4F6 Signal Detection + Estimation 2013

(Solutions)

Q1) The Neyman-Fisher factorization theorem is useful in estimation theory since it helps decide what the "sufficient statistics" are for estimating parameters of a model for data.

The theorem states that if the likelihood $p(x|\theta)$ can be written as

$$p(x|\theta) = g(T(x), \theta)\, h(x)$$

where x is data and θ the parameter of the model, then $T(x)$ is the sufficient statistic for θ.

The theorem can be proved in two ways.

1) For an unbiased, efficient estimator we can write:

1 cont)

$$\frac{\partial}{\partial\theta} \ln p(x|\theta) = k(\theta)\left(\hat{\theta}(x) - \theta\right)$$

$$\therefore \ln p(x|\theta) = \int k(\theta')\left(\hat{\theta}(x) - \theta'\right) d\theta' + \ln h(x).$$

$$\therefore \underline{p(x|\theta) = h(x)\; g\left(T(x), \theta\right)}.$$ [20%]

The second way uses Bayes theorem.

i) assume $p(x|\theta) = g(T(x), \theta)\; h(x)$

$$\therefore \quad p(\theta|x) = \frac{p(\theta)\; p(x|\theta)}{p(x)}$$

$$= \frac{g(T(x),\theta)\; \cancel{h(x)}\; p(\theta)}{\cancel{h(x)} \int p(\theta)\, g(T(x),\theta)\, d\theta}.$$

$$= p(\theta | T(x))$$

ie $T(x)$ is the only way the data enters into the estimation and is $\therefore$ sufficient.

QED

1 cont) The likelihood for the data $\underline{x}$ is

$$p(\underline{x}|\theta) = \prod_{n=0}^{N-1} \exp\left(A(\theta)\, B(x_n) + C(x_n) + D(\theta)\right)$$

$$= \exp\left\{ A(\theta) \sum B(x_n) + \sum C(x_n) + N\, D(\theta) \right\}.$$

$$= \exp\left\{ A(\theta) \sum B(x_n) + N\, D(\theta) \right\} \exp\left\{ \sum C(x_n) \right\}.$$

$\therefore$ using the N-F theorem, the sufficient statistic $T(\underline{x})$ is

$$T(\underline{x}) = \sum_{n=0}^{N-1} B(x_n).$$

[40%]

---

a) Gaussian case (+ likewise for the exponential)

$$p(x|\mu) = \frac{1}{\sqrt{2\pi}} \exp\left(-\tfrac{1}{2}(x-\mu)^2\right) \quad \text{for } \sigma = 1$$

$$= \exp\left\{ x\mu - \tfrac{1}{2}x^2 + \left(-\tfrac{1}{2}\mu^2 + \ln \tfrac{1}{\sqrt{2\pi}}\right) \right\}$$

$$\therefore \; T(\underline{x}) = \sum_{n=0}^{N-1} x_n$$

[40%]

Q2/ first part is book work. 35%.

—//—

parameterizing A, and B as $\theta = \begin{bmatrix} A \\ B \end{bmatrix}$,

the 2×2 Fisher information is

$$I(\theta) = \begin{bmatrix} -E\dfrac{\partial^2 \ln p(d|\theta)}{\partial A^2} & -E\dfrac{\partial^2 \ln p(d|\theta)}{\partial A \partial B} \\ -E\dfrac{\partial^2 \ln p(d|\theta)}{\partial B \partial A} & -E\dfrac{\partial^2 \ln p(d|\theta)}{\partial B^2} \end{bmatrix}.$$

and $p(d|\theta) = \dfrac{1}{(2\pi\sigma^2)^{N/2}} \exp\left(-\dfrac{1}{2\sigma^2}\sum_{n=0}^{N-1}(d(n) - A - Bn)^2\right).$

$$\frac{\partial \ln p(d|\theta)}{\partial A} = \frac{1}{\sigma^2}\sum_{n=0}^{N-1}(d(n) - A - Bn) \; ; \; \frac{\partial^2 \ln p(d|\theta)}{\partial A^2} = -\frac{N}{\sigma^2}$$

$$\frac{\partial \ln p(d|\theta)}{\partial B} = \frac{1}{\sigma^2}\sum_{n=0}^{N-1}(d(n) - A - Bn)\,n \; ; \; \frac{\partial^2 \ln p(d|\theta)}{\partial B^2} = -\frac{1}{\sigma^2}\sum_{n=0}^{N-1} n^2$$

$$\frac{\partial \ln p(d|\theta)}{\partial A \partial B} = -\frac{1}{\sigma^2}\sum_{n=0}^{N-1} n$$

? cont)

$$I(\theta) = \frac{1}{\sigma^2}\begin{pmatrix} N & \sum n \\ \sum n & \sum n^2 \end{pmatrix}$$

$$= \frac{1}{\sigma^2}\begin{pmatrix} N & \frac{N(N-1)}{2} \\ \frac{N(N-1)}{2} & \frac{N(N-1)(2N-1)}{6} \end{pmatrix}$$

[30%]

$$I^{-1}(\theta) = \sigma^2 \begin{pmatrix} \frac{2(2N-1)}{N(N+1)} & -\frac{6}{N(N+1)} \\ \frac{-6}{N(N+1)} & \frac{12}{N(N^2-1)} \end{pmatrix}$$

$$\therefore \mathrm{var}(\hat{A}) \geq \frac{2(2N-1)}{N(N+1)}\sigma^2$$

$$\mathrm{Var}(\hat{B}) \geq \frac{12\sigma^2}{N(N^2-1)}.$$

$$\therefore \frac{\mathrm{Var}(\hat{A})}{\mathrm{Var}(\hat{B})} = \frac{(2N-1)(N-1)}{6}$$ which is > 1 for $N \geq 3$

$\therefore$ easier to estimate B [35%]

Q3) First part is book work – see notes. [20%]

b) For the two hypotheses H_0 and H_1, the likelihoods are

$$p(\underline{y}|H_0) = \frac{1}{(2\pi\sigma^2)^{N/2}} \exp\left(-\frac{1}{2\sigma^2} \underline{y}^T \underline{y}\right)$$

&

$$p(\underline{y}|H_1) = \frac{1}{(2\pi\sigma^2)^{N/2}} \exp \frac{-1}{2\sigma^2} (\underline{y}-\underline{s})^T(\underline{y}-\underline{s})$$

The likelihood ratio test is

$$L(\underline{y}) = \frac{p(\underline{y}|H_1)}{p(\underline{y}|H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} k$$

where k depends upon the detection criteria used i.e. MAP, Bayes or Neyman–Pearson.

Q 3 cont)

Therefore

$$L(\underline{y}) = \exp \frac{1}{2\sigma^2}\left(2\underline{y}^T\underline{s} - \underline{s}^T\underline{s}\right)$$

taking logs,

$$\therefore \quad \underline{y}^T\underline{s} \underset{H_0}{\overset{H_1}{\gtrless}} \frac{1}{2}\underline{s}^T\underline{s} + \sigma^2 \ln k. \quad [60\%]$$

c) for the case of coloured noise with covariance matrix C the detector becomes

$$y^T C^{-1} S \underset{H_0}{\overset{H_1}{\gtrless}} \frac{1}{2} S^T C^{-1} S + \ln k.$$

using decomposition methods for C ie (cholesky etc

we get $\quad y^{T'} S' \underset{H_0}{\overset{H_1}{\gtrless}} \frac{1}{2} S^{T'} S' + \ln k.$

where $y^{T'}$ and s' are prewhitened versions of y + s

[20%]

Q4 First part is book work.
If we wish to compare the two following hypotheses,

$$p(y|H_0) = \frac{1}{(2\pi\sigma^2)^{1/2}} e^{-\frac{1}{2\sigma^2}y^2}$$

and

$$p(y|H_1) = \frac{1}{(2\pi\sigma^2)^{1/2}} e^{-\frac{1}{2\sigma^2}(y-A)^2}.$$

Then we have

$$\frac{p(y|H_1)}{p(y|H_0)} = e^{-\frac{1}{2\sigma^2}(-2Ay+A^2)} \underset{H_0}{\overset{H_1}{\gtrless}} \lambda.$$

Where λ is a threshold.
The N-P hypothesis test requires that the false alarm probability be

$$P(D_1|H_0) = \int_{y\in R_1} P(y|H_0)\,dy = \alpha$$

$$\int_{y_T}^{\infty} \frac{1}{(2\pi\sigma^2)^{1/2}} e^{-y^2/2\sigma^2}\,dy = \alpha.$$

Q4 cont) Let $\frac{y}{\sqrt{2}\sigma} = u$.

$$\therefore \alpha = \int_{u_T}^{\infty} e^{-u^2} du.$$

where y_T is given by the equality of (1).

$$\therefore \alpha = \frac{1}{2} \operatorname{erfc}\left(\frac{1}{\sqrt{2}}\left(\frac{1}{2}\left(\frac{A}{\sigma}\right) + \frac{\sigma}{A}\log\lambda\right)\right)$$

The detection and false alarm probabilities are

$$P(D_1 | H_1) = \int_{y_T}^{\infty} p(y | H_1)\, dy \qquad \text{(det. prob)}$$

$$P(D_1 | H_0) = \int_{y_T}^{\infty} p(y | H_0)\, dy \qquad \text{(F.A. prob.)}$$

Slope of ROC curve is

$$\text{slope} = \frac{dP(D_1 | H_1)}{dP(D_1 | H_0)} = \frac{dP(D_1 | H_1)}{dy_T} \cdot \frac{dy_T}{dP(D_1 | H_0)}$$

but $\frac{dP(D_1 | H_1)}{dy_T} = -p(y_T | H_1)$, $\frac{dP(D_1 | H_0)}{dy_T} = -p(y_T | H_0)$

[30%]

Q4 cont)

$$\frac{dP(D_1 | H_1)}{dP(D_1 | H_0)} = \frac{p(y_T | H_1)}{p(y_T | H_0)} = \lambda.$$

$\therefore \lambda$ is the slope of the ROC curve at the required F.A. prob α.

—— // ——

$$P_1(y) = \frac{1}{(2\pi\sigma^2)^{N/2}} \exp\left(-\frac{1}{2}\sum_{i=1}^{n}\left(\frac{y_i - 1}{\sigma}\right)^2\right)$$

$$P_0(y) = \frac{1}{(2\pi\sigma^2)^{N/2}} \exp\left(-\frac{1}{2}\sum\left(\frac{y_i}{\sigma}\right)^2\right).$$

$\therefore$ decision rule is: choose H_1 if

$$\exp\left(\frac{1}{2}\sum_{i=1}^{N}\left(\frac{2y_i - 1}{\sigma}\right)\right) \geq \lambda.$$

[40%]