CHE384, From Data to Decisions: Measurement, Uncertainty, Analysis, and Modeling

Lecture 74
Bayesian Regression, part 1

Chris A. Mack
Adjunct Associate Professor

http://www.lithoguru.com/scientist/statistics/

Maximum Likelihood Regression

- For conventional regression, we try to maximize the likelihood function:
  - Likelihood = $\mathrm{P}(\boldsymbol{y}|\hat{\boldsymbol{\theta}})$
  - $\boldsymbol{y}$ = experimental data set
  - $\boldsymbol{\theta}$ = parameters of the model, $\hat{\boldsymbol{\theta}}$ = estimate of $\boldsymbol{\theta}$
  - These unknown parameters are thought of as constants that must be estimated
- But should we instead maximize $\mathrm{P}(\boldsymbol{\theta}|\boldsymbol{y})$?
  - Here, the unknown parameters are thought of as random variables that have pdf's

Bayes' Theorem

- From the definition of conditional probability we can write $\mathrm{P}(A \cap B)$ two ways

$$\mathrm{P}(A \cap B) = \mathrm{P}(A|B)\mathrm{P}(B)$$
$$\mathrm{P}(A \cap B) = \mathrm{P}(B|A)\mathrm{P}(A)$$

- Combining, $\mathrm{P}(A|B) = \dfrac{\mathrm{P}(B|A)\mathrm{P}(A)}{\mathrm{P}(B)}$
  - Bayes theorem allows us to flip what is given

Bayes' Theorem Example

- Consider a lab test for a disease
  - It is 95% effective at detecting the disease
  - It has a false positive rate of 1%
  - The rate of occurrence of the disease in the general population is 0.5%
- I take the screening test and get a positive result. What is the likelihood I have the disease?

Bayes' Theorem Example (2)

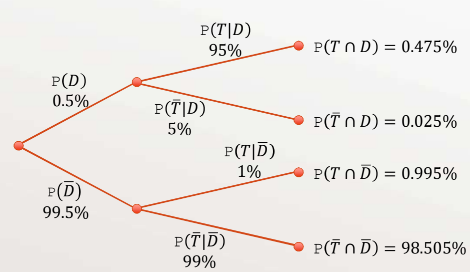
- Carefully define our variables:
  - Let T = I tested positive
  - Let D = I have the disease
  - We want $\mathrm{P}(D|T)$

$$\mathrm{P}(T|D) = 0.95$$
$$\mathrm{P}(T|\bar{D}) = 0.01$$
$$\mathrm{P}(D) = 0.005$$

$$\mathrm{P}(D|T) = \frac{\mathrm{P}(T|D)\mathrm{P}(D)}{\mathrm{P}(T)} = \frac{\mathrm{P}(T|D)\mathrm{P}(D)}{\mathrm{P}(T|D)\mathrm{P}(D) + \mathrm{P}(T|\bar{D})\mathrm{P}(\bar{D})}$$

$$\mathrm{P}(D|T) = \frac{(0.95)(0.005)}{(0.95)(0.005) + (0.01)(0.995)} = 0.32$$

Probability Tree Diagram



Bayesian Statistics

- We can use Bayes theorem to systematically combine what we learned from *this* experiment with what we knew *before* we began the experiment
- Our previous example:
  – What we knew before: effectiveness of test, false positive rate, prevalence of the disease in the population
  – What we learned from this experiment: the results of the test on one patient (me)

Bayesian Regression

Posterior Distribution, Likelihood Function, Prior Distribution, Normalizing constant

$$P(\theta|\mathbf{y}) = \frac{P(\mathbf{y}|\theta)P(\theta)}{P(\mathbf{y})}$$

- We already know how to calculate the likelihood function
- The term $P(\mathbf{y})$ is a constant (independent of θ) and used to normalize the posterior pdf
- We need to provide the prior distribution $P(\theta)$ in order to calculate the posterior distribution
- How do we interpret these distributions?
  – To a Bayesian, all probabilities are subjective

Bayesians vs. Frequentists

- Frequentists
  – Probability is the limiting case of repeated (possibly hypothetical) measurements
  – Model parameters are unknown constants and data (results of experiments) are random variables
- Bayesians
  – Probability is subjective, based on my degree of certainty in the event (my degree of knowledge)
  – Model parameters are random variables and data are constants

Lecture 74: What have we learned?

- Be able to explain and use Bayes' Rule
- How are data and model parameter estimates treated in maximum likelihood estimation as compared to Bayesian regression?
- What is the meaning of the prior distribution?
- What is the meaning of the posterior distribution?
- What is the difference between frequentist and Bayesian views of statistics?