17th April 2024

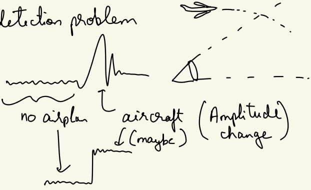
Detection Theory

H_0 (Null hypothesis): Regular state of world

H_1 (Alternate hyp)

Given $X_1, \dots X_n$, is H_0 or H_1 true?

Eg: Consider airplane detection problem



H_0: no airplane

H_1: There is an airplane

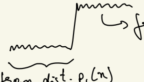
Mean change example.

$$H_0 : x[n] = \varepsilon_n \quad \varepsilon_n \sim \mathcal{N}(0, \sigma^2)$$

$$H_1 : x[n] = \theta + \varepsilon_n$$

Given data, we check mean value of the signal and if its an anomaly.

Change point detection



We want to check if the data has changed the distribution from p_1 to p_2.

H_0: does x continue to follow P_1

H_1: does x change.

Classic Example:

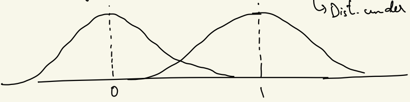
$$X_1, \dots X_N \sim \mathcal{N}(\mu, 1)$$

$H_0: \mu = 0$

$H_1: \mu = 1$ (could be something else)

$P_{H_0}(x) = \mathcal{N}(0,1)$ $\quad$ $P_{H_1}(x) = \mathcal{N}(1,1)$

↳ Dist. under alternative



If $X = -4$, $P_{H_0}(X = -4) = P_{H_1}(X = -4)$.

Can also look at it like likelihood is greater

Consider a more general example:

$$X_1, \dots X_N \overset{iid}{\sim} \mathcal{N}(\mu, 1)$$

$$L_{H_0} = \prod_{i=1}^{N} \mathcal{N}(x_i | 0, 1) \leftarrow p(x_i | \mu, \sigma^2)$$

$$L_{H_1} = \prod_{i=1}^{N} \mathcal{N}(x_i | 1, 1)$$

Gaussian properties

$$P(x|\mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

(left tail mass)

$$F(a) = P(x \leq a) = \int_{-\infty}^{a} P(x|\mu, \sigma^2)\, dx. = \Phi_{\mu,\sigma^2}(a)$$

Special notation for gaussian CDF

$$P(x > a) = \int_{a}^{+\infty} P(x|\mu, \sigma^2)\, dx = Q_{\mu,\sigma^2}(a)$$

(right tail mass)

If $X \sim N(0,1)$ & $y = aX + b$ then $Y \sim (b, a^2)$

Closed under linear transformation

Given any $P(x) \sim N(\mu, \sigma^2)$ and $\Phi_{\mu,\sigma^2}(a) = P_r(X \leq a)$

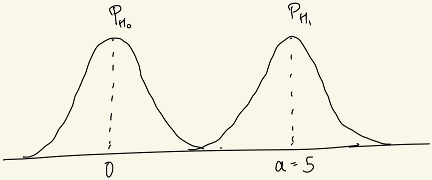
we can write $X = \mu + \sigma z$, where $z \sim N(0,1)$

$$\therefore P_r(X \leq a) = P(\mu + \sigma z \leq a)$$

$$= P\left(z \leq \frac{a-\mu}{z}\right)$$

Neyman - Pearson : Relates to likelihood ratio test

Consider $H_0 : \mu = 0$
$H_1 : \mu = a$



Define $r = \dfrac{P_{H_0}(x)}{P_{H_1}(x)}$

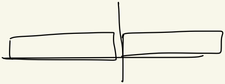
If H_0 is more likely $r > 1$
Else $r < 1$

More interesting when there is larger overlap.
So we set a threshold γ and check $r \gtreqless \gamma$
If $r > \gamma$, H_0 accept $\qquad \gamma \in (0, \infty)$

If γ is low, we will end up accepting H_0 in cases where we shouldn't. (revise this)

Neamen Pearson Lemma: How to select γ

$$P(X \geq a \mid H_0) = \alpha$$



$$\int_{r(x) > \gamma} p(x \mid H_0)\, dx = \alpha = \text{area}$$

a will be given ⇒ Doctor says in a normal state (H_0) a will never be greater than 5. ⇒ We find γ using a

Think of r as a func^n. $r(x)$ = Ratio of likelihoods.

r divides the domain of x into 2 disjoint regions.

Considering $H_0: \mu = 0$, $H_1: \mu = 1$

$$r(x) = \frac{\exp\left(-\frac{x^2}{2}\right)}{\exp\left(-\frac{(x-1)^2}{2}\right)}$$

$$= \exp\left(\frac{(x-1)^2}{2} - \frac{x^2}{2}\right)$$

$$= \exp\left(-\frac{1}{2}(2x-1)\right)$$

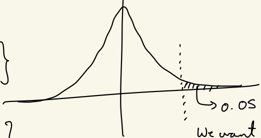
$= \exp\left(-x + \frac{1}{2}\right)$

We want to find γ for which.

$$P(X \geq a \mid H_0) = 0.05 \quad \text{is true}$$

$\{X : r(x) > \gamma\}$

$\Rightarrow \left\{x : \exp\left(-x+\frac{1}{2}\right) > \gamma\right\}$



$\Rightarrow \left\{x : -x + \frac{1}{2} > \log\gamma\right\}$

$\Rightarrow \left\{x : x < -\log\gamma + \frac{1}{2}\right\}$

$\Downarrow$

$$\int_{\{x < -\log\gamma + \frac{1}{2}\}} p(x)\,dx = 0.05$$

$$\Rightarrow \int_0^{-\log\gamma + 1/2} p(x)\,dx = 0.05 \Rightarrow \Phi_{0,1}\left(-\log\gamma + \frac{1}{2}\right) = 0.05.$$

$\Downarrow$

$-\log\gamma + \frac{1}{2} = 2.96$ (from table of CDFs)

$\Rightarrow \gamma = e^{-2.46}$

Example 2:

$$H_0 : \sigma^2 = 1 \qquad N(0,1)$$
$$H_1 : \sigma^2 = 4 \qquad N(0,4)$$ we have N points

$$r(x_{1:N}) = \frac{\prod_{i=1}^{N} P_{H_0}(x_i)}{\prod_{i=1}^{N} P_{H_1}(x_i)}$$

$$= \frac{\prod_{i=1}^{N} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right)}{\prod_{i=1}^{N} \frac{1}{\sqrt{8\pi}} \exp\left(-\frac{x^2}{8}\right)}$$

$$= \frac{(2\pi)^{-N/2} \exp\left(-\frac{1}{2}\sum x_i^2\right)}{(8\pi)^{-N/2} \exp\left(-\frac{1}{8}\sum x_i^2\right)}$$

$$= 4^{N/2} \cdot \exp\left(-\frac{1}{2}\sum x_i^2 \left(1 - \frac{1}{4}\right)\right)$$

$$= 2^N \exp\left(-\frac{3}{8}\sum x_i^2\right)$$

If H_0 was true $\sigma^2 = 1 \Rightarrow 99\%$ data lies in 3 s.d

$\therefore (-3, 3)$

$\big\downarrow$ Because of x^2

$(0, 27)$

If H_0 was rejected $\sigma^2 = 4 \Rightarrow 99\%$ data $\in (-6, 6)$

$\downarrow$ on squaring

(?) why $(0, -)$? $(0, 36)$.

whats the point?

$$\left\{ x_{1:N} : 2^N \exp\left(\frac{-3}{8}\sum x_i^2\right) > \gamma \right\}$$

check.

$$\Rightarrow \quad \sum x_i^2 < \frac{8}{3} N \log(2) - \frac{8}{3} \log(\gamma)$$

$$N \frac{\sum x_i^2}{N} < \text{———''———}$$

$$N\bar{x} < \text{———''———}$$

$$\bar{x} < \frac{8}{3}\log 2 - \frac{8}{3N}\log(\gamma).$$

χ^2_D dist D is degrees of freedom. If X_i are normally dist &

$Y = \sum x_i^2$ then $Y \sim \chi^2_D$

If distribution has a sufficient stat, LLRT will be a func^n of the S.S

Check: sufficient stat & likelihood ratio test.

(simple hyp test)

In the more general case $H_0: \mu = 0$, $H_1: \mu \neq 0$.

$$P_{H_1} = \begin{cases} 0.001 \\ 0.002 \\ \vdots \\ \infty \text{ choices for } \mu \end{cases}$$

→ choose most likely mean and use that in ratio test.

Do likelihood

$$\hat{\mu} = \sum_{i=1}^{N} \log \mathcal{N}(x_i \mid \mu, 1)$$

$$\hat{\mu} = \frac{\sum x_i}{N}$$

$$r(X_{1:N}) = \frac{\prod P_0(x_i)}{\prod \mathcal{N}(x_i \mid \hat{\mu}, 1)}$$

* Inverse CDF

CDF exist only in 1D.

If $\Phi_{\mu,\sigma^2}(x) = \alpha$, $\Phi^{-1}_{\mu,\sigma}(\alpha) = x$.