ECE801
Monitoring and Estimation

Fundamentals

Instructor: Christos Panayiotou

Outline

- Hypothesis Testing
  - Maximum A Posteriori Probability (MAP) criterion
  - Bayes criterion
  - Neyman-Pearson (NP) criterion
- Estimator properties
- Estimation
  - Maximum A Posteriori Probability (MAP) criterion
  - Bayes criterion
  - Maximum Likelihood (ML)

Hypothesis Testing

Problem

- Assume a "system" with
- Take a measurement of the output y corrupted by noise

$$y = s_i + n, \quad i = 0,1$$

- Decide which was the **true** output of the system

Hypothesis

- Make two *hypothesis*: H_0 which corresponds to the event that s_0 is the correct output and H_1 which corresponds to s_1
- Define $\Pr[H_0|y]$ and $\Pr[H_1|y]$ and decide that the output was
  - s_0
  - s_1

Maximum A Posteriori Probability (MAP) Criterion

- Assume we know the *prior* probabilities $\pi_0 = \Pr[H_0]$ and $\pi_1 = \Pr[H_1]$
- We can use Bayes' Theorem
- Therefore, the decision rule
  - s_0 if $\Pr[H_0|y] > \Pr[H_1|y]$ or
  - s_1 if $\Pr[H_1|y] > \Pr[H_0|y]$
- Define *Decisions* D_0 and D_1 we can write

Likelihood Ratio

$$f(y \mid H_0)\pi_0 \underset{D_1}{\overset{D_0}{\gtrless}} f(y \mid H_1)\pi_1$$

- Rearrange terms to get

- Define the *Likelihood Ratio*

$$L(y) = \frac{f(y \mid H_1)}{f(y \mid H_0)}$$

- The MAP criterion

Example

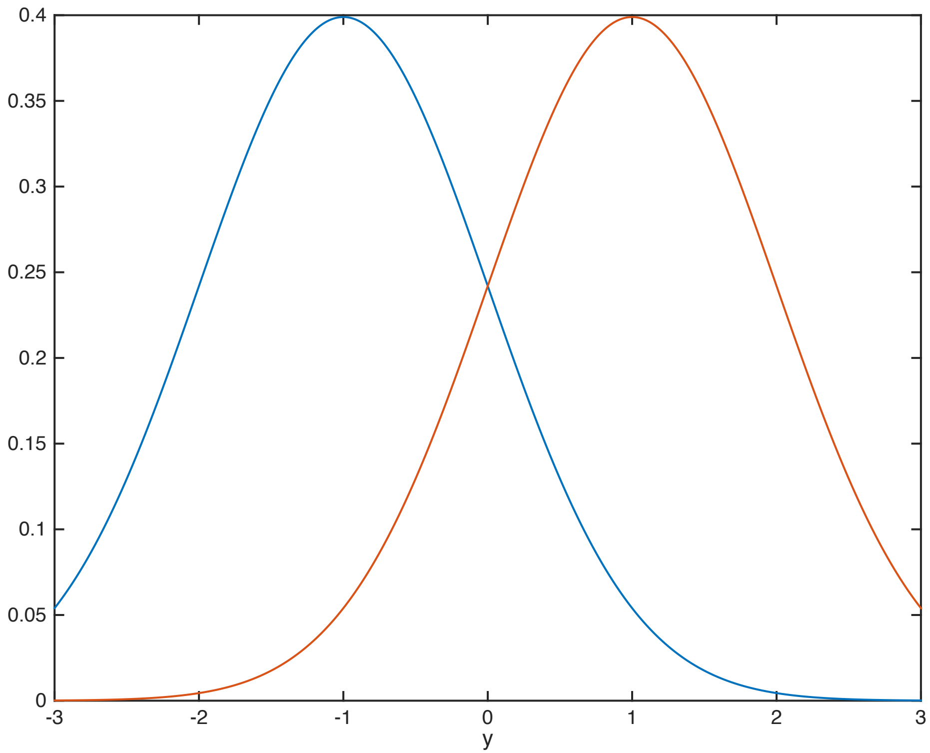
- Assume $s_0 = -a$ and $s_1 = a$.
- A priori probabilities $\pi_0 = 0.2$ and $\pi_1 = 0.8$
- Zero-mean Gaussian white noise

Solution

Example

Example

- If $\pi_0 = \pi_1 = 0.5$,
- If $\pi_0 = 0.2,\ \pi_1 = 0.8$,



Types of Errors

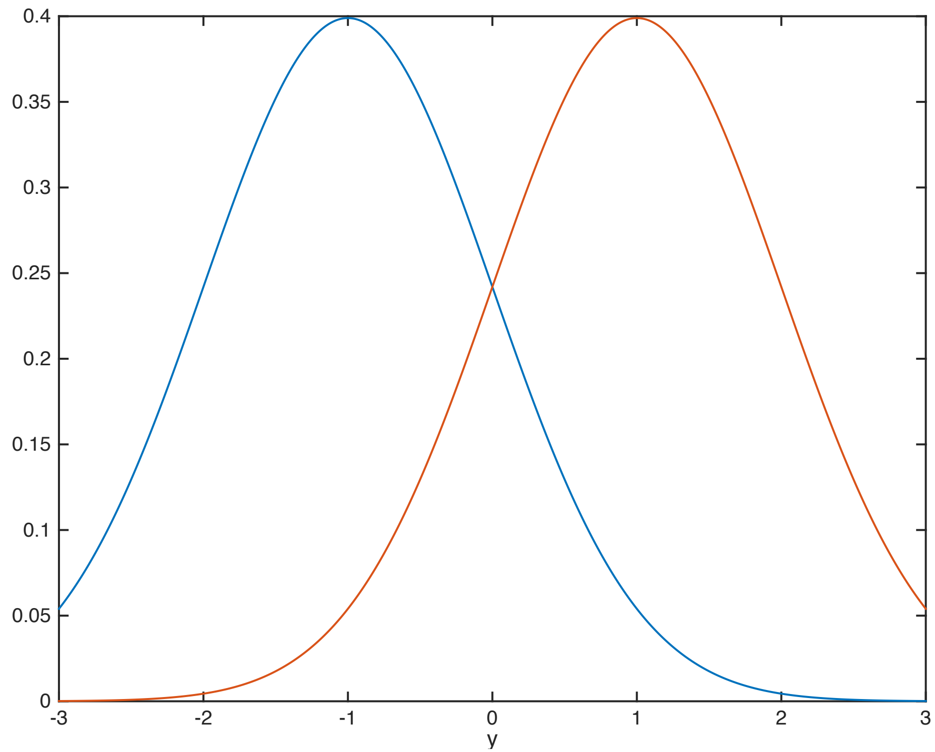
S\D	D_0	D_1
s_0		
s_1		

$$\Pr[s_0, D_0] = \int_{-\infty}^{\tau'} f(y \mid H_0)\, dy$$

$$\Pr[s_1, D_1] = \int_{\tau'}^{\infty} f(y \mid H_1)\, dy$$

$$\Pr[s_0, D_1] = \int_{\tau'}^{\infty} f(y \mid H_0)\, dy$$

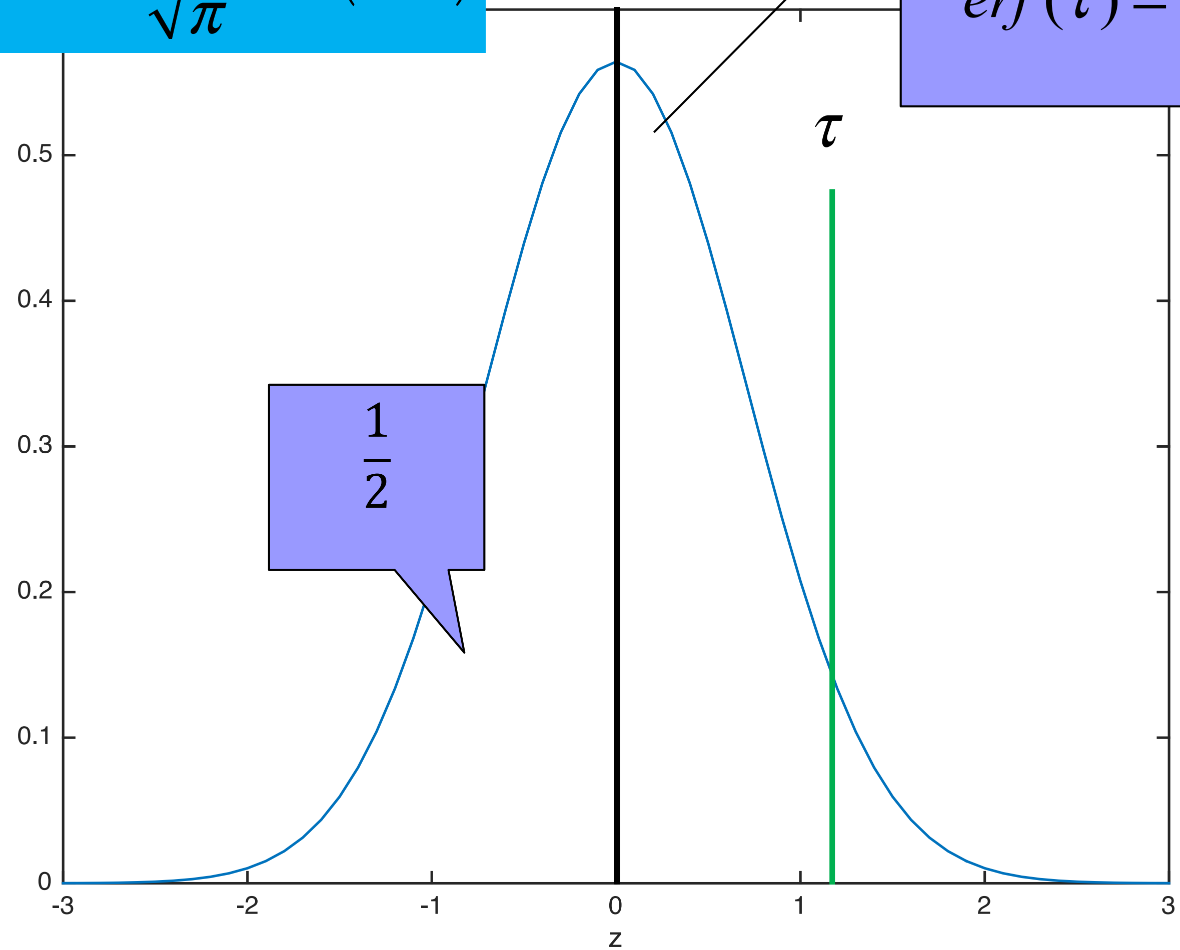
$$\Pr[s_1, D_0] = \int_{-\infty}^{\tau'} f(y \mid H_1)\, dy$$



Gaussian Integrals

$$f(z) = \frac{1}{\sqrt{\pi}} \exp\left(-z^2\right)$$

$$erf(\tau) = \frac{2}{\sqrt{\pi}} \int_0^{\tau} e^{-z^2}\, dz$$

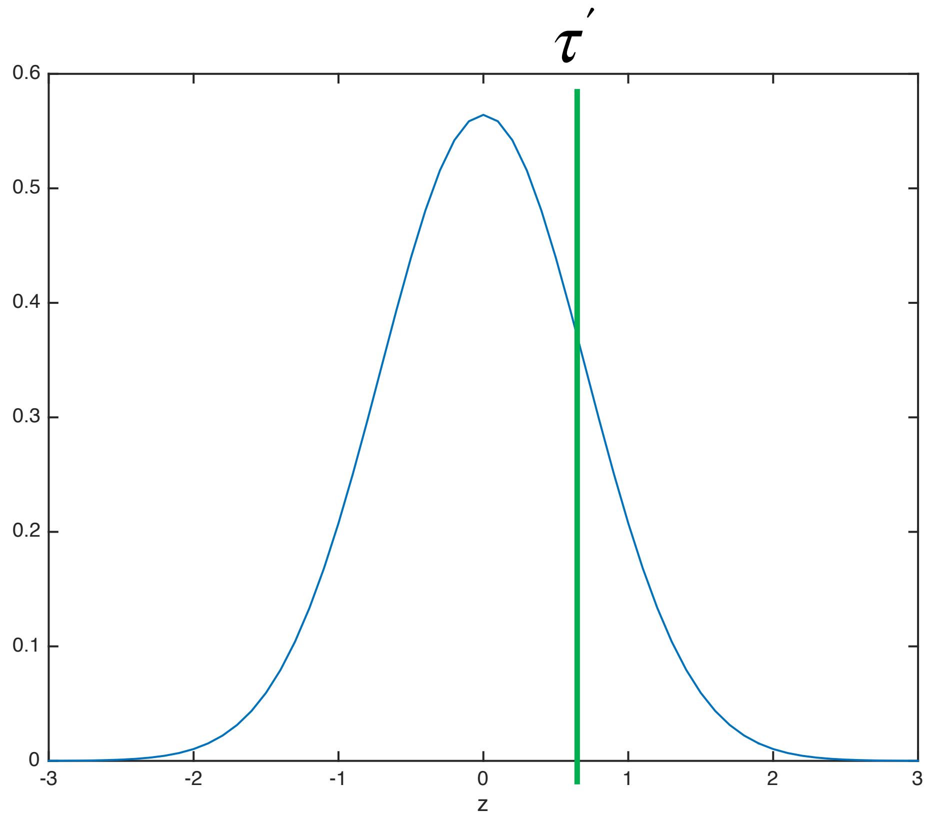


Gaussian Integrals

- Let $Y \sim N(\mu, \sigma^2)$ and define the random variable $Z = \frac{Y-\mu}{\sqrt{2}\sigma}$. Then $Z \sim N\left(0, \frac{1}{2}\right)$.

$$\Pr[s_0, D_1] = \int_{\tau'}^{\infty} f(y \mid H_0)\, dy$$

$$\Pr[s_1, D_0] = \int_{-\infty}^{\tau'} f(y \mid H_1)\, dy$$



Bayes Criterion

- Some errors may be more important than others!
- Assume we know the *cost* associated with every decision
- Assume we know the *prior* probabilities $\pi_0 = \Pr[H_0]$ and $\pi_1 = \Pr[H_1]$
- We can define the Bayes' risk (or cost)
- Or
-

Bayes Criterion

- After some computations, we arrive at the criterion that we decide D_0 if
- or
- Therefore the Bayes Decision is given by

Example

- Assume $s_0 = -a$ and $s_1 = a$.
- A priori probabilities $\pi_0 = 0.2$ and $\pi_1 = 0.8$
- Costs: $C_{00} = C_{11} = 0$ and $C_{01} = 1$, $C_{10} = 2$
- Zero-mean Gaussian white noise

Neyman-Pearson (NP) Criterion

- What if neither costs nor prior probabilities are known?
- **NP Criterion**: Keep the False Alarm probability below some level α_f
- and **maximize** the detection probability
- Constrained optimization problem:
- where we can obtain

Detection vs Estimation

- **Detection theory** involves the selection among a finite number of possible hypotheses
- **Estimation theory** involves the selection among a continuum of "hypotheses"
  - As the number of hypotheses in detection theory grows larger, the distinction between detection and estimation becomes blurred.
-

Estimator Properties

- Suppose that we want to **estimate** the value of a parameter α using the observations $y_1, \ldots, y_n$ using an estimator $\hat{\alpha}_n$ which is a function of the observations. Then, it may be desirable that the estimator (which is a random variable) may have the following properties
- **Unbiased**
- **Consistent**
- **Invariant** under transformation. Let the function $g(\alpha)$, then

Estimator Properties

- **Sufficient:** Intuitively, this property states that the estimator utilizes all available information.
- **Minimum Variance**:
  - The smaller the variance, the better the quality of the estimator.
  - **Cramer-Rao lower bound**

where $F(\alpha)$ is the Fisher Information

$$F(\alpha) = -E\left[\frac{\partial^2}{\partial\alpha^2}\ln f(y_1,\cdots,y_n;\alpha)\right]$$

Estimator Properties

- **Efficient** estimators, let two unbiased estimators $\hat{\alpha}_n^0$ and $\hat{\alpha}_n^1$ with $\hat{\alpha}_n^0$ being the one with the lowest variance. Then efficiency is defined as

- **Asymptotically Efficient**

- **Asymptotically Normal**
  - $\hat{\alpha}_n$ approaches a normal distribution as n goes to infinity

Maximum A Posteriori (MAP) Estimation

- We want to **estimate** the value of a parameter α using the observations $y_1, \dots, y_n$ and the a priori distribution $f(\alpha)$.
- **MAP Estimator:** *Maximize* the pdf $f(\alpha|\mathbf{y})$ where $\mathbf{y} = [y_1, \dots, y_n]$.
- Using Bayes' rule
- Thus

Example

- Assume that
  - the observations $y_1, \ldots, y_n$ are i.i.d. taken from a Gaussian distribution with an *unknown mean* μ and known variance σ^2, $y_i \sim N(\mu, \sigma^2)$, $i = 1, \ldots, n$.
  - The mean μ is also a random variable $\mu \sim N(m_1, \beta^2)$
- **MAP Estimator:**

Example

- Set the derivative equal to 0

$$\frac{d}{d\mu}\left\{\frac{1}{\left(2\pi\sigma^2\right)^{n/2}}\exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\left(y_i-\mu\right)^2-\frac{1}{2\beta^2}(\mu-m_1)^2\right)\right\}=0$$

Example

- Estimator properties
  - If μ is held constant, then
  - But the expected value of the sample mean $E[\bar{y}|\mu] = \mu$
  - Therefore the estimator is ***biased***, but ***asymptotically unbiased***

Bayes' Estimator

- We want to **estimate** the value of a parameter α using
  - the observations $\mathbf{y} = [y_1, \ldots, y_n]$
  - the a priori distribution $f(\alpha)$.
  - the Bayes' cost (loss) which is a function of the error $a_e = \alpha - \hat{\alpha}$
- **Bayes' Estimator:**
- Various cost functions

$$C(\hat{\alpha},\alpha)=\begin{cases} 0 & \text{if } |a_e| \leq \Delta/2 \\ 1 & \text{if } |a_e| > \Delta/2 \end{cases}$$

Bayes' Estimator

- Mean Square Error (MSE)
  - $C(\hat{\alpha}, \alpha) = \alpha_e^2 = (\alpha - \hat{\alpha})^2$

$$E\left[C\left(\hat{\alpha},\alpha\right)\right] = \int_{-\infty}^{\infty}\left(\alpha-\hat{\alpha}\right)^2 f(\alpha\,|\,\mathbf{y})\,d\alpha$$

- Differentiate with respect to $\hat{\alpha}$

Bayes' Estimator

$$\hat{\alpha}_{MSE} = \int_{-\infty}^{\infty} \alpha f(\alpha \,|\, \mathbf{y})\, d\alpha = E\left[\alpha \,|\, \mathbf{y}\right]$$

- Using Bayes' Rule

$$f(\alpha \,|\, \mathbf{y}) = \frac{f(\mathbf{y} \,|\, \alpha) f(\alpha)}{f(\mathbf{y})} = \frac{f(\mathbf{y} \,|\, \alpha) f(\alpha)}{\int_{-\infty}^{\infty} f(\mathbf{y} \,|\, \alpha) f(\alpha)\, d\alpha}$$

- Which results to

Example

- Assume that
  - the observations $y_1, \ldots, y_n$ are i.i.d. taken from a Gaussian distribution with an *unknown mean* μ and known variance σ^2, $y_i \sim N(\mu, \sigma^2)$, $i = 1, \ldots, n$.
  - The mean μ is also a random variable $\mu \sim N(m_1, \beta^2)$
- **MSE Estimator:**

$$\hat{\mu}_{MSE} = \int_{-\infty}^{\infty} \mu f(\mu \mid \mathbf{y}) d\mu = E[\mu \mid \mathbf{y}]$$

with

$$f(\mathbf{y} \mid \mu) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (y_i - \mu)^2 \right)$$

Example

- Using Bayes' rule again we obtain

$$f\left(\mu \mid \mathbf{y}\right) = \frac{1}{\sqrt{2\pi\gamma^2}} \exp\left(-\frac{(\mu - \gamma^2\omega)^2}{2\gamma^2} \right)$$

Where

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$$

so

$$\Rightarrow \hat{\mu}_{MSE} = \frac{\beta^2 \bar{y} + \sigma^2 m_1 / n}{\beta^2 + \sigma^2 / n}$$

Maximum Likelihood (ML) Estimator

- We want to **estimate** the value of a parameter α using
  - the observations $\mathbf{y} = [y_1, \ldots, y_n]$
  - **NO** a priori distribution $f(\alpha)$ and **NO** cost function are available.
- **ML Estimator:** Maximize the likelihood distribution
- Assuming independent observations each with pmf $f(y_i|\alpha)$

Relation between ML and MAP Estimator

- Again use Bayes' rule and taking logarithms

$$\ln f(\alpha \mid \mathbf{y}) = \ln f(\mathbf{y} \mid \alpha) + \ln f(\alpha) - \ln f(\mathbf{y})$$

- For the minimization, take derivatives with respect to α

Example

- Assume that
  - the observations $y_1, \dots, y_n$ are i.i.d. taken from a Gaussian distribution with an *unknown mean* μ and known variance σ^2, $y_i \sim N(\mu, \sigma^2)$, $i = 1, \dots, n$.
  - The mean μ is also a random variable $\mu \sim N(m_1, \beta^2)$
- **ML Estimator:**

$$f(\mathbf{y} \mid \mu) = \frac{1}{\left(2\pi\sigma^2\right)^{n/2}} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^{n} \left(y_i - \mu\right)^2 \right)$$

Set the derivative with respect to μ equal to 0.