Testing

Fair coins example to show power.

Def: The true value of θ is the true <u>state of nature</u>.

Def: The hypotheses of a testing problem are the:

a) <u>null hypothesis</u> $H_0: \theta \in \omega$

b) <u>alternative hypothesis</u> H_1 or H_A or Alt : $\theta \in \Theta - \omega$

These hypotheses are gleaned from the experimentor (usually painfully)

Ex: Mean time to cure of the common cold is 14 days. AAA Pharm claims their drug ZZZ is effective in speeding the cure of the common cold.

Let θ = mean time to cure using ZZZ

$H_0: \theta \geq 14$ everything else

$H_1: \theta < 14$ effective

We run the experiment and then decide

a) H_1 is probably true

b) H_0 " "

c) we can't tell which is true because we need more info

Traditionally we pool "can't tell" with "probably H_0"

Our strongest result/action is to "reject H_0" or "accept H_1"

Our other choice is "don't reject H_0"

We never, ever, "accept H_0"

Thus we set up our tests so that H_1 is what we want to confirm. Watch out for non-dichotomous outcomes.

Ex.

	Action: Rej H_0	Do not rej H_0
H_0 True	Type I error (α)	
H_0 False	power $= 1-\beta$	Type II error (β)

$$\alpha = P(\text{rej } H_0 \mid H_0 \text{ true}) = \int_\omega f_0(\vec{x})\,d\vec{x}$$

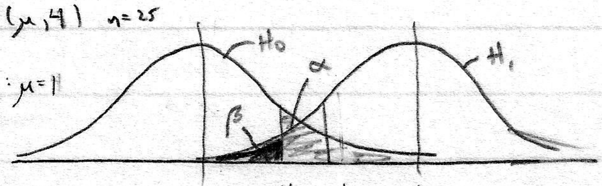
$$\beta = P(\text{not rej } H_0 \mid H_0 \text{ false}) = \int_{\Theta-\omega} f_1(\vec{x})\,d\vec{x} = 1 - \int_\omega f_1(\vec{x})\,d\vec{x}$$

$$\text{Power} = 1-\beta = P(\text{rej } H_0 \mid H_0 \text{ false}) = \int_\omega f_1(\vec{x})\,d\vec{x}$$

Def: A hypothesis (either H_0 or H_1) is a <u>simple hypoth</u> if it contains a single element of Θ. Otherwise it is a <u>composite hypoth</u>.

Ex: $X \overset{iid}{\sim} N(\mu, 4)$ $n=25$

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



$$\alpha = P(\bar{X} > k \mid \mu=0) = 1 - \Phi\left(\frac{k-0}{2/\sqrt{25}}\right)$$

$$\beta = P(\bar{X} < k \mid \mu=1) = \Phi\left(\frac{k-1}{2/\sqrt{25}}\right)$$

For $k=.4$, $\alpha = .1587$ $\beta = .0668$ power $= .9332$

$k=.6$, $\alpha = .0668$ $\beta = .1587$ power $= .8413$

Def: Let $X \sim f(x, \theta)$ for $\theta \in \Theta$. Consider $H_0: \theta \in \omega$ vs $H_1: \theta \in \Theta - \omega$.

A test is a statistic $\phi(X)$ where $0 \le \phi(x) \le 1 \;\forall x \in \mathcal{X}$

When $\phi(x) = p$ then a bernoulli trial ($B(p)$), Y, is performed with $P(Y=1) = p$. If $Y=1$ then H_0 is rejected, else H_0 is not rejected.

Note: Usually $\phi(X)$ is defined such that $\phi \in \{0,1\}$, in which case

Def: The critical region for a test ϕ is $\mathcal{C} = \{x \mid \phi(x) = 1\}$

Def: Consider a test $\phi(X)$ for which there is an $x \in \mathcal{X} \ni 0 < \phi(x) < 1$

Then $\phi(X)$ is a randomized test.

Thm: $E_\theta(\phi(X)) = P(\text{rej } H_0 \mid \theta)$

Def: $E_\theta(\phi(X)) = 1 - \beta(\theta)$ is the power of the test $\phi(X)$.

Ex: "Flipping" "fair" coin 10 times. $X = \#$ of heads

$H_0: p = .5 \qquad H_1: p > .5$

Spose we want $\alpha = .05$.

1-pbinom(7,10,.5) $\quad P(X > 7 \mid p = .5) = 1 - P(X \le 7) = .0547 > .05$

1-pbinom(8,10,.5) $\quad P(X > 8 \mid p = .5) = 1 - P(X \le 8) = .0107 < .05$

So reject H_0 when $X > 8$

"flip" a coin with $P(\text{rej}) = p^*$ when $X = 8$

dbinom(8,10,.5) $\quad$ do not rej H_0 when $X \le 7$

$P(\text{rej } H_0 \mid p = .5) = 0 \cdot P(X \le 7) + .8952\, \overset{.0439}{P(X=8)} + 1\, \overset{.0107}{P(X > 8)}$

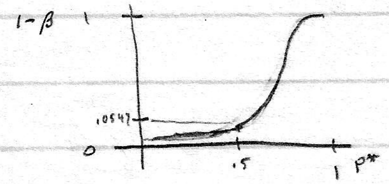
$= 0 + .0393 + .0107 = .05$

So randomizer coin has $p^* = .8952$ of causing rejection

S'pose we choose $\mathcal{C} = \{8, 9, 10\}$

Then $\alpha = P(\text{rej } H_0 \mid p = .5) = .0547$

p^*	0	.1	.2	.3	.4	.5	.6	.7	.8	.9	1
$1-\beta = P(\text{rej } H_0 \mid p = p^*)$	0	0	.0001	.0016	.0123	.0547	.1673	.3828	.6778	.9298	1



Def: A test φ is of <u>size α</u> if $E_\theta \varphi(X) \leq \alpha \quad \forall \theta \in \omega$.

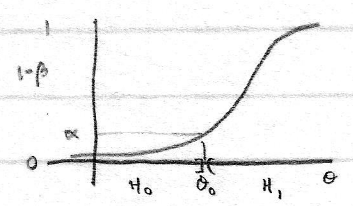
(Fisher gave us historical $\alpha = .05$)

Def: A test φ is <u>similar size α</u> if $E_\theta \varphi(X) = \alpha \quad \forall \theta \in \omega$.

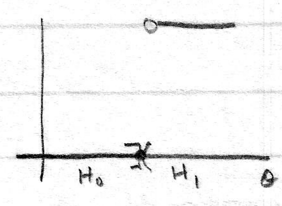
Simple H_0 may generate similar size α, but composite won't.

Def: A test φ is <u>exact size α</u> if $E_\theta \varphi(X) \leq \alpha \quad \forall \theta \in \omega$

and $\exists\, \theta^* \in \omega$ for which $E_{\theta^*} \varphi(X) = \alpha$



Typical



Optimal

$H_1: \theta > \theta_0$

$H_0: \theta \leq \theta_0$