LECTURE 21

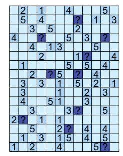
- **Readings:** Sections 8.1–8.2

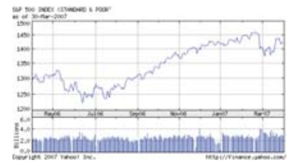
Lecture outline

- Statistical inference
  - Contrast with probability theory
  - Bayesian vs. classical
- Bayesian inference
  - Four primary forms of Bayes' rule
  - Types of problems/outputs
  - MAP estimation

Statistics

- Drawing infererences from limited and imperfect data
- Design and interpretation of experiments
  - polling, census, medical/pharmaceutical trials
- Netflix competition • Finance

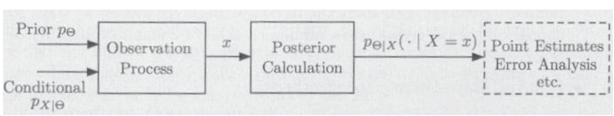




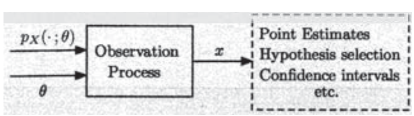
- Signal processing
  - Tracking, detection, speaker identification, forensics

Bayesian vs. classical inference

- Want to make inferences about *parameter(s)* θ
- Bayesian: θ is a realization of random variable Θ



- Classical: θ is unknown but not random



Four versions of Bayes' rule

- Θ discrete, X discrete: $$p_{\Theta|X}(\theta \mid x) = \frac{p_\Theta(\theta)\, p_{X|\Theta}(x \mid \theta)}{\sum_k p_\Theta(k)\, p_{X|\Theta}(x \mid k)}$$
- Θ discrete, X cont.: $$p_{\Theta|X}(\theta \mid x) = \frac{p_\Theta(\theta)\, f_{X|\Theta}(x \mid \theta)}{\sum_k p_\Theta(k)\, f_{X|\Theta}(x \mid k)}$$
- Θ cont., X discrete: $$f_{\Theta|X}(\theta \mid x) = \frac{f_\Theta(\theta)\, p_{X|\Theta}(x \mid \theta)}{\int f_\Theta(t)\, p_{X|\Theta}(x \mid t)\, dt}$$
- Θ cont., X cont.: $$f_{\Theta|X}(\theta \mid x) = \frac{f_\Theta(\theta)\, f_{X|\Theta}(x \mid \theta)}{\int f_\Theta(t)\, f_{X|\Theta}(x \mid t)\, dt}$$

Parameter of a coin

- Suppose a coin has probability of heads θ. What do we believe about θ after observing X heads in n tosses?
- What is the Bayesian approach?
- $f_{\Theta|X}(\theta \mid k) =$
- Beta(α, β) prior:

$$f_\Theta(\theta) = \begin{cases} \frac{1}{B(\alpha,\beta)}\theta^{\alpha-1}(1-\theta)^{\beta-1}, & \text{if } 0 < \theta < 1; \\ 0, & \text{otherwise} \end{cases}$$

where

$$B(\alpha, \beta) = \frac{(\alpha-1)!(\beta-1)!}{(\alpha+\beta-1)!}$$

Common mean of normal random variables

- Suppose normal $X_1, X_2, \ldots, X_n$ have an unknown common mean θ and known variances $\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2$
- If Θ is normal and X_is are conditionally independent given Θ, then $f_{\Theta|X}$ is normal
  - [Algebraic details in Example 8.3]
- Every observed $X_i \Rightarrow$ posterior update within normal class
  - Only mean update and variance update
- Important in engineering applications

Questions asked about θ

- **Binary hypothesis testing**:
  Choose between two possibilities for θ
- m-**ary hypothesis testing**:
  Choose between m possibilities for θ
- **Estimation**:
  Pick a number $\hat{\theta}$ that approximates θ
  - **Estimator**: random variable $\widehat{\Theta} = g(X)$ for some g
  - **Estimate**: value $\hat{\theta}$ of an estimator
    (determined by realization x of X)

Maximum a posteriori probability (MAP) rule

- Posterior distribution: PMF $p_{\Theta|X}(\cdot \mid x)$ or PDF $f_{\Theta|X}(\cdot \mid x)$

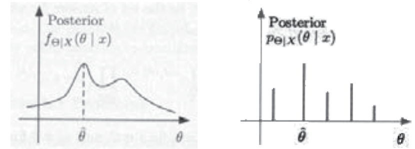


Figure 8.3. Illustration of the MAP rule for inference of a continuous parameter (left figure) and a discrete parameter (right figure).

- Pick $\hat{\theta}$ such that

$$p_{\Theta|X}(\hat{\theta} \mid x) = \max_{\theta} p_{\Theta|X}(\theta \mid x)$$

$$f_{\Theta|X}(\hat{\theta} \mid x) = \max_{\theta} f_{\Theta|X}(\theta \mid x)$$