Lecture 19 - Hypothesis Testing

Defn: Hypothesis

a hypothesis is a statement about a parameter

$$\underbrace{H_0 : \theta \in \Theta_0}_{\text{null hypothesis}} \quad v. \quad \underbrace{H_a : \theta \in \Theta_a}_{\text{alternative hypothesis}}$$

Constraint:

1. $\Theta_0 \cap \Theta_a = \emptyset$
2. $\Theta = \Theta_0 \cup \Theta_a$

Θ_0 and Θ_a partition Θ.

equiv. $\Theta_0 = \Theta \setminus \Theta_a$ or $\Theta_a = \Theta \setminus \Theta_0$

Ex. let θ is the proportion of defective items so that $\theta \in \Theta = [0, 1]$.

Might test

$$\rightarrow H_0 : \theta \leq .1 \quad v. \quad H_a : \theta > .1$$

$$\left[\Theta_0 = [0, .1] \;,\; \Theta_a = (.1, 1] \right]$$

Ex. Let θ denote the change in blood pressure after taking some medicine.

Maybe we test

$$H_0 : \theta = 0 \; ; \; H_a : \theta \neq 0$$

$$\left[\Theta = \mathbb{R} \; ; \; \Theta_0 = \{0\} \; ; \; \Theta_a = \mathbb{R} \setminus \{0\} \right]$$

If θ is a 1-dimensional parameter, then a test of the form

(1)

$$H_0 : \theta \leq c \quad \text{v.} \quad H_a : \theta > c$$

$$H_0 : \theta < c \quad \text{v.} \quad H_a : \theta \geq c$$

$$H_a : \theta \geq c \quad \text{v.} \quad H_a : \theta < c$$

$$H_a : \theta > c \quad \text{v.} \quad H_a : \theta \leq c$$

is called a <u>one-sided test</u>.

(2)

$$H_0 : \theta = c \quad \text{v.} \quad H_a : \theta \neq c$$

$$H_0: \theta \neq c \quad v. \quad H_a: \theta = c$$

is called a <u>two-sided</u> test

(3)

$$H_0: \theta = c_0 \quad v. \quad H_a: \theta = c_a$$

$[\Theta = \{c_0, c_a\}]$

is called a <u>simple hypothesis</u>.

<u>Defn: A hypothesis testing procedure</u>

A rule that determines for which samples $\underset{\sim}{X}$ we prefer H_0 and for which $\underset{\sim}{X}$ we prefer H_a.

Let $\mathcal{X}$ be the support of possible values of $\underset{\sim}{X}$
[typically $\mathcal{X} \subset \mathbb{R}^N$]

A HT procedure is simply a rule that partitions $\mathcal{X}$ into

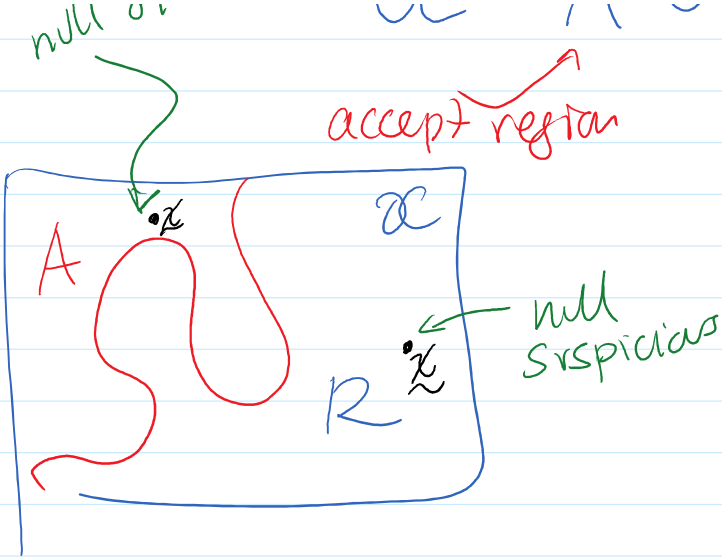
null ok

$$\mathcal{X} = A \cup R$$

reject

accept region

reject region



and we "reject H_0" if $\underset{\sim}{x} \in R$ and "accept H_0" if $\underset{\sim}{x} \in A$.

Often, we can define R (equiv. A) through a <u>test statistic</u> so that our HT is

$$\rightarrow \quad R = \{ \underset{\sim}{x} \mid T(\underset{\sim}{x}) \in C \}$$

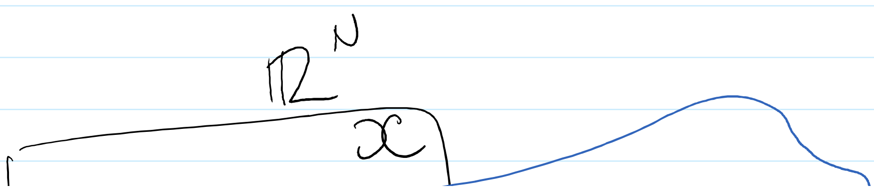
test statistic

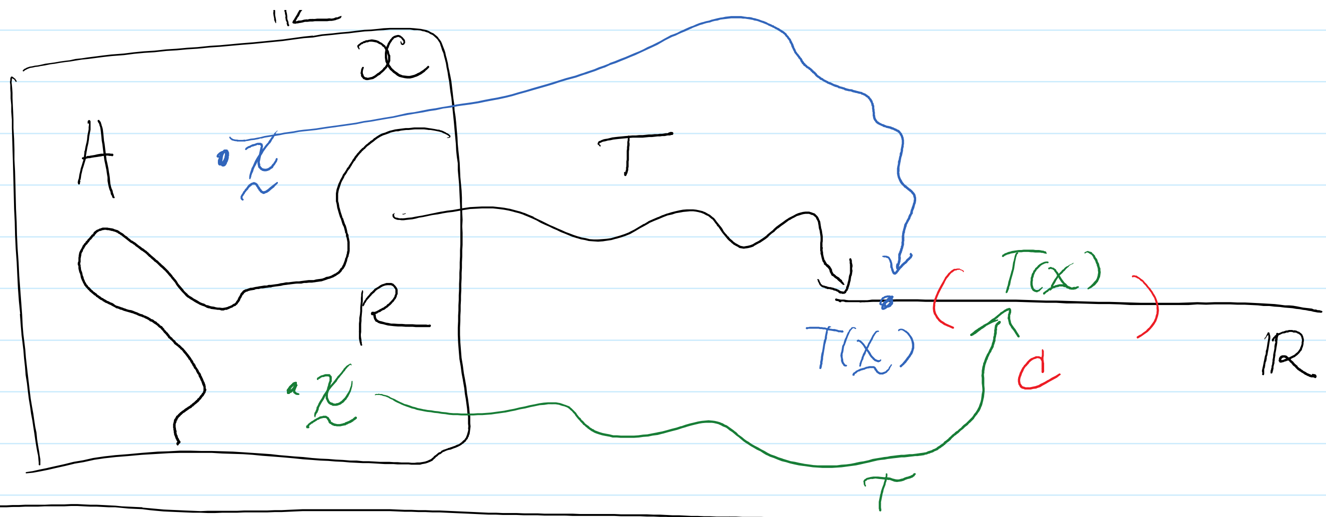
critical region

$C \subset \mathbb{R}$ (typically)

<u>alt.</u>

$$\rightarrow \quad A = \{ \underset{\sim}{x} \mid T(\underset{\sim}{x}) \notin C \}$$





Ex. $T = \overline{X}$ and $C = (z, \infty)$

then my HT

$$R = \{\underset{\sim}{x} \mid \overline{X} > z\} \; ; \; A = \{\underset{\sim}{x} \mid \overline{X} \le z\}$$

Defn! Test Function

The test function φ is

$$\varphi(\underset{\sim}{x}) = \mathbb{1}(\underset{\sim}{X} \in R) = \begin{cases} 1, & \underset{\sim}{x} \in R \text{ (reject } H_0) \\ 0, & \underset{\sim}{x} \notin R \text{ (accept } H_0) \end{cases}$$

notice that

$$\mathbb{E}[\varphi(\underset{\sim}{X})] = \mathbb{E}[\mathbb{1}(\underset{\sim}{X} \in R)] = P(\underset{\sim}{X} \in R)$$

$$\boxed{\mathbb{E}[\mathbb{1}(\underset{\sim}{X} \in A) = \mathbb{P}(\underset{\sim}{X} \in A)]}$$

Defn: Type I and Type II errors

Truth		Accept H_0 ($\underset{\sim}{x} \in A$)	Reject H_0 ($\underset{\sim}{x} \in R$)
Null True	$\theta \in \Theta_0$	Correct Decision	Type I error
Null False	$\theta \in \Theta_a$	Type II error	Correct Decision

Goal: create a HT procedure that minimizes type I and II error probs.

Defn: Power Function

For any $\theta \in \Theta$ the power function β is defined as

$$\beta(\theta) = \mathbb{E}_\theta[\varphi(\underset{\sim}{X})]$$
$$= \mathbb{P}_\theta(\underset{\sim}{X} \in R)$$

for true θ this is the prob. we reject.

→ For $\theta \in \Theta_0$ then the prob. I make a type I error (reject H_0 when $\theta \in \Theta_0$) is

$$\beta(\theta) = P_\theta(\underline{X} \in R)$$

→ If $\theta \in \Theta_a$ then the prob. of a type II error (accept H_0 when $\theta \in \Theta_a$)

$$P_\theta(\underline{X} \in A) = 1 - P_\theta(\underline{X} \in R) = 1 - \beta(\theta)$$

Ex, Let $X_1, \ldots, X_5 \overset{iid}{\sim} \text{Bernoulli}(p)$

→ $H_0: p \leq 1/2$ v. $H_a: p > 1/2$

$$\left[\Theta = [0,1] \; ; \; \Theta_0 = [0, 1/2] \; ; \; \Theta_a = (1/2, 1] \right]$$

Need a test.

$$R = \{(1,1,1,1,1)\}$$

we can define this in terms of a test stat.

$$T(\underline{X}) = T = \sum_{n=1}^{5} X_n \sim \text{Bin}(5, p)$$

then $R = \{ \underset{\sim}{x} \mid T(\underset{\sim}{x}) = 5 \}$

critical region $C = \{5\}$

<u>what is $\beta(p)$?</u>

$$\beta(p) = E[\varphi(\underline{X})] = P(\underline{X} \in R)$$

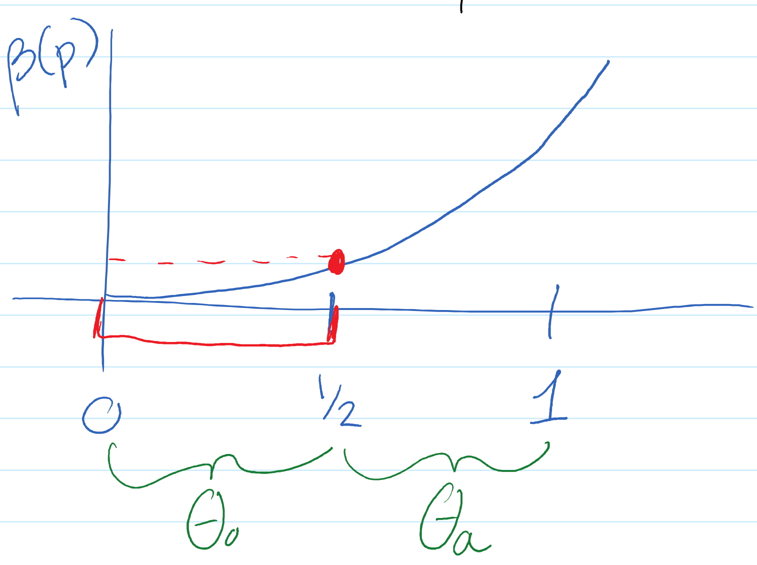
$$= P(T \in C)$$

$$= P(T = 5)$$

$T \sim \text{Bin}(5, p)$

$$= \binom{5}{5} p^5 (1-p)^{5-5}$$

$$= p^5$$



① What is the max prob. I make a type I error?

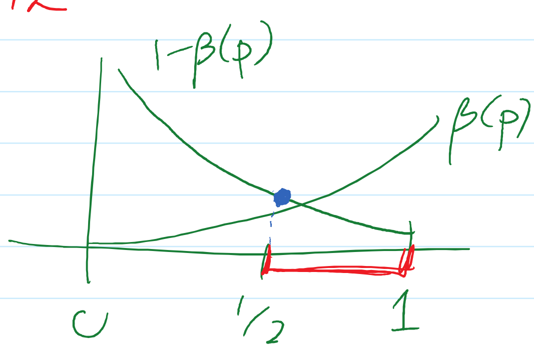
for $\theta \in \Theta_0$, $\beta(\theta)$ = prob. of making type I error

$$\max_{\theta \in \Theta_0} \beta(\theta) = \max_{p \le 1/2} \beta(p) = \beta(1/2) = (1/2)^5 = 1/32$$

② What is the max prob of a type II error?

$\theta \in \Theta_a$, $1-\beta(\theta)$ = prob. of type II

$$\max_{\theta \in \Theta_a} 1-\beta(\theta) = \max_{p > 1/2} 1-\beta(p) = 1-\beta(1/2) = 31/32.$$



Ex. Continue setup. $H_0: p \le 1/2$ v. $H_a: p > 1/2$

But consider a different test.

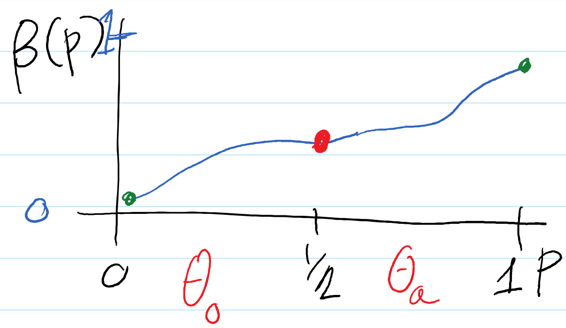
$$R = \{ \underset{\sim}{x} \mid T(\underset{\sim}{x}) \ge 3 \}$$

In this case

$$\beta(p) = \mathbb{P}(\underline{X} \in R) \qquad T \sim Bin(5,p)$$
$$= \mathbb{P}(T \geq 3)$$
$$= \mathbb{P}(T=3) + \mathbb{P}(T=4) + \mathbb{P}(T=5)$$
$$= \binom{5}{3} p^3 (1-p)^2 + \binom{5}{4} p^4 (1-p) + p^5$$
$$= \cdots \text{algebra} \cdots$$
$$= \underline{p^3(6p^2 - 15p + 10)}$$

Q: what is max prob of a type I/II error?

$$\frac{\partial \beta}{\partial p} = 30p^2(p-1)^2 > 0$$



max type I error

$$\underline{\max_{p \leq 1/2} \beta(p) = \beta(1/2)}$$

max type II error

$$\underline{\max_{p > 1/2} 1 - \beta(p) = 1 - \beta(1/2)}$$

Defn: Size and Level of a Test.

We say a test is size $\alpha \in [0,1]$ if

$$\alpha = \max_{\theta \in \Theta_0} \beta(\theta) = \text{max type I error prob.}$$

we say a test is <u>level</u> $\alpha \in [0,1]$ if

$$\text{max type I prob} = \max_{\theta \in \Theta_0} \beta(\theta) \leq \alpha$$
