- Classical estimation and detection
  - Sufficient statistics, factorization theorem,
  - Fisher information, Cramér-Rao bound, efficient estimators, information inequality theorem,
  - Maximum likelihood (ML) estimators, ML invariance principle, asymptotic properties of ML estimators.
  - Neyman-Pearson, uniformly most powerful (UMP), and generalized likelihood ratio (GLR) tests.

- Bayesian estimation and detection
  - Conjugate priors,
  - Bayes' rule: computing
    - posterior and
    - marginal distributions (i.e. marginalization using the proportionality trick),
  - Bayesian MMSE and MAP estimation,
  - LMMSE estimation, orthogonality principle,
  - Bayesian detection (binary and multihypothesis).

Notation: $i_A(x)$ denotes the indicator function:

$$i_A(x) = \begin{cases} 1, & x \in A, \\ 0, & \text{otherwise} \end{cases} .$$

A Few Examples

Example 1. Our measurements $X[0]$ and $X[1]$ are conditionally independent given the parameter $\Theta = \theta$, following

$$\begin{aligned} f_{X[0] \,|\, \Theta}(x[0] \,|\, \theta) &= \text{Poisson}(x[0] \,|\, \theta) \\ f_{X[1] \,|\, \Theta}(x[1] \,|\, \theta) &= \text{Poisson}(x[1] \,|\, \theta^2). \end{aligned}$$

(a) Find the minimal sufficient statistic for θ based on $x[0]$ and $x[1]$.

(b) Does an efficient estimator exist for θ? Does an efficient estimator exist for an invertible function of θ?

Example 2. Suppose that $\theta > 0$ is a parameter of interest and that the observations $X[0], X[1], \ldots, X[N-1]$ are conditionally independent, identically distributed (i.i.d.) given θ, with the conditional distribution $X[n] \,|\, \theta$ described by the following cumulative distribution function (cdf):

$$\Pr_{X[n]\,|\,\Theta}\{X[n] \le x \,|\, \theta\} = [F(x)]^{1/\theta}, \quad x \in (-\infty, +\infty)$$

for $n = 0, 1, \ldots, N-1$. Here, $F(z)$ is a known cdf of a continuous random variable Z with pdf $f(z) = dF(z)/dz$.

(a) Compute the likelihood function $f_{\boldsymbol{X}\,|\,\Theta}(\boldsymbol{x} \,|\, \theta)$ where

$$\boldsymbol{x} = [x[1], x[2], \ldots, x[N-1]]^T.$$

Is $f_{\boldsymbol{X}\,|\,\Theta}(\boldsymbol{x} \,|\, \theta)$ a member of the exponential family of distributions?

(b) Show that

$$\widehat{\theta}_{\text{eff}} = -\frac{1}{N} \sum_{n=0}^{N-1} \ln F(x[n])$$

is an *efficient* estimator of θ.

(c) Suppose now that we assign a scaled inverted χ^2 prior pdf for Θ:

$$\pi(\theta) = \text{Inv-}\chi^2(\theta \,|\, \nu_0, \theta_0) \propto \theta^{-(\nu_0/2+1)} \cdot e^{-\nu_0\theta_0/(2\theta)} \cdot i_{[0,+\infty)}(\theta)$$

where $\theta_0 > 0$ and $\nu_0 > 0$ are known prior scale and degrees of freedom, see the distribution table handout. Find the posterior pdf

$$f_{\Theta \,|\, \boldsymbol{X}}(\theta \,|\, \boldsymbol{x}).$$

and compute the Bayesian minimum mean-square error (MMSE) and maximum *a posteriori* (MAP) estimates of θ, respectively.

Example 3. Camera measurement. The measurement X from a camera can be expressed as

$$X = A\,\Theta + W$$

where

- Θ is the object position with mean μ_Θ and variance τ_Θ^2,
- A is the *occlusion indicator function*, equal to 1 (if the camera can see the object) with probability p, and 0 (if the camera cannot see the object) with probability $(1-p)$, and
- W is noise with mean 0 and variance σ_W^2.

Assume that Θ, A, and W are

- real-valued and
- mutually independent and, therefore, uncorrelated.

Find the linear minimum mean-square error (LMMSE) estimate of Θ based on the camera measurement X. Your answer should be in terms of $\mu_\Theta, \tau_\Theta^2, \sigma_W^2$, and p.

Example 4. We observe N time samples $X[n]$, modeled as

$$X[n] = \sum_{i=1}^{p} A_i \, g_i(n - \tau_i) + w[n], \quad n = 0, 1, \ldots, N-1,$$

where

- $W[n]$ is zero-mean white Gaussian noise with known variance σ_W^2,
- A_i, $i = 1, 2, \ldots, p$ are p i.i.d. zero-mean Gaussian random variables with known variance σ_A^2, and
- $g_i(u)$, $i = 1, 2, \ldots, p$ are known time functions over $u \in (-\infty, +\infty)$.

The A_i and $W[n]$ are independent and p is known. Define the parameter vector $\boldsymbol{\tau} = [\tau_1, \tau_2, \ldots, \tau_p]^T$ and $K(\boldsymbol{\tau})$ as the $p \times p$ matrix of inner products of the g_i's, i.e. K has entries $K_{i,k}(\boldsymbol{\tau}) = \sum_{n=0}^{N-1} g_i(n - \tau_i) g_k(n - \tau_k)$.

Show that the ML estimator of the τ_i's involves maximizing a quadratic form

$$[\boldsymbol{y}(\boldsymbol{\tau})]^T \, [I + \rho \, K(\boldsymbol{\tau})]^{-1} \, \boldsymbol{y}(\boldsymbol{\tau}) - b(\boldsymbol{\tau})$$

where $\boldsymbol{y}(\boldsymbol{\tau}) = [y_1(\tau_1), y_2(\tau_2), \ldots, y_p(\tau_p)]^T$ is a vector of p correlator outputs

$$y_i(\tau_i) = \sum_{n=0}^{N-1} x[n]\, g_i(n - \tau_i), \quad i = 1, 2, \ldots, p$$

$b(\boldsymbol{\tau})$ is an observation-independent bias term, and $\rho = \sigma_A^2/\sigma_W^2$ is the SNR.

Hints: Express the log-likelihood function in the vector-matrix form and use the matrix inversion lemma. These idenitites may be useful:

$$|A\,B| = |B\,A|, |I + A\,B| = |I + B\,A|$$

and matrix inversion lemma:

$$(I + S\,U)^{-1} = I - S\,(I + U\,S)^{-1}\,U.$$

Example 5. Consider the signal $\Theta \in \{-\sqrt{2}, \sqrt{2}\}$ with prior probability mass function (pmf)

$$p_\Theta(-\sqrt{2}) = p_\Theta(\sqrt{2}) = \tfrac{1}{2}.$$

We send Θ over a cascade of two additive Gaussian noise channels, modeled as

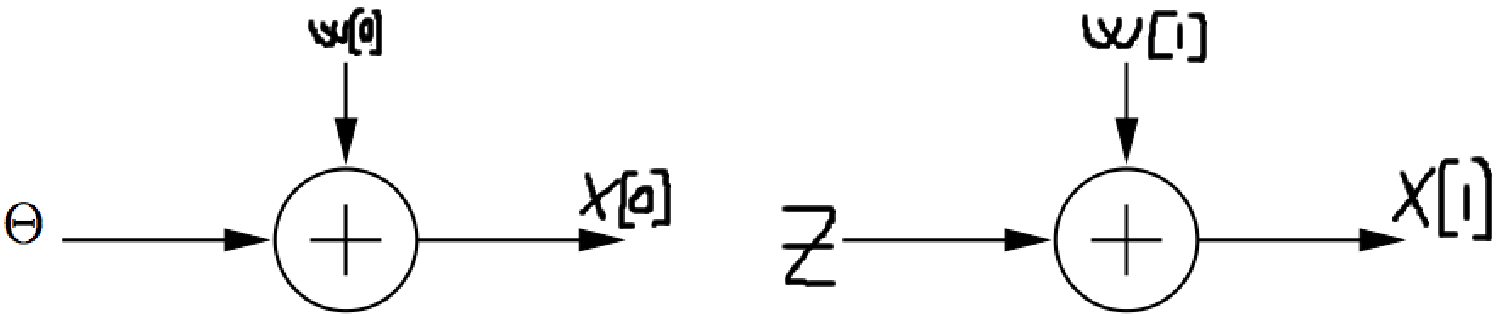
$$X[0] = \Theta + W[0]$$

and

$$X[1] = Z + W[1]$$

where $\Theta, W[0]$, and $W[1]$ are independent and

$$W[0] \sim \mathcal{N}(0,1), \quad W[1] \sim \mathcal{N}(0,2).$$



A communication engineer wishes to decide between the following two relaying schemes:

Decode-and-forward. Upon receiving $X[0]$, find the MAP estimate of Θ, $\widehat{\theta}_{\mathrm{MAP,DF},1}(X[0])$ and then send

$$Z = \widehat{\theta}_{\mathrm{MAP,DF},1}(X[0])$$

over the second channel. The MAP estimate $\widehat{\theta}_{\mathrm{MAP,DF},1}(X[0]) \in \{-\sqrt{2}, \sqrt{2}\}$ can be viewed as a detector (decoder) corresponding to the Bayes' rule for 0-1 loss. Upon receiving $Y[1]$, apply the Bayes' rule for 0-1 loss to detect Z and use the obtained detection result as a detector for Θ.

Amplify-and-forward. Upon receiving $X[0]$, send

$$Z = \sqrt{\frac{2}{3}}\, X[0]$$

over the second channel. Here, the multiplicative term $\sqrt{\frac{2}{3}}$ is used to normalize the power of Z to 2, which is equal to the power of Z in the decode-and-forward case. Upon receiving $X[1]$, apply the Bayes' rule for 0-1 loss to detect Θ.

Which scheme do you recommend?

Example 6. We model the real-valued observations $X[0], X[1], \ldots, X[N-1]$ as

$$X[n] = \theta\, S[n] + W[n], \quad n = 1, 2, \ldots, N$$

where

- $S[n],\ n = 0, 1, \ldots, N-1$ are i.i.d. Bernoulli random variables taking values ± 1 with equal probability, i.e.

$$\Pr\{S[n] = 1\} = \Pr\{S[n] = -1\} = \tfrac{1}{2}$$

  and

- $W[n]$ is zero-mean additive white Gaussian noise with unit variance, i.e. $\mathrm{E}_{\boldsymbol{W}}[\boldsymbol{W}] = \boldsymbol{0}$, $\mathrm{E}_{\boldsymbol{W}}[\boldsymbol{W}\,\boldsymbol{W}^T] = I$, where $\boldsymbol{w} = \begin{bmatrix} W[0] \\ W[1] \\ \vdots \\ W[N-1] \end{bmatrix}$ and I denotes the identity matrix.

(a) Derive the Bayes' decision rule for testing

$$\begin{aligned} \mathcal{H}_0 &: \quad \Theta = 0 \quad \text{versus} \\ \mathcal{H}_1 &: \quad \Theta = \theta_1 \end{aligned}$$

where θ_1 is a known constant. Assume piecewise-constant loss functions and set the losses $\mathrm{L}(1\,|\,1)$ and $\mathrm{L}(0\,|\,0)$ due to correct decisions to zero (i.e. $\mathrm{L}(1\,|\,1) = \mathrm{L}(0\,|\,0) = 0$) and describe the prior probability mass function (pmf) for the hypotheses using

$$\pi_0 = \Pr\{\Theta = 0\}, \quad \pi_1 = \Pr\{\Theta = \theta_1\} = 1 - \pi_0.$$

(b) Specialize the detector proposed in (a) to the case of 0-1 loss $[\mathrm{L}(1\,|\,1) = \mathrm{L}(0\,|\,0) = 0$ and $\mathrm{L}(0\,|\,1) = \mathrm{L}(1\,|\,0) = 1]$ and *a priori* equiprobable hypotheses ($\Pr\{\Theta = 0\} = \Pr\{\Theta = \theta_1\} = \frac{1}{2}$).

Hint: To express your solution in a compact manner, note that the hyperbolic cosine is defined as

$$\cosh x = \tfrac{1}{2}\,(e^x + e^{-x}).$$