# Operational Subjective Statistical Methods: a primer on subjective probability and statistics in the mode of Bruno de Finetti

Frank Lad

University of Canterbury

May 4, 2012

# Topics for this primer:

- Quantities
- Coherent prevision
- Coherent conditional prevision
- The Fundamental Theorem of Prevision
- Exchangeable assertions
- Sequential forecasting
- Proper Scoring Rules

# Basic Concepts and Definitions:

Observable quantities ... $X$

Realm of $X$ ... $\Re(X) = \{x_1, x_2, ..., x_K\}$

Your prevision for X ... by asserting $P(X)$
you avow your indifference to $\$X$ or $\$P(X)$
scaling down the monetary units

Events ... When $\Re(E) = \{0, 1\}$ ...
$P(E)$ is also called your probability for $E$

Partition vector ...

$$\mathbf{Q}_K \equiv [(X = x_1), (X = x_2), ..., (X = x_K)]^T$$

# Coherency

Your $P(X)$ is coherent ... as long as

No $s$ pos or neg for which

$s(x - P(X)) < 0$ for every $x \in \Re(X)$.

Otherwise you are a "sure loser" !

# Definitions for a Vector of Quantities:

A vector of quantities ... $\mathbf{X}_N \in \Re(\mathbf{X}_N) \quad = \quad \{\mathbf{x}_{N1}, \mathbf{x}_{N2}, ..., \mathbf{x}_{NK}\}$

Constituents of the partition it generates ...

specify the logical relations among its components:

$$\mathbf{Q}_K = [(\mathbf{X}_N = \mathbf{x}_{N1}),\ (\mathbf{X}_N = \mathbf{x}_{N2}),\ ...,\ (\mathbf{X}_N = \mathbf{x}_{NK})]^T$$

Note: $\quad \mathbf{X}_N \quad = \quad \mathbf{R}(\mathbf{X}_N)\ \mathbf{Q}_K$

where $\mathbf{R}(\mathbf{X}_N)$ is an $(N \times K)$ "realm matrix"

# Coherence of $P(\mathbf{X}_N$

Your prevision for $\mathbf{X}_N$ ... a vector of such prices

The assertion $P(\mathbf{X}_N)$ is coherent ...

there is no vector $\mathbf{s}_N$ for which

$$\mathbf{s}_N^T[\mathbf{x}_N - P(\mathbf{X}_N)] < 0$$

for every $\mathbf{x}_N \in \Re(\mathbf{X}_N)$.

Coherent $P(\mathbf{X}_N)$ must lie within the convex hull of $\Re(\mathbf{X}_N)$

since there is no hyperplane that can
separate $P(\mathbf{X}_N)$ from $\Re(\mathbf{X}_N)$

Coherent $P(\cdot)$ a linear operator:
your $P(\mathbf{s}_N^T\mathbf{X}_N) \ = \ \mathbf{s}_N^T P(\mathbf{X}_N)$ for any coefficient vector $\mathbf{s}_N$ .

# Fundamental Theorem of Prevision

Whatever you assert about any N quantities whatsoever,

there are computable NASC bounds on your prevision for any further quantity $X_{N+1}$

if it is to cohere with your asserted $P(\mathbf{X}_N)$.

Think why.

This theorem includes all Prob inequalities as special cases, including Chebyshev and Kolmogorov

# Coherent Conditional Prevision and Conditional Quantities

Your asserting $P(X|E)$ attests your indifference between

$$\$P(X|E) \quad \text{and} \quad \$(XE + (1-E)P(X|E))$$

Having asserted $P(X|E)$, define $(X|E) \equiv XE + (1-E)P(X|E)$

Coherency requires $P(XE) = P(X|E)P(E)$.

Conditional probability has nothing at all to do with "before and after observing E".

# Exchangeability and Inference

You regard $E_1, E_2, ..., E_N$ exchangeably if ...

equivalently ... if you regard the sum and N as sufficient statistics!

and exchangeably extendible to $N + K$ if ...

General extensions via numbers of sufficient stats.

# More

If you regard the sequence as infinitely exchangeably extendible

then your $P(S_N = a)$ can be represented as

$\int_0^1 \theta^a \, (1-\theta)^{(N-a)} \, dF(\theta)$ for some distribution function $F(.)$

# and more

... and $P(E_{N+1}) \; = \; \int_0^1 \theta \; dF(\theta | S_N = a)$
This is a representation of ...

# Proper Scoring Rules

$S(\mathbf{p}_N, X = x^o)$ is a scoring rule for $\mathbf{p}_N$ on observing that $X = x^o$

The rule is *proper* if

$$P\{S(\mathbf{p}_N, X = x^o)\} \quad \geq \quad P\{S(\mathbf{q}_N, X = x^o)\}$$

for any other $\mathbf{q}_N$ in $\mathbf{S}^{N-1}$, when $P\{\cdot\}$ is assessed with respect to $\mathbf{p}_N$.

# Uniqueness of the Logarithmic Scoring rule:

$$S_{log}(\mathbf{p}_N, X = x^o) \equiv \sum_{i=1}^{N} E_i \; log(p_i) = log(p^o)$$

relative to other proper rules, such as :

$$S_{Quadratic}(\mathbf{p}_N, X = x^o) \equiv -[\, \sum_{i=1}^{N} E_i(1-p_i)^2 + \sum_{i=1}^{N} \tilde{E}_i(0-p_i)^2 \,]$$

$$= -[\, \sum_{i=1}^{N} p_i^2 \; - 2p^o \; + 1 \,]$$

or

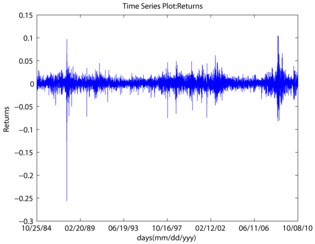
$$S_{Spherical}(\mathbf{p}_N, X = x^o) \equiv \sum_{i=1}^{N} E_i \; p_i \; / \; [\, \sum_{i=1}^{N} p_i^2 \,]^{1/2}$$

... "the only proper score that is a function only of the actual observation, not values of $X$ that were not observed."

... "like the likelihood principle" ... Bernardo (1979) ... WOOPS!

Daily Dow-Jones returns

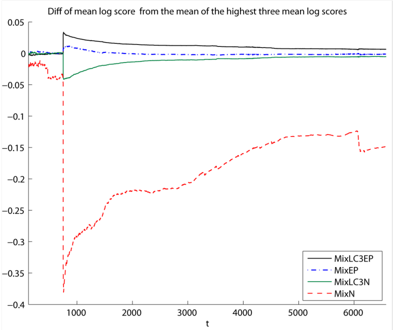
25/10/84 through 20/08/10



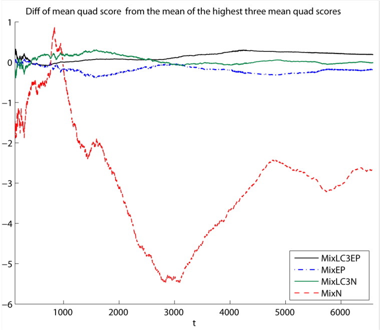

Notice the day 19/10/87 ...

# Comparative Mean Accumulating Scores

## Comparative Mean Log Scores




## Comparative Mean Quadratic Scores




Notice that larger parametric mixture structures
do NOT ensure better scores !
MixN and MixLC3N have long periods of good Quad scoring