Statistical Analysis of CMB Anisotropy Data

Eric Gawiser

Bayesian and Frequentist Approaches

Ground rule: there is no “right” approach to statistics. Bayesian and Frequentist approaches ask different questions.

$$p(d,t) = p(d|t)p(t) = p(t|d)p(d) \ , \tag{1}$$

$$p(t|d) = \frac{p(d|t)p(t)}{p(d)} \ , \ p(d) = \int_t p(d|t)p(t) \ . \tag{2}$$

Assumptions, Basics, and Correlations

$$d_i = s_i + n_i \ . \tag{3}$$

$$< n_i n_j >= C^N_{ij} \ . \tag{4}$$

Assumption 1: The instrument noise obeys a multivariate Gaussian distribution described by $< n_i >= 0$ and the correlation matrix C^N.

If a theory specified all s_i directly,

$$L(d|t) = \frac{1}{(2\pi)^{\frac{n}{2}}|\mathbf{C}^N|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(\mathbf{d}-\mathbf{s})^T(\mathbf{C}^N)^{-1}(\mathbf{d}-\mathbf{s})\right) \ . \tag{5}$$

If the instrument noise were uncorrelated between pixels, $\mathbf{C}^N$ would be diagonal:

$$L(d|t) = \frac{1}{(2\pi)^{n/2}\prod_i \sigma_i} \exp\left(-\frac{\chi^2}{2}\right) \ . \tag{6}$$

$$\chi^2 = \sum_i (d_i - s_i)^2/\sigma_i^2 \ . \tag{7}$$

Assumption 2: The sky temperatures obey a multivariate Gaussian distribution given by $< s_i >= 0$ and the correlation matrix C^S.

$$C_{ij}^S =< s_i s_j >= \frac{1}{4\pi} \sum_\ell (2\ell+1) C_\ell W_\ell P_\ell(\theta_{ij}) \ , \tag{8}$$

Assumption 3: The combination of pixel convolution, beam convolution, and observing strategy are azimuthally symmetric and can be represented as a window function W_ℓ.

For zero instrument noise,

$$L(d|t) = \frac{1}{(2\pi)^{\frac{n}{2}} |\mathbf{C}^S|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}\mathbf{d}^T (\mathbf{C}^S)^{-1}\mathbf{d}\right) \ . \tag{9}$$

Assumption 4: The instrument noise is uncorrelated with the temperature fluctuations on the sky.

$$C_{ij}^D =< d_i d_j >=< (s_i+n_i)(s_j+n_j) >=< s_i s_j > + < n_i n_j > + < s_i n_j > + < n_i s_j >= C_{ij}^S + C_{ij}^N \ . \tag{10}$$

$$\begin{aligned} L(d|t) &= \int_{\mathbf{s}} \int_{\mathbf{n}} p(\mathbf{s}|t) p(\mathbf{n}) \delta(\mathbf{d} - \mathbf{s} - \mathbf{n}) \\ &= \frac{1}{(2\pi)^{\frac{n}{2}} |\mathbf{C}^D|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}\mathbf{d}^T (\mathbf{C}^D)^{-1}\mathbf{d}\right) \ , \end{aligned} \tag{11}$$

Deduction: The data obeys a multivariate Gaussian distribution.

Calibration Uncertainty

Reported temperature fluctuations are really in units of $c\ uK$ where $< c >= 1$ with uncertainty σ_c.

$$
\begin{aligned}
L(d|t) &= \int_c \int_{\mathbf{s}} \int_{\mathbf{n}} p(\mathbf{s}|t) p(\mathbf{n}) p(c) \delta(\mathbf{d} - c\mathbf{s} - \mathbf{n}) \\
&= \int dc \frac{1}{(2\pi)^{\frac{n+1}{2}} |c^2 \mathbf{C}^S + \mathbf{C}^N|^{\frac{1}{2}} \sigma_c} \exp\left\{ -\frac{1}{2} \left(\mathbf{d}^T (c^2 \mathbf{C}^S + \mathbf{C}^N)^{-1} \mathbf{d} + \frac{(c-1)^2}{\sigma_c^2} \right) \right\} . \qquad (12)
\end{aligned}
$$

$$
\chi_c^2 = \mathbf{d}^T (c^2 \mathbf{C}^S + \mathbf{C}^N)^{-1} \mathbf{d} + \frac{(c-1)^2}{\sigma_c^2} . \qquad (13)
$$

Beam Uncertainty

Best-fit Gaussian beam has standard deviation b^* with uncertainty σ_b.

$$
W_\ell = \exp[-\ell(\ell+1) b^2] . \qquad (14)
$$

$$
\begin{aligned}
L(d|t) &= \int_c \int_b \int_{\mathbf{s}} \int_{\mathbf{n}} p(\mathbf{s}|t) p(\mathbf{n}) p(b) p(c) \delta(\mathbf{d} - c\mathbf{s}(b,t) - \mathbf{n}) \\
&= \int dc \int db \frac{1}{(2\pi)^{\frac{n+2}{2}} |c^2 \mathbf{C}^S(b,t) + \mathbf{C}^N|^{\frac{1}{2}} \sigma_c \sigma_b} \exp\left\{ -\frac{1}{2} \left(\mathbf{d}^T (c^2 \mathbf{C}^S(b,t) + \mathbf{C}^N)^{-1} \mathbf{d} + \frac{(c-1)^2}{\sigma_c^2} + \frac{(b-b^*)^2}{\sigma_b^2} \right) \right\} (15)
\end{aligned}
$$

Maximizing versus Marginalizing over Nuisance Parameters

$$
p(d|t) = \int du\, p(d|u,t) p(u) . \qquad (16)
$$

Maximization corresponds to **retroactively assuming a delta-function prior** in the nuisance parameters u at their optimal values u^* for each theory:

$$\int du\, p(d|u,t)\delta(u-u^*(t)) = p(d|u^*(t),t)\ . \tag{17}$$

The same result occurs for

$$\int du\, p(d|u^*,t)\exp\left(-\frac{(u-u^*(t))^2}{2\sigma_u^2}\right) = p(d|u^*(t),t)\ , \tag{18}$$

which requires u to be **independently** Gaussianly distributed:

$$\frac{\partial^2 L}{\partial u \partial t_i} \propto \delta_{ut_i} \tag{19}$$

Impressive difference for BOOMERANG!

Dangers of Assuming a Gaussian Shape to Set Credible Regions

Melchiorri et al., sometimes Lange et al., and maybe Balbi et al. assume that 68% (or 95%) of area is found down to $L/L_{max} = \exp(-1/2)$ (or $\exp(-2)$) as would be the case for a one-dimensional Gaussian.

1. Best-fit L_{max} fluctuates more than an integral does.

2. n parameters may only use up m effective degrees of freedom in the data.

3. Varying the prior is guaranteed to vary the shape that you're integrating over.

How Should I Set the Credible Region?

Determine the threshold value of the likelihood for which the volume under it comprises 68% of the data, **even if this region is not simply connected**. Quite different from the 16%, 50%, and 84% integrals used by Jaffe et al.

Data Compression into Bandpower Estimates

$$L(d|t) = \frac{1}{(2\pi)^{\frac{n_B}{2}} |\mathbf{C}^{BB'}|^{\frac{1}{2}}} \exp\left\{ -\frac{1}{2} C_B^T(t) (\mathbf{C}^{BB'})^{-1} C_B(t) \right\} \; . \tag{20}$$

Only an approximation that all theories which predict a given C_B are equally likely.

The Offset Lognormal Approximation

Bond, Jaffe, and Knox [astro-ph/9808264] look at curvature at likelihood max:

$$-\frac{\partial^2 \log L(d|t)}{\partial C_\ell \partial C_{\ell'}} = \frac{2\ell+1}{2} (C_\ell + N_\ell / B_\ell^2)^{-2} \delta_{\ell\ell'} \; . \tag{21}$$

$$Z_\ell \equiv \log(C_\ell + N_\ell / B_\ell^2) \; . \tag{22}$$

Then curvature at Z_ℓ^* is constant so this looks Gaussian, and if we're going to do a χ^2 analysis Z_ℓ is the best variable to use. Should probably go ahead and use a better approximation to the likelihood function and not do a χ^2 analysis.