ECE531 Lecture 3b: Neyman-Pearson Hypothesis Testing (Infinite Number of Possible Observations)

D. Richard Brown III

Worcester Polytechnic Institute

26-January-2011

Neyman Pearson Hypothesis Testing

Recall, when we had a finite number of possible observations, our approach was:

- Sort the likelihood ratio $L_\ell = \frac{P_{\ell,1}}{P_{\ell,0}}$ by observation index in descending order. The order of L's with the same value doesn't matter.
- Now pick the threshold v to be the smallest value such that

$$P_{\mathsf{fp}} = \sum_{\ell : L_\ell > v} P_{\ell,0} \leq \alpha$$

- If necessary, compute the randomization coefficient γ so that $P_{\mathsf{fp}} = \alpha$.
- The N-P decision rule is then (binary HT notation)

$$\rho^{\mathsf{NP}}(y_\ell) = \begin{cases} 1 & L_\ell > v \\ \gamma & L_\ell = v \\ 0 & L_\ell < v \end{cases}$$

Can same intuition that we developed for the discrete observation case be applied in the continuous observation case?

The Neyman-Pearson Lemma: Part 1 of 3: Optimality

Recall $p_j(y)$ for $j \in \{0,1\}$ and $y \in \mathcal{Y}$ is the conditional pmf or pdf of the observation y given that the state is x_j .

Lemma

Let ρ *be any decision rule satisfying* $P_{\mathit{fp}}(\rho) \le \alpha$ *and let* ρ' *be any decision rule of the form*

$$\rho'(y) = \begin{cases} 1 & \text{if } p_1(y) > vp_0(y) \\ \gamma(y) & \text{if } p_1(y) = vp_0(y) \\ 0 & \text{if } p_1(y) < vp_0(y) \end{cases}$$

where $v \ge 0$ *and* $0 \le \gamma(y) \le 1$ *are such that* $P_{\mathit{fp}}(\rho') = \alpha$*. Then* $P_D(\rho') \ge P_D(\rho)$*.*

The Neyman-Pearson Lemma: Part 1 of 3: Optimality

Proof.

By the definitions of ρ and ρ', we always have $[\rho'(y) - \rho(y)][p_1(y) - vp_0(y)] \geq 0$. Hence

$$\int_{\mathcal{Y}} [\rho'(y) - \rho(y)][p_1(y) - vp_0(y)]\, dy \geq 0$$

Rearranging terms, we can write

$$\begin{aligned}
\int_{\mathcal{Y}} \rho'(y)p_1(y)\, dy - \int_{\mathcal{Y}} \rho(y)p_1(y)\, dy &\geq v\left[\int_{\mathcal{Y}} \rho'(y)p_0(y)\, dy - \int_{\mathcal{Y}} \rho(y)p_0(y)\, dy\right] \\
P_D(\rho') - P_D(\rho) &\geq v\left[P_{\mathsf{fp}}(\rho') - P_{\mathsf{fp}}(\rho)\right] \\
P_D(\rho') - P_D(\rho) &\geq v\left[\alpha - P_{\mathsf{fp}}(\rho)\right]
\end{aligned}$$

But $v \geq 0$ and $P_{\mathsf{fp}}(\rho) \leq \alpha$ implies that the RHS is non-negative. Hence

$$P_D(\rho') \geq P_D(\rho).$$

□

The Neyman-Pearson Lemma: Part 2 of 3: Existence

Lemma

For every $\alpha \in [0,1]$ *there exists a decision rule* ρ^{NP} *of the form*

$$\rho^{NP}(y) = \begin{cases} 1 & \textit{if } p_1(y) > vp_0(y) \\ \gamma(y) & \textit{if } p_1(y) = vp_0(y) \\ 0 & \textit{if } p_1(y) < vp_0(y) \end{cases}$$

where $v \geq 0$ *and* $\gamma(y) = \gamma \in [0,1]$ *(a constant) such that* $P_{fp}(\rho^{NP}) = \alpha$.

The Neyman-Pearson Lemma: Part 2 of 3: Existence

Proof by construction.

Let $\nu \geq 0$, $\mathcal{Y}_\nu = \{y \in \mathcal{Y} : p_1(y) > \nu p_0(y)\}$ and $\mathcal{Z}_\nu = \{y \in \mathcal{Y} : p_1(y) = \nu p_0(y)\}$. For $\nu_2 \geq \nu_1$, $\mathcal{Y}_{\nu_2} \subseteq \mathcal{Y}_{\nu_1}$ and $\int_{\mathcal{Y}_{\nu_2}} p_0(y)\,dy \leq \int_{\mathcal{Y}_{\nu_1}} p_0(y)\,dy$.

Let v be the smallest value of ν such that

$$\int_{\mathcal{Y}_v} p_0(y)\,dy \leq \alpha$$

Choose

$$\gamma = \begin{cases} \frac{\alpha - \int_{\mathcal{Y}_v} p_0(y)\,dy}{\int_{\mathcal{Z}_v} p_0(y)\,dy} & \text{if } \int_{\mathcal{Y}_v} p_0(y)\,dy < \alpha \\ \text{any arbitrary number in } [0,1] & \text{otherwise} \end{cases}$$

Then

$$P_{\mathsf{fp}}(\rho^{\mathsf{NP}}) = \int_{\mathcal{Y}_v} p_0(y)\,dy + \gamma \int_{\mathcal{Z}_v} p_0(y)\,dy = \alpha$$

□

The Neyman-Pearson Lemma: Part 3 of 3: Uniqueness

Lemma

Suppose that $\rho''(y)$ is any N-P decision rule for $\mathcal{H}_0$ versus $\mathcal{H}_1$ with significance level α. Then $\rho''(y)$ must be of the same form a $\rho^{NP}(y)$ except possibly on a subset of $\mathcal{Y}$ having zero probability under $\mathcal{H}_0$ and $\mathcal{H}_1$.

The Neyman-Pearson Lemma: Part 3 of 3: Uniqueness

Proof.

If ρ'' is a N-P decision rule with significance level α, then it must be true that $P_D(\rho'') = P_D(\rho^{\text{NP}})$. From part 1 of the Lemma, we know that

$$P_D(\rho^{\text{NP}}) - P_D(\rho'') \quad \geq \quad v\left[\alpha - P_{\text{fp}}(\rho'')\right]$$

which implies that $P_{\text{fp}}(\rho'') = \alpha$ since the LHS of the inequality is zero. So $P_D(\rho'') = P_D(\rho^{\text{NP}})$ and $P_{\text{fp}}(\rho'') = P_{\text{fp}}(\rho^{\text{NP}})$. We can work the proof of part 1 of the Lemma back to write

$$\int_{\mathcal{Y}} [\rho^{\text{NP}}(y) - \rho''(y)][p_1(y) - vp_0(y)]\, dy = 0$$

Note that the integrand here is non-negative. This implies that $\rho^{\text{NP}}(y)$ and $\rho''(y)$ can differ only on the set $\mathcal{Z}_v = \{y \in \mathcal{Y} : p_1(y) = vp_0(y)\}$. This then implies that $\rho^{\text{NP}}(y)$ and $\rho''(y)$ must have the same form and can differ only in the choice of γ.

From part 2 of the lemma, we know that γ is arbitrary when $\int_{\mathcal{Z}_v} p_0(y)\, dy = 0$. Otherwise, if $\int_{\mathcal{Z}_v} p_0(y)\, dy > 0$, $\rho^{\text{NP}}(y)$ and $\rho''(y)$ must share the same value of γ. □

Example: Coherent Detection of BPSK

Suppose a transmitter sends one of two scalar signals a_0 or a_1 and the signals arrive at a receiver corrupted by zero-mean additive white Gaussian noise (AWGN) with variance σ^2.

We want to use N-P hypothesis testing to maximize

$$P_D = \mathrm{Prob}(\text{decide } \mathcal{H}_1 \,|\, a_1 \text{ was sent})$$

subject to the constraint

$$P_{\mathsf{fp}} = \mathrm{Prob}(\text{decide } \mathcal{H}_1 \,|\, a_0 \text{ was sent}) \leq \alpha.$$

Signal model conditioned on state x_j:

$$Y = a_j + \eta$$

where a_j is the scalar signal and $\eta \sim \mathcal{N}(0, \sigma^2)$. Hence

$$p_j(y) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(\frac{-(y - a_j)^2}{2\sigma^2}\right)$$

Example: Coherent Detection of BPSK

How should we approach this problem? We know from the N-P Lemma that the optimum decision rule will be of the form

$$\rho^{\text{NP}}(y) = \begin{cases} 1 & \text{if } p_1(y) > vp_0(y) \\ \gamma & \text{if } p_1(y) = vp_0(y) \\ 0 & \text{if } p_1(y) < vp_0(y) \end{cases}$$

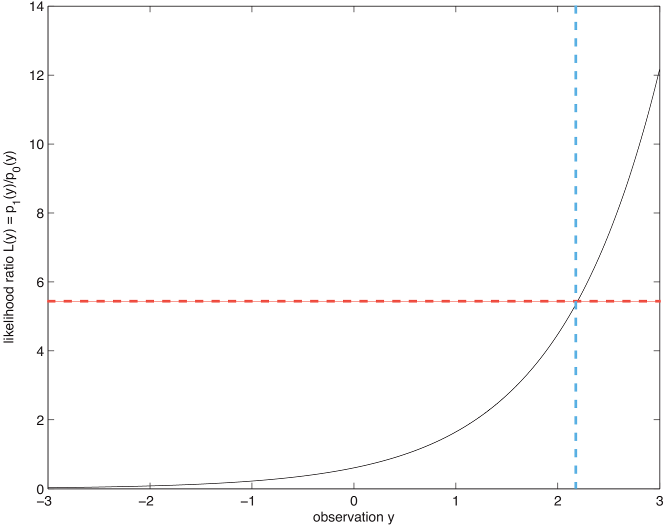
where $v \geq 0$ and $0 \leq \gamma(y) \leq 1$ are such that $P_{\text{fp}}(\rho^{\text{NP}}) = \alpha$. How should we choose our threshold v?

We need to find the smallest v such that

$$\int_{\mathcal{Y}_v} p_0(y)\, dy \leq \alpha$$

where $\mathcal{Y}_v = \{y \in \mathcal{Y} : p_1(y) > vp_0(y)\}$.

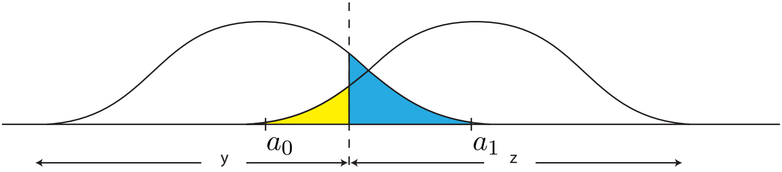
Example: Likelihood Ratio for $a_0 = 0$, $a_1 = 1$, $\sigma^2 = 1$



Example: Coherent Detection of BPSK

Note that, since $a_1 > a_0$, the likelihood ratio $L(y) = \frac{p_1(y)}{p_0(y)}$ is monotonically increasing. This means that finding v is equivalent to finding a threshold τ so that

$$\int_{\tau}^{\infty} p_0(y)\,dy \leq \alpha \quad \Leftrightarrow \quad Q\left(\frac{\tau - a_0}{\sigma}\right) \leq \alpha$$



How are τ and v related? Once we find τ, we can determine v by computing

$$v \;=\; L(\tau) = \frac{p_1(\tau)}{p_0(\tau)}.$$

Example: Coherent Detection of BPSK: Finding τ

Unfortunately, no "closed form" solution exists to exactly solve the inverse of a Q function. We can use Matlab's `qfunc` and `qfuncinv` to numerically solve these sorts of problems, however.

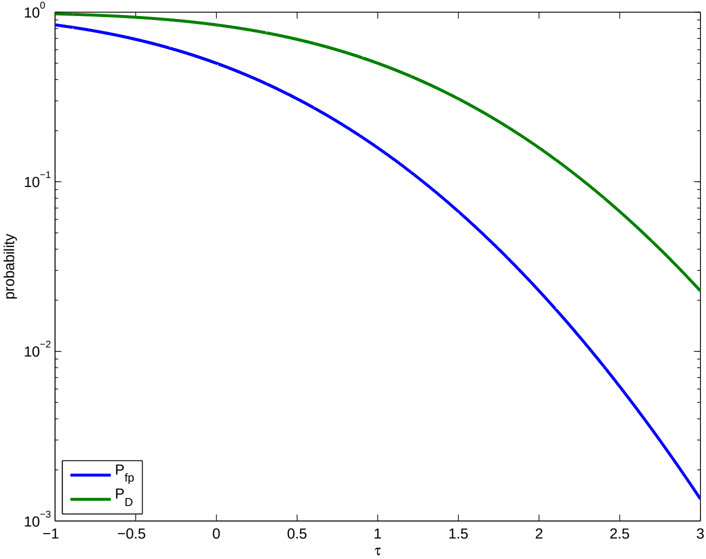
It is not difficult to see that we are going to always be able to find a value of τ such that

$$Q\left(\frac{\tau - a_0}{\sigma}\right) = \alpha$$

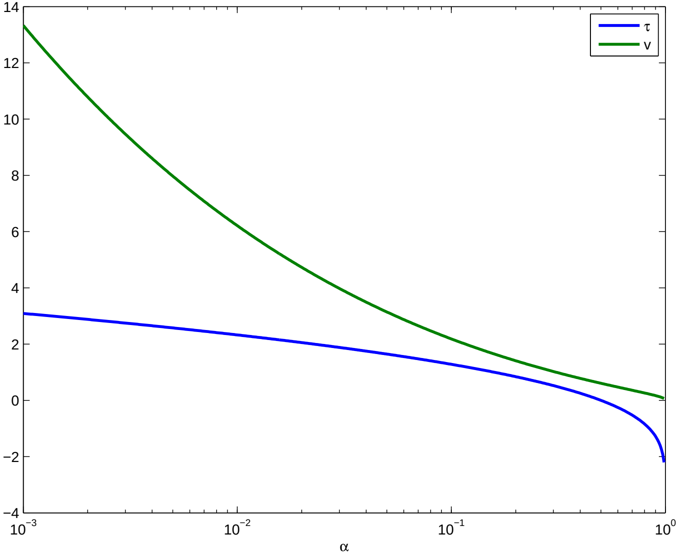
so we won't have to worry about randomization here. Explicitly,

$$\tau = \sigma Q^{-1}(\alpha) + a_0.$$

Example: Coherent Detection of BPSK



Example: Coherent Detection of BPSK



Final Comments on Neyman-Pearson Hypothesis Testing

1. N-P decision rules are useful in asymmetric risk scenarios or in scenarios where one has to guarantee a certain probability of false detection.
2. N-P decision rules are always based on simple likelihood ratio comparisons. The comparison threshold is chosen to satisfy the significance level constraint.
3. Randomization may be necessary for N-P decision rules. Without randomization, the power of the test may not be maximized for the significance level constraint.
4. The original N-P paper: "On the Problem of the Most Efficient Tests of Statistical Hypotheses," J. Neyman and E.S. Pearson, *Philosophical Transactions of the Royal Society of London, Series A, Containing Papers of a Mathematical or Physical Character*, Vol. 231 (1933), pp. 289-337. Available on jstor.org.