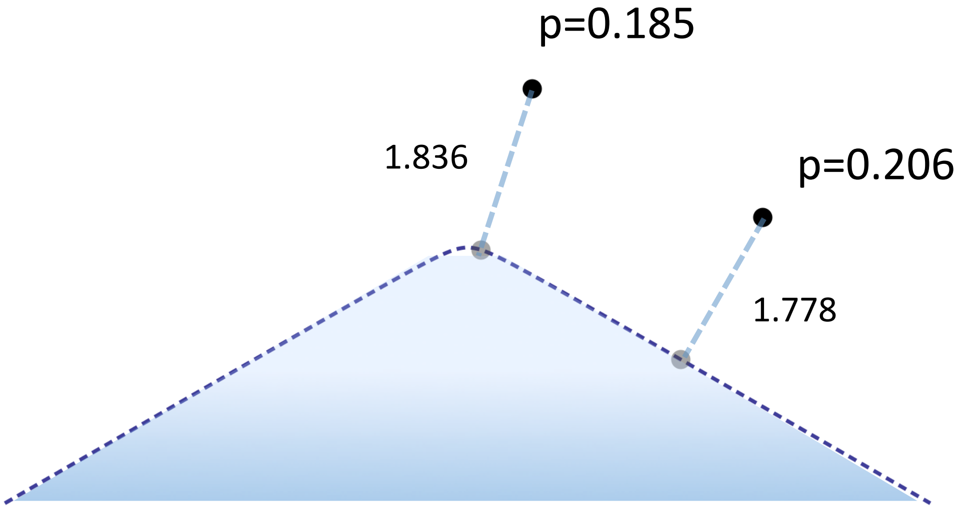
$$H = \left\{ (u, v) \mid v \leq -h(u), u \in \mathbb{R}^q \right\}$$

region: $H = \mathcal{R}(h)$ boundary surface: $\partial H = \mathcal{B}(h)$

Null hypothesis: $\mu \in H$ v.s. Alternative hypothesis: $\mu \notin H$

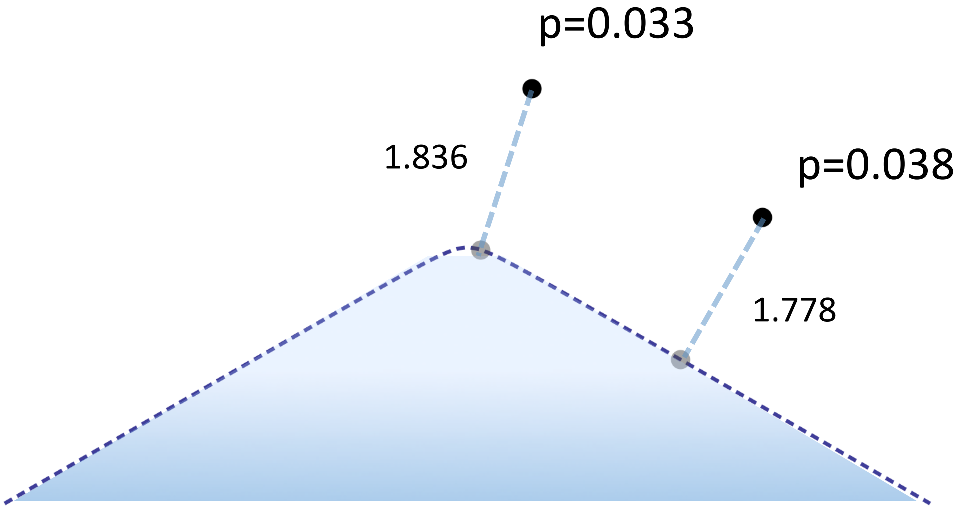
Chi-square test (very conservative)

$\text{distance}^2 \sim$ chi-square distribution with df=2



Normal test (rejecting too much)

distance ~ N(0,1)



Bootstrap probability (=Bayesian PP)

$$Y^* \sim N_{q+1}(y, \sigma^2 I_{q+1}) \qquad \sigma^2 = 1$$

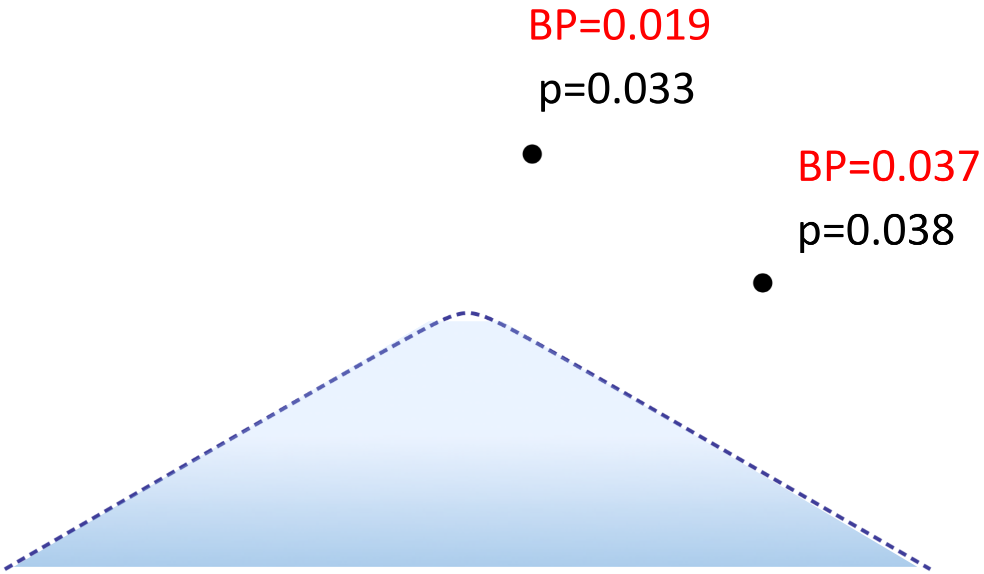
$Y^{*1}, \ldots, Y^{*B}$ ➡ $\widehat{\mathrm{BP}}_{\sigma^2}(H|y) = \dfrac{\#\{Y^{*b} \in H, b = 1, \ldots, B\}}{B}$

$$\mathrm{BP}_{\sigma^2}(H|y) = P(Y^* \in H|y)$$

BP is interpreted as the Bayesian posterior probability of H
if the prior distribution of mu is uniform.

Efron and Tibshirani (1998)

BP is even worse



Double bootstrap probability

Projection of y onto the boundary surface:

$$\hat{\mu}(H|y) = \underset{\mu \in \partial H}{\arg\min} \|y - \mu\|$$

Adjusting BP using resampling from the projection

$$Y^{+} \sim N_{q+1}(\hat{\mu}(H|y), \tau^2 I_{q+1}) \qquad \tau^2 = 1$$

$$\mathrm{DBP}_{\sigma^2,\tau^2}(H|y) = P\Big[\mathrm{BP}_{\sigma^2}(H|Y^{+}) \le \mathrm{BP}_{\sigma^2}(H|y) \mid \hat{\mu}(H|y)\Big]$$

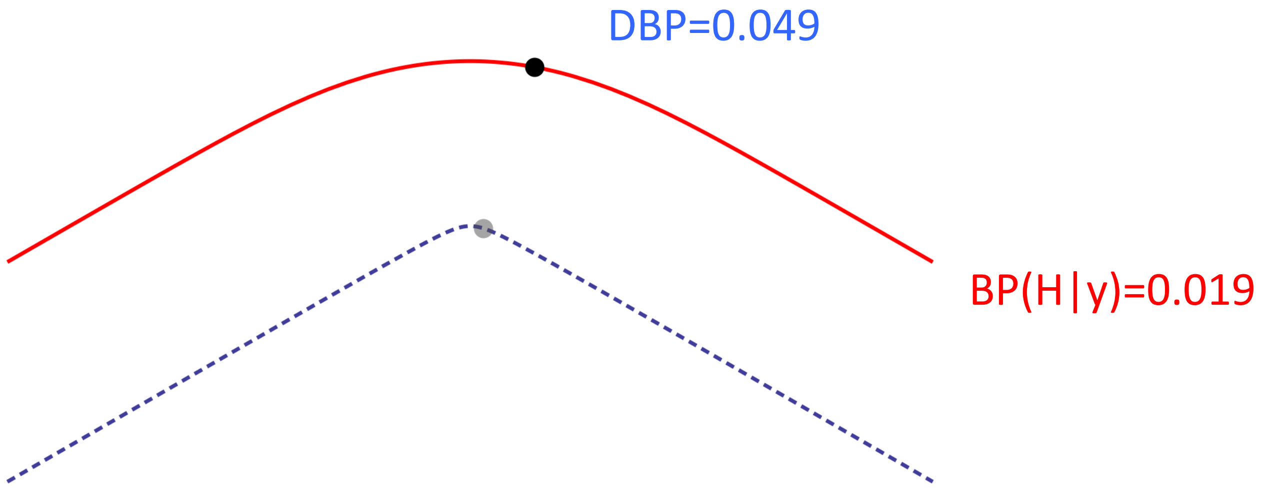
Hall (1992), Efron and Tibshirani (1998)

contour surface of BP=0.019

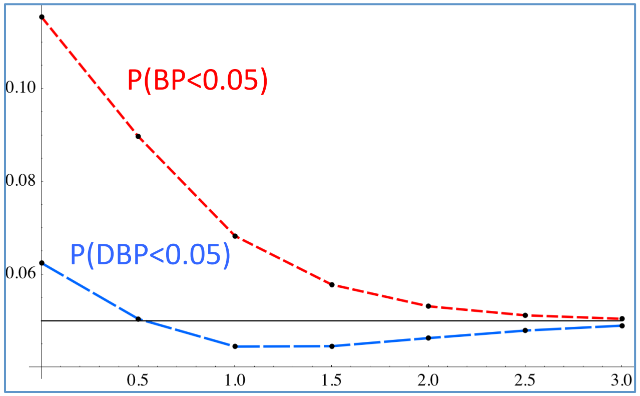
Computing DBP:

$$Y^{+} \sim N_2(\hat{\mu}, I_2)$$

$$\mathrm{DBP} = P\Big[\mathrm{BP}(H|Y^{+}) \leq 0.019 \mid \hat{\mu}\Big]$$

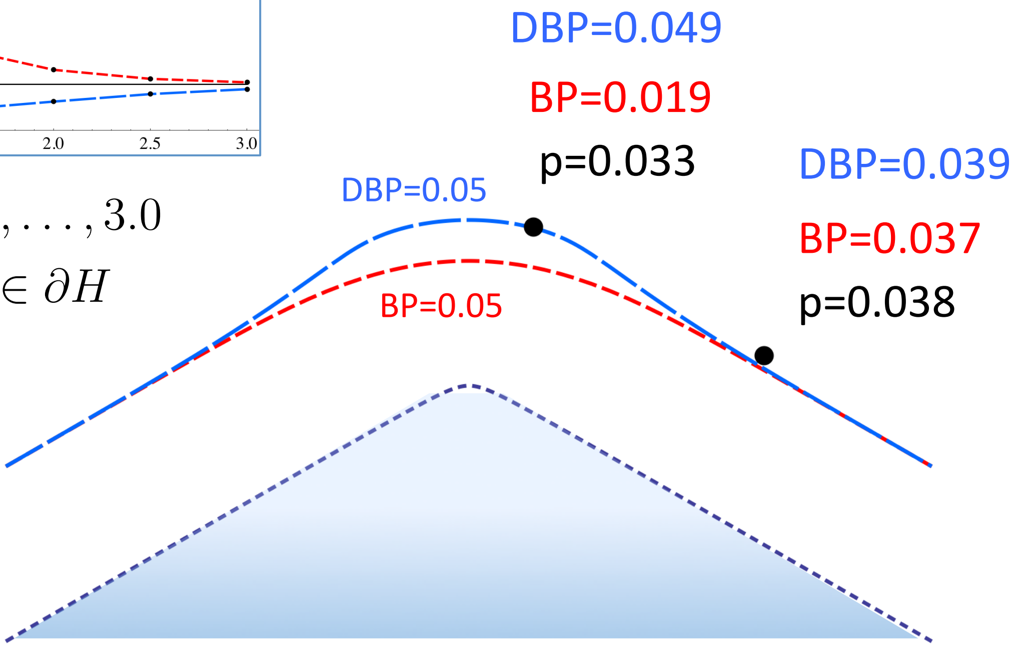


DBP adjusts the bias of BP



$$\theta = 0.0, 0.5, 1.0, \ldots, 3.0$$

$$\mu = (\theta, -h(\theta)) \in \partial H$$



Approximately unbiased p-values via Multiscale bootstrap

m out of n bootstrap : Politis and Romano (1994), Bickel et al. (1997)

$$\mathcal{X} = \{x_1, \ldots, x_n\} \quad \Rightarrow \quad \mathcal{X}^* = \{x_1^*, \ldots, x_m^*\}$$

The idea of multiscale bootstrap : Shimodaira (2002, 2004, 2008)

$$\sigma^2 = \frac{n}{m}$$

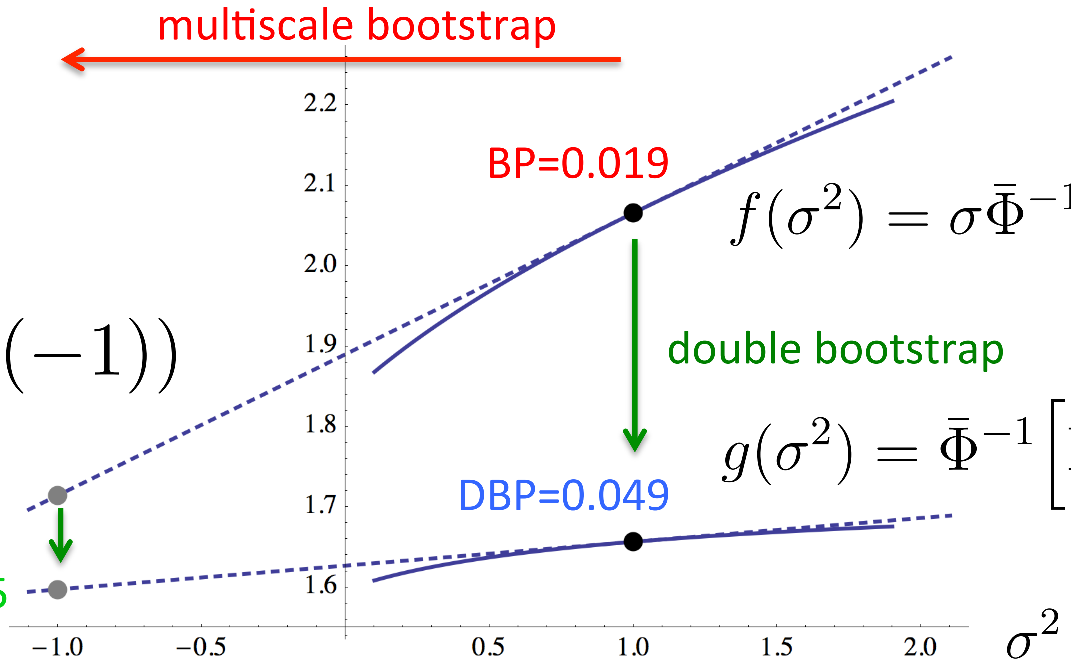
We compute BP for $\sigma_1^2, \ldots, \sigma_S^2$ and extrapolate BP to $\sigma^2 = -1$

(equivalently $m = -n$)

The BP with m = -n is denoted as AU (= Approximately Unbiased)

Extrapolation to sigma^2=-1

$$\bar{\Phi}(z) = 1 - \Phi(z) = \int_z^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\, dx$$



$$f(\sigma^2) = \sigma \bar{\Phi}^{-1}\left[\mathrm{BP}_{\sigma^2}(H|y)\right]$$

$$g(\sigma^2) = \bar{\Phi}^{-1}\left[\mathrm{DBP}_{\sigma^2,1}(H|y)\right]$$

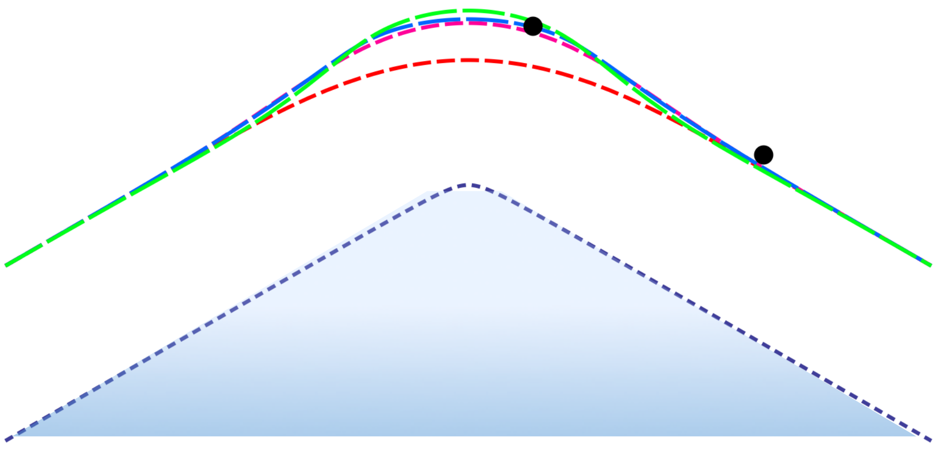
$$\mathrm{AU} = \bar{\Phi}(f(-1))$$

$$\mathrm{DAU} = \bar{\Phi}(g(-1))$$

We apply the multiscale bootstrap to DBP for getting DAU (THIS PRESENTATION)
Equivalently, we could say applying double bootstrap to AU for getting DAU

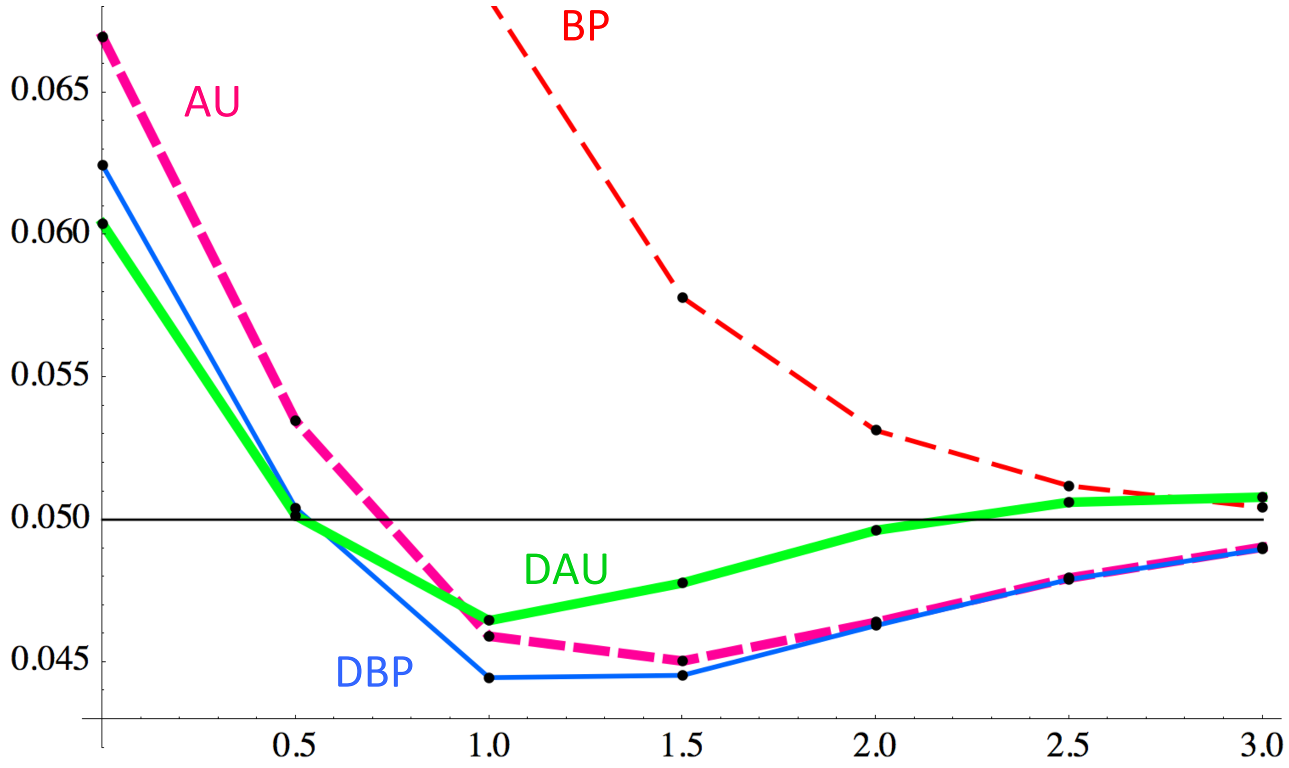
contour curves of p=0.05

BP, AU, DBP, DAU



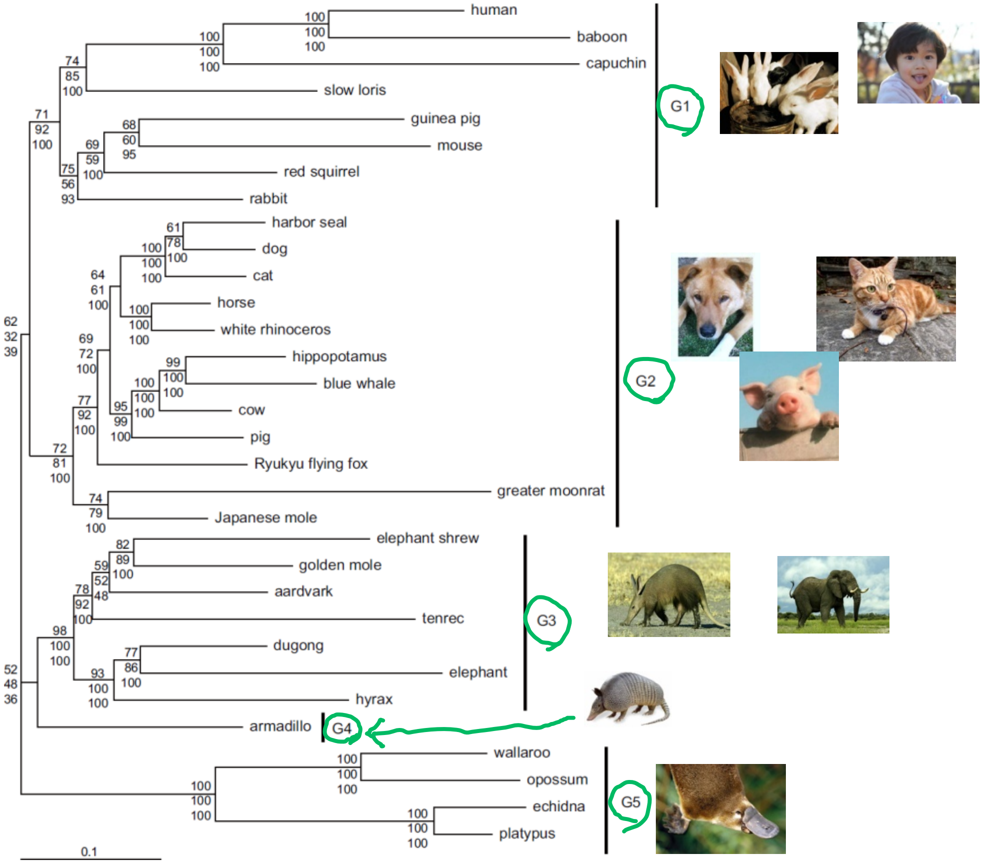
Rejection probabilities $P(p<0.05)$

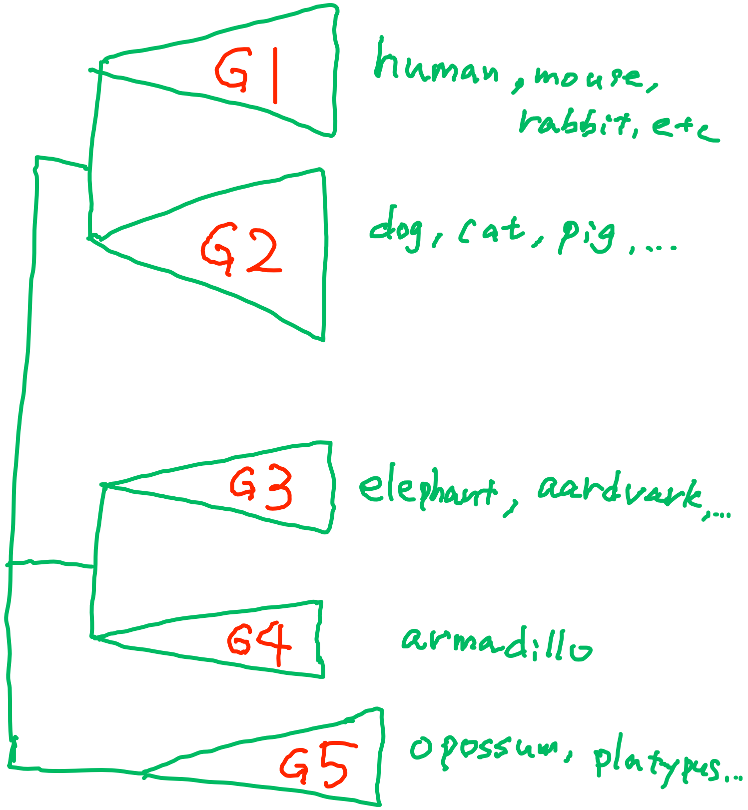
Error: DAU < {DBP, AU} < BP



$\theta = 0.0, 0.5, 1.0, \ldots, 3.0$ $\mu = (\theta, -h(\theta)) \in \partial H$

Evolution of mammal species





ML tree topology: ((G1,G2), (G3,G4), G5)

Fig 1 of Shimodaira and Hasegawa (2005) from the book (ed. Nielsen)
Data: mt protein sequences of n=3392 amino acids for s=32 species

Comparing 15 trees (cont.)

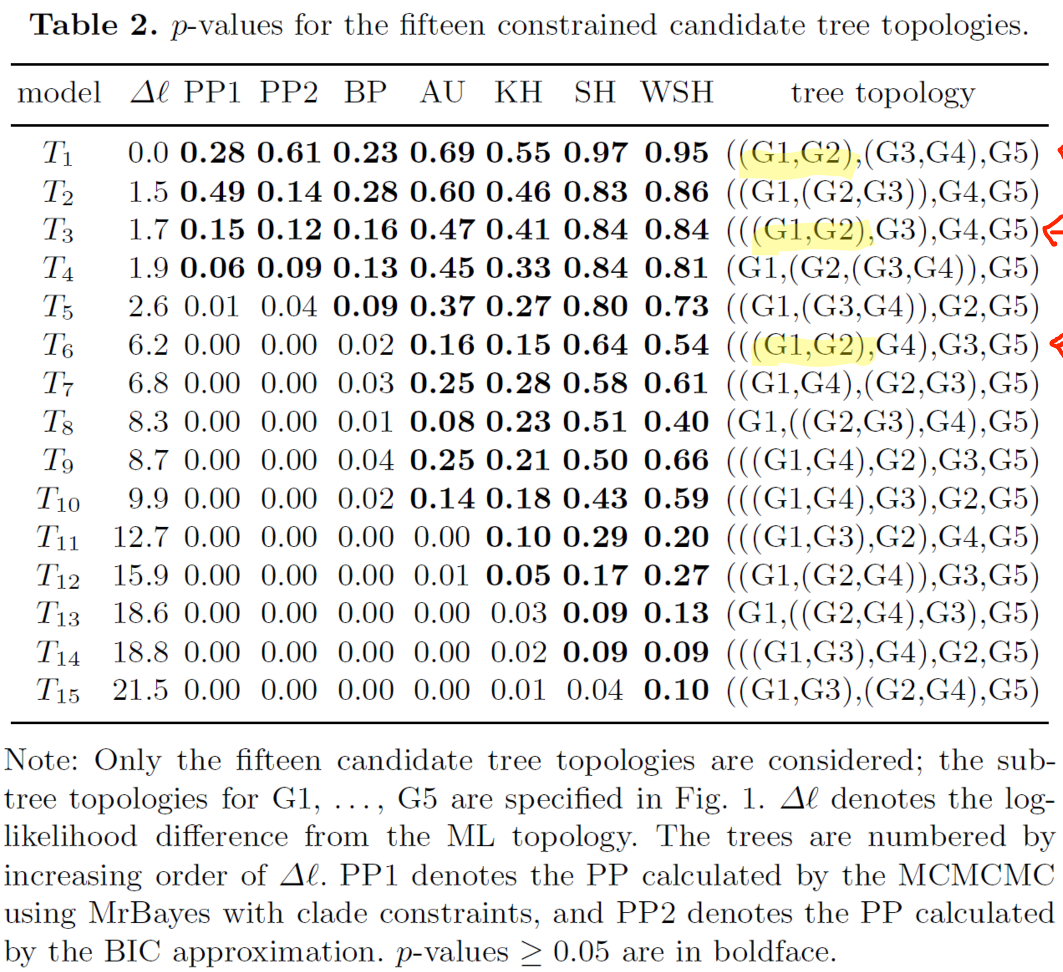


Table 2. *p*-values for the fifteen constrained candidate tree topologies.

model	$\Delta\ell$	PP1	PP2	BP	AU	KH	SH	WSH	tree topology
T_1	0.0	**0.28**	**0.61**	**0.23**	**0.69**	**0.55**	**0.97**	**0.95**	((G1,G2),(G3,G4),G5) ← ③
T_2	1.5	**0.49**	**0.14**	**0.28**	**0.60**	**0.46**	**0.83**	**0.86**	((G1,(G2,G3)),G4,G5)
T_3	1.7	**0.15**	**0.12**	**0.16**	**0.47**	**0.41**	**0.84**	**0.84**	(((G1,G2),G3),G4,G5) ← ②
T_4	1.9	**0.06**	**0.09**	**0.13**	**0.45**	**0.33**	**0.84**	**0.81**	(G1,(G2,(G3,G4)),G5)
T_5	2.6	0.01	0.04	**0.09**	**0.37**	**0.27**	**0.80**	**0.73**	((G1,(G3,G4)),G2,G5)
T_6	6.2	0.00	0.00	0.02	**0.16**	**0.15**	**0.64**	**0.54**	(((G1,G2),G4),G3,G5) ← ①
T_7	6.8	0.00	0.00	0.03	**0.25**	**0.28**	**0.58**	**0.61**	((G1,G4),(G2,G3),G5)
T_8	8.3	0.00	0.00	0.01	**0.08**	**0.23**	**0.51**	**0.40**	(G1,((G2,G3),G4),G5)
T_9	8.7	0.00	0.00	0.04	**0.25**	**0.21**	**0.50**	**0.66**	(((G1,G4),G2),G3,G5)
T_{10}	9.9	0.00	0.00	0.02	**0.14**	**0.18**	**0.43**	**0.59**	(((G1,G4),G3),G2,G5)
T_{11}	12.7	0.00	0.00	0.00	0.00	**0.10**	**0.29**	**0.20**	(((G1,G3),G2),G4,G5)
T_{12}	15.9	0.00	0.00	0.00	0.01	**0.05**	**0.17**	**0.27**	((G1,(G2,G4)),G3,G5)
T_{13}	18.6	0.00	0.00	0.00	0.00	0.03	**0.09**	**0.13**	(G1,((G2,G4),G3),G5)
T_{14}	18.8	0.00	0.00	0.00	0.00	0.02	**0.09**	**0.09**	(((G1,G3),G4),G2,G5)
T_{15}	21.5	0.00	0.00	0.00	0.00	0.01	0.04	**0.10**	((G1,G3),(G2,G4),G5)

Note: Only the fifteen candidate tree topologies are considered; the subtree topologies for G1, …, G5 are specified in Fig. 1. $\Delta\ell$ denotes the log-likelihood difference from the ML topology. The trees are numbered by increasing order of $\Delta\ell$. PP1 denotes the PP calculated by the MCMCMC using MrBayes with clade constraints, and PP2 denotes the PP calculated by the BIC approximation. *p*-values ≥ 0.05 are in boldface.

Recent Results

Nishihara et al (2008)

Genome-Scale analysis of 1 Mbp

Asymptotic theory of 4th order accuracy

$$h(u) = \sum_{i=1}^{q}\sum_{j=1}^{q} h_{ij}u_iu_j + \sum_{i=1}^{q}\sum_{j=1}^{q}\sum_{k=1}^{q} h_{ijk}u_iu_ju_k + \cdots$$

$$h_{ij} = \frac{1}{2}\frac{\partial^2 h(u)}{\partial u_i \partial u_j}\Big|_0 = O(n^{-1/2}) \qquad h_{ijk} = \frac{1}{6}\frac{\partial^3 h(u)}{\partial u_i \partial u_j \partial u_k}\Big|_0 = O(n^{-1})$$

The k-th order derivatives are $O(n^{-(k-1)/2})$ for $k \geq 1$, because

the coordinates $u_1, \ldots, u_q$ as well as $h(u)$ are scaled by the factor $\sqrt{n}$

(Class $\mathcal{S}$)

We take care of terms up to $O(n^{-3/2})$ ignoring $O(n^{-2})$

$$h(u) \simeq h_0 + h_iu_i + h_{ij}u_iu_j + h_{ijk}u_iu_ju_k + h_{ijkl}u_iu_ju_ku_l$$

$$h_0 = O(1),\ h_i = O(n^{-1}),\ h_{ij} = O(n^{-1/2}),\ h_{ijk} = O(n^{-1}),\ h_{ijkl} = O(n^{-3/2})$$

Thm: Asymptotic expansion of BP1

$$\mathrm{BP}_1(H|y) \simeq 1 - \Phi(\beta_0 + \beta_1 + \beta_2)$$

Proved by a simple argument of Taylor expansion and integration.

data point $y = (0, \lambda_0 - h_0)$

signed distance $\beta_0 = \lambda_0 = O(1)$

curvature + ... $\beta_1 = \underline{\gamma_1} - \lambda_0\gamma_2 + \frac{4}{3}\lambda_0^2\gamma_3 = O(n^{-1/2})$

fourth-order terms

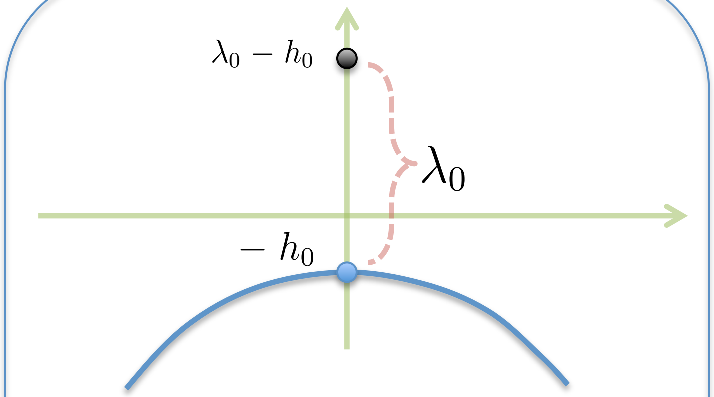
$$\left\{\begin{array}{l} \beta_2 = 3\gamma_4 - \gamma_1\gamma_2 - \frac{4}{3}\gamma_3 = O(n^{-3/2}) \\ \beta_3 = 6\gamma_4 - 2\gamma_1\gamma_2 - 4\gamma_3 = O(n^{-3/2}) \end{array}\right.$$

$$\gamma_1 = \frac{1}{2}\frac{\partial^2 h(u)}{\partial u_i \partial u_i}\bigg|_0$$

mean curvature of the surface

$$\beta_3 = \frac{1}{2}\frac{\partial^2 \gamma_1(h,u)}{\partial u_i \partial u_i}\bigg|_0$$

mean curvature of the mean curvature



$\underline{\gamma_1 = h_{ii} = O(n^{-1/2})}$

$\gamma_2 = h_{ij}h_{ij} = O(n^{-1})$

$\gamma_3 = h_{ij}h_{jk}h_{ki} = O(n^{-3/2})$

$\gamma_4 = h_{iijj} = O(n^{-3/2})$

Thm: scaling law of BP

$$\mathrm{BP}_{\sigma^2}(H|y) \simeq 1 - \Phi\left[\beta_0\sigma^{-1} + \beta_1\sigma + \beta_2\sigma^3\right]$$

$$\sigma\bar{\Phi}^{-1}\left[\mathrm{BP}_{\sigma^2}(H|y)\right] \simeq \beta_0 + \beta_1\sigma^2 + \beta_2\sigma^4$$

Proved by a simple rescaling argument.

$$\mathrm{BP}_{\sigma^2}(H|y) = \mathrm{BP}_1(H/\sigma|y/\sigma)$$

$$\beta_0 \to \beta_0\sigma^{-1}, \quad \beta_1 \to \beta_1\sigma, \quad \beta_2 \to \beta_2\sigma^3$$

$$\lambda_0 \to \lambda_0/\sigma$$

$$h_0 \to h_0/\sigma, h_i \to h_i, h_{ij} \to \sigma h_{ij}, h_{ijk} \to \sigma^2 h_{ijk}, h_{ijkl} \to \sigma^3 h_{ijkl}$$

$$\gamma_1 \to \sigma\gamma_1, \gamma_2 \to \sigma^2\gamma_2, \gamma_3 \to \sigma^3\gamma_3, \gamma_4 \to \sigma^3\gamma_4$$

Thm: unbiased p-value

Def: k-th order accurate p-values should satisfy

$$P\Big[\mathrm{PV}(H|Y) < \alpha \mid \mu\Big] = \alpha + \underset{\text{error}}{\underline{O(n^{-k/2})}}, \quad \mu \in \partial H$$

Thm: fourth-order accuracy (k=4) is achieved by

$$\mathrm{PV}(H|y) \simeq 1 - \Phi\Big[\underset{\underset{O(1)}{\uparrow}}{\beta_0} - \underset{\underset{O(n^{-1/2})}{\uparrow}}{\beta_1} - \underset{\underset{O(n^{-3/2})}{\uparrow}}{\beta_2} + \underset{\uparrow}{\beta_3}\Big]$$

Corollary: BP is first-order accurate (k=1), AU is third-order accurate (k=3)

$$\sigma\bar{\Phi}^{-1}\Big[\mathrm{BP}_{\sigma^2}(H|y)\Big] \simeq \beta_0 + \beta_1\sigma^2 + \beta_2\sigma^4$$

$$\mathrm{BP} = \mathrm{PV} + O(n^{-1/2}) \qquad \mathrm{AU} = \mathrm{PV} + O(n^{-3/2})$$

Rejection probabilities of BP and AU

BP is first-order accurate (k=1)

$$P\Big(\mathrm{BP}(H|Y) < \alpha\Big) = \Phi(z_\alpha + \underline{2\gamma_1}) + O(n^{-1}) = \alpha + O(n^{-1/2})$$

$$\boxed{z_\alpha = \Phi^{-1}(\alpha)}$$

AU is third-order accurate (k=3)

$$P\Big(\mathrm{AU}(H|Y) < \alpha\Big) \simeq \Phi(z_\alpha + \underline{\tfrac{4}{3}\gamma_3}) = \alpha + O(n^{-3/2})$$

Using q x q hessian matrix $D = \left(\left.\frac{\partial^2 h(u)}{\partial u_i \partial u_j}\right|_0 : i,j = 1,\dots,q \right)$

$$\gamma_1 = h_{ii} = \tfrac{1}{2}\,\mathrm{tr}(D) \qquad \gamma_3 = h_{ij}h_{jk}h_{ki} = \tfrac{1}{8}\,\mathrm{tr}(D^3)$$

sketch of the proof for PV

Contour surface of BP $s = \mathcal{L}_{\sigma^2}(h, a)$

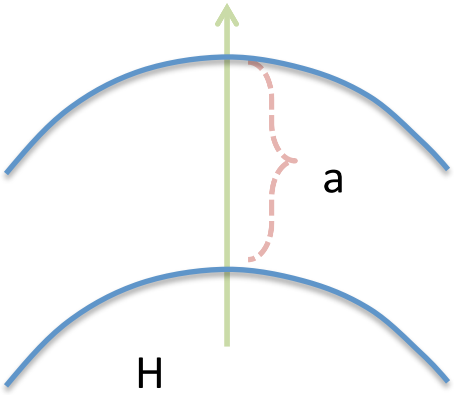
is defined by $\mathrm{BP}_{\sigma^2}(H|y) = \text{constant for any } y \in \mathcal{B}(s)$

additivity $\mathcal{L}_{\sigma_2^2}(\mathcal{L}_{\sigma_1^2}(H, a_1), a_2) \doteq \mathcal{L}_{\sigma_1^2+\sigma_2^2}(H, a_1 + a_2)$

identity $h = \mathcal{L}_0(h, 0)$

inverse $\mathcal{L}_{\sigma^2}^{-1}(\cdot, a) \doteq \mathcal{L}_{-\sigma^2}(\cdot, -a)$

$\mathcal{B}(s) = \{(u, v) : v = -s(u)\}$



Consider contour surface of PV: $\mathrm{PV}(H|y) = \text{constant for any } y \in \mathcal{B}(s)$

then unbiasedness requires $\mathrm{BP}_1(\mathcal{R}(s)|\mu) = \text{constant for any } \mu \in \mathcal{B}(h)$

$$h = \mathcal{L}_1(s, -\lambda_0) \Rightarrow s = \mathcal{L}_{-1}(h, \lambda_0)$$

Thm: scaling-law of DBP

$$\mathrm{DBP}_{\tau^2,\sigma^2}(H|y) \simeq 1 - \Phi\left[\beta_0\tau^{-1} - \beta_1\tau - \beta_2\tau^3 - \beta_3\tau\sigma^2\right]$$

$$\bar{\Phi}^{-1}\left[\mathrm{DBP}_{1,\sigma^2}(H|y)\right] \simeq (\beta_0 - \beta_1 - \beta_2) \underline{- \beta_3\sigma^2}$$

Corollary: DBP is third-order accurate (k=3), DAU is fourth-order accurate (k=4)

$$\mathrm{PV}(H|y) \simeq 1 - \Phi\left[\beta_0 - \beta_1 - \beta_2 + \beta_3\right]$$

$$\mathrm{DBP} = \mathrm{PV} + O(n^{-3/2}) \qquad \mathrm{DAU} = \mathrm{PV} + O(n^{-2})$$

sketch of the proof for DBP

$s = \mathcal{L}_{\sigma^2}(h, \lambda_0)$ contour surface of BP

$$\widetilde{\mathrm{DBP}}_{\tau^2,\sigma^2}(H|y) = 1 - \mathrm{BP}_{\tau^2}(\mathcal{R}(s)|\tilde{\mu})$$

The proof completes by applying the asymptotic expansion of BP to R(s)

Rejection probabilities of DBP and DAU

$$P\Big(\mathrm{DBP}_{1,\sigma^2}(H|Y) < \alpha\Big) \simeq \Phi\Big[z_\alpha - \underline{(1+\sigma^2)\beta_3}\Big]$$

DBP is third-order accurate (k=3)

$$P\Big(\mathrm{DBP}(H|Y) < \alpha\Big) \simeq \Phi(z_\alpha - \underline{2\beta_3}) = \alpha + O(n^{-3/2})$$

DAU is fourth-order accurate (k=4)

$$P\Big(\mathrm{DAU}(H|Y) < \alpha\Big) \simeq \Phi(z_\alpha) = \alpha$$

$$\gamma_1 = \frac{1}{2}\frac{\partial^2 h(u)}{\partial u_i \partial u_i}\Big|_0$$

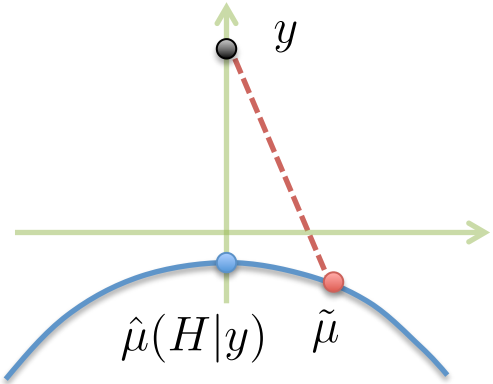
mean curvature of the surface

$$\beta_3 = \frac{1}{2}\frac{\partial^2 \gamma_1(h,u)}{\partial u_i \partial u_i}\Big|_0$$

mean curvature of the mean curvature

Robustness to projection error

If $\hat{\mu}(H|y) = (0, -h(0)) \in \partial H$ is replaced by $\tilde{\mu} = (\theta, -h(\theta)) \in \partial H$



DBP becomes

$$\widetilde{\mathrm{DBP}}_{\tau^2,\sigma^2}(H|y) \simeq 1 - \Phi\left[\beta_0\tau^{-1} - \beta_1\tau - \beta_2\tau^3 - \beta_3\tau\sigma^2 \underline{- \tau^{-1}(\tau^2+\sigma^2)\tilde{\Delta}(\theta)}\right]$$

error = $O(n^{-1})$

$$\tilde{\Delta}(\theta) = (3h_{mmi} - 6\lambda_0 h_{ml}h_{mli})\theta_i + (6h_{mmij} - 2\gamma_1 h_{mi}h_{mj} - 4h_{ml}h_{mi}h_{lj})\theta_i\theta_j = O(n^{-1})$$

Corollary: DBP becomes only second-order accurate (k=2), but
DAU keeps fourth-order accuracy (k=4)

$$\widetilde{\mathrm{DBP}} = \mathrm{DBP} + O(n^{-1}) \qquad \widetilde{\mathrm{DAU}} = \mathrm{DAU} + O(n^{-2})$$

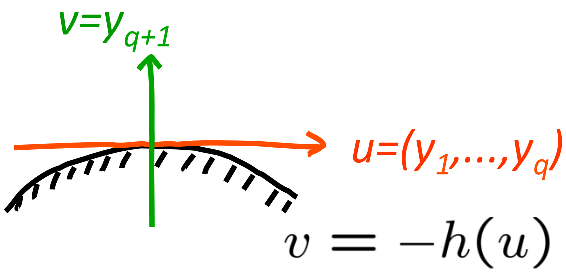
Asymptotic Theories of approaching flat surfaces

Traditional: $n \to \infty$ (sample size)

$$h(0) = \frac{\partial h}{\partial u}\bigg|_0 = 0$$

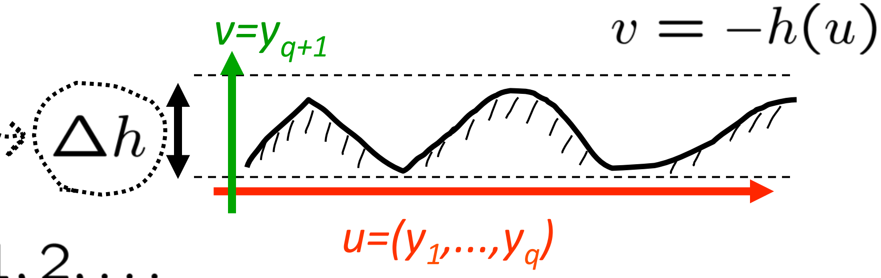
$$\frac{\partial^j h}{\partial u^j}\bigg|_0 = O(n^{-\frac{j-1}{2}}), \quad j = 2, 3, \ldots$$

Higher order derivatives disappear faster



New proposal (Nearly Flat Surface): $\tau \to 0$ (an artificial order parameter)

$$\Delta h = O(\tau)$$



This is interpreted as

$$\frac{\partial^j h}{\partial u^j}\bigg|_0 = O(\tau), \quad j = 0, 1, 2, \ldots$$

All order derivatives disappear at the same rate

Nearly Flat Surfaces (Shimodaira 2008)

Three conditions

1. $\|h\|_\infty = \sup_{u \in \mathbb{R}^m} |h(u)| = O(\tau), \quad \tau \to 0$

   (i.e., approaches a flat surface)

2. $\|h\|_1 = \int_{\mathbb{R}^m} |h(u)|\, du < \infty$

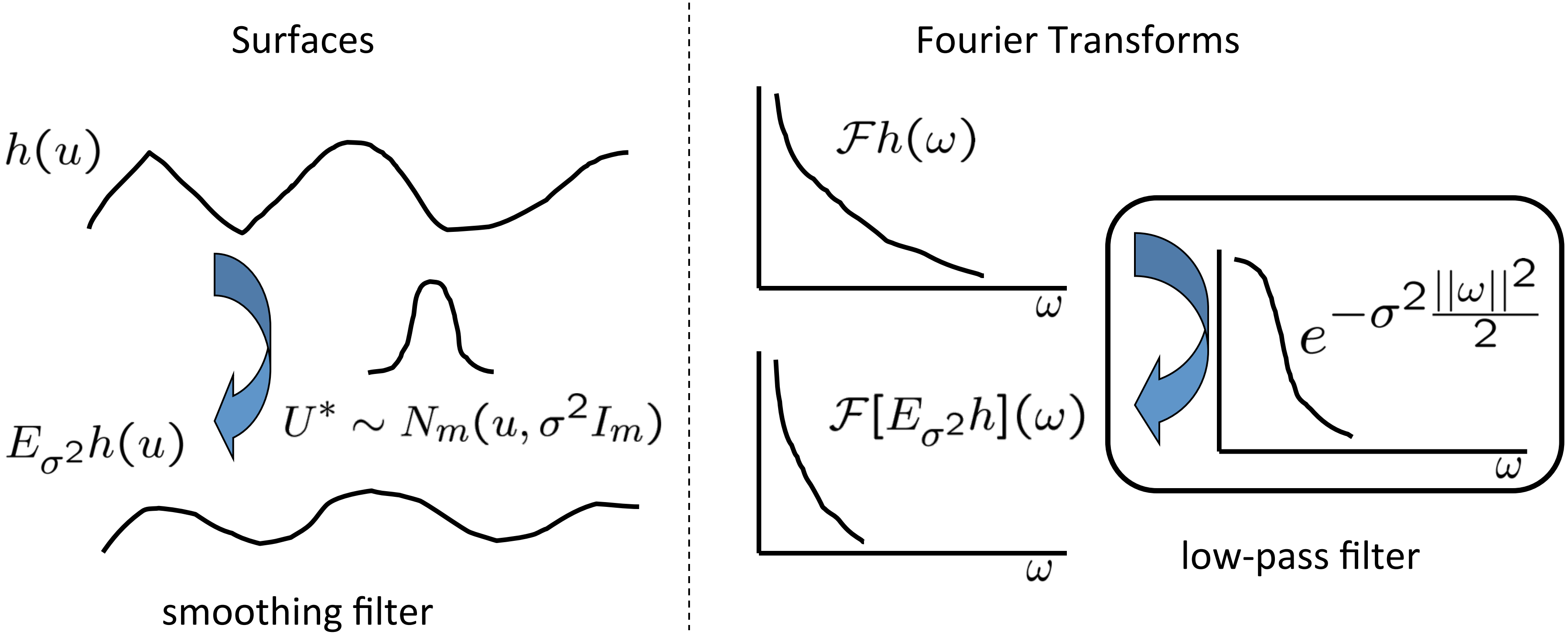
   Fourier transform: $\tilde{h}(\omega) = \mathcal{F}h(\omega) = \int_{\mathbb{R}^m} e^{-i\omega \cdot u} h(u)\, du$

   for $\omega = (\omega_1, \ldots, \omega_m)$

3. $\|\tilde{h}\|_1 < \infty$

Expectation Operator (Gaussian Smoothing)

$$\boxed{E_{\sigma^2} h(u) := E_{\sigma^2}(h(U^*)|u)}$$



smoothing filter

low-pass filter

Bridging Bayesian to Frequentist

$$\sigma\bar{\Phi}^{-1}\Big[\mathrm{BP}_{\sigma^2}(H|y)\Big] = \beta_0 + \beta_1\sigma^2 + \beta_2\sigma^4 + \beta_3\sigma^6 + \cdots$$

$$\bar{\Phi}^{-1}\Big[\mathrm{PV}(H|y)\Big] = \beta_0 - \beta_1 + \beta_2 - \beta_3 + \cdots$$

$\sigma^2 = 1$ gives Bayesian posterior probability

$\sigma^2 = -1$ gives unbiased p-value

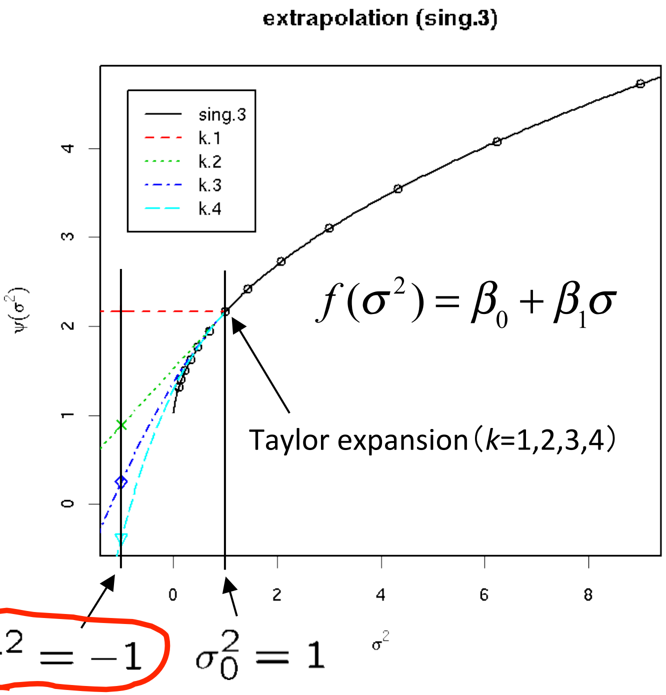
Shown for smooth “nearly flat surfaces” in Shimodaira (2008)

Taylor expansion using *k* terms

Shimodaira (2008)

$$nBP_k(\sigma^2) = \bar{\Phi}\left[\sum_{j=0}^{k-1} \frac{(\sigma^2-\sigma_0^2)^j}{j!} \frac{\partial^j f(\sigma^2)}{\partial(\sigma^2)^j}\bigg|_{\sigma_0^2}\right] \qquad f(\sigma^2) = \sigma\bar{\Phi}^{-1}\left[\mathrm{BP}_{\sigma^2}(H|y)\right]$$

$$\mathrm{AU}_k = nBP_k(-1)$$



$f(\sigma^2) = \beta_0 + \beta_1\sigma$

Taylor expansion（*k*=1,2,3,4）

$\sigma^2 = -1$ $\quad \sigma_0^2 = 1$

Our Method and Generalization

Our corrected *p*-values are represented as:

$$q_k(u,v) = v + \sum_{j=0}^{k-1} \frac{(-1-\sigma_0^2)^j}{j!} \frac{\partial^j}{\partial(\sigma^2)^j}\Big|_{\sigma_0^2} \mathcal{F}^{-1}\left[\tilde{h}(\omega) e^{-\sigma^2 \frac{\|\omega\|^2}{2}}\right](u) + O(\tau^2)$$

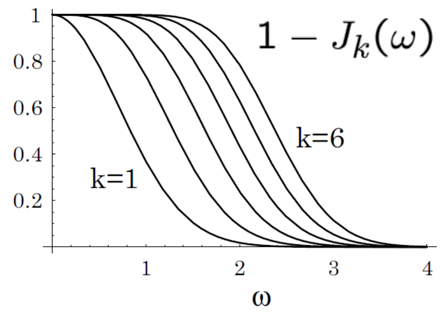
$$\boxed{= v + \mathcal{F}^{-1}\left[\tilde{h}(\omega) e^{\frac{\|\omega\|^2}{2}} (1 - J_k(\omega))\right](u) + O(\tau^2), \quad \ldots(*)}$$

Generalization: (*) defines a new *p*-value from a given $J_k(w)$

For our method, $J_k(w)$ is defined by

$$J_k(\omega) = 1 - e^{-(1+\sigma_0^2)\frac{\|\omega\|^2}{2}} \sum_{j=0}^{k-1} \frac{(1+\sigma_0^2)^j}{j!} \left(\frac{\|\omega\|^2}{2}\right)^j$$

$$= \frac{\gamma(k, (1+\sigma_0^2)\frac{\|\omega\|^2}{2})}{\Gamma(k)} = \sum_{j=k}^{\infty} \frac{(-1)^{j-k}(1+\sigma_0^2)^j \|\omega\|^{2j}}{(k-1)!(j-k)!j2^j}.$$



Bootstrap Iteration

Another example satisfying conditions (i)-(iv).

$$q_1(u,v) = z_1(u,v)$$
$$q_{k+1}(u,v) = \Phi^{-1}\Big\{P_1(q_k(U^*,V^*) \le q_k(u,v) \mid \widehat{\theta}(u,v), -h(\widehat{\theta}(u,v)))\Big\}.$$

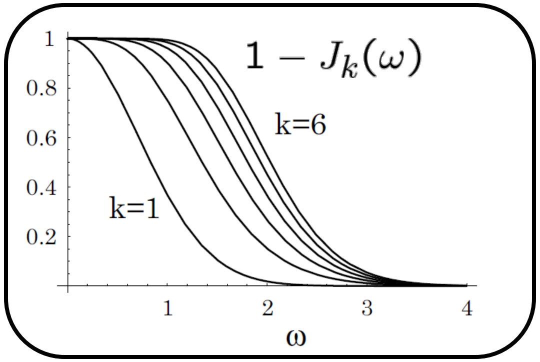
Disadvantages:
1. computation requires *O(B^k)* steps; *B*=10,000.
2. requires resampling from "projection" instead of data.

For bootstrap iteration, *J$_k$*(*w*) is defined by

$$J_k(\omega) = (1 + e^{-\frac{\|\omega\|^2}{2}})(1 - e^{-\frac{\|\omega\|^2}{2}})^k$$
$$= (-1)^k k! \sum_{j=k}^{\infty} (S2(j,k) + S2(j+1,k+1)) \frac{(-1)^j \|\omega\|^{2j}}{2^j j!}$$

$$S2(j,k) = \sum_{i=0}^{k} (-1)^{k-i} i^j / i!(k-i)!$$

(Stirling numbers of the second kind)



Summary and other issues

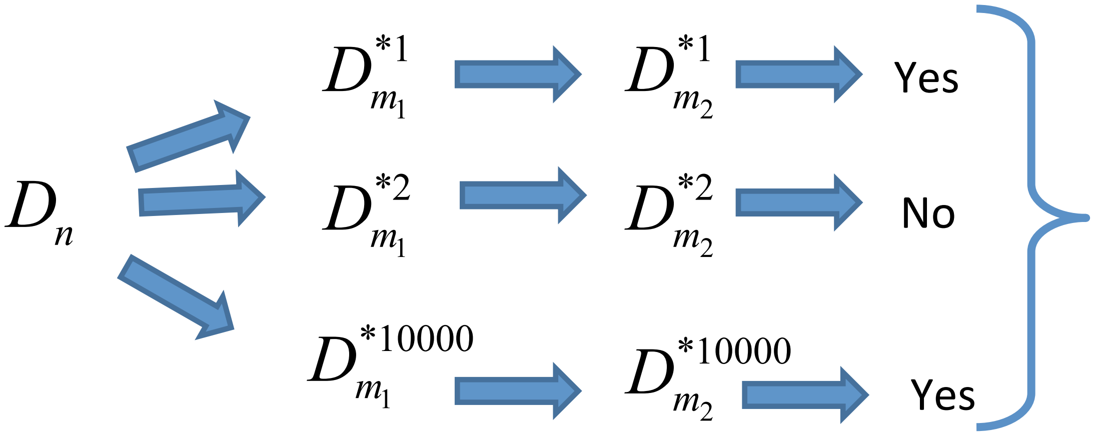
- DAU = "DBP with m=-n" is proposed
- The accuracy of BP is first order (k=1), AU is third-order (k=3), DBP is third-order (k=3)
- DAU is fourth-order accurate (k=4)
- DAU is robust to the projection error (surprisingly, k=4)
- Geometry of surfaces played important roles

- Shimodaira (2008) showed another theory of AU using unusual asymptotic theory of "nearly flat surfaces"

- Shimodaira (2004) discussed deviation from the multivariate normal model, and results for exponential family distributions are given there for multistep-AU

- Future topics may be DAU for nearly flat surfaces, or for exponential family distributions

Estimating the skewness term *A*



D_n

$D_{m_1}^{*1} \rightarrow D_{m_2}^{*1} \rightarrow$ Yes

$D_{m_1}^{*2} \rightarrow D_{m_2}^{*2} \rightarrow$ No

$D_{m_1}^{*10000} \rightarrow D_{m_2}^{*10000} \rightarrow$ Yes

$$BP(\sigma_1^2, \sigma_2^2) = \frac{\#\{Yes\}}{10000}$$

$$= BP(\sigma^2) + O_p(n^{-1/2})$$

with $\sigma^2 = \sigma_1^2 + \sigma_2^2$

$$-\sigma\Phi^{-1}(BP(\sigma_1^2, \sigma_2^2)) = -\sigma\Phi^{-1}(BP(\sigma^2)) + n^{-1/2} A\sigma_1^2\sigma_2^2\sigma^{-4}(\beta_0 - \sigma^2) + O_p(n^{-1})$$

Two-step multiscale bootstrap of Shimodaira (2004)