Quantifying Uncertainty in Inverse Problems

Philip B. Stark
Department of Statistics
University of California, Berkeley

`stark@stat.berkeley.edu`

Mathematical Geophysics & Uncertainty in Earth Models
14-25 June 2004

Some work joint w/
Steve Evans (Berkeley)
David Freedman (Berkeley)
Ben Hansen (Michigan)
Chris Genovese (Carnegie-Mellon)
Chad Schafer (Berkeley/Carnegie-Mellon)

Ref. for basic Prob&Stat

Online intro text with self-test exercises, dynamic examples, applets to illustrate key concepts, *etc.*:

`www.stat.berkeley.edu/∼stark/SticiGui`

More detail on some of the topics Dr. Malinverno discussed.

Part I:
Comparing two Groups

Prof. Young compared heights of statisticians and geophysicists, based on a (presumed random) sample.

Tested the hypothesis that the average height of statisticians and the average height of geophysicists are equal, using a confidence interval based on Student's t distribution.

The assumptions of the test are hard to justify: null hypothesis says populations of heights have normal distributions with same mean and same variance.

What would an anal-retentive statistican do?

Part I:
Comparing two Groups, contd

Data

Yesterday's data:

sort	no.	mean height (in)	sd(height)
Geophysicist	20	69.716	2.447
Math/Stat	9	68.562	3.416
Pooled	29	69.358	2.774

Idea: if there is no difference in heights of the two groups, it is as if labeling someone as a mathematician/statistician or geophysicist is determined by tossing a coin—as if we decided who is a geophysicist by drawing 20 people at random from the 29 in the pooled group. Equivalently, as if the 29 heights were assigned at random to the 29 people.

Part I:
Comparing two Groups, contd

Permutation Test & t-test

Null hypothesis for permutation test: height independent of occupation, as if height or occupation assigned at random among people in sample.

Null hypothesis for t-test: groups are independent random samples from normal distributions with same mean μ and same variance σ^2.

When do the tests reject?

MATLAB code

For permutation test, in principle, calculate diff. of sample means for all $\binom{29}{9} = 10,015,005$ assignments. Can approximate by simulation.

```
geo = [ 72 71 70 68 1.75 1.8 1.61 1.73 73 1.8 ...
       1.79 1.69 1.83 66 1.8 1.8 1.74 1.75 73 71 ]';
math = [ 1.86 71 69 68 1.85 1.7 1.65 1.73 1.6 ]';

geo(geo <= 4) = geo(geo <= 4) * 100/2.54;
math(math <= 4) = math(math <= 4) * 100/2.54;
allSorts = [geo; math];

nGeo = length(geo);
nMath = length(math);
N = nGeo + nMath;

gBar = mean(geo)
mBar = mean(math)
```

MATLAB code, contd

```
% Student's t-test
sHat = std(allSorts);
seHat = sHat* sqrt(1/nMath + 1/nGeo);
t = (gBar - mBar)/seHat;
tPval = 2*(1-tcdf(abs(t),nMath+nGeo-1))

% Permutation test
iter = 1000;
sims = zeros(iter,1);
for j=1:iter,
    dat = allSorts(randperm(N));
    sims(j) = abs(mean(dat(1:nGeo)) - ...
                  mean(dat(nGeo+1:N)));
end;
pPval = sum(sims >= abs(gBar - mBar))/iter
```

Results

P-value for t-test: 0.3088

P-value for permutation test: 0.3020.

In this case, not much difference.

General case: differ, but distribution of permutation test when null is true is asymptotically Student's t as both groups grow (in number, not height!).

Aside: sampling distributions

Applet time!

Part II:
What's the Chance of an Earthquake?
(Joint work with D.A. Freedman)

USGS Forecast (1999):

What's the chance that an earthquake of magnitude 6.7 or greater will occur by the year 2030 in the San Francisco Bay Area?

USGS says 0.7 ± 0.1.

What does that *mean*?

Very hard to make sense of the number.

Ingredients of USGS Forecast

- geological maps
- rules of thumb
- expert opinion
- physical models
- stochastic models
- numerical simulations
- geodetic, seismic, and paleoseismic data.

Probability hard to define in this context.

Interpreting Probability

Two aspects:

Formal mathematical theory, axiomatized by *Kolmogorov* (1956).

Informal theory connects math to the world—defines what ‘probability’ means for real events.

Simple example: tossing a coin.
What does $\mathbb{P}(\text{heads}) = 1/2$ mean?

3 common interpretations: symmetry, relative frequency, and strength of belief. (Others, too.)

Symmetry Theory

Earliest interpretation of probability: ‘equally likely outcomes,’ from the study of gambling.

If the n possible outcomes of a chance experiment are judged equally likely—for instance, on the basis of symmetry—each must have probability $1/n$.

For example, for a coin toss, $n = 2$, so chance of heads is $1/2$

Difficulties
Tossing 2 coins: total number of heads can be 0, 1 or 2—not equally likely (for purposes of gambling).

Earthquake forecasting (like many problems) has no obvious symmetry.

Frequency Theory

$\mathbb{P}(A)$ is limit of the relative frequency with which A occurs, in repeated trials under the same conditions.

If we toss a coin repeatedly under the same conditions, the fraction of tosses that result in heads will converge to $1/2$: that's why the chance of heads is $1/2$.

Difficulties

To interpret USGS forecast, need to imagine repeating the years 2000–2030 over and over again. ????

Bayesian approach

Probability is degree of belief, on scale from 0 to 1.

$\mathbb{P}(\text{impossible event}) = 0$; $\mathbb{P}(\text{certain event}) = 1$.

Different observers need not have the same beliefs; differences don't imply that anyone is wrong.

Difficulties

Changes the topic: probability is a summary of an opinion, not something inherent in the system being studied.

'There is chance 0.7 of at least one earthquake with magnitude 6.7 or greater in the Bay Area between 2000 and 2030' is just the USGS reporting its corporate state of mind, and may not mean anything about tectonics and seismicity.

Not clear why one observer should care about the opinion of another.

Principle of Insufficient Reason

Bayesians—and some frequentists—often make probability assignments using Laplace's principle of insufficient reason:

> If there is no reason to believe that outcomes are not equally likely, take them to be equally likely.

Difficulties

Not believed to be unequal is one thing; known to be equal is another. Moreover, all outcomes cannot be equally likely, so Laplace's prescription is ambiguous.

Principle of Insufficient Reason, contd

Example: Thermodynamics

Gas of n particles, each of which can be in any of r quantum states.

State of the gas defined by a ‘state vector’.

Three conventional models for such a gas; different definitions of the state vector.

In each model, all possible values of the state vector equally likely.

Principle of Insufficient Reason, contd

1. Maxwell-Boltzman. State vector specifies state of each particle; there are

$$r^n$$

possible values of the state vector.

2. Bose-Einstein. State vector specifies #particles in each state. There are

$$\binom{n+r-1}{n}$$

possible values of the state vector.

3. Fermi-Dirac. State vector specifies #particles in each state—but at most one perticle in each state. There are

$$\binom{r}{n}$$

possible values of the state vector.

Principle of Insufficient Reason, contd

Maxwell-Boltzman stats model coins, but not gas.

Bose-Einstein stats model *bosons*—particles with spin angular momentum integer multiple of $\hbar$. (E.g., photons and He^4 atoms.)

Fermi-Dirac stats model *fermions*, particles with spin angular momentum a half-integer multiple of $\hbar$. (E.g., electrons and He^3 atoms.)

Bose-Einstein condensates, low temperature gases with all atoms in 1 quantum state, observed at CU Boulder by *Anderson et al.* (1995). Occur for bosons, not fermions.

Outcomes of an experiment can be defined in quite different ways. Seldom clear *a priori* which—if any—are equally likely.
Principle of Insufficient Reason is insufficient.

Earthquake & Weather Forecasts

Earthquake forecasts look similar to weather forecasts.

How do meteorologists interpret 'the chance of rain tomorrow is 0.7'?

Standard interpretation applies frequentist ideas to forecasts: the chance of rain tomorrow is 0.7 means 70% of such forecasts are followed by rain the next day.

Meteorology *not* like earthquake prediction:

Large regional earthquakes recur on scale of centuries.

Weather forecasters have a *much* shorter time horizon.

Weather prediction not good analogue for earthquake prediction.

Kolmogorov's Axioms of Probability

For most statisticians, Kolmogorov's axioms are the basis for probability theory—no matter the interpretation.

Σ a σ-algebra of subsets of a set S.

$\mathbb{P}$ a real-valued function on Σ. $\mathbb{P}$ is a probability if

- $\mathbb{P}(A) \geq 0$ for every $A \in \Sigma$ (nonnegative);

- $\mathbb{P}(S) = 1$ (scale);

- if $A_j \in \Sigma$ for $j = 1, 2, \ldots,$ and $A_j \cap A_k = \emptyset$ whenever $j \neq k$, then
$$\mathbb{P}\left(\bigcup_{j=1}^{\infty} A_j\right) = \sum_{j=1}^{\infty} \mathbb{P}(A_j) \text{ (additivity).} \qquad (1)$$

Model-based interpretation

Another interpretation of probability: probability is just a property of a mathematical model intended to describe features of the world.

To be useful, the model must be shown to correspond well with the system it describes.

That's where the statistics meets the science.

Probability models, contd

Illustration: A coin will be tossed n times.

2^n possible sequences of heads and tails.

Model: all 2^n sequences are equally likely. Each has probability $1/2^n$ (probability $1/2$ of heads on each toss and independence of tosses).

Observational consequences can be used to test model validity.

For example, model says the distribution of total number X of heads in n tosses is binomial:

$$P(X = k) = \binom{n}{k}\frac{1}{2^n}.$$

If model is right, when n is large should see around $n/2$ heads, with an error on the order of $\sqrt{n}$.

Probability models, contd

Model also gives probability distributions for number of runs, lengths of runs, *etc.*, which can be checked against data.

The model is good, but imperfect: with many thousands of tosses, the difference between a real coin and the model coin is likely to be detectable.

The probability of heads will not be exactly 1/2, and there may be some correlation between successive tosses.

Probability models, contd

Probability Model interpretation—that probability is a property of a mathematical model and has meaning for the world only by analogy—seems best for earthquake forecasts.

To apply the interpretation, posit stochastic model for earthquakes; interpret a number calculated from the model to be the probability of an earthquake in some time interval.

Problem: models in forecasts not tested against relevant data.

Models can't be tested on human time scale.

Little reason to believe the probability estimates.

The USGS Forecast

USGS Forecast (1999):

2 stages. Stage 1 major steps:

1. Determine regional constraints on aggregate fault motions from geodetic measurements.

2. Map faults and fault segments; identify fault segments with slip rates $\geq$ 1 mm/y. Estimate slip on each fault segment principally from paleoseismic data, occasionally augmented by geodetic and other data. Determine (by expert opinion) for each segment a 'slip factor', the extent to which long-term slip is accommodated aseismically. Represent uncertainty in fault segment lengths, widths, and slip factors as independent Gaussian random variables

with mean 0 (and sometimes variance 0). Draw a set of fault segment dimensions and slip factors at random from that probability distribution.

3. Identify (by expert opinion) ways segments of each fault can rupture separately and together. Each such combination of segments is a ‘seismic source’.

4. Determine (by expert opinion) how much long-term fault slip is accommodated by rupture of each combination of segments for each fault.

5. Choose at random (with probabilities of 0.2, 0.2, and 0.6 respectively) one of three generic relationships between fault area and moment release to characterize magnitudes of events that each

combination of fault segments supports. Represent uncertainty in the generic relationship as Gaussian with zero mean and standard deviation 0.12, independent of fault area.

6. Using the chosen relationship and the assumed probability distribution for its parameters, determine mean event magnitude for each source by Monte Carlo simulation.

7. Combine seismic sources along each fault 'in such a way as to honor their relative likelihood as specified by the expert groups'; adjust the relative frequencies of events on each source so that every fault segment matches its geologic slip rate—as estimated from paleoseismic and geodetic data. Discard the combination of sources if it violates a regional slip constraint.

8. Repeat the previous steps until 2,000 regional models meet the slip constraint. Treat the 2,000 models as equally likely to estimate magnitudes, rates, and uncertainties.

9. Steps 1-8 model events on seven identified fault systems, but there are background events not associated with those faults. Estimate the background rate of seismicity as follows. Use an (unspecified) Bayesian procedure to categorize historical events from three catalogs either as associated or not associated with the seven fault systems. Fit a generic Gutenberg-Richter magnitude-frequency relation $N(M) = 10^{a-bM}$ to the events deemed not to be associated with the seven fault systems. Model this background seismicity as a marked Poisson process. Extrapolate the Poisson model to $M \geq 6.7$, which gives a probability of 0.09 of at least one event.

2nd Stage of USGS Forecast

First stage gives 2,000 models & estimates long-term seismicity rates as a function of magnitude for each source.

Second stage fits 3 types of stochastic models for earthquake recurrence—Poisson, Brownian passage time and ‘time-predictable’—to long-term seismicity rates estimated in 1st stage. Stochastic models combined to estimate probability of large earthquake. Fitting models requires more assumptions; data are noisy and scarce (e.g., estimating fault loading over time).

Outputs of the 3 stochastic models for each segment weighted according to opinions of panel of 15 experts.

The USGS Forecast

$\exists$ no straightforward interpretation of USGS probability forecast.

Many steps involve models that are largely untestable; modeling choices often arbitrary.

Frequencies are equated with probabilities, fiducial distributions are used, outcomes are assumed to be equally likely, and subjective probabilities are used in ways that violate Bayes' rule.

USGS Uncertainty Estimate

USGS forecast is 0.7 ± 0.1, where 0.1 is an uncertainty estimate.

The 2,000 regional models produced in stage 1 give estimate of long-term seismicity rate for each source (linked fault segments), and estimate of uncertainty in each rate.

Those uncertainties propagated through stage 2 to estimate uncertainty of estimated probability of a large earthquake.

0.1 is a gross underestimate . . .

USGS Uncertainty Estimate, contd

Overlooked errors include:

1. Errors in fault maps and identification of fault segments.

2. Errors in geodetic measurements, paleoseismic data, and viscoelastic models used to estimate fault loading and sub-surface slip from surface data.

3. Errors in estimated fraction of stress relieved aseismically through creep in each fault segment and errors in relative amount of slip assumed to be accommodated by each source.

4. Errors in estimated magnitudes, moments, and locations of historical earthquakes.

5. Errors in relationships between fault area and seismic moment.

6. Errors in models for fault loading.

7. Errors in models for fault interactions.

8. Errors in generic Gutenberg-Richter relationships, not only in parameter values but also in functional form.

9. Errors in estimated probability of an earthquake not associated with any of fault in model.

10. Errors in form of probability models for earthquake recurrence and in estimated parameters of those models.

A Mathematician's View from the Past

Littlewood (1953) wrote:

Mathematics (by which I shall mean pure mathematics) has no grip on the real world; if probability is to deal with the real world it must contain elements outside mathematics; the *meaning* of 'probability' must relate to the real world, and there must be one or more 'primitive' propositions about the real world, from which we can then proceed deductively (i.e. mathematically). We will suppose (as we may by lumping several primitive propositions together) that there is just one primitive proposition, the 'probability axiom,' and we will call it A for short. Although it has got to be *true*, A is by the nature of the case incapable of deductive proof, for the sufficient reason that it is about the real world . . . .

Littlewood, contd

There are 2 schools. One, which I will call mathematical, stays inside mathematics, with results that I shall consider later. We will begin with the other school, which I will call philosophical. This attacks directly the 'real' probability problem; what are the axiom A and the meaning of 'probability' to be, and how can we justify A? It will be instructive to consider the attempt called the 'frequency theory'. It is natural to believe that if (with the natural reservations) an act like throwing a die is repeated n times the proportion of 6's will, *with certainty*, tend to a limit, p say, as $n \to \infty$. (Attempts are made to sublimate the limit into some Pickwickian sense—'limit' in inverted commas. But either you *mean* the ordinary limit, or else you have the problem of explaining how 'limit' behaves, and you are no further. You do not make an illegitimate conception legitimate by putting it into inverted commas.)

Littlewood, contd

If we take this proposition as 'A' we can at least settle off-hand the other problem, of the *meaning* of probability; we define its measure for the event in question to be the number p. But for the rest this A takes us nowhere. Suppose we throw 1000 times and wish to know what to expect. Is 1000 large enough for the convergence to have got under way, and how far? A does not say. We have, then, to add to it something about the rate of convergence. Now an A cannot assert a *certainty* about a particular number n of throws, such as 'the proportion of 6's will *certainly* be within $p \pm \epsilon$ for large enough n (the largeness depending on ϵ)'. It can only say 'the proportion will lie between $p \pm \epsilon$ *with at least such and such probability (depending on* ϵ *and* n_0*) whenever* $n > n_0$'. The vicious circle is apparent. We have not merely failed to *justify* a workable A; we have failed even to *state* one which would work if its truth were granted. It is generally agreed that the frequency theory won't work. But whatever the theory it is clear that the vicious circle is very deep-seated: certainty being impossible, whatever A is made to state can be stated only in terms of 'probability'.

Summary

Making sense of earthquake forecasts is hard: standard interpretations of probability fail.

A model-based interpretation is better, but lacks empirical justification.

Probability models are only part of forecasting machinery.

USGS San Francisco Bay Area forecast for 2000–2030 involves geological mapping, geodetic mapping, viscoelastic loading calculations, paleoseismic observations, extrapolating rules of thumb across geography and magnitude, simulation, and many appeals to expert opinion.

The numerical probability values seem rather arbitrary. Uncertainty estimates are shaky, too.

Part II: Reducing Uncertainty using Physical Constraints

Examples of Physical Constraints

Nonnegativity:
Counts, energies, densities; monotonicity
seismic velocity in Earth's core (w/ R.L. Parker et al.)
aftershock probability density (w/ N. Hengartner)

Bounds on functionals:
Energy in geomagnetic field,
rotation rate in solar interior (w/ C. Genovese)

Parametrizations:
Power law for CMB spectrum (w/ L. Tenorio, C. Lineweaver)

How can we use such constraints to reduce uncertainty?

Most texts don't treat constraints.

Examples.

Models and Parameters

Model specifies prob. distribution of data X.
Models indexed by θ; call θ and $\mathbb{P}_\theta$ “the model.”

Physical Constraint: Know *a priori* that $\theta \in \Theta$

Parameter: image $f(\theta)$ of θ under a mapping f.

Building block: Bounded Normal Mean
Model $X \sim N(\theta, 1)$, with $\theta \in \Theta = [-\tau, \tau]$
Seek to estimate $f(\theta) = \theta$.

Confidence Sets and Coverage

Confidence Set:
Random set S of parameter values: depends on X

Coverage probability $(1 - \alpha)$:
Minimum chance $S(X)$ contains true $f(\theta)$.

$$\mathbb{P}_\theta\{S(X) \ni f(\theta)\} \geq 1 - \alpha, \quad \forall \theta \in \Theta. \tag{2}$$

Duality between testing and confidence sets:
Invert family of level α tests $\Rightarrow$ $1 - \alpha$ conf. set–
all $\theta' \in \Theta$ that aren't rejected.

Standard 95% conf. interval for normal mean:
$\mathcal{I} = [X - 1.96, X + 1.96]$.
Doesn't use the constraint $\theta \in [-\tau, \tau]$.

Procrustean (truncated) interval:
$\mathcal{I}_T = [X - 1.96, X + 1.96] \cap [-\tau, \tau]$.
Uses constraint, but is it best?

Common ways to add constraints

Ignore them.

Ad hoc; procrustean: Make unconstrained estimate; force it to be in the constraint set.

Bayesian: Use prior π that assigns probability 1 to the constraint set.

$$\mathbb{P}_\pi\{\theta \in \Theta\} = 1. \tag{3}$$

Frequentist minimax: Use estimator that (within some class) is minimax for some loss over the constraint set

Shortcomings

Ad hoc; procrustean: Can do “better.”

Bayesian: Where does the prior come from? No such thing as uninformative prior
Sensitive to loss and assumptions
Usual approach doesn’t have frequentist coverage (should we care?)

Frequentist minimax: Driven by worst case
Sensitive to loss and assumptions

Defining “best:” small expected size

Want a precise answer:
Minimize expected size of the confidence set.
Expected size depends on true value of θ: tradeoff.

Bayesian: minimize average (for prior π) expected size

Frequentist Minimax: minimize max expected size for $\theta \in \Theta$

Bayes/Minimax duality:
Minimax is Bayes for *least favorable prior*

Minimax expected size CI for BNM

τ	standard $\mathcal{I}$		truncated standard $\mathcal{I} \cap [-\tau, \tau]$		Best meas. fixed-width[a] $\mathcal{I}_{\mathrm{N}}$		Opt. meas.[b] $\mathcal{I}_{\mathrm{OPT}}$
1.75	3.9	49%	2.9	10%	3.3	25%	2.6
2.00	3.9	38%	3.2	11%	3.3	16%	2.8
2.25	3.9	31%	3.4	13%	3.3	10%	3.0
2.50	3.9	26%	3.6	14%	3.3	6%	3.1
2.75	3.9	22%	3.7	15%	3.3	3%	3.2
3.00	3.9	21%	3.8	16%	3.3	1%	3.3
3.25	3.9	19%	3.8	16%	3.3	0%	3.3
3.50	3.9	18%	3.9	16%	3.5	5%	3.3 [c]
3.75	3.9	16%	3.9	15%	3.6	6%	3.4
4.00	3.9	14%	3.9	13%	3.6	5%	3.4

[a]These have form $[\hat{\theta}(X) - e, \hat{\theta}(X) + e]$, with $\hat{\theta}(\cdot)$ measurable and e constant.

[b]These have form $\{\theta \in \Theta : (\theta, X) \in S\}$, where $S \subseteq \Theta \times \mathcal{X}$ is product-measurable.

[c]Truncated Pratt interval $\mathcal{I}_{\mathrm{TP}}$ is optimal for $\tau \leq 3.29$. The entries in the rightmost column for $\tau = 3.50, 3.75$, and 4.00 are approximated numerically.

Size isn't everything: how you use it matters.

Cost of including $\theta' \neq \theta$ can depend on θ:
E.g., might sacrifice length to include values with one sign only.

Curvature of the universe?

Can allow size measure to depend on θ.

Numerics:
Approximate least-favorable π by Monte Carlo
Well suited to distributed/parallel computing.
Uses importance sampling.
Don't need closed-form likelihood.
Solve 2-player matrix game iteratively.

Example: Microwave cosmology–relic of the Big Bang

Model for Power Spectrum

$$\mathbf{X} \sim N\left(\mathbf{0},\ \mathbf{N} + \sum_{\ell=1}^{\infty}\left(\frac{2\ell+1}{4\pi}\right)C_\ell(\theta)B_\ell^2\,\mathbf{P}_\ell\right)$$

$\mathbf{P}_\ell$: the ℓ^{th} Legendre polynomial matrix

$\mathbf{N}$: data noise covariance matrix

$\{C_\ell(\theta)\}$: power spectrum for the set of cosmological parameters θ

B_ℓ: transfer function of the observing filter.

Complicated relationship between cosmological "parameters" and spectrum $\{C_\ell\}$: nonlinear PDE.

Constraints on Parameters of Cosmological Model

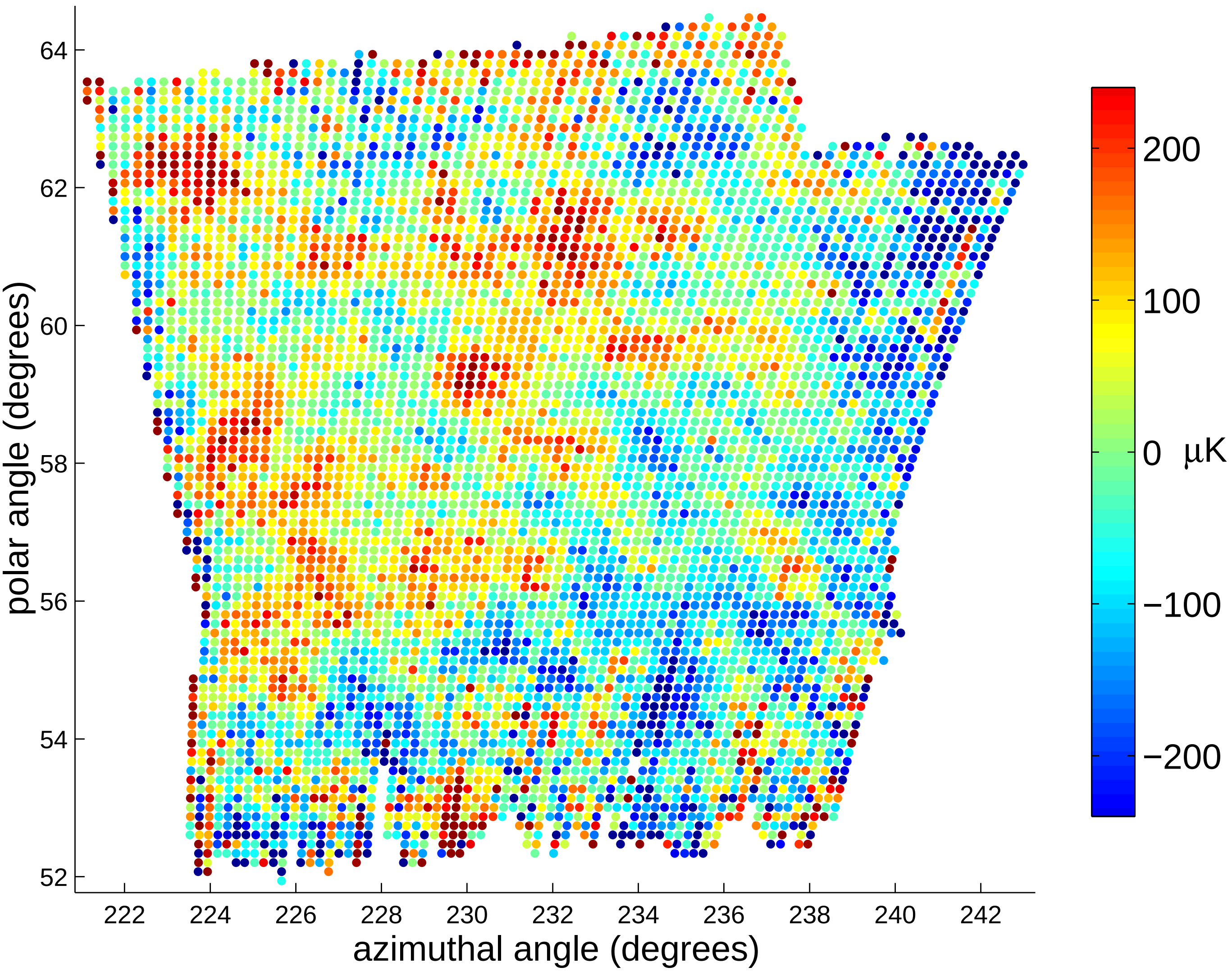
Parameter		**Lower**	**Upper**
Total Matter†	Ω_m	0.05	1.00
Baryonic Matter†	Ω_b	0.005	0.15
Cosmological Constant†	Ω_Λ	0.0	1.0
Hubble Constant (km s^{-1} Mpc^{-1})	H_0	40.0	90.0
Scalar Spectral Index	n_s	0.6	1.5
Optical Depth	τ	0.0	0.5

† Relative to critical density.

Monte-Carlo chooses model uniformly, subject to

$$\Omega_b \leq \Omega_m \text{ and } 0.6 \leq \Omega_m + \Omega_\Lambda \leq 1.4. \qquad (4)$$

MAXIMA-1 Data



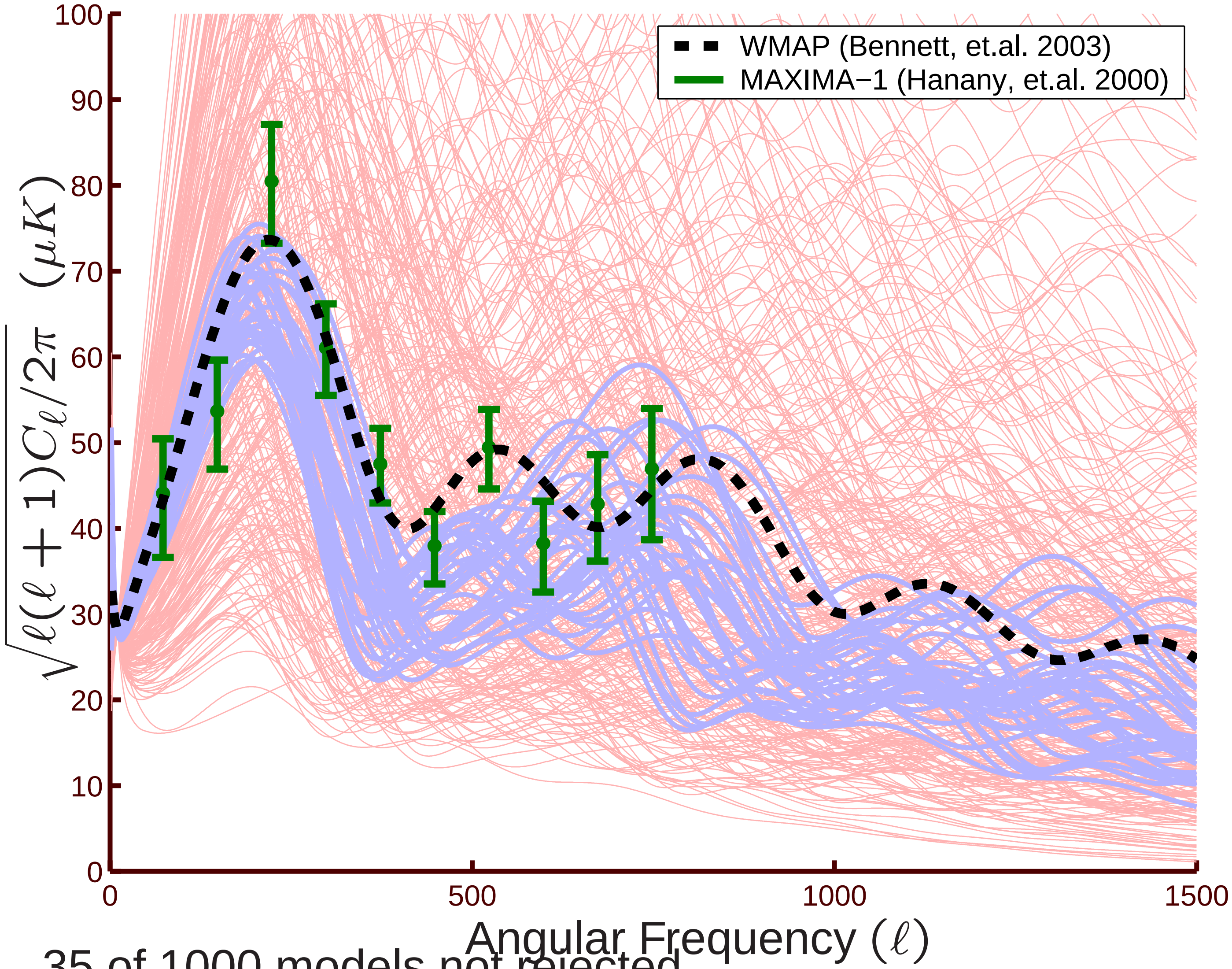
Data reduction/compression

Form 20,000 linear combinations of 5972 pixel temps.

Select 5972 linear combinations–eigenvectors of covariance matrix for reference model.

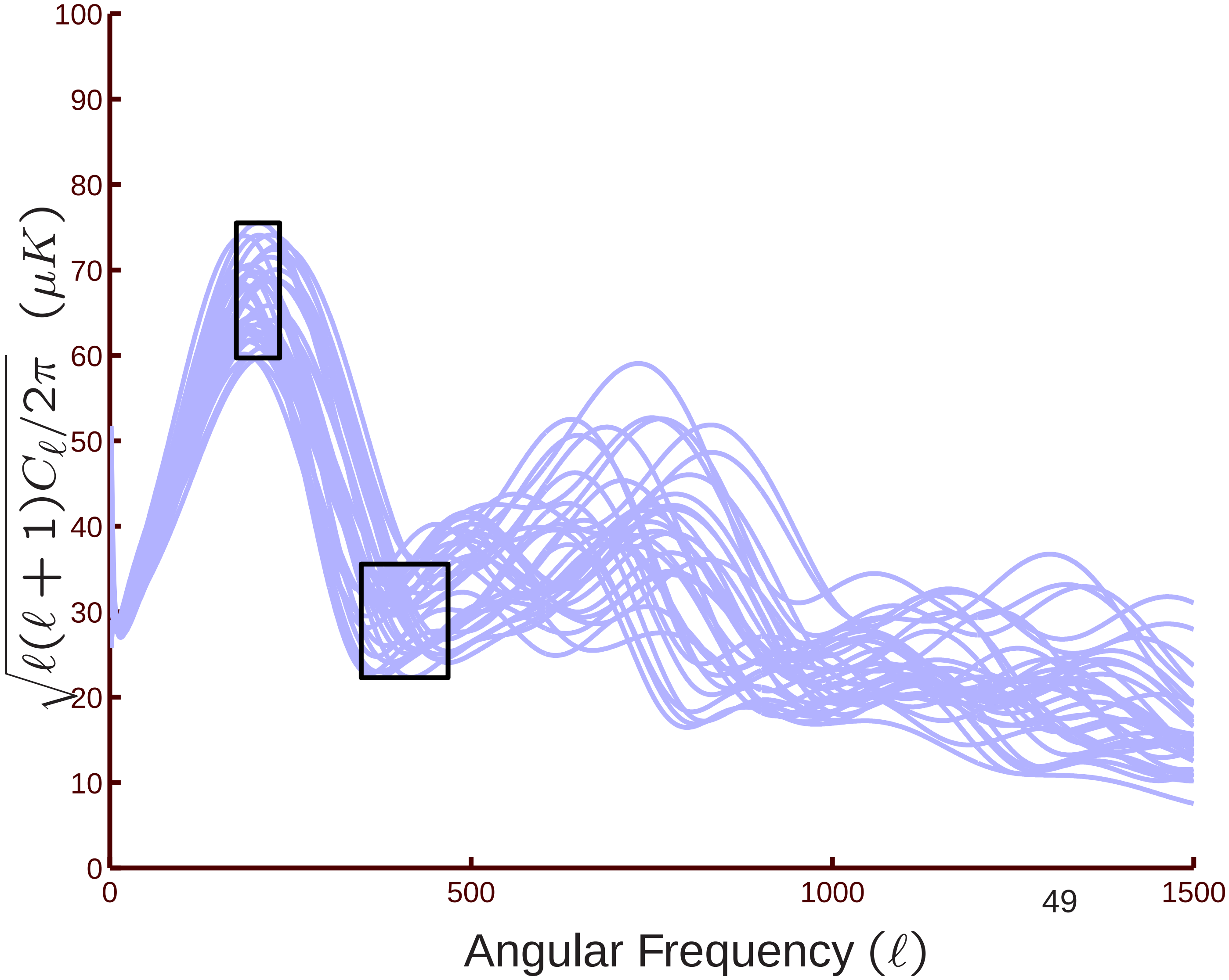
Select 2000 of those using estimated Kullbach-Leibler divergence to approximate false coverage prob.

Results



35 of 1000 models not rejected.
Spectra of 300 rejected models shown.

Results



100
90
80
70
60
50
40
30
20
10
0
0
500
1000
1500
$\sqrt{\ell(\ell+1)C_\ell/2\pi}\quad(\mu K)$
Angular Frequency (ℓ)

Results (contd)

Location of main peak (ℓ) : $(175, 235)$.
Height of main peak (μK) : $(59.7, 75.5)$.

Location of first valley (ℓ) : $(348, 468)$.
Height of first valley (μK) : $(22.3, 35.6)$.

Results (contd)

Some Accepted Models:

	Ω_b	Ω_m	Ω_Λ	τ	H_0	n_s	Ω
	0.042	0.674	0.241	0.317	77.00	1.117	0.915
	0.091	0.620	0.319	0.346	65.93	1.095	0.939
	0.104	0.684	0.316	0.152	53.00	0.960	1.000
	0.081	0.540	0.526	0.000	77.15	0.809	1.066
	0.072	0.754	0.328	0.116	86.55	0.970	1.082
	0.134	0.940	0.161	0.466	66.83	1.038	1.101
	0.085	0.301	0.815	0.058	62.21	0.729	1.116
	0.096	0.479	0.730	0.428	87.64	1.043	1.209
	0.096	0.555	0.693	0.260	81.28	0.923	1.248
	0.093	0.708	0.551	0.000	76.96	0.855	1.259
	0.139	0.667	0.623	0.269	61.15	0.954	1.290
	0.133	0.692	0.642	0.068	41.47	0.846	1.334
Min incl	0.011	0.058	0.082	0.000	41.47	0.729	0.915
Max incl	0.139	0.994	0.988	0.466	89.24	1.151	1.334
Min poss	0.005	0.000	0.000	0.000	40.00	0.600	0.600
Max poss	0.150	1.000	1.000	0.500	90.00	1.500	1.400

References

Balbi, A., et al., 2000. Constraints on Cosmological Parameters from MAXIMA-1, *Astrophys. J.*, *545*, L1–L4

Bennett, C.L., et al., 2003. First Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Maps and Basic Results, *Ap. J.*, in press.

Evans, S.N., B. Hansen, and P.B. Stark, 2002. Minimax Expected Measure Confidence Sets for Restricted Location Parameters. Tech. Rept. 617, Dept. Statistics Univ. Calif Berkeley.

Hanany, et al., 2000. MAXIMA-1: A Measurement of the Cosmic Microwave Background Anisotropy on angular scales of 10 arcminutes to 5 degrees, *Astrophys. J.*, *545*, L5.

Jaffe, et al., 2001. Cosmology from Maxima-1, Boomerang and COBE/DMR CMB Observations, *Phys. Rev. Lett.*, *86*, 3475–3479.

Kempthorne, P.J., 1987. Numerical specification of discrete least favorable prior distributions, *SIAM J. sci. statis. Comp.*, *8*, 171–184.

Nelson, W., 1966. Minimax solution of statistical decision problems by iteration, *Ann. math. Statis.*, *37*, 1643–1657.

Pratt, J.W., 1961. Length of Confidence Intervals, *J. amer. statis. Assoc.*, *56*, 549–567.

Rao, C.R., 1962. Efficient estimates and optimum inference procedures, *J. roy. statis. Soc.*, *24 (B)*, 46–72.

Robinson, J., 1951. An iterative method for solving a game, *Ann. Math.*, *54*, 296–301.