Bayesian Data Analysis (PHY/CSI/INF 451/551)
HW#5w

In class we discussed using a Taylor series approximation of the logarithm of a probability density function (PDF) to obtain an estimate of the uncertainty about the mode (most probable solution).

Basically, the Log (base-e) of the PDF, denoted $L(x) = \ln(P(x))$, is approximated at the mode $\hat{x}$ by:

$$L(x) \approx L(\hat{x}) + \left.\frac{dL}{dx}\right|_{x=\hat{x}} (x-\hat{x}) + \frac{1}{2}\left.\frac{d^2L}{dx^2}\right|_{x=\hat{x}} (x-\hat{x})^2 + \cdots \tag{1}$$

where $\left.\frac{dL}{dx}\right|_{x=\hat{x}} = 0$, since $x = \hat{x}$ is a maximum. This leaves

$$L(x) \approx L(\hat{x}) + \frac{1}{2}\left.\frac{d^2L}{dx^2}\right|_{x=\hat{x}} (x-\hat{x})^2 + \cdots \tag{2}$$

so that taking the exponential of both sides we have an approximation to the original PDF,

$$P(x) \approx P(\hat{x})\, Exp\left[\frac{1}{2}\left.\frac{d^2L}{dx^2}\right|_{x=\hat{x}} (x-\hat{x})^2\right], \tag{3}$$

which is essentially a Gaussian since at the peak $x = \hat{x}$, the second derivative is negative $\left.\frac{d^2L}{dx^2}\right|_{x=\hat{x}} < 0$. We usually write the exponent of a Gaussian as

$$Exp\left[-\frac{(x-\hat{x})^2}{2\,\sigma^2}\right], \tag{4}$$

which, using the two equations above, allows us to write

$$\sigma^2 = \left(-\left.\frac{d^2L}{dx^2}\right|_{x=\hat{x}}\right)^{-1}, \tag{5}$$

resulting in an objectively-determined measure of uncertainty

$$\sigma = \left(-\left.\frac{d^2L}{dx^2}\right|_{x=\hat{x}}\right)^{-\frac{1}{2}} \tag{6}$$

that comes from approximating the peak of the distribution $P(x)$ with a Gaussian distribution.

1. **Consider that you have solved a Bayesian estimation problem for a model parameter x, that has resulted in a posterior probability density of the form**

   $$P(x) = \frac{1}{\Gamma(k)\,\theta^k}\, x^{k-1}\, e^{-\frac{x}{\theta}}$$

   With $k = 9$ and $\theta = \frac{1}{2}$ and Γ is the Gamma function, where $\Gamma(9) = 40230$.

   A. Plot the function $P(x)$ for $0 \le x \le 20$, and verify that it has a peak within that range.

   B. Take the first derivative of the log probability to find the most probable value of x, written $\hat{x}$, and verify that this result agrees with your plot.

   C. Take the second derivative of the log probability and find the uncertainty, σ, of this estimate $x = \hat{x}$. Write the most probable solution as $x = \hat{x} \pm \sigma$.

   D. Take your plot of P(x) and overlay a plot of the Gaussian approximation as in (3) and (4) using a different color to verify that the Gaussian approximation is a good description of the peak of P(x).

Bayesian Data Analysis (PHY/CSI/INF 451/551)
HW#5w

2. The Mean

Consider a problem where you are interested in the distance x between two cities. You would like to come up with a "best" estimate, $\hat{x}$, but all you know about x, is the probability density over all the possible distances $p(x \mid I)$.

Here we consider a cost function that will penalize you if you do not get the answer right. This cost function depends on the square of the difference between your estimate $\hat{x}$ and the correct, but unknown, distance x. That is, the cost function is $(x - \hat{x})^2$.
The idea is to minimize the cost $(x - \hat{x})^2$. However, all you know about the distance x is the probability density $p(x \mid \mathrm{I})$. Therefore, the best you can do is minimize the expected value of the cost function.

A. Given that the expectation value of a given function $f(x)$ is
$\langle f(x) \rangle = \int f(x)\, p(x \mid \mathrm{I})\, dx$,
write an expression for the expectation value of the cost function $(x - \hat{x})^2$.
Don't worry about solving it directly, I haven't told you $p(x \mid I)$.

B. Minimize the expectation of the cost function with respect to your estimate $\hat{x}$, and show that the result is the expectation value of x: $\langle x \rangle$, which is called the mean value of x, commonly denoted as $\bar{x} = \langle x \rangle$.
Hint: Remember that you can find maxima and minima of functions by taking the derivative.

3. The Median

Now that you have had some experience with cost functions, try minimizing the expectation value of this cost function: $|x - \hat{x}|$. This is not as nice, since there is an absolute value, which is discontinuous. Show that the result of minimizing the expected value of $|x - \hat{x}|$ with respect to your estimate $\hat{x}$ is the median.
Hint: Break the integral into two pieces using the definition of the absolute value.
Hint 2: You don't have to show that one-half of the integral is 50%. You only need to show that the probability to the left of the median is equal to the probability to the right of the median.

4. The Mode

Show that the mode results from minimizing the expectation value of a cost function defined by the delta function: $-\delta(x - \hat{x})$. This cost function is very strict. You lose unless you get the answer absolutely right!
Hint: Recall that the delta function is zero except when its argument is 0, and that it integrates to 1.
Hint 2: Recall that $\int f(x)\, \delta(x - \hat{x})\, dx = f(\hat{x})$.
Hint 3: You don't need to take derivatives to solve this problem. Just do the integral and think about what condition will minimize the result.

These are the implicit concepts that lie behind choosing the mean, the median or the mode to summarize your results. Pretty arbitrary huh?