Note: No lecture on
Wed Oct 12.

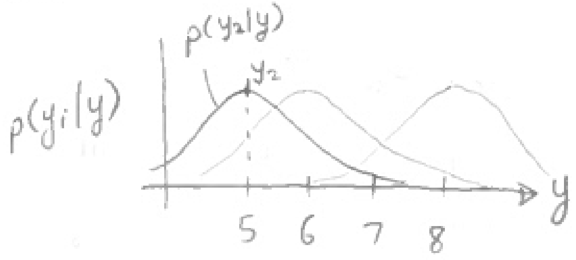
On Fri Oct 14 I will do more blackboard + lab, but I am going to try preparing a notebook ahead of time with the overall structure but with places where you have to fill in code.

We will do:

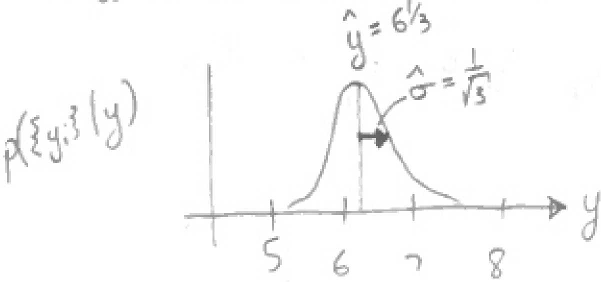
- handling outliers via foreground-background modeling
- intro to Markov Chain Monte Carlo
- using Julia's nonlinear optimizer to find solutions to this model

Review & Comments from last time

Likelihoods of our observations



Likelihood of getting the 3 observations we got
= product of 3 individual likelihoods (Gaussian)
= also a Gaussian (← magical!)



Then we used Bayes' rule to take

$$p(\{y_i\}|y) \quad \text{to} \quad p(y|\{y_i\})$$

which is shorthand for $p(y|$ observation <u>and</u> all our assumptions$)$

Comments

- probabilities : "beliefs", "knowledge", "uncertainty"

Bayes' rule : math for updating our statement of our knowledge.

You take a new obs. y_{new}

$$p(y \mid y_{new}) = \frac{p(y_{new} \mid y)\, p(y)}{p(y_{new})}$$

Implicitly, Include "everything you already know", E :

$$p(y \mid y_{new}, E) = \frac{p(y_{new} \mid y, E)\, p(y, E)}{p(y_{new}, E)}$$

"E" can include previous measurements!

$$p(y \mid y_{new}) \propto p(y_{new} \mid y) \cdot N(\hat{y}, \hat{\sigma}^2)$$

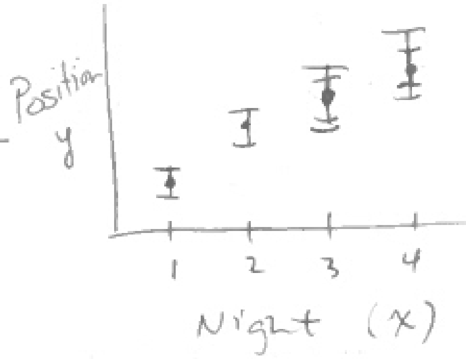
Updates our statement of what we know / believe about the value of y.

- imagine you know y. Then can plot:

"you don't know where your observation sits on this curve"

Fitting a line

Night	Pos	σ
1	14	1
2	26	1
3	37	2
4	43	2



Model:

$$y = mx + b$$

Assumptions:

- uncertainties on $\underline{x}$ are negligible
- measured $y_i = m x_i + b + e_i$
- e_i are iid. drawn from $N(0, \sigma_i^2)$

As before, can write likelihood of the observations given parameters:

$$\underbrace{p(\{y_i\}|m,b)}_{\mathcal{L}} = \prod_i \frac{1}{\sqrt{2\pi}\,\sigma_i} \exp\left(-\frac{1}{2}\frac{(mx_i+b-y_i)^2}{\sigma_i^2}\right)$$

$$\log \mathcal{L} = c - \frac{1}{2}\sum \frac{(mx_i+b-y_i)^2}{\sigma_i^2}$$

Could take derivatives to find m, b as before, but that gets messy.

Instead, take the χ^2 term

$$\chi^2 = \sum_i \frac{(mx_i+b-y_i)^2}{\sigma_i^2}$$

and write as matrix math

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \end{bmatrix} \quad A = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \end{bmatrix} \quad W = \begin{bmatrix} \frac{1}{\sigma_1^2} & & 0 \\ & \frac{1}{\sigma_2^2} & \\ 0 & & \ddots \end{bmatrix}$$

$$X = \begin{bmatrix} b \\ m \end{bmatrix}$$

awkward change of convention on "x"!

$$\chi^2 = (AX-Y)^T W (AX-Y)$$

$$\chi^2 = \left[W^{1/2}(AX-Y)\right]^T \left[W^{1/2}(AX-Y)\right]$$

This is a matrix least-squares problem!

Julia has a built-in "\" operator to solve this:

$$\hat{X} = \underset{x^T x}{\text{argmin}} \left(r = (W^{1/2}A)\,X - (W^{1/2}Y) \right)$$

$$\hat{X} = (W^{1/2}A) \backslash (W^{1/2}Y)$$

~~We want to find a "least squares" solution!~~

$\cancel{\chi =}$

$$\text{cov} = (A^T W A)^{-1}$$

What this solves is

$$\hat{X} = (A^T W A)^{-1} (A^T W Y)$$

Lab exercises

- set up this matrix equation using our data set, and solve it <u>with</u> and <u>without</u> the weight matrix W.

- plot the data points (with error bars) and the weighted and unweighted fit lines.

- compute the covariance of the maximum likelihood m and b values

- plot, in the m,b plane, the best-fit (weighted and unweighted) values, and a sampling of values from the covariance matrix.

- plot, in x,y space, a sampling of lines drawn from the best-fit m,b plus covariance.

- optional: fit a quadratic

- write the log-likelihood function for our data set given values of m, b.
- evaluate this function on a grid of m, b values
- plot this log-likehood surface (eg, as a heatmap)