Tuesday Feb 11 (Part B)

HYPOTHESIS TESTING

Ex Taste test

Model $X_1 \dots X_n \overset{iid}{\sim} \text{Bernoulli}(\theta)$

$\theta = \frac{1}{2}$ It doesn't matter

Null hypothesis $H_0 : \theta = \frac{1}{2}$ $\quad$ $H_1 : \theta \neq \frac{1}{2}$

Alternative hypothesis

Ex Educational Program

$$\left.\begin{array}{l} X_1 \dots X_{n_1} \overset{iid}{\sim} N(\mu_1, \sigma_1^2) \\ Y_1 \dots Y_{n_2} \sim N(\mu_2, \sigma_2^2) \end{array}\right\} ind$$

$H_0 : \mu_1 = \mu_2$ $\quad$ $H_1 : \mu_1 \neq \mu_2$

Ex Wait times for surgery: Govt mandate expected wait time cannot exceed 2 weeks.

$X_1 \dots X_n \overset{iid}{\sim} \text{Exp}(\lambda)$ $\quad$ $H_0 : E(X_i) = \mu \leq 2$

$\iff \lambda \geq \frac{1}{2}$

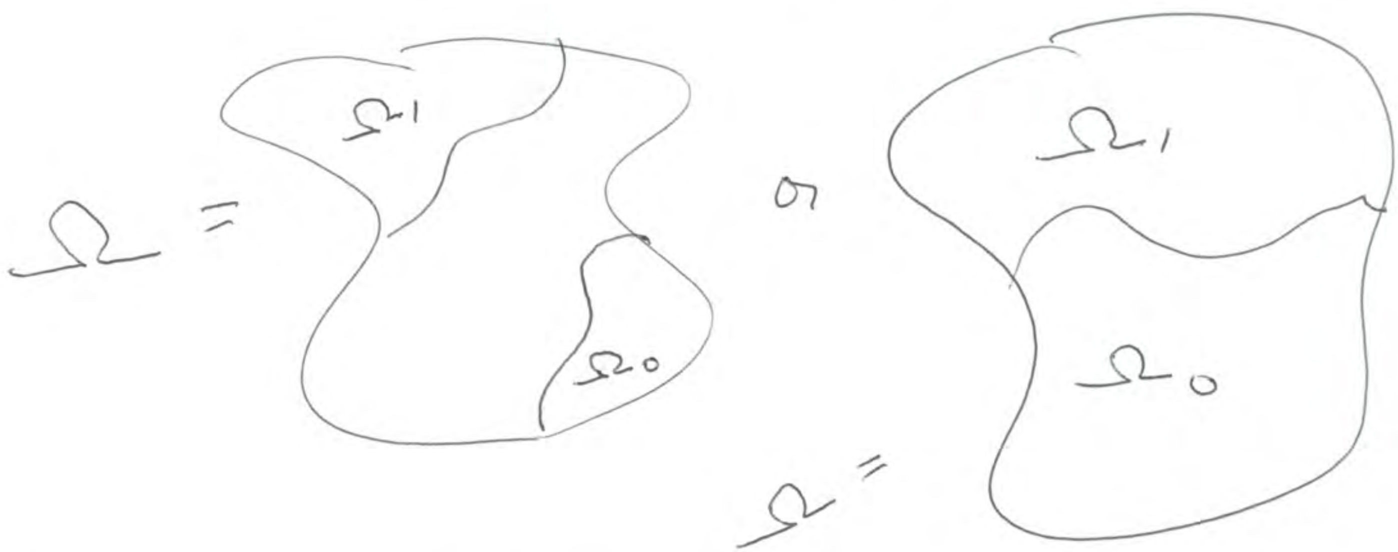
$H_1 : \lambda < \frac{1}{2}$

Def A statistical hypothesis is an assertion that θ is in some subset of parameter space Ω, like

$H_0: \theta \in \Omega_0$ vs $H_1: \theta \in \Omega$,

where $\Omega_0 \subset \Omega$, $\Omega_1 \subset \Omega$

Example $\Omega = (0,1)$ $\theta \in \Omega$



Ex $X_1 \dots X_n \overset{iid}{\sim} B(\theta)$

$\Omega_0 = \{\frac{1}{2}\}$, $\Omega_1 = (0, \frac{1}{2}) \cup (\frac{1}{2}, 1)$

Strategy: Divide sample space $S = \mathbb{R}^n$ into 2 Regions: C & C^c

usually ↑

$$S = C \cup C^c, \ \&$$

Reject H_0 if observed data $s \in C$

C is called the CRITICAL REGION

If $s \in C$, Reject H_0 & Accept H_1

If $s \in C^c$, Accept H_0, Reject H_1

C is usually defined in terms of a ~~TEST~~ TEST STATISTIC. For example

$$C = \left\{ x_1 \ldots x_n = s : \underbrace{\left| \frac{\sqrt{n}\,(\bar{X}_n - \frac{1}{2})}{\sqrt{\frac{1}{2}(1-\frac{1}{2})}} \right|}_{= |Z|} > 1.96 \right\}$$

		TRUTH	
		$\theta \in \Omega_0$	$\theta \in \Omega_1$
Decision Conclusion	$\theta \in \Omega_0$ $s \in C^c$	correct	TYPE II error
	$\theta \in \Omega_1$ $s \in C$	TYPE I error	correct

Wish we could minimize P(TYPE I) & P(TYPE II) at same time.

Can't do both

<u>DUMB TEST ONE</u>: $C = \emptyset$

<u>DUMB TEST TWO</u>: $C = S$

Solution (Fisher, later Neyman-Pearson)
Hold MAX prob of TYPE I error to some small value, & seek a test (Critical region) with SMALLEST POSSIBLE P(TYPE II error)

Missing true findings

TYPICALLY, Hold P(TYPE I) to $\alpha = 0.05$ or 0.01