BASIC STATISTICS III

CHI SQUARED AND LIKELIHOOD

Individual values of "Chi squared" – comparison of observed residuals from a model to expected residuals (=uncertainties)

$$\chi^2 = \sum_i \frac{(O_i - E_i)^2}{\sigma_i^2}$$

These values are drawn from the Chi squared probability distribution if the residuals follow normal distributions.

The likelihood of a model given a data set is *proportional to* $e^{-\chi^2/2}$ if we the residuals follow normal distributions around each individual data point. (Verify for yourself that multiplying the individual probability distributions then gives an overall probability distribution proportional to $e^{-\chi^2/2}$.)

TRADITIONAL MAXIMUM LIKELIHOOD

- seek "best fit" models/parameters
- typically assume likelihood L proportional to $e^{-\chi^2/2}$
- min χ^2 ➔ max likelihood
- "maximum likelihood estimators" (MLEs) of parameters α_i of a model are usually found by $\frac{\partial L}{\partial \alpha_i} = 0$ or equivalently $\frac{\partial \ln(L)}{\partial \alpha_i} = 0$
- assuming all residuals follow same normal distribution (i.e. have same σ), called "ordinary least-squares" (OLS) fitting or "minimizing the rms" (root mean square deviations)

TRADITIONAL MAXIMUM LIKELIHOOD

- Example: model $y = \alpha X + \beta$ with equal Gaussian errors σ
- $\chi^2 = \sum_i \frac{(Y_i - (\alpha X_i + \beta))^2}{\sigma^2}$ ➔ max likelihood
- $\frac{\partial \ln(L)}{\partial \alpha} = 0$ ➔ $\frac{\partial \ln\left(e^{-\frac{\chi^2}{2}}\right)}{\partial \alpha} = 0$ ➔ $\sum_i \frac{\left(Y_i - (\alpha X_i + \beta)\right) X_i}{\sigma^2} = 0$
- $\frac{\partial \ln(L)}{\partial \beta} = 0$ ➔ $\frac{\partial \ln\left(e^{-\frac{\chi^2}{2}}\right)}{\partial \beta} = 0$ ➔ $\sum_i \frac{\left(Y_i - (\alpha X_i + \beta)\right)}{\sigma^2} = 0$
- two eqns, two unknowns – solve to get result in tutorial:

  $\alpha = \frac{\bar{X}\bar{Y} - \overline{XY}}{(\bar{X})^2 - \overline{X^2}}$ and $\beta = \bar{Y} - \bar{X}\alpha$

  (so for this simple case, no numerical χ^2 minimization is needed; but harder for more parameters or different σ_i)

TRADITIONAL MAXIMUM LIKELIHOOD

- uncertainties on MLE params estimated by 1/E(-H) = inverse of expectation of negative "Hessian matrix"
- Hessian matrix example: $y = \alpha X + \beta$

$$\text{Hessian}(\alpha,\beta) = \begin{bmatrix} \frac{\partial^2}{\partial\alpha^2}\log L(\alpha,\beta) & \frac{\partial^2}{\partial\alpha\partial\beta}\log L(\alpha,\beta) \\ \frac{\partial^2}{\partial\alpha\partial\beta}\log L(\alpha,\beta) & \frac{\partial^2}{\partial\beta^2}\log L(\alpha,\beta) \end{bmatrix}$$

note covariance terms!

- complicated to compute Hessians, often done numerically
- fully worked Hessian for least squares case at http://mathworld.wolfram.com/LeastSquaresFitting.html - note errors on parameters generally decrease as $\frac{1}{\sqrt{N}}$