Lecture 11.

Subject: No: Date:

Error Curve & Neyman–Pearson Rule. (1)

Binary Hypothesis Testing: $X=0$ or $X=1$

Given a test $A \subset \mathcal{Y}$ and $X=0$, error is made ~~wherever~~ wherever $y \in A$ (A is in favor of $X=1$). If $X=1$, an error is made if $y \in A^c$.

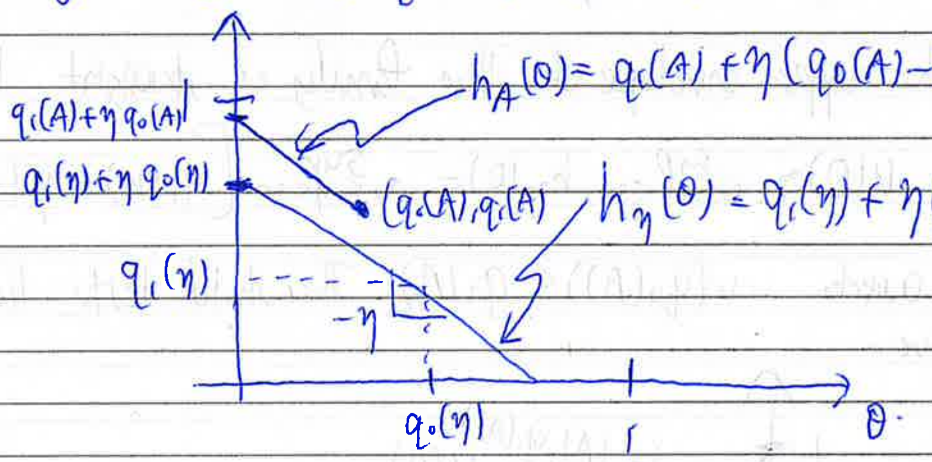
$q_0(A) = Pr(Y \in A \mid X=0)$, $\quad q_1(A) = Pr(Y \in A^c \mid X=1)$.

Threshold tests are important. $A = \left\{ y : \frac{f_{Y|X}(y|1)}{f_{Y|X}(y|0)} \geq \eta \right\}$.

We write $q_0(\eta) = Pr(Y \in A_\eta \mid X=0) = Pr(e_\eta \mid X=0)$

$q_1(\eta) = Pr(Y \in A_\eta^c \mid X=1) = Pr(e_\eta \mid X=1)$

Thm: Consider a 2-D plot in which $(q_0(A), q_1(A))$ is plotted for each test A. Then for each threshold test η where $0 \leq \eta < \infty$, and each arbitrary test A, the point $(q_0(A), q_1(A))$ lies in the closed halfspace to the top-right of a straight line of slope $-\eta$ passing through ~~$q_0(\eta)$~~ $(q_0(\eta), q_1(\eta))$.



Pf: For any η, consider the prior probabilities given by $\eta = p_0/p_1$. Overall (Bayesian) error probability for these prior probabilities is

$$Pr(e(A)) = p_0 q_0(A) + p_1 q_1(A) = p_1[q_1(A) + \eta q_0(A)].$$

Overall (Bayesian) error probabilities for the threshold test & these priors

$$Pr(e(\eta)) = p_0 q_0(\eta) + p_1 q_1(\eta) = p_1[q_1(\eta) + \eta q_0(\eta)].$$

By the optimality of threshold tests for Bayesian hypothesis testing,

$$q_1(\eta) + \eta q_0(\eta) \leq q_1(A) + \eta q_0(\eta).$$

$q_1(\eta) + \eta q_0(\eta)$: point at which $h_\eta(\theta) = q_1(\eta) + \eta(q_0(\eta) - \theta)$ touches the y-axis.

$q_1(A) + \eta q_0(A)$: point at which $h_A(\theta) = q_1(A) + \eta(q_0(A) - \theta)$ touches the y-axis.

~~All points on the~~ The lines $h_\eta(\theta)$ & $h_A(\theta)$ have slope $-\eta$ and pass through $(q_0(\eta), q_1(\eta))$ and $(q_0(A), q_1(A))$ resp.

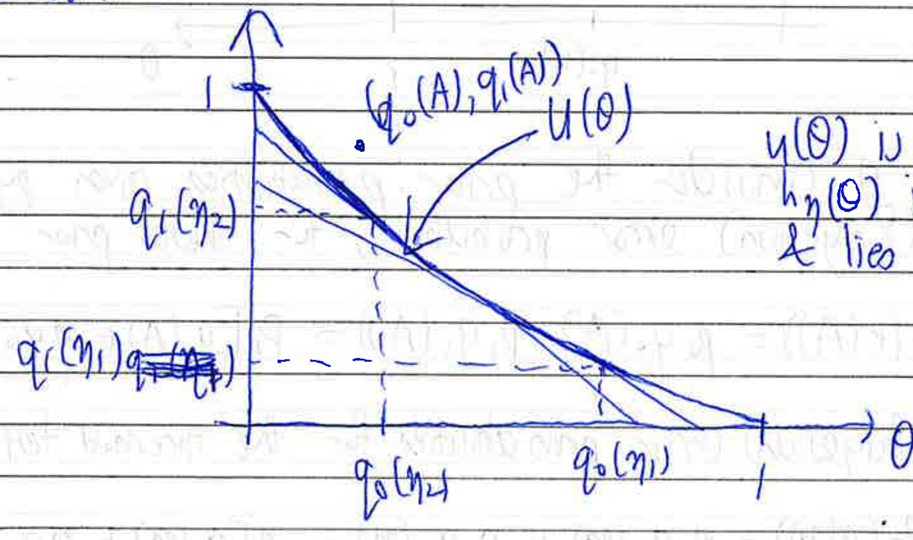
Thus all points on $h_A(\theta)$ lie in the closed halfplane to the top right of all points on the first line $h_\eta(\theta)$. ■

The point $(q_0(A), q_1(A))$ for an arb. test A lies to the top-right of the straight line $h_\eta(\theta)$ for all $0 \leq \eta < \infty$.

Define the upper envelope of this family of straight lines as

$$u(\theta) = \sup_{0 \leq \eta < \infty} h_\eta(\theta) = \sup_{0 \leq \eta < \infty} \left[q_1(\eta) + \eta(q_0(\eta) - \theta) \right]$$

Lemma asserts $u(q_0(A)) \leq q_1(A)$. Threshold tests lie on the curve.



$u(\theta)$ is convex as $h_\eta(\theta)$ is a tangent to it & lies below $u(\theta)$.

Subject:

No:

Date:

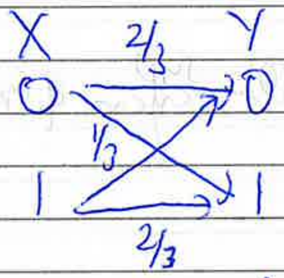
Error curve gives a tradeoff between prob. of error give $X=0$ and $X=1$.

Threshold tests are optimal points for this tradeoff; they lie on the curve.

Are all optimal tests (deterministic) threshold tests? No!

Thm: Consider a BHT problem. The error curve is convex, all threshold tests lie on the error curve and all other tests lie on or above the error curve.

Eg:



$P_{Y|X}(0|0) = P_{Y|X}(1|1) = 2/3.$

$P_{Y|X}(1|0) = P_{Y|X}(0|1) = 1/3.$

$$\Lambda(y) = \frac{P_{Y|X}(y|1)}{P_{Y|X}(y|0)} \overset{'1'}{\underset{'0'}{\gtrless}} \eta$$

Note that y can only take on 2 values $y \in \{0,1\}$.

When $y=1$, $\Lambda(y) = \Lambda(1) = \frac{2/3}{1/3} = 2$

When $y=0$, $\Lambda(y) = \Lambda(0) = \frac{1/3}{2/3} = 1/2$.

Under ML, $\eta = 1$, we would choose $\hat{x}(y) = y$.

If $\eta \le 1/2$, $\hat{x}(y) = 1$ for both $y=0,1$.

If $\eta > 2$, $\hat{x}(y) = 0$ —— " ——

$$\Rightarrow \quad \hat{x}(y) = \begin{cases} 1 & \eta \le \frac{1}{2} \\ y & 1/2 < \eta \le 2 \\ 0. & \eta > 2. \end{cases}$$

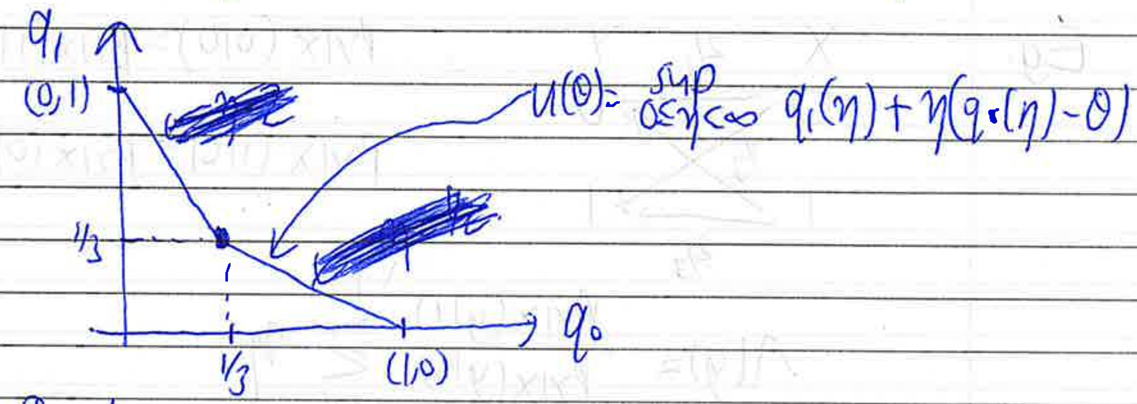
Subject: No: Date:

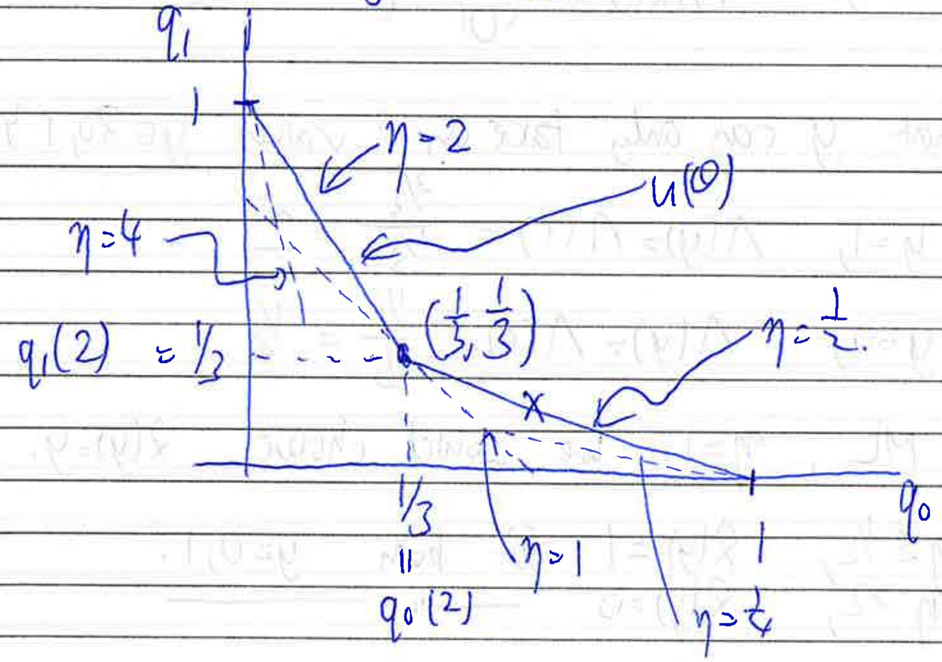
For $\eta \le \frac{1}{2}$, $\hat{x}(y)=1$ so for both $y=0$ & $y=1$, error is made for $X=0$ but not $X=1$ $\Rightarrow$ $q_0(\eta)=1$, $q_0(\eta)=0$ for $\eta \le \frac{1}{2}$.

For $\frac{1}{2} < \eta \le 2$, $\hat{x}(y)=y$, $q_0(\eta) = \frac{1}{3}$, $q_1(\eta)=\frac{1}{3}$.

For $\eta > 2$, $\hat{x}(y)=0$ so for any observation $y=0$ or $y=1$, error is made for $X=1$ but not $X=0$ $\Rightarrow$ $q_0(\eta)=0$ & $q_1(\eta)=1$.

$$q_0(\eta) = \begin{cases} 1 & \eta \le \frac{1}{2} \\ \frac{1}{3} & \frac{1}{2} < \eta \le 2 \\ 0 & \eta > 2 \end{cases} \qquad q_1(\eta) = \begin{cases} 0 & \eta \le \frac{1}{2} \\ \frac{1}{3} & \frac{1}{2} < \eta \le 2. \\ 1 & \eta > 2. \end{cases}$$





However, no test lies in an interior point between $(\frac{1}{3}, \frac{1}{3})$ and $(1,0)$, i.e., the point indicated by X.

Subject: No: Date:

How to achieve the best (smallest) $q_1(A)$ for a given $q_0(A) \le \frac{1}{2}$ (say)?

Randomize between two tests $\hat{x}(y) = y$ and $\hat{x}(y) = 1$. The former yields $(q_0(\eta), q_1(\eta)) = (1/3, 1/3)$ while the latter yields $(1, 0)$ on the error curve.

For $q_0(A) = \frac{1}{2} = \alpha \cdot 1 + (1-\alpha)\, 1/3 = 1/3 + 2/3\,\alpha \Rightarrow \alpha = 1/4$

Our decision $\hat{x}(y) = \begin{cases} 1 & y = 1 \\ 1 & y = 0 \;\&\; z = 1 \\ 0 & y = 0 \;\&\; z = 0 \end{cases}$ $\quad z \sim \text{Bern}(\alpha = 1/4)$, $\Pr(z=1) = 1/4$, $\Pr(z=0) = 3/4$

Why? $\quad z = 1 \qquad\qquad z = 0.$

$\hat{x}(y) = 1 \;\forall y \qquad\qquad \hat{x}(y) = y,\; y \in \{0, 1\}.$

$$\Rightarrow \hat{x}(y) = \begin{cases} 1 & y = 1 \\ 1 & y = 0 \quad z = 1 \\ 0 & y = 0 \quad z = 0. \end{cases}$$

Eg. H_0: $q = q_0 = 1/2$, $\qquad H_1$: $q = q_1 = 1/4$.

Likelihood function. $P_{K|H}(k|H_j) = \begin{cases} q_0(1-q_0)^k, & k \in \mathbb{N} \cup \{0\},\; j = 0 \\ q_1(1-q_1)^k, & j = 1 \end{cases}$

1. If $p_0 = p_1$, find the min prob. of error rule:

$$L(k) = \frac{q_1(1-q_1)^k}{q_0(1-q_0)^k} \underset{\text{'}H_0\text{'}}{\overset{\text{'}H_1\text{'}}{\gtrless}} 1 \Rightarrow k \underset{\text{'}H_0\text{'}}{\overset{\text{'}H_1\text{'}}{\gtrless}} \frac{\log(q_0/q_1)}{\log\left(\frac{1-q_1}{1-q_0}\right)} = 1.71$$

Decide in favor of H_1 if $k \ge 2$.
$\qquad\qquad$ — " — $\qquad H_0$ if $k \le 1$

(6) $P_D = 1 - q_1(A)$.
↑ notation in book

Subject: No: Date:

2. Plot the (P_F, P_D) curve. Note that $P_D = 1 - P_{MD}$

$$P_F = \Pr(K \geq \gamma \mid H = H_0) = \sum_{i=\gamma}^{\infty} q_0(1-q_0)^i = \left(\frac{1}{2}\right)^\gamma$$

$$P_D = \Pr(K \geq \gamma \mid H = H_1) = \sum_{i=\gamma}^{\infty} q_1(1-q_1)^i = \left(\frac{3}{4}\right)^\gamma.$$

The ROC (set of all (P_F, P_D) points) is a countable set of pairs of points.

(P_F, P_D): $(1,1)$, $(\frac{1}{2}, \frac{3}{4})$, $(\frac{1}{4}, \frac{9}{16})$, $(\frac{1}{8}, \frac{27}{64})$, ...
γ: 0, 1, 2, 3, ...

3. Find the maximum P_D s.t. $P_F \leq \frac{1}{128}$. What's the decision rule?

If $P_F \leq \frac{1}{128}$, we choose $\gamma = 7$. Thus, if $K \geq 7$, we declare H_1 is true, otherwise declare H_0 is true.

4. Find the maximum P_D s.t. $P_F \leq \frac{1}{100}$. What's the decision rule?

Note that $0.01 \in (\frac{1}{128}, \frac{1}{64}) = (1/2^7, 1/2^6)$. Need to perform a randomized test between these 2 operating points on the ROC.

$$P_F = 0.01 = \alpha \frac{1}{128} + (1-\alpha)\frac{1}{64} \Rightarrow \alpha = 0.72.$$

$$\text{Max } P_D = \alpha \left(\frac{3}{4}\right)^7 + (1-\alpha)\left(\frac{3}{4}\right)^6 = 0.1459.$$

$$\hat{H} = \begin{cases} H_1 & K \geq 7 \quad Z = 1 \\ H_0 & K < 7 \quad Z = 1 \\ H_1 & K \geq 6 \quad Z = 0 \\ H_0 & K < 6 \quad Z = 0 \end{cases} \Rightarrow \hat{H} = \begin{cases} H_1 & K \geq 7 \\ H_0 & K < 6 \\ H_1 & K = 6 \text{ w.p. } 1-\alpha \\ H_0 & K = 6 \text{ w.p. } \alpha \end{cases}$$

Subject: No: Date:

Eg: Zero-mean Gaussians with one of two positive variances σ_0^2 & $\sigma_1^2 > 0$. (say $\sigma_1^2 > \sigma_0^2$).

$$f_{Y|X}(y|0) = \frac{1}{\sqrt{2\pi\sigma_0^2}} \exp\left(-\frac{y^2}{2\sigma_0^2}\right).$$

$$f_{Y|X}(y|0) = \frac{1}{\sqrt{2\pi\sigma_1^2}} \exp\left(-\frac{y^2}{2\sigma_1^2}\right).$$

$$\text{LRT}: \quad \Lambda(y) = \frac{\frac{1}{\sqrt{2\pi\sigma_1^2}} \exp\left(-\frac{y^2}{2\sigma_1^2}\right)}{\frac{1}{\sqrt{2\pi\sigma_0^2}} \exp\left(-\frac{y^2}{2\sigma_0^2}\right)} \underset{\hat{x}(y)=0}{\overset{\hat{x}(y)=1}{\gtrless}} \eta = \frac{p_0}{p_1}$$

Straightforward to show that

$$\frac{\sigma_0}{\sigma_1} \exp\left(-\frac{y^2}{2\sigma_1^2} + \frac{y^2}{2\sigma_0^2}\right) \gtrless \eta.$$

$\eta > \frac{\sigma_0}{\sigma_1}$

$$\exp\left(-y^2\left(\frac{1}{2\sigma_1^2} - \frac{1}{2\sigma_0^2}\right)\right) \gtrless \eta \frac{\sigma_1}{\sigma_0}.$$

$$y^2\left(\frac{1}{2\sigma_0^2} - \frac{1}{2\sigma_1^2}\right) \gtrless \ln\left(\eta \frac{\sigma_1}{\sigma_0}\right).$$

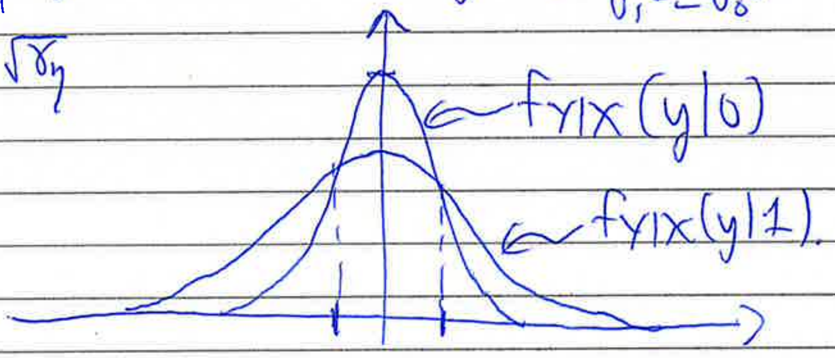
'X=1'

$|y| > \sqrt{\gamma_\eta}$

$$\Leftarrow \quad y^2 \gtrless \frac{2\sigma_0^2\sigma_1^2}{\sigma_1^2 - \sigma_0^2} \ln\left(\eta \frac{\sigma_1}{\sigma_0}\right) = \gamma_\eta.$$

'X=0'

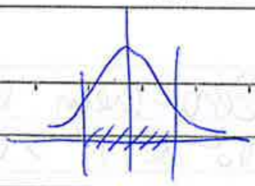
$|y| \le \sqrt{\gamma_\eta}$



$$P_{FA} = \Pr(|Y| > \sqrt{\gamma_\eta} \mid X=0) = \Pr(Y > \sqrt{\gamma_\eta} \mid X=0) + \cancel{\Pr(Z \le \sqrt{\gamma_\eta} \mid X=0)}$$
$$+ \Pr(Y < \sqrt{\gamma_\eta} \mid X=0)$$

$$= 2Q\left(\frac{\sqrt{\gamma_\eta}}{\sigma_0}\right)$$

Subject: No: Date:



$$P_{MD} = P_r(|Y| \leq \sqrt{\gamma_\eta} \mid X=1).$$

$$= 2 P_r(Y \in [0, \sqrt{\gamma_\eta}] \mid X=1)$$

$$= 2\left(\frac{1}{2} - Q\left(\frac{\sqrt{\gamma_\eta}}{\sigma_1}\right)\right) = 1 - 2Q\left(\sqrt{\frac{\gamma_\eta}{\sigma_1^2}}\right).$$

$$P_D = 2Q\left(\frac{\sqrt{\gamma_\eta}}{\sigma_1}\right).$$