ELC 5396: Digital Communications

Liang Dong

Electrical and Computer Engineering
Baylor University

liang_dong@baylor.edu

September 6, 2016

Probability Theory

Conditional Probability

$$\mathbb{P}[A|B] = \frac{\mathbb{P}[A \cap B]}{\mathbb{P}[B]} = \frac{\mathbb{P}[A, B]}{\mathbb{P}[B]}$$

Bayes' Rule

$$\mathbb{P}[B|A] = \frac{\mathbb{P}[A|B]\mathbb{P}[B]}{\mathbb{P}[A]}$$

Independence

$$\mathbb{P}[B|A] = \mathbb{P}[B], \;\; \mathbb{P}[A|B] = \mathbb{P}[A], \;\; \mathbb{P}[A \cap B] = \mathbb{P}[A]\mathbb{P}[B]$$

Probability Theory

Conditional pdf

$$f_Y(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)}$$

Independence of Random Variables

$$f_{X,Y}(x,y) = f_X(x)f_Y(y)$$

Sum of Independent Random Variables

$$Z = X + Y, \;\; f_Z(z) = \int_{-\infty}^{\infty} f_X(x)f_Y(z-x)dx$$

Probability Theory

Expectation: $\mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y]$
$\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$ (X and Y are independent)

Function of a random variable: $Y = g(X)$

nth moment: $\mathbb{E}[X^n]$

Variance: $\mathrm{var}[X] = \sigma_X^2 = \mathbb{E}[X^2] - \mu_X^2$

Covariance: $\mathrm{cov}[XY] = \mathbb{E}[XY] - \mu_X \mu_Y$

X and Y are uncorrelated: $\mathrm{cov}[XY] = 0$

X and Y are orthogonal: $\mathbb{E}[XY] = 0$

Distributions

Gaussian distribution:

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad x \sim \mathcal{N}(\mu, \sigma^2)$$

Exponential distribution:

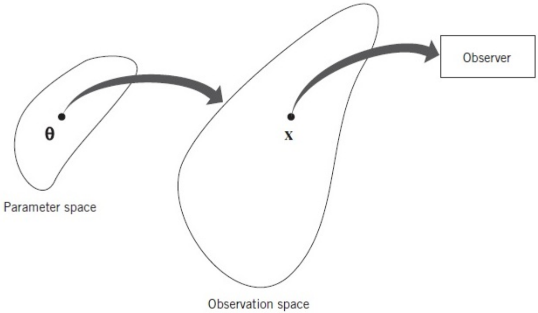
$$f_X(x; \lambda) = \begin{cases} \lambda e^{-\lambda x} & , x \geq 0 \\ 0 & , x < 0 \end{cases}, \quad \mathbb{E}[X] = \frac{1}{\lambda}, \mathrm{Var}[X] = \frac{1}{\lambda^2}$$

Rayleigh distribution:

$$f_X(x; \sigma) = \begin{cases} \frac{x}{\sigma^2} e^{-\frac{x^2}{2\sigma^2}} & , x \geq 0 \\ 0 & , x < 0 \end{cases}$$

Statistical communication theory

Two finite-dimensional space: a parameter space and an observation space.



Bayesian Inference

- Probabilistic modeling: Formulate $f_{\mathbf{X}|\mathbf{\Theta}}(\mathbf{x}|\theta)$
  Characterize the future behavior of $\mathbf{x}$ conditioned on θ.
- Statistical analysis: Find $f_{\mathbf{\Theta}|\mathbf{X}}(\theta|\mathbf{x})$
  Make inference about θ given $\mathbf{x}$.

Bayesian Inference

By Bayes' theorem,

$$f_{\mathbf{\Theta}|\mathbf{X}}(\theta|\mathbf{x}) = \frac{f_{\mathbf{X}|\mathbf{\Theta}}(\mathbf{x}|\theta) f_{\mathbf{\Theta}}(\theta)}{f_{\mathbf{X}}(\mathbf{x})}$$

where $f_{\mathbf{X}}(\mathbf{x})$ is a marginal density.

$$f_{\mathbf{X}}(\mathbf{x}) = \int_{\mathbf{\Theta}} f_{\mathbf{X}|\mathbf{\Theta}}(\mathbf{x}|\theta) f_{\mathbf{\Theta}}(\theta) d\theta = \int_{\mathbf{\Theta}} f_{\mathbf{X}|\mathbf{\Theta}}(\mathbf{x}|\theta) f_{\mathbf{\Theta}}(\theta) d\theta$$

$$\text{posterior} = \frac{\text{observation density} \times \text{prior}}{\text{evidence}}$$

Likelihood

Likelihood function:

$$l(\theta|\mathbf{x}) = f_{\mathbf{X}|\mathbf{\Theta}}(\mathbf{x}|\theta)$$

The likelihood $l(\theta|\mathbf{x})$ is not a distribution; rather, it is a function of the parameter vector θ, given $\mathbf{x}$.

$$\text{posterior} = \frac{\text{likelihood} \times \text{prior}}{\text{evidence}}$$

Parameter Estimation

The maximum a posteriori (MAP) estimate of θ

$$\hat{\theta}_{\mathrm{MAP}} = \arg\max_{\theta} f_{\boldsymbol{\Theta}|\mathbf{X}}(\theta|\mathbf{x}) = \arg\max_{\theta} l(\theta|\mathbf{x}) f_{\boldsymbol{\Theta}}(\theta)$$

The maximum a posteriori estimate $\hat{\theta}_{\mathrm{MAP}}$ of the unknown parameter vector θ is the globally optimal solution to the parameter-estimation problem, in the sense that there is no other estimator that can do better.

Parameter Estimation

The maximum likelihood estimation

$$\hat{\theta}_{\mathrm{ML}} = \arg\sup_{\theta} l(\theta|\mathbf{x})$$

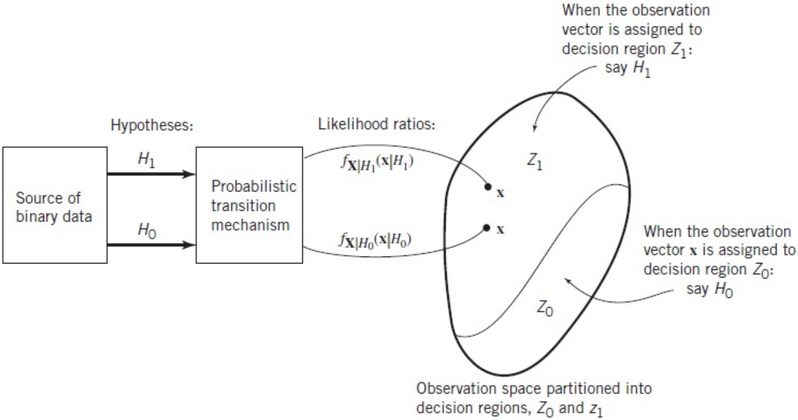
The maximum likelihood estimate ignores the prior.

Nevertheless, maximum likelihood estimation is widely used in the literature on statistical communication theory, largely because in ignoring the prior $f_{\mathbf{\Theta}}(\theta)$, it is less demanding than maximum posterior estimation in computational complexity.

Hypothesis Testing

Hypothesis testing is basic to signal detection in digital communications.

Binary Hypotheses:



Likelihood Receiver

Likelihood functions: $f_{\mathbf{X}|H_0}(\mathbf{x}|H_0)$, $f_{\mathbf{X}|H_1}(\mathbf{x}|H_1)$

Priors of hypotheses H_0 and H_1 are π_0 and π_1, respectively.

In processing observation vector $\mathbf{x}$, there are two kinds of errors that can be made by the decision rule:

1. Error of the first kind. This arises when hypothesis H_0 is true but the rule makes a decision in favor of H_1.
2. Error of the second kind. This arises when hypothesis H_1 is true but the rule makes a decision in favor of H_0.

Likelihood Receiver

The conditional probability of an error of the first kind:

$$\int_{Z_1} f_{\mathbf{X}|H_0}(\mathbf{x}|H_0)d\mathbf{x}$$

The conditional probability of an error of the second kind:

$$\int_{Z_0} f_{\mathbf{X}|H_1}(\mathbf{x}|H_1)d\mathbf{x}$$

The cost function in digital communications is the average probability of symbol error (Bayes risk):

$$\mathcal{R} = \pi_0 \int_{Z_1} f_{\mathbf{X}|H_0}(\mathbf{x}|H_0)d\mathbf{x} + \pi_1 \int_{Z_0} f_{\mathbf{X}|H_1}(\mathbf{x}|H_1)d\mathbf{x}$$

Likelihood Receiver

Because Z_0 and Z_1 are the complement of each other,

$$\mathcal{R} = \pi_0 + \int_{Z_0} \left[\pi_1 f_{\mathbf{X}|H_1}(\mathbf{x}|H_1) - \pi_0 f_{\mathbf{X}|H_0}(\mathbf{x}|H_0)\right] d\mathbf{x}$$

where π_0 is a fixed cost.

The optimum decision rule:

1. If $\pi_1 f_{\mathbf{X}|H_1}(\mathbf{x}|H_1) > \pi_0 f_{\mathbf{X}|H_0}(\mathbf{x}|H_0)$, the observation vector $\mathbf{x}$ should be assigned to Z_1.
2. If $\pi_1 f_{\mathbf{X}|H_1}(\mathbf{x}|H_1) < \pi_0 f_{\mathbf{X}|H_0}(\mathbf{x}|H_0)$, the observation vector $\mathbf{x}$ should be assigned to Z_0.

Likelihood Receiver

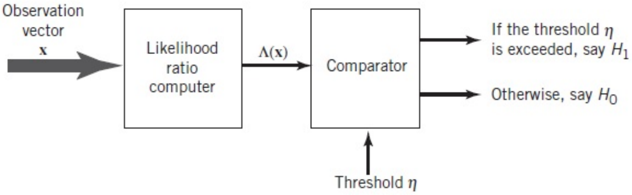
Likelihood ratio:

$$\Lambda(\mathbf{x}) = \frac{f_{\mathbf{X}|H_1}(\mathbf{x}|H_1)}{f_{\mathbf{X}|H_0}(\mathbf{x}|H_0)}$$

Threshold of the test:

$$\eta = \frac{\pi_0}{\pi_1}$$

Compare $\Lambda(\mathbf{x})$ and η — Likelihood ratio receiver



Log-Likelihood Receiver

Log-Likelihood ratio:

$$\ln \Lambda(\mathbf{x}) = \ln \frac{f_{\mathbf{X}|H_1}(\mathbf{x}|H_1)}{f_{\mathbf{X}|H_0}(\mathbf{x}|H_0)}$$

Threshold of the test:

$$\ln \eta = \ln \frac{\pi_0}{\pi_1}$$

Compare $\ln \Lambda(\mathbf{x})$ and $\ln \eta$ — Log-likelihood ratio receiver