Hypothesis Tests and Threshold Crossing Probabilities in Random Walks.

→ The MAP Test and the Sequential Test

→ Main objective: Bounds on probability of error in each test

→ Tools from Random Walks one of which is the Chernoff Bound.

i) We have a discrete signal H taking values $\{0, 1\}$, transmitted over a communication channel.

We receive noisy observations of H.

Based on data we receive, we want to decide what H is.

ii) Random Variable $Y \sim P(\cdot | H)$.

We want to test for what value H actually is. $H \in \{0, 1\}$.

Based on observations of Y, we calculate a certain statistic

Maximum Apriori Probability Test

Let H be a discrete random variable with support $\{0, 1\}$. We assign
$P_0 = P(H=0)$, $P_1 = P(H=1)$.
$\Rightarrow P_0 + P_1 = 1$.

→ We want to perform a test on H where $H_0: H=0$ vs. $H_1: H=1$.

To perform this test, we make observations $Y_1, Y_2, \ldots, Y_n$ of H such that $Y_1, Y_2, \ldots, Y_n$ are iid conditional on $H=0$ are iid conditional $H=1$.

Define $\vec{y} = (y_1, y_2, \ldots, y_n)$ and let $f(\cdot|\cdot)$ be the conditional density of Y given H.

$$\Rightarrow f(\vec{y}\,|\,h) = \prod_{i=1}^{n} f(y_i\,|\,h) \quad \text{where}$$

$h \in \{0,1\}$.

From Bayes' Theorem, we know that

$$P(H=0\,|\,\vec{y}) = \frac{p_0\, f(\vec{y}\,|\,0)}{\sum_{i=0}^{1} p_i\, f(\vec{y}\,|\,i)}.$$

$$P(H=1\,|\,\vec{y}) = \frac{p_1\, f(\vec{y}\,|\,1)}{\sum_{i=0}^{1} p_i\, f(\vec{y}\,|\,i)}$$

$$\frac{P(H=0\,|\,\vec{y})}{P(H=1\,|\,\vec{y})} = \frac{p_0\, f(\vec{y}\,|\,0)}{p_1\, f(\vec{y}\,|\,1)}.$$

$$= \frac{P_0}{P_1} \prod_{i=1}^{n} \frac{f(y_i \mid 0)}{f(y_i \mid 1)} = \frac{P_0}{P_1} \lambda(\vec{y})$$

$\lambda(\vec{y}) \downarrow$ Likelihood ratio.

The MAP test states that:

if $\lambda(\vec{y}) > \frac{P_1}{P_0}$, choose $H=0$

if $\lambda(\vec{y}) \leq \frac{P_1}{P_0}$, choose $H=1$

We want to find bounds on the probability of error in the MAP test.

Why is it important to find a bound on probability of error?

$\longrightarrow$ Gives us an estimate of how useful our test is.

Define $Z_1 = \ln \dfrac{f(y_1 \mid 0)}{f(y_1 \mid 1)}$ and we

let $\gamma_1(r) = \ln \underbrace{g_{Z_1 \mid H=1}(r)}_{\text{mgf of } Z_1 \text{ conditional on } H=1.}$

Theorem 1: Let $a = \frac{1}{n} \ln \frac{p_1}{p_0}$ and

let $\exists\, r_1$ such that $\gamma_1'(r_1) = a$.

Then

i) $P(\text{choosing } H=0 \mid H=1) = P(\text{error} \mid H=1)$

$$\leq e^{n[\gamma_1(r_1) - r_1 a]}$$

ii) $P(\text{choosing } H=1 \mid H=0) = P(\text{error} \mid H=0)$

$$\leq e^{n[\gamma_1(r_1) + (1-r_1)a]}.$$

Outline of Proof for Theorem 1:

Chernoff Bound: for any random walk $S_n = X_1 + X_2 + \cdots + X_n$ where $\{X_i\}_{i \geq 1}$ are iid, let $\gamma_x(r) = \ln g_x(r)$. If $\exists\, s$ such that $\gamma_x'(s) = a$, then

$$P(S_n \geq na) \leq e^{n[\gamma_x(s) - sa]}.$$

This is a sharp bound.

i) $P(\text{choosing } H=0 \mid H=1)$

$$= P\left(\lambda(\vec{y}) > \frac{p_1}{p_0} \;\middle|\; H=1\right) \quad (*)$$

We let $S_n = \ln(\lambda(\vec{y}))$

$$\Rightarrow S_n = \sum_{i=1}^{n} \ln \frac{f(y_i \mid 0)}{f(y_i \mid 1)} = \sum z_i$$

S_n is a random walk because

$Y_1, Y_2 -- Y_n$ are iid.

This is stated as Proposition 1 in the notes.

$$(*) = P(S_n > na \mid H=1) \quad \text{where } a = \frac{1}{n} \ln \frac{P_1}{P_0}$$

From the Chernoff Bound, we have directly that

$$P(S_n > na \mid H=1) \le e^{n[\gamma_1(r_1) - r_1 a]}.$$

This proves (i).

For (ii)

$P(\text{choosing } H=1 \mid H=0)$

$$= P\left(\lambda(\vec{y}) \le \frac{P_1}{P_0} \mid H=0\right)$$

$$= P\left(S_n \le \ln \frac{P_1}{P_0} \mid H=0\right)$$

$$= P\left(-S_n \ge -\ln \frac{P_1}{P_0} \mid H=0\right) \quad (**)$$

We define $T_n \equiv -S_n$.

$$\Rightarrow T_n = \sum_{i=1}^{n} X_i \quad \text{s.t. } X_i = -z_i$$

(**)

$= P(T_n \geq -na \mid H=0)$

We use the Chernoff bound to see

that $P(T_n \geq -na \mid H=0)$

$$\leq e^{n[\gamma_0(r_0) - r_0(-a)]}$$

$$= e^{n[\gamma_0(r_0) + r_0 a]} \quad (***).$$

where $\gamma_0 = \ln g_{X \mid H=0}(r)$

and r_0 is such that $\gamma_0'(r_0) = -a$.

<u>Claim</u> (proved in the notes)

i) $\gamma_0(r) = \gamma_1(1-r)$ for r that lies in domain of $\gamma_1(r)$.

ii) $r_0 + r_1 = 1$

iii) $\gamma_0(r_0) = \gamma_1(r_1)$.

We substitute (i), (ii) and (iii)

in (***) to get

$P(T_n \geq -na \mid H=0)$

$$\leq e^{n[\gamma_1(r_1) + (1-r_1)a]}.$$

This proves (ii)

$\square$.

Chernoff Bound (Proposition 2)

For any random walk $S_n = X_1 + X_2 + \cdots + X_n$ where X_i are iid, let $\gamma_X(r) = \ln g_X(r)$

If there exists s such that $\gamma_X'(s) = a$

then $P(S_n \geq na) \leq e^{n[\gamma_X(s) - sa]}$.

and this is a sharp bound.

Outline of the proof of Chernoff Bound

$$P(S_n \geq na) = P(e^{rS_n} \geq e^{rna})$$

for $r > 0$.

$$P(e^{rS_n} \geq e^{rna}) \leq \frac{E[e^{rS_n}]}{e^{rna}} \quad \text{[Markov Inequality]}.$$

$$= [g_X(r)]^n e^{-rna}.$$

$$\Longrightarrow P(S_n \geq na) \leq [g_X(r)]^n e^{-rna}$$

$$= e^{n[\gamma_X(r) - ra]}.$$

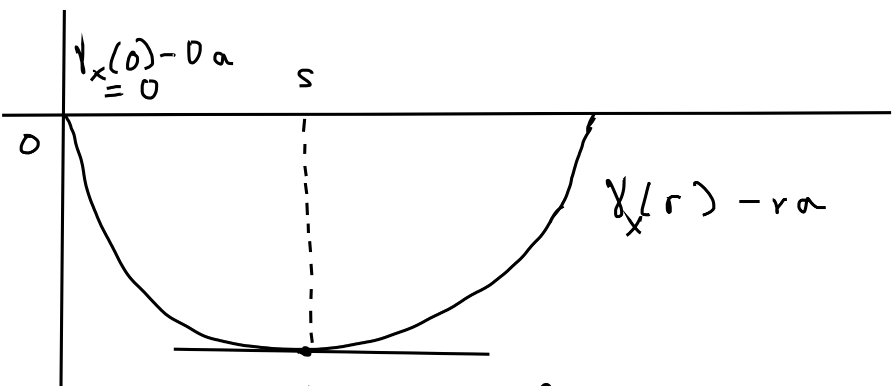
for $r > 0$.

$$\min_{r>0} \gamma_X(r) - ra$$

In theorem 3 of the notes, we use Hölder's Inequality to prove that $\gamma_X(r)$ is a convex function.

If $\gamma_X(r)$ is convex, then a local minimum is also a global minimum.

$\gamma_X'(s) = a.$



$\gamma_X'(s) - a = 0.$

The smallest value of $\gamma_X(r) - ra$ is at $r = s.$

$\Rightarrow \quad P(S_n \geq na) \leq e^{n[\gamma_X(s) - sa]}$

is the going to be the smallest bound on $P(S_n \geq na)$. $\square$.

<u>Sequential test</u>

→ No fixed n.

→ $\alpha > 0, \quad \beta < 0$.

J is a stopping time such that J is the smallest n for which $S_n \geq \alpha$ or $S_n \leq \beta$.

For $n \in \mathbb{N}$, the test states that if

i) If $S_n \in (\beta, \alpha)$, then record observation Y_{n+1}

ii) If $S_n \geq \alpha$ i.e $S_J \geq \alpha$, then we stop testing and we select $H=0$

iii) If $S_n \leq \beta$ i.e $S_J \leq \beta$, then stop testing and we select $H=1$.

Theorem 2:

i) $P(\text{choosing } H=0 \mid H=1) = P(\text{error} \mid H=1)$
$$\le e^{-\alpha}$$

ii) $P(\text{choosing } H=1 \mid H=0) = P(\text{error} \mid H=0)$
$$\le e^{\beta}$$

Outline of Proof of Theorem 2

i) $P(\text{error} \mid H=1) = P(S_J \ge \alpha \mid H=1)$

$\le e^{-r^* \alpha}$ where

$\gamma_1(r^*) = 0$.

Recall that $\gamma_1(r) = \ln g_{Z|H=1}(r)$.

In the notes I have shown that

$\gamma_1(1) = 0. \Longrightarrow r^* = 1$

$\Longrightarrow P(S_J \ge \alpha \mid H=1) \le e^{-\alpha}$.

This proves (i).

ii) $P(\text{error} \mid H=0) = P(S_J \le \beta \mid H=1)$

$$\le e^{r^* \beta} = e^{\beta}$$

$\square$.

In section 4, I have discussed the power of the MAP test and why the sequential test is advantageous compared to MAP test.