ECE 830 Fall 2010 Statistical Signal Processing

instructor: R. Nowak , **scribe:** A. Pasha Hosseinbor

Lecture 6: Detection Theory

1 Neyman-Pearson Lemma

Consider two densities where $H_o : X\ p_o(x)$ and $H_1 : X\ p_1(x)$.To maximize a **probability of detection** (true positive) P_D for a given **false alarm** (false positive or type 1 error) $P_{FA} = \alpha$, decide according to

$$\Lambda(x) = \frac{p(x|H_1)}{p(x|H_o)} \frac{P_o c_{00} - P_o c_{10}}{P_1 c_{11} - P_1 c_{01}} \underset{H_0}{\overset{H_1}{\gtrless}} \gamma \tag{1}$$

The Neyman-Pearson theorem is a constrained optimazation problem, and hence one way to prove it is via Lagrange multipliers.

1.1 Method of Lagrange Multipliers

In the method of Lagrange multipliers, the problem at hand is of the form $max\ f(x)$ such that $g(x) \leq c$.

Theorem: Let λ be a fixed non-negative number and let $x_o(\lambda)$ be a maximizer of

$$M(x, \lambda) = f(x) - \lambda\, g(x) \tag{2}$$

Then $x_o(\lambda)$ maximizes f(x) over all x such that $g(x) \leq g(x_o(\lambda))$.

Proof: We assume $\lambda \geq 0$ and that $x_o = x_o(\lambda)$ satisfies

$$f(x_o) - \lambda\, g(x_o) \geq f(x) - \lambda\, g(x) \tag{3}$$

Then

$$f(x_o) \geq f(x) - \lambda\, (g(x) - g(x_o)) \tag{4}$$

Now let $S = \{x : g(x) \leq g(x_o)\}$. Thus, for all $x \in S$, $g(x) \leq g(x_o)$. Since λ is non-negative, we conclude

$$f(x_o \geq f(x)\ \forall\ x \in S \tag{5}$$

1.2 Proof of Neyman-Pearson Theorem

The problem at hand is *max* $P_D(\gamma)$ such that $P_{FA}(\gamma) \leq \alpha$. The Lagrangian is

$$\begin{aligned} M(\gamma, \lambda) &= P_D(\gamma) - \lambda P_{FA}(\gamma) \\ &= \int_{R_1(\gamma)} p(x|H_1)\, dx - \lambda \int_{R_1(\gamma)} p(x|H_o)\, dx \\ &= \int_{R_1(\gamma)} [p(x|H_1) - \lambda p(x|H_o)]\, dx, \end{aligned} \tag{6}$$

where $R_1(\gamma) = \{x : p(x|H_1) > \lambda\, p(x|H_o)\}$. The likelihood ratio is

$$\Lambda(x) = \frac{p(x|H_1)}{p(x|H_o)} \leq \lambda. \tag{7}$$

Now determine $\lambda = \gamma$ as value that satisfies

$$P_{FA}(\gamma) = \alpha \tag{8}$$

Thus,

$$\int_{\gamma}^{\infty} p(\Lambda|H_o)\, dx = \int_{x:\Lambda(x)>\gamma} p(x|H_o)\, dx = \alpha \tag{9}$$

1.3 Example: DC Level in Additive White Gaussian Noise (AWGN)

Consider independent random variables x_i for $i = 1, ..., n$:

$$H_o : x_i \sim N(0, \sigma^2) \qquad H_1 : x_i \sim N(\mu, \sigma^2) \tag{10}$$

According to likelihood ratio test (LRT)

$$\Lambda(x) = \frac{p(x|H_1)}{p(x|H_o)} = \frac{\frac{1}{(2\pi\sigma^2)^{n/2}} e^{\frac{-1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2}}{\frac{1}{(2\pi\sigma^2)^{n/2}} e^{\frac{-1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2}} \underset{H_0}{\overset{H_1}{\gtrless}} \gamma \tag{11}$$

Lets take the natural logarithm of the likelihood ratio:

$$\ln(\Lambda(x)) = \frac{-1}{2\sigma^2}(-2\mu\sum_{i=1}^{n} x_i + n\mu^2) \underset{H_0}{\overset{H_1}{\gtrless}} \ln(\gamma) \tag{12}$$

Assuming $\mu > 0$,

$$\sum_{i=1}^{n} x_i \underset{H_0}{\overset{H_1}{\gtrless}} \frac{\sigma^2}{\mu}\ln\gamma + \frac{n\mu}{2} \equiv \nu, \tag{13}$$

where ν is the threshold. Note that $y \equiv \sum_{i=1}^{n} x_i$ is simply the sufficient statastic for μ of a normal distribution of unknown mean. Lets rewrite our hypotheses test in terms of the sufficient statistic:

$$H_o : y \sim N(0, n\sigma^2) \qquad H_1 : y \sim N(n\mu, n\sigma^2) \tag{14}$$

Lets now determine P_{FA} and P_D.

$$P_{FA} = P(pick\ H_1|given\ H_o) = \int_{\nu}^{\infty} \frac{1}{\sqrt{2n\pi\sigma^2}} e^{-t^2/2n\sigma^2}\, dt = Q(\frac{\nu}{\sqrt{n\sigma^2}}) \tag{15}$$

$$P_D = P(pick\ H_1|given\ H_1) = Q(\frac{\nu - n\mu}{\sqrt{n\mu^2}}) \tag{16}$$

Here Q is the complementary error function. Noting that $\nu = \sqrt{n\sigma^2}Q^{-1}(P_{FA})$, we can rewrite P_D as

$$P_D = Q(Q^{-1}(P_{FA}) - \sqrt{\frac{n\mu^2}{\sigma^2}}), \tag{17}$$

where $\sqrt{\frac{n\mu^2}{\sigma^2}}$ is simply the signal-to-noise ratio (SNR).

1.4 Example: Change in Variance

Consider independent random variables x_i for $i = 1, ..., n$:

$$H_o : x_i \sim N(0, \sigma_o^2) \qquad H_1 : x_i \sim N(\mu, \sigma_1^2) \tag{18}$$

Assume $\sigma_1^2 > \sigma_o^2$. Lets apply LRT, taking natural log of both sides:

$$\frac{n}{2}\ln(\frac{\sigma_o^2}{\sigma_1^2} + .5(\frac{1}{\sigma_o^2} - \frac{1}{\sigma_1^2})\sum_{i=1}^{n} x_i^2 \underset{H_0}{\overset{H_1}{\gtrless}} \ln(\gamma) \tag{19}$$

After doing some algebra, we obtain

$$\sum_{i=1}^{n} x_i^2 \underset{H_0}{\overset{H_1}{\gtrless}} 2(\frac{\sigma_1^2\sigma_o^2}{\sigma_1^2 - \sigma_o^2})(\ln(\gamma) + n\ln(\frac{\sigma_1}{\sigma_o}) \equiv \nu \tag{20}$$

Note that $y \equiv \sum_{i=1}^{n} x_i^2$ is simply the sufficient statistic for variance of a normal distribution of unknown variance.

Now recall that if $x_1, ..., x_n$ are *iid* $N(0, 1)$, then $\sum_{i=1}^{n} x_i^2 \sim \chi_n^2$ (chi-square of degree n). Lets rewrite our null hypothesis test using the sufficient statistic:

$$H_o : y = \sum_{i=1}^{n} \frac{x_i^2}{\sigma_o^2} \sim \chi_n^2 \tag{21}$$

Then, the probability of false alarm is

$$\begin{aligned} P_{FA} &= P(pick\ H_1|given\ H_o) \\ &= \int_{\nu}^{\infty} p(y|H_o)\,dy \\ &= P(y > \nu) \\ &= P(\frac{y}{\sigma_o^2} > \frac{\nu}{\sigma_o^2}) \\ &= P(\chi_n^2 > \frac{\nu}{\sigma_o^2}) \end{aligned} \tag{22}$$

We have to compute the variance numerically. For example, if we have $n = 20$ realizations of x_i and $P_{FA} = 0.01$, then we can numerically compute the threshold to be $\nu = 37.57\sigma_o^2$.

1.5 Neyman-Pearson Lemma: A Second Look

Here is an alternate proof of the Neyman-Pearson Lemma. Consider a binary hypothesis test and LRT:

$$\Lambda(x) = \frac{p_1(x)}{p_o(x)} \underset{H_0}{\overset{H_1}{\gtrless}} \lambda \tag{23}$$

$$P_{FA} = P(\Lambda(x) \geq |H_o) = \alpha \tag{24}$$

There does not exist another test with $P_{FA} = \alpha$ and a detection problem larger than $P(\Lambda(x) \geq |H_o)$. That is, the LRT is the **most powerful test** with $P_{FA} = \alpha$.

Proof: The region where the LRT decides H_1 is

$$R_{np} = \{x : \frac{p_1(x)}{p_o(x)} \geq \lambda\} \tag{25}$$

Let R_T denote the region where some other test describes H_1. Define for any region R

$$P_i(R) = \int_R p_i(x)\,dx, \tag{26}$$

which is simply the probability of $x \in R$ under hypothesis H_i. By assumption both tests have $P_{FA} = \alpha$:

$$\alpha = P_o(R_{np}) = P_o(R_T). \tag{27}$$

Next observe that

$$P_i(R_{NP}) = P_i(R_{NP} \cap R_T) + P_i(R_{NP} \cap R_T^c) \tag{28}$$

$$P_i(R_T) = P_i(R_{NP} \cap R_T) + P_i(R_{NP}^c \cap R_T) \tag{29}$$

Therefore from *Eq.* (27), we conclude that

$$P_o(R_{NP} \cap R_T^c) = P_o(R_{NP}^c \cap R_T) \tag{30}$$

Now, we want to show

$$P_1(R_{NP}) \geq P_1(R') \tag{31}$$

which from $Eq.$ $(28 - 29)$ holds if

$$P_1(R_{NP} \cap R_T^c) \geq P_1(R_{NP}^c \cap R_T) \tag{32}$$

Note

$$\begin{aligned} P_1(R_{NP} \cap R_T^c) &= \int_{R_{NP} \cap R_T^c} p_1(x)\,dx \\ &\geq \lambda \int_{R_{NP} \cap R_T'} p_o(x)\,dx \\ &= \lambda P_o(R_{NP} \cap R_T^c) \\ &= \lambda P_o(R_{NP}^c \cap R_T) \\ &= \lambda \int_{R_{NP}^c \cap R_T} p_o(x)\,dx \\ &\geq \int_{R_{NP}^c \cap R_T} p_1(x)\,dx \\ &= P_1(R_{NP}^c \cap R_T). \end{aligned} \tag{33}$$

Thus, from $Eq.$ (33) we see that at λ increases, R_{np} decreases, and hence P_{FA} decreases. In other words, if $\lambda_1 \geq \lambda_2$, then $R_{np}(\lambda_1) \cap R_{np}(\lambda_2)$, and hence $\alpha_1 \leq \alpha_2$.