Minimax Hypothesis Testing

○ Def.

Suppose that a binary hypothesis testing has
(i) Costs $C_{0,0}$, $C_{1,0}$, $C_{0,1}$, $C_{1,1}$ and
(ii)

H_0 : $\underline{Y} \sim f_{\underline{Y};\theta=0}(\underline{y})$ or $\underline{Y} \sim f_{\underline{Y}|\Theta=0}(\underline{y})$
vs
H_1 : $\underline{Y} \sim f_{\underline{Y};\theta=1}(\underline{y})$ or $\underline{Y} \sim f_{\underline{Y}|\Theta=1}(\underline{y})$

Source $\xrightarrow{\Theta / \theta}$ prob. mapping rule $\xrightarrow{\underline{Y}}$ Detector w/ $C_{i,j}$ $\xrightarrow{\hat{\theta}}$ SINK

$\theta \backslash \hat{\theta}$	0	1
0	$C_{0,0}$	$C_{1,0}$
1	$C_{0,1}$	$C_{1,1}$

$\theta \backslash \hat{\theta}$	0	1
0		FA
1	M	D

In radar engineering target not present / target present are the conventions

We want to find a detector ($\delta_{minimax}$) s.t.

$$\delta_{minimax} = \arg\min_{\delta} \max\{ R(\delta|\theta=0), R(\delta|\Theta=1) \}$$

This detector is called **the minimax detector**.

where the conditional risks are given by

$$R(\delta|\Theta=0) = C_{0,0} \Pr(\hat{\Theta}=0|\Theta=0) + C_{1,0} \Pr(\hat{\Theta}=1|\Theta=0)$$
$$\left(= C_{0,0}(1-P_{FA}) + C_{1,0} P_{FA}\right)$$

$$R(\delta|\Theta=1) = C_{0,1} \Pr(\hat{\Theta}=0|\Theta=1) + C_{1,1} \Pr(\hat{\Theta}=1|\Theta=1)$$
$$= C_{0,1} P_M + C_{1,1}(1-P_M)$$

○ Notes

(i) Uniform cost assignment.

$$R(\delta|\theta=0) = 0 \cdot \Pr(\hat{\Theta}=0|\theta=0) + 1 \cdot \Pr(\hat{\Theta}=1|\theta=0) \triangleq P_{FA}$$

$$R(\delta|\theta=1) = 1 \cdot \Pr(\hat{\Theta}=0|\theta=1) + 0 \cdot \Pr(\hat{\Theta}=1|\theta=1) \triangleq P_M$$

$$\therefore \quad \delta_{minimax} = \arg\min_{\delta} \max\{P_{FA}, P_M\}$$

(ii) The minimax criterion can be applied to both problems w/ Bayesian & non-Bayesian observation models.

(iii) See Poor II.C.

○ Robust detector for Bayesian test

Cancer detector (H_0: no cancer, H_1: yes cancer)

$C_{0,0} = -\$20$

We design a $C_{1,1} = -\$20$

to be sold to $C_{1,0} = \$10{,}000$ → **false alarm**

4 different countries. $C_{0,1} = \$1{,}000{,}000$ → **miss**

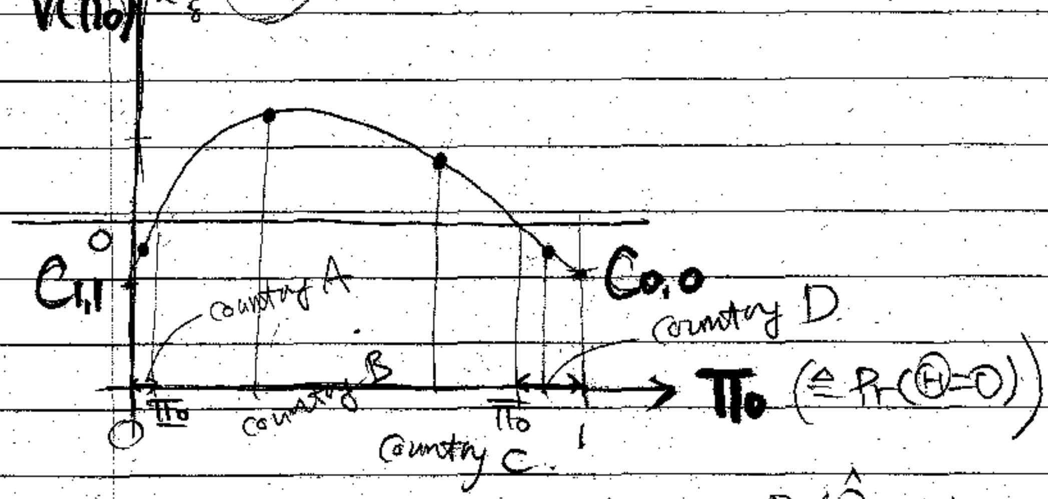
- To maximize the revenue, $\min_{\delta} R(\delta)$ (Bayes detector)

If $P_r(H_0)$ is known

← Bayesian detector

Bayes risk

$\min_{\delta} R(\delta)$ is a function of $\pi_0 \triangleq V(\pi_0)$



If $\pi_0 = 0$, then $Pr(\Theta = 1) = 1$. Hence, $Pr(\hat{\Theta} = 1 \mid Y = y) = 1 \ \forall y$ is optimum. In this case, $R(\delta) = \underbrace{\pi_0}_{=0} R(\delta \mid \Theta = 0) + \underbrace{\pi_1}_{1} R(\delta \mid \Theta = 1) = R(\delta \mid \Theta = 1) = C_{1,1}$

Similarly, if $\pi_0 = 1$, $R(\delta) = C_{0,0}$.

In this case, if $\pi_0 < \underline{\pi_0}$ or $\pi_0 > \overline{\pi_0}$ then the revenue is positive.

What if it is known that the a priori probabilities $Pr(\Theta=0)$ and $Pr(\Theta=1)$ exist but are not known? or incorrectly known?

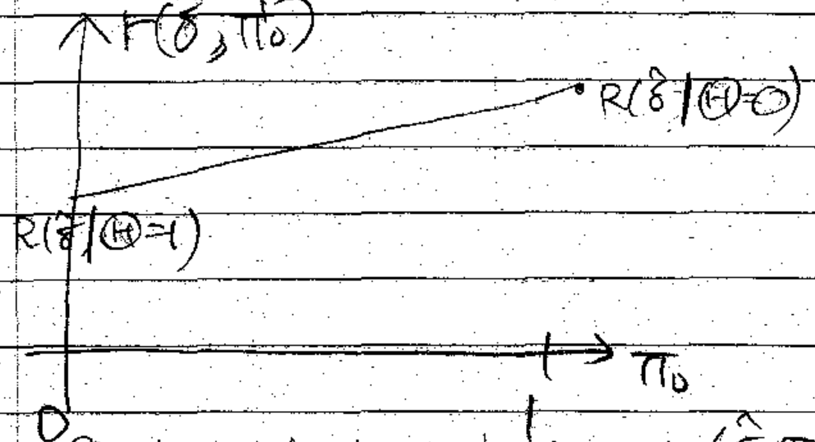
→ need a robust detector.

○ Def.

- Given a fixed $\hat{\delta}$, $\mathbf{r(\hat{\delta}, \pi_0)}$ is defined as **the Bayes risk of $\hat{\delta}$**, i.e.,
$$r(\hat{\delta}, \pi_0) \triangleq R(\hat{\delta} \mid \Theta = 0)\, \pi_0 + R(\hat{\delta} \mid \Theta = 1)(1 - \pi_0)$$

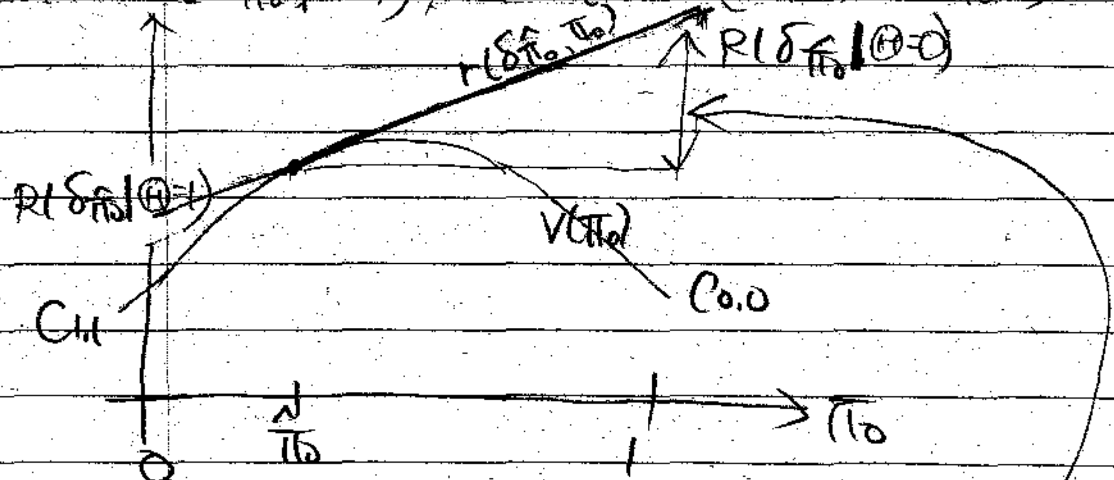
- **As a function of π_0**, the graph $\{(\pi_0, r(\hat{\delta}, \pi_0)) \mid 0 \le \pi_0 \le 1\}$ is given by a straight line crossing

$\left(0,\ R(\hat{\delta}\,|\,\Theta=1)\right)$ and $\left(1,\ R(\hat{\delta}\,|\,\Theta=0)\right)$



* Every fixed detector has $r(\hat{\delta}, \pi_0)$ a straight line.

ex) For a Bayesian detector $\delta_{\hat{\pi}_0}$, $r(\delta_{\hat{\pi}_0}, \pi_0)$ is a straight line crossing $\left(0,\ R(\delta_{\hat{\pi}_0}\,|\,\Theta=1)\right)$, $\left(1,\ R(\delta_{\hat{\pi}_0}\,|\,\Theta=0)\right)$, and $\left(\hat{\pi}_0,\ V(\hat{\pi}_0)\right)$



For a Bayesian detector and other detectors, a **mismatch** between $\hat{\pi}_0$ and $\pi_0 = \Pr(\Theta=0)$ could lead to a very high Bayes risk. $\rightarrow$ The detector

is not **robust** to mismatch. We want to have a detector that is robust or **insensitive** to

the mismatch. $\rightarrow$ the minimax detector.
We will see that
The minimax detector is not only (i) a
non-Bayesian detector just defined as

$$\hat{\delta}_{minimax} = \arg\min_{\delta} \max\left(R(\delta|\Theta=0), R(\delta|\Theta=1)\right)$$

but also (ii) **a robust detector for Bayesian Test**
that is insensitive to mismatch in π_0
& whose performance is guaranteed & uniform.

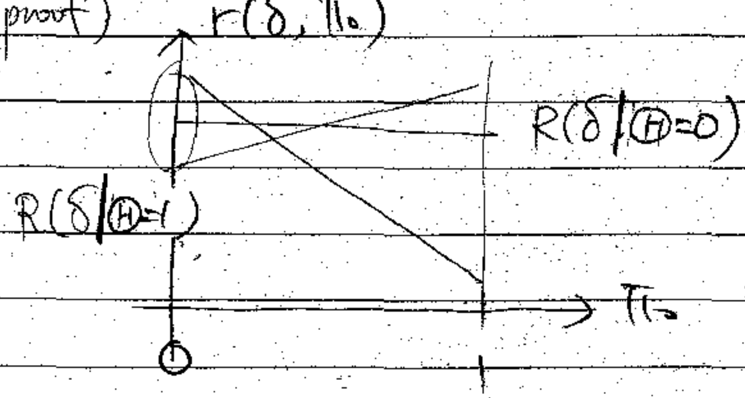
relation to minimax & Bayesian

○ A Twist.

(i) $\max\{R(\delta|\Theta=0), R(\delta|\Theta=1)\}$

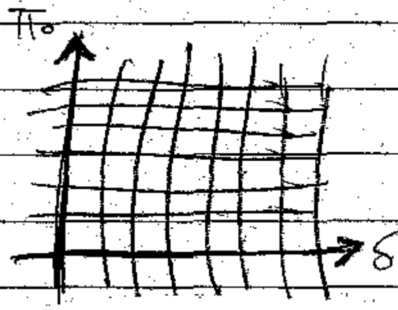
$$= \max_{0\le\pi_0\le1} r(\delta,\pi_0)$$

proof)



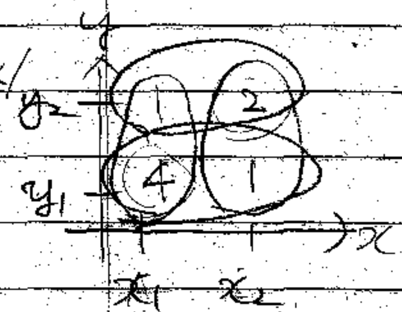
(ii) $\min_{\delta} \max \{ R(\delta|\Theta=0), R(\delta|\Theta=1) \}$

$= \min_{\delta} \max_{0\le\pi_0\le 1} r(\delta,\pi_0)$ by (i)



$\geq \max_{0\le\pi_0\le 1} \min_{\delta} r(\delta,\pi_0)$

ex/



$\min_{x} \max_{y} f(x,y) = 2.$

$\max_{y} \min_{x} f(x,y) = 1$

$\cancel{\min_{\delta} \max \{ R(\delta|\Theta=0), R(\delta|\Theta=1) \}}$

(iii) $= \min_{\delta} \max_{0\le\pi_0\le 1} \underbrace{R(\delta|\Theta=0)\pi_0 + R(\delta|\Theta=1)(1-\pi_0)}_{r(\delta,\pi_0)}$

$\geq \max_{0\le\pi_0\le 1} \min_{\delta} r(\delta,\pi_0)$ by (ii)

Finding the Bayesian detector δ for $Pr(\Theta=0)=\pi_0$

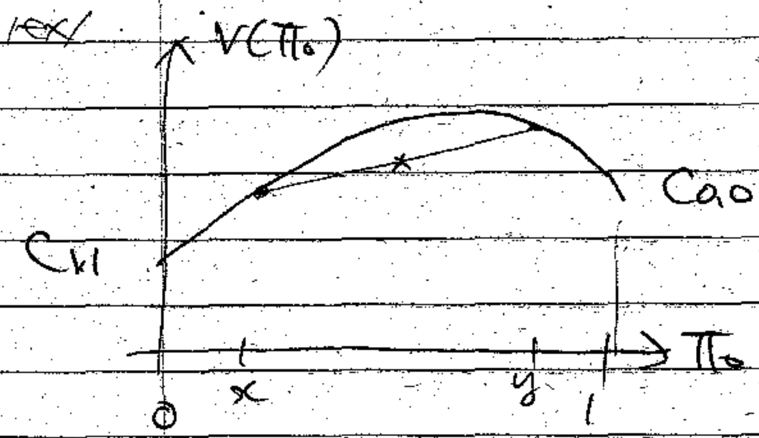
$= \max_{0\le\pi_0\le 1} r(\delta_{\pi_0},\pi_0)$

$\triangleq \max_{0\le\pi_0\le 1} V(\pi_0)$. We have found **a lower bound** on the objective function(al).

○ $V(\pi_0)$

(i) $V(\pi_0)$ is a concave function.

~~ex~~



Def. of concavity : for all $x, y \in [0,1]$.

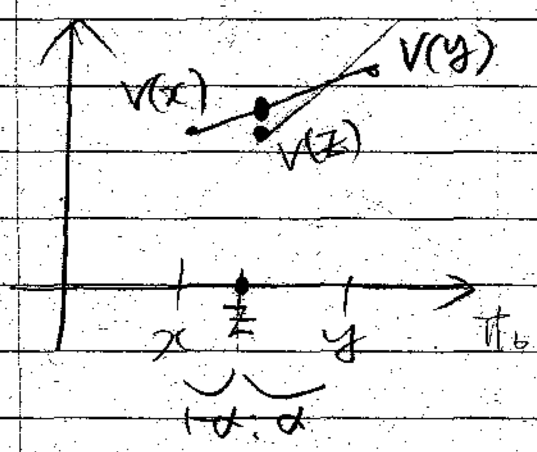
$V(\alpha x+(1-\alpha)y) \geq \alpha V(x)+(1-\alpha)V(y)$, for **all** $0 \leq \alpha \leq 1$

(proof) We use proof by contradiction.

Suppose ~~that~~ there exist x & $y \in [0,1]$ such that for **some** $\alpha \in [0,1]$,

$$V(\alpha x+(1-\alpha)y) < \alpha V(x)+(1-\alpha)V(y). \quad \cdots (*)$$

i.e.,



Define $\delta_z = \arg\min_{\delta} r(\delta, z)$.

where $z = \alpha x + (1-\alpha) y$.

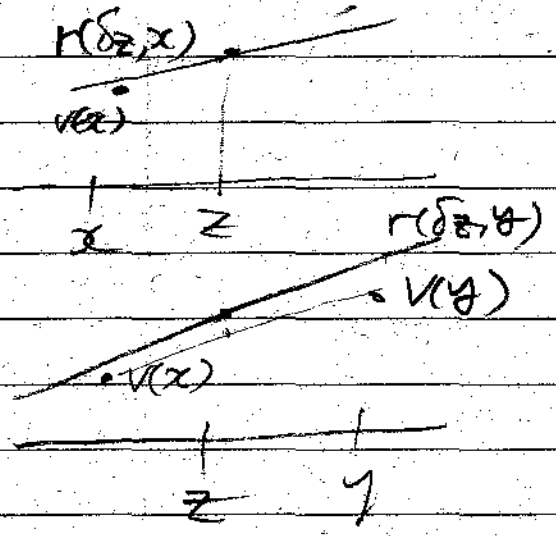
Note. $\min r(\delta, z) = V(z)$

Then,

$V(x) \le r(\delta_z, x)$

& $V(y) \le r(\delta_z, y)$

by definition of $V(\pi_0)$.



which implies

$V(\alpha x + (1-\alpha) y) \ge$
$\alpha V(x) + (1-\alpha) V(y)$

∴ concave.

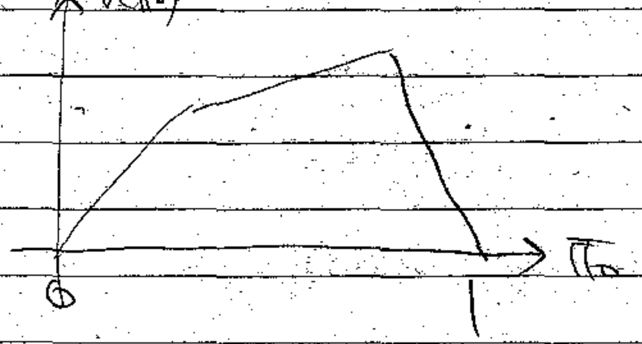
(i) $r(\delta, \pi_0) \ge V(\pi_0)$ $\forall \delta, \forall \pi_0$!

(ii) $V(\pi_0)$ is continuous.

(proof) Any convex or concave function defined on $[0,1]$ is continuous.

(iii) Not necessarily differentiable.

(ex)



iv) $V(\pi_0)$ is **upper** & lower **bounded**

$$\min V(\pi_0) = \min (C_{11}, C_{00})$$

$$\max V(\pi_0) < \max_{i,j} \{C_{ij}\} < \infty$$

$\Rightarrow$ $V(\pi_0)$ has minimum & maximum on $[0, 1]$

v) $$\underset{\pi_0}{\operatorname{argmax}} V(\pi_0) = \begin{cases} 0, \\ 1, \\ \underline{\text{something}} \in (0,1) \end{cases}$$

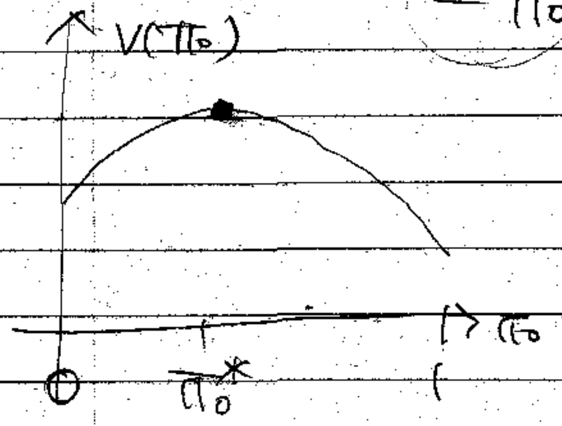
we only consider this case

w/ unique...

The minimax detector

· When $0 < \underset{0 \le \pi_0 \le 1}{\arg\max}\ V(\pi_0) < 1$

$= \pi_0^*$ **the least favorable prior** for Bayesian detection.



(i) If $V(\pi_0)$ is differentiable at π_0^*, then
$V'(\pi_0) = 0$.
$r(\delta_{\pi_0^*}, \pi_0)$ must be $V(\pi_0^*)$ $\forall \pi_0$.

Hence, by adopting $\delta_{\pi_0^*}$, we have

$$\min_{\delta} \max_{\pi_0} r(\delta, \pi_0) \le \max_{\pi_0} r(\delta_{\pi_0^*}, \pi_0)$$

RHS

$$= V(\pi_0^*) = \max_{\pi_0} \overbrace{\min_{\delta} r(\delta, \pi_0)}^{V(\pi_0)} \le \text{LHS}$$

/ex/

$\min_x f(x)$
s.t. $x \in \Omega$

(i) $f(x) \ge 5$ $\forall x \in \Omega$ is shown.
(ii) We find $f(1) = 5$.
Then $x^* = 1$

Hence, $\delta_{\pi_0^*}$ achieves the lower bound.
Therefore, $\delta_{\pi_0^*}$ is the minimax detector.

Note that $R(\delta_{\pi_0^*} | \Theta = 0) = R(\delta_{\pi_0^*} | \Theta = 1)$.
So called **an equalizer rule**.

For differentiable $V(\pi_0)$, the equalizer rule is unique.

$\therefore$ minimax rule $\Leftrightarrow$ equalizer rule

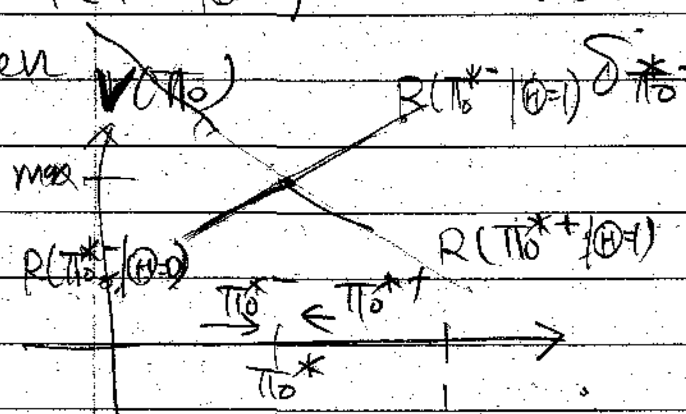
for $V(\pi_0)$ differentiable

(ii) If $V(\pi_0)$ is not differentiable at π_0^* but still have $V'(\pi_0^{*-}) > 0$ and $V'(\pi_0^{*+}) < 0$

left (hand) derivative — right (hand) derivative

$R(\pi_0^{*+} | \Theta=0)$

then



Let's construct a **randomized detector** δ s.t.

$$Pr(\delta = \delta_{\pi_0^{*-}}) = \beta \quad \&$$
$$Pr(\delta = \delta_{\pi_0^{*+}}) = 1-\beta .$$

where β is chosen to have

equalizer rule

$$R(\delta|\Theta=0) = \beta R(\delta_{\pi_0^{*-}}|\Theta=0) + (1-\beta) R(\delta_{\pi_0^{*+}}|\Theta=0)$$
$$= R(\delta|\Theta=1) = \beta R(\delta_{\pi_0^{*+}}|\Theta=1) + (1-\beta) R(\delta_{\pi_0^{*+}}|\Theta=1)$$

, i.e.,

$$q = \frac{R(\delta_{\pi_0^{*+}}|\Theta=0) - R(\delta_{\pi_0^{*+}}|\Theta=1)}{R(\delta_{\pi_0^{*+}}|\Theta=0) - R(\delta_{\pi_0^{*+}}|\Theta=1) + R(\delta_{\pi_0^{*-}}|\Theta=1) - R(\delta_{\pi_0^{*-}}|\Theta=0)} > 0$$

$$= \frac{V'(\pi_0^{*+})}{V'(\pi_0^{*+}) - V'(\pi_0^{*-})}$$

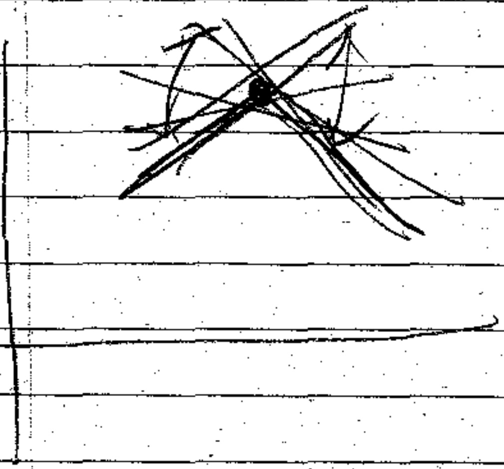
Note that this detector δ has $r(\delta) = V(\pi_0^*)$. (Hence, it is a Bayes detector)

Therefore,

Similarly, δ becomes the minimax detector.

Again, it's the **unique equalizer rule**

Note that by using different q, we can have



All these rules have the **same** $r(\delta) = V(\pi_0^*)$. **Bayes' risk**

However, their **conditional risks** are **different**!

Neyman-Pearson

$\therefore$ randomization is important in Non Bayesian detection such as N-P & minimax.

There are some cases missing in this course.