TSKS04 Digital Communication
Continuation Course

Lecture 3

Introduction to Estimation

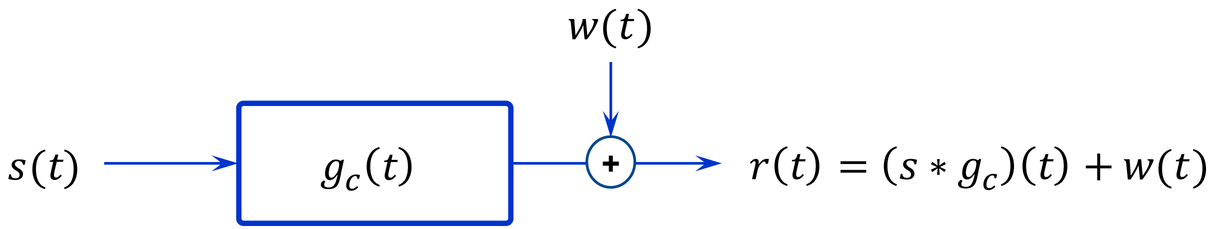
Emil Björnson

Department of Electrical Engineering (ISY)
Division of Communication Systems

COMMUNICATION SYSTEMS

LINKÖPING UNIVERSITY

Classification of Communication Channels

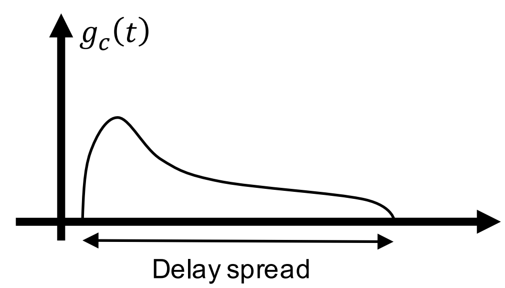


Classification of channel filter:

- Non-dispersive: $g_c(t) = \delta(t)$
- Time-dispersive: Spreads signal in time

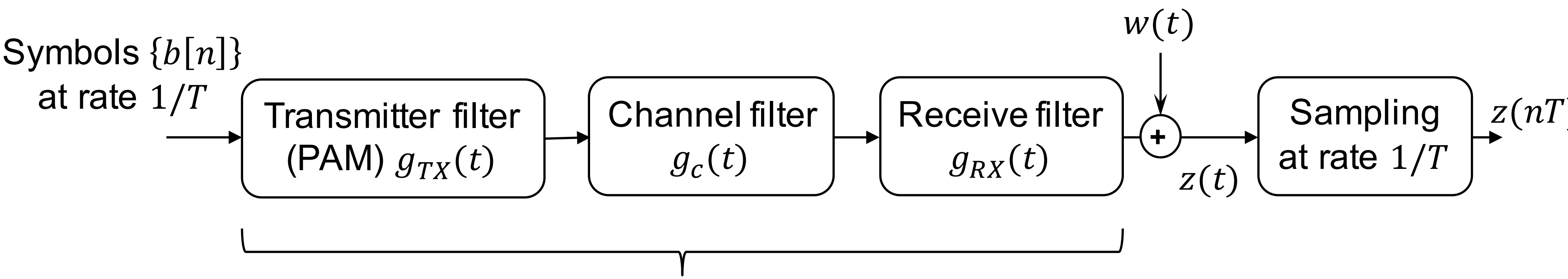
Delay spread:

Time interval where $g_c(t) \neq 0$



Approximately non-dispersive: Delay spread small compared to symbol time

Nyquist Criterion at Receiver (1/2)



Effective pulse shape:
$p(t) = (g_{TX} * g_c * g_{RX})(t)$

Signal model:

- $z(t) = \sum_k b[k]p(t - kT) + w(t)$

ISI-free communication if

- $z(nT) = \sum_k b[k]p(nT - kT) + w(nT) = p(0)b[n] + w(nT)$

Achieved when $p(nT) = K\delta[n] \quad \Leftrightarrow \quad \frac{1}{T}\sum_{m=-\infty}^{\infty} P\left(f - \frac{m}{T}\right) = K$

Nyquist ISI criterion

Nyquist Criterion at Receiver (2/2)

- Can we get rid of ISI?
  - Select $g_{TX}(t)$ and $g_{RX}(t)$ based on $g_c(t)$ to make:

$$p(nT) = (g_{TX} * g_c * g_{RX})(nT) = K\delta[n]$$

  - Needed: Knowledge of $g_c(t)$, tweaking $g_{TX}(t), g_{RX}(t)$ if channel changes

More realistic:

- Make T large compared to delay spread
- Detect multiple symbols jointly

- Discrete-time model with ISI:

$$z(nT) = \sum_k b[k] \underbrace{p(nT - kT)} + w(nT)$$

We want to learn this!

How to estimate them?

Estimation: Learning Unknown Parameters

Basic notation:

- Unknown parameter: θ
- Observation: $\boldsymbol{y} = (y[1], \ldots y[K])^T$ (realization of stochastic variable $\boldsymbol{Y}$)
- Relation: $p(\boldsymbol{y}|\theta)$ (conditional PDF)

Other name: Likelihood function

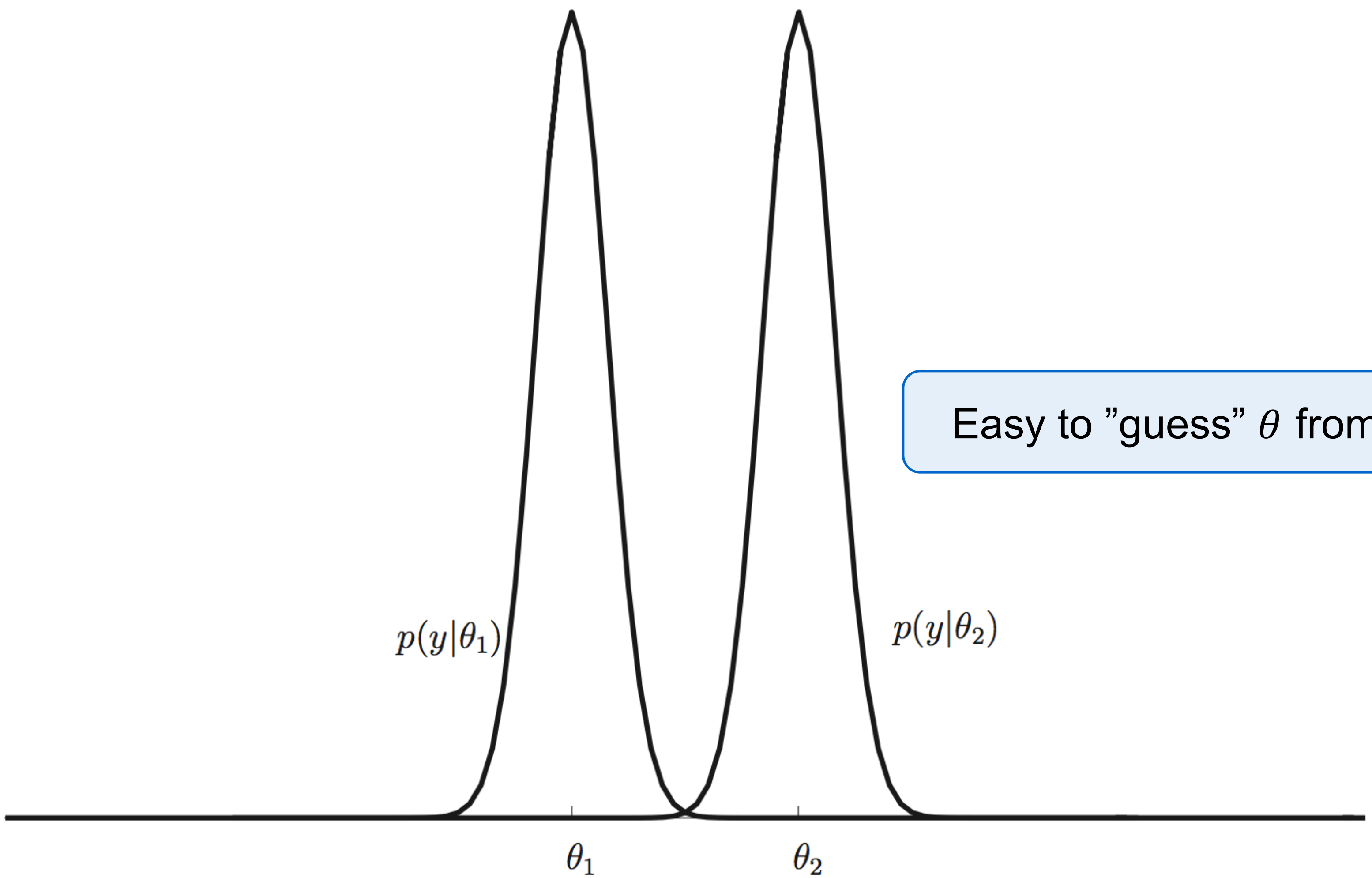
Goal: Find an estimation function $\hat{\theta}(\boldsymbol{y})$

Two paradigms

- Classical estimation: θ is deterministic
- Bayesian estimation: θ is realization of a stochastic variable Θ

Main topic today

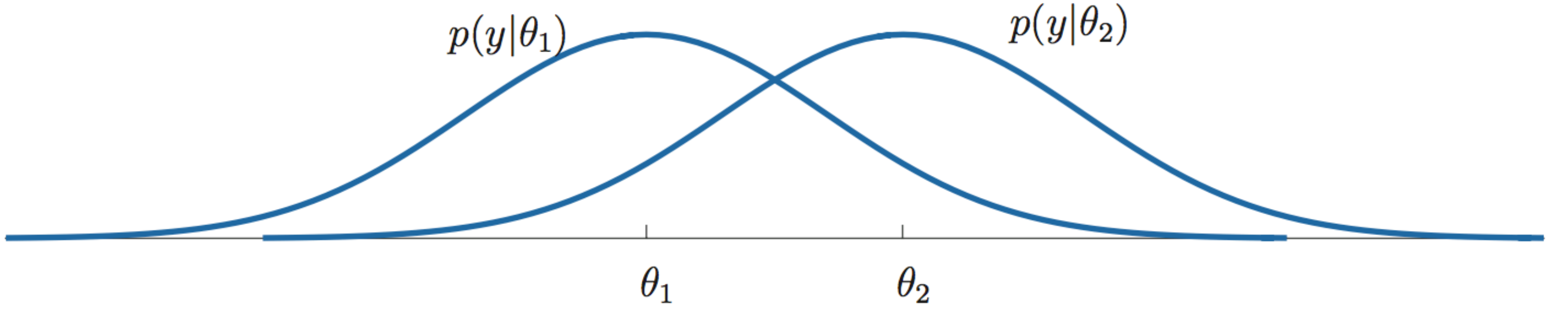
Observation Depends Strongly on θ



Easy to ”guess” θ from y

Observation Depends Weakly on θ

Hard to "guess" θ from y



Properties of a Good Estimator

What properties should $\hat{\theta}(\boldsymbol{y})$ have?

1. Unbiased:

$$E\{\hat{\theta}(\boldsymbol{y})\} = \theta$$

2. Low estimation error variance:

$$E\left\{\Big(\underbrace{\hat{\theta}(\boldsymbol{y}) - \theta}_{\text{Estimation error}}\Big)^2\right\}$$

Best choice: *Minimum variance unbiased (MVU)* estimator

Bounding the Estimation Error Variance (1/2)

Theorem 1 (Cramer-Rao Lower Bound (CRLB)) *If the regularity condition* (B.3) *holds, then the estimation error variance of any unbiased estimator* $\hat{\theta}(\mathbf{y})$ *must satisfy*

$$\mathrm{E}\left\{\left(\hat{\theta}(\mathbf{y})-\theta\right)^2\right\} \geq \frac{1}{-\mathrm{E}\left\{\frac{\partial^2 \ln p(\mathbf{y}|\theta)}{\partial \theta^2}\right\}} \tag{B.6}$$

where the expectations are taken with respect to $p(\mathbf{y}|\theta)$.

Regularity condition:

$$\mathrm{E}\left\{\frac{\partial \ln p(\mathbf{y}|\theta)}{\partial \theta}\right\} = 0, \tag{B.3}$$

Corresponds to: Order of derivative and integral can be interchanged

Bounding the Estimation Error Variance (2/2)

Theorem 1 (Cramer-Rao Lower Bound (CRLB)) *If the regularity condition* (B.3) *holds, then the estimation error variance of any unbiased estimator* $\hat{\theta}(\mathbf{y})$ *must satisfy*

$$\mathrm{E}\left\{\left(\hat{\theta}(\mathbf{y})-\theta\right)^2\right\} \geq \frac{1}{-\mathrm{E}\left\{\frac{\partial^2 \ln p(\mathbf{y}|\theta)}{\partial \theta^2}\right\}} \qquad \text{(B.6)}$$

where the expectations are taken with respect to $p(\mathbf{y}|\theta)$.

An unbiased estimator that attains the CRLB can be found, for all θ, *if and only if the first derivative of* $\ln p(\mathbf{y}|\theta)$ *can be decomposed as*

$$\frac{\partial \ln p(\mathbf{y}|\theta)}{\partial \theta} = I(\theta)\left(g(\mathbf{y})-\theta\right) \qquad \text{(B.7)}$$

for $I(\theta) = -\mathrm{E}\left\{\frac{\partial^2 \ln p(\mathbf{y}|\theta)}{\partial \theta^2}\right\}$ *and some function* $g(\mathbf{y})$. *Then the MVU estimator is* $\hat{\theta}(\mathbf{y}) = g(\mathbf{y})$ *and has the estimation error variance* $1/I(\theta)$.

Fisher Information $I(\theta)$

Another way to write the CRLB:

$$E\left\{\left(\hat{\theta}(\boldsymbol{y}) - \theta\right)^2\right\} \geq \frac{1}{I(\theta)}$$

- Inversely to the Fisher information

$$I(\theta) = -E\left\{\frac{\partial^2 \ln p(\boldsymbol{y}|\theta)}{\partial \theta^2}\right\}$$

- Describes how much information $\boldsymbol{y}$ provides about θ
- Second derivative: Sharpness in the curvature of the function

The CRLB cannot always be attained with equality

Other Estimators (1/2)

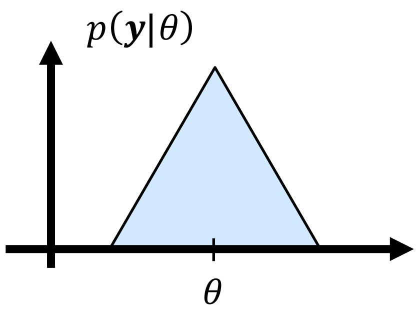
Maximum Likelihood (ML) estimator

- Find value of θ that maximizes likelihood function $p(\mathbf{y}|\theta)$

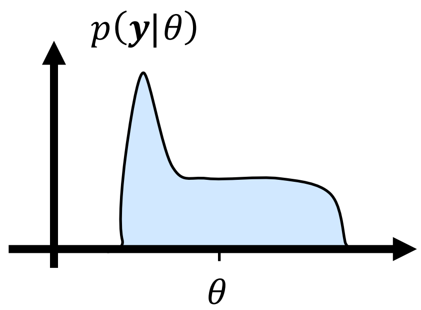
- Mathematically: $\hat{\theta}(\mathbf{y}) = \arg\max_{\theta} \, p(\mathbf{y}|\theta) = \arg\max_{\theta} \, \ln p(\mathbf{y}|\theta)$

- Found by solving: $\frac{\partial}{\partial\theta} \ln p(\mathbf{y}|\theta) = 0$

Can use any increasing function



Suitable for ML estimation



Unsuitable for ML estimation

ML estimators attain CRLB asymptotically as $K \to \infty$

Other Estimators (2/2)

Suppose: θ is realization of a stochastic variable Θ

- Known prior distribution: $p(\theta)$
- Utilized to obtain:

$$p(\theta|\boldsymbol{y}) = \frac{p(\boldsymbol{y}[\theta)p(\theta)}{p(\boldsymbol{y})}$$

Which θ is most likely to have caused $\boldsymbol{y}$?

Maximum A Posteriori (MAP) estimator

- Choose θ most likely to give $\boldsymbol{y}$
- Mathematically: $\hat{\theta}(\mathbf{y}) = \arg\max_{\theta} p(\theta|\mathbf{y}) = \arg\max_{\theta} p(\mathbf{y}|\theta)p(\theta)$

This is a Bayesian estimator

LINKÖPING UNIVERSITY

www.liu.se