

Lecture 5

◦ Twin pillars of statistical inference

Classical / Frequentist

Probabilities refer to the frequency of outcomes in repeated experiments

Bayesian

Probabilities refer to subjective belief conditioned on data and prior knowledge.

⇒ Priors in Bayesian statistics act as boundaries on our inference.

◦ Maximum Likelihood Estimation

- ① Formulate a model with some parameters, θ .
- ② Find the parameters θ^* that best fit the observed data.
- ③ Estimate the uncertainty on those parameters.
- ④ Try other models, and see which fits best.

• Likelihood \rightarrow fitting/objective function.

e.g let's assume the noise in our data is Gaussian.

$$\text{data} = \text{signal} + \text{noise}$$
$$d = s + n$$

$$p(n_i) = N(0, \sigma_i)$$

$$\therefore p(d_i) = N(s_i, \sigma_i)$$

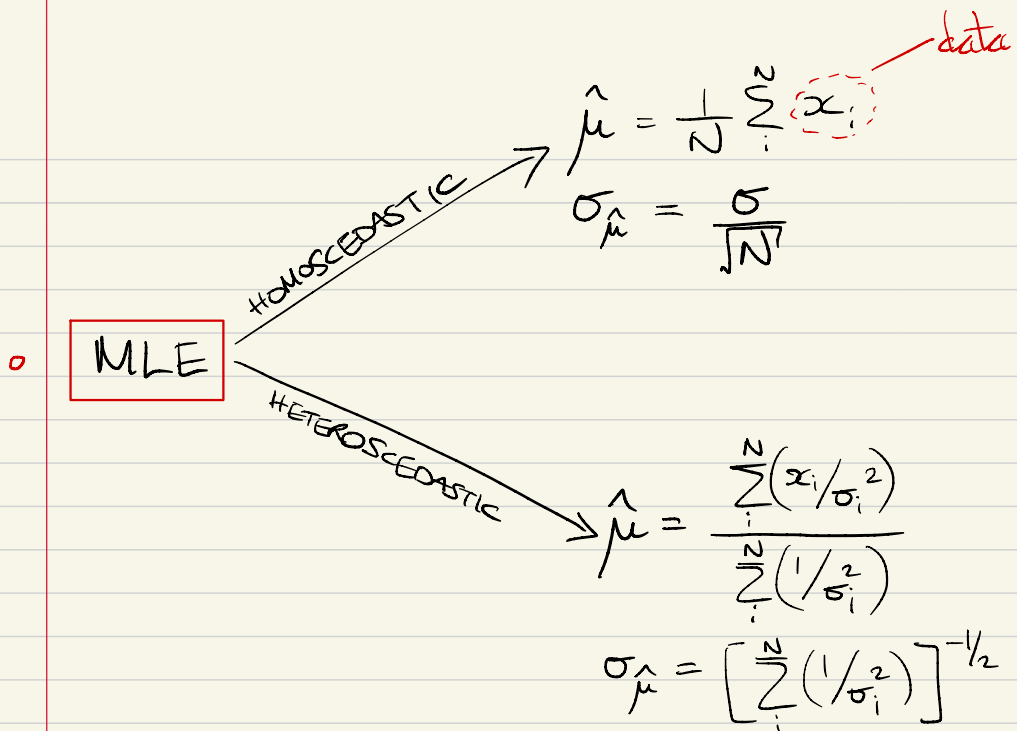
$$\text{LIKELIHOOD}_i = \frac{1}{\sqrt{2\pi\sigma_i^2}} e^{-\frac{1}{2} \frac{(d_i - s_i)^2}{\sigma_i^2}}$$

PROBABILITY OF
DATA GIVEN PARAMETERS

$$\Rightarrow (\text{LIKELIHOOD OF ALL DATA}) = \prod p(d_i)$$
$$= \prod \frac{1}{\sqrt{2\pi\sigma_i^2}} e^{-\frac{1}{2} \sum_i \frac{(d_i - s_i)^2}{\sigma_i^2}}$$

NOTE $\chi^2 = \sum_i \frac{(d_i - s_i)^2}{\sigma_i^2}$

--- MAXIMIZING the (\log) -likelihood is the SAME
as MINIMIZING the χ^2 value.



Quantifying Uncertainty on Model Parameters

* With > 1 parameter, you can have correlated measurement uncertainties,

$$\sigma_{jk} = ([F^{-1}]_{jk})^{1/2}$$

where $F_{jk} = - \frac{\partial^2 \ln L}{\partial \theta_j \partial \theta_k} \Big|_{\theta = \theta_{MLE}}$

FISHER INFORMATION MATRIX

e.g. Measuring the mean μ of Gaussian-distributed data with homoscedastic uncertainties $\sigma_{\hat{\mu}} = \frac{\sigma}{\sqrt{N}}$