Fiducial prediction

Peter McCullagh

Department of Statistics
University of Chicago

Royal Holloway College
London, June 22, 2010

Outline

Statistical models

Components of a statistical model?

- A parameter space Θ (non-empty)
- Space of r.vs. $\mathcal{S} = \mathcal{R}^\infty$ (with σ-field)
- $P\colon \theta \mapsto P_\theta$ (prob distn on $\mathcal{S}$)

Population, sample and observation (structure of $\mathcal{S}$)

- Population (index set) $\mathbb{N} = \{1, 2, \ldots\}$
- Random variable $Y = (Y_1, Y_2, \ldots) \quad Y\colon \mathbb{N} \to \mathcal{R}$
- Sample units $S = \{u_1, \ldots, u_n\} \subset \mathbb{N}$ finite
- Sample values $Y[S] = \{(u_i, Y(u_i)) : 1 \le i \le n\}$
  or $\{(u_i, x(u_i), Y(u_i)) : 1 \le i \le n\}$ if covariates present

Generating mechanism:
For some $\theta \in \Theta, \quad Y \sim P_\theta, \quad Y[S] \sim P_{n,\theta}$

Inferential goals, parametric and non-parametric

Observation $\mathbf{y} = Y[S]$ on sample $S \subset \mathbb{N}$

Parametric inference:

Statement about $\theta \in \Theta$

principles: likelihood, sufficiency, conditionality,...

Non-parametric (sample-space) inference:

Statements about Y such as

prequential: $\mathrm{pr}(Y(u') > 7.1 \mid Y[S] = \mathbf{y}) = ??$

test sample predictions (classification)

tail events $\mathrm{pr}(\bar{Y}_\infty < 6.8 \mid \mathbf{y}) = ??$

p-value calculations?

Techniques for inference

Inference concerned with events E, in principle observable

Root problem: $P_\theta(E) \neq P_{\theta'}(E)$

Solution: replace $\{P_\theta\}$ with a single compromise distn

Standard Bayes solution: (three steps)

(i) Extend space to $\mathcal{S} \times \Theta$

(ii) Use joint distn $P_\pi(d\theta\, dy) = P_\theta(dy)\pi(d\theta)$

... or use mixture $P_\pi(dy) = \int P_\theta(dy)\,\pi(d\theta)$ on $\mathcal{S}$

(iii) Compute conditional distribution $P_\pi(E \mid \mathbf{y})$

Q: Are there any principled alternatives to Bayes?

(i) Fiducial?

(ii) Improper priors?

The fiducial compromise: my interpretation

Ancillary events: $P_\theta(A) = P_{\theta'}(A)$ for $\theta, \theta' \in \Theta$
Full agreement on the value of $P(A)$ for $A \in \mathcal{A}$

Problems: ....

Suppose in addition that $\mathcal{A}$ is a σ-field

Fiducial/pivotal compromise:
replace set $\{(\mathcal{F}, P_\theta)\}$ with distn $(\mathcal{A}, P)$ by restriction to $A \subset \mathcal{F}$
Bayes compromise:
replace set $\{(\mathcal{F}, P_\theta)\}$ with the average $(\mathcal{F}, P_\pi)$

What useful inferences can we draw using the fiducial model?
We can compute $P(E \mid Y = y)$ for $E \in \mathcal{A} \quad (E \subset \mathcal{S})$
Is this potentially useful or is it a dead end?
Where are the parameters?

The standard location-scale example

$\mathcal{S} = \mathcal{R}^\infty = \{(y_1, y_2, \ldots) : y_i \in \mathcal{R}\}$
Location-scale group elements $g = [a, b]$
Multplication $[a, b][a', b'] = [a + ba', bb']$
$\mathcal{G}$ acts on $\mathcal{R}^\infty$ component-wise $gy_i = a + by_i$

$\mathcal{G}$-invariant events: $gE = E$ for each $g \in \mathcal{G}$
Examples: $\{y \in \mathcal{R}^\infty : (y_n - y_1)/|y_2 - y_1| < 1\}$
$\{y \in \mathcal{R}^\infty : (y_{n+1} - \bar{y}_n)/s_n < 1.4\}$
$\{y \in \mathcal{R}^\infty : (\bar{y}_{n+1:n+m} - \bar{y}_n)/s_n < 1.4\}$
$\{y \in \mathcal{R}^\infty : (\bar{y}_\infty - \bar{y}_n)/s_n < 1.4\}$

$\mathcal{G}$-invariant events are ancillary wrt *any* location-scale model
Set $\mathcal{A} = \mathcal{B}/\mathcal{G}$ of $\mathcal{G}$-invariant Borel events is a σ-field
$\mathcal{G}$-invariant events in $\mathcal{R}^n$: $\mathcal{A}_n = \mathcal{A} \cap \mathcal{B}_n$

But there may be ancillary events that are not in $\mathcal{A}$

The location-scale model contd.

Fiducial compromise:

Replace $\{(\mathcal{R}^\infty, \mathcal{B}, P_\theta)\}$ with restriction $(\mathcal{R}^\infty, \mathcal{A}, P)$

... rather than average $(\mathcal{R}^\infty, \mathcal{B}, P_\pi)$

Given an obs $Y^{(n)}: (\mathcal{R}^\infty, \mathcal{A}) \to (\mathcal{R}^n, \mathcal{A}_n)$

Fiducial inference $P(A \,|\, \mathcal{A}_n)$ for events A

Fundamental problem I:

No parameters, nothing to estimate!

Fundamental problem II:

Discussion limited to invariant (ancillary) events

$P(Y_{n+1} > 4 \,|\, Y^{(n)} = \mathbf{y})$ meaningless

But we can talk about events in a relative sense

$P(Y_{n+1} > \bar{Y}_n + 4s_n \,|\, \mathbf{y})$

Digression: Quotient-space distributions

New symbol: $Y \sim N_n(\mathcal{K}, \mu, \Sigma)$; $\mathcal{K} \subset \mathcal{R}^n$

Prob distn on $(\mathcal{R}^n, \mathcal{B}_n/\mathcal{K})$

if $\ker(L) = \mathcal{X}$, then $LY \sim N(L\mu, L\Sigma L') \iff Y \sim N_n(\mathcal{K}, \mu, \Sigma)$

Many reasons for working directly with Y rather than LY

directness of prediction (Kriging)

relative to what is observed (equivariant, not absolute)

Log likelihood based on $Y \sim N_n(\mathcal{K}, \mu, \Sigma)$

$$l(\mu, \Sigma) = \tfrac{1}{2} \log \mathrm{Det}(WQ) - \tfrac{1}{2}(y - \mu)' WQ(y - \mu)$$

$W = \Sigma^{-1}$, $\quad Q = I - K(K'WK)^{-1}K'W$

same as REML if $\mu \in \mathcal{K}$

Gaussian processes

Brownian motion: $Y \sim GP(0, K)$ where
$K(x, x') = -|x - x'|/2)$
What does this mean?
$\Rightarrow Y(x) - Y(x') \sim N(0, K(x) + K(x') - 2K(x, x')) = N(0, |x - x'|)$
$Y(x_1), \ldots, Y(x_n) \sim N_n(\mathbf{1}, 0, K)$: $\mathcal{F} = \mathcal{B}_n/\mathbf{1}$

Gaussian spline model: $Y \sim GP(0, \delta_{x,x'} - |x - x'|^3)$
Kernel $\mathcal{K} = span\{1, x\}$ dimension 2
meaning depends on $\mathcal{K}$ and the implied σ-field

Conditional expectation is such that

$$E(Y(x^*) - \hat{\mu}(x^*) \mid Y(x_1), \ldots, Y(x_n)) = \text{cubic spline in } x^*$$

Kriging = fiducial prediction = equivariant prediction

Gaussian families and Gosset families

Parameter space $\Theta = \mathcal{R} \times \mathcal{R}^{+} = \{(\mu, \sigma) : \sigma > 0\}$

Gaussian family: $N_\infty(\mu, \sigma^2)$ iid sequences

Gosset family: $G_\infty(\mu, \sigma)$

$$Y_1 = \mu, \quad Y_2 = \mu \pm \sigma \qquad Y_{n+1} = \bar{Y}_n + s_n \epsilon_n \sqrt{1 + 1/n}$$

(i) Both families defined on Borel sets
(ii) Mathematically, as dissimilar as two families can be
(iii) Inguishable from observation on one sequence
(iv) Identical on $\mathcal{G}$-invariant events
(v) Likelihood ratio always selects Gosset
(vi) Identical as point processes in $\mathcal{R}^\infty$ $(d\mu\, d\sigma/\sigma)$
(vii) Gosset process gives predictions in a relative sense

Gauss and Gosset contd

Why are these families essentially indistinguishable?
Transformation properties:

$X = (X_1, X_2, \ldots) \sim N_\infty(\mu, \sigma^2)$
Define $Y_i = (X_i - X_1)/|X_2 - X_1|$
Then $Y \sim G_\infty(0, 1)$

$Y = (Y_1, Y_2, \ldots) \sim G_\infty(\mu', \sigma')$
Define $X_i = (Y_i - \bar{Y}_\infty)/s_\infty$
Then $X \sim N_\infty(0, 1)$

For every invariant event $N_\infty(\mu, \sigma^2)(A) = G_\infty(\mu, \sigma)(A)$

Gosset process is a Borel extension of the fiducial process giving the same predictions.

Inversion: the fatal fiducial flaw?

The Fisherian argument:

location model: P_μ such that $Y_1, \ldots \sim N(\mu, 1)$

$$P_\mu(\bar{Y}_n - \mu \in B) = N(0, 1/n)(B) \quad B \subset \mathcal{R}$$

$$\text{pr}(\bar{Y}_n - \mu < z/\sqrt{n} \,|\, \mu) = \Phi(z) \text{ same thing!}$$

$$\text{pr}(\mu > \bar{Y}_n - z/\sqrt{n} \,|\, \mathbf{y}^{(n)}) = \Phi(z) \text{ false?}$$

But ...

The fiducial compromise

Let F be the restriction of P_μ to $\mathcal{A} = \mathcal{B}/\mathbf{1}$

F is the inferential compromise distn/process
$B \subset \mathcal{R}$ Borel set
$A = \{y : \bar{y}_n - \bar{y}_\infty \in B\}$ is invariant event in $\mathcal{A}$

$F(\{y : \bar{y}_n - \bar{y}_\infty \in B\}) = N(0, 1/n)(B)$ correct!

$F(\{y : \bar{y}_n - \bar{y}_\infty < z/\sqrt{n}\}) = \Phi(z)$ correct!

$F(\bar{Y}_\infty > \bar{Y}_n - z/\sqrt{n} \,|\, \mathcal{A}_n) = \Phi(z)$ also correct!
$F(\bar{Y}_\infty > \bar{Y}_n - z/\sqrt{n} \,|\, y^{(n)}) = \Phi(z)$ same thing!

Where does that leave Student, Fisher, Barnard, Fraser,...?

Extensions

Behrens-Fisher type models

Non-Gaussian models
Fisher (1934)

Linear regression:

Sequential prediction

Computational issues
McC, V.V., I.N., D.D. and A.G. (2009)