Hypothesis testing framework

Recap: Constructing a test

$$H_0 : \theta \in \Theta_0 \qquad H_A : \theta \in \Theta_1$$

Observe data $X_1, \ldots, X_n$.

- Calculate a test statistic $T_n = T(X_1, \ldots, X_n)$
- Choose a rejection region $R = \{(x_1, \ldots, x_n) : \text{ reject } H_0\}$
- Reject H_0 if $(X_1, \ldots, X_n) \in R$

Goal: minimize probability of a type II error
while controlling probability of a type I error

P(type I error) = P(reject H_0 | H_0 is true)
P(type II error) = 1 − P(reject H_0 | H_A is true)

More specifically: we want P(reject H_0 | true parameter $= \theta$)

Power function

Suppose we reject H_0 when $(X_1, \ldots, X_n) \in R$. The **power function** $\beta(\theta)$ is

$$\underset{\text{function of } \theta}{\beta(\theta)} = \underbrace{P_\theta((X_1, \ldots, X_n) \in R)}_{\text{probability of rejecting } H_0 \text{ if the parameter is } \theta}$$

Goal: want $\beta(\theta)$ small when $\theta \in \Theta_0$
and large when $\theta \in \Theta_1$

Formally: ① Fix $\alpha \in [0,1]$
② Try to maximize $\beta(\theta)$ for $\theta \in \Theta_1$, subject to
$\beta(\theta) \leq \alpha$ for $\theta \in \Theta_0$

- if $\sup_{\theta \in \Theta_0} \beta(\theta) = \alpha$, our test is size α
- if $\sup_{\theta \in \Theta_0} \beta(\theta) \leq \alpha$, our test is level α

Φ = denote $N(0,1)$ cdf

Example

$X_1, \ldots, X_n$ iid from a population with mean μ and variance σ^2.

$$H_0 : \mu = \mu_0 \qquad H_A : \mu > \mu_0$$

reject H_0 when $Z_n = \dfrac{\bar{X}_n - \mu_0}{\sigma/\sqrt{n}} > c$

Power function:

$$\beta(\mu) = P_\mu\left(\frac{\bar{X}_n - \mu_0}{\sigma/\sqrt{n}} > c\right)$$

$$= P_\mu\left(\frac{\bar{X}_n - \mu + \mu - \mu_0}{\sigma/\sqrt{n}} > c\right)$$

$$= P_\mu\left(\underbrace{\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}}}_{\approx N(0,1)} > c - \frac{(\mu - \mu_0)}{\sigma/\sqrt{n}}\right)$$

$$\approx 1 - \Phi\left(c - \frac{(\mu - \mu_0)}{\sigma/\sqrt{n}}\right)$$

For $\mu = \mu_0$:

$$\beta(\mu_0) = 1 - \Phi(c)$$

To get $\beta(\mu_0) = \alpha$

$$1 - \Phi(c) = \alpha$$

$$\Rightarrow \quad c = \Phi^{-1}(1-\alpha)$$

$$= z_\alpha$$

Class activity

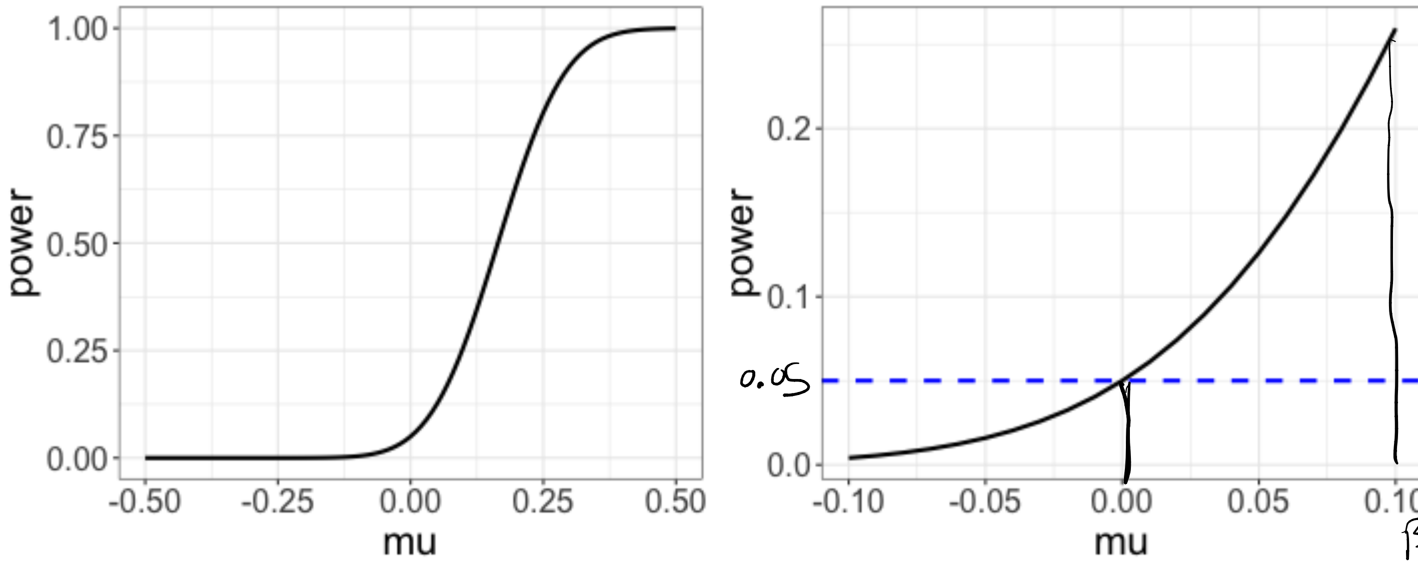
$$\beta(\mu) \approx 1 - \Phi\left(z_\alpha - \frac{(\mu - \mu_0)}{\sigma/\sqrt{n}}\right)$$

- Suppose that $\mu_0 = 0, n = 100$, and $\sigma = 1$. Make a plot of $\beta(\mu)$ vs. μ for $\alpha = 0.05$.
- Now consider testing $H_0 : \mu \leq \mu_0$ vs. $H_A : \mu > \mu_0$. Will this change our rejection region if we want a size α test?

$$\beta(\mu) = 1 - \Phi\left(z_\alpha - \frac{(\mu-\mu_0)}{\sigma/\sqrt{n}}\right)$$

Class activity

if $\mu < \mu_0$, $z_\alpha - \frac{(\mu-\mu_0)}{\sigma/\sqrt{n}} > z_\alpha$



$\beta(0.1) \approx 0.25$

$\beta(\mu)$ increases as μ increases (easier to reject H_0)

$\alpha = 0.05$, $\beta(\mu_0) = 0.05$

$\sup_{\mu \le \mu_0} \beta(\mu) = \beta(\mu_0)$

$H_0: \mu \le \mu_0$ vs. $H_A: \mu > \mu_0$

$H_0: \mu = \mu_0$ vs. $H_A: \mu > \mu_0$

same rejection region for a size α test

$$\beta(\mu) = 1 - \Phi\left(z_\alpha - \frac{(\mu - \mu_0)}{\sigma/\sqrt{n}}\right)$$

if $\mu < \mu_0$, $\quad z_\alpha - \dfrac{(\mu-\mu_0)}{\sigma/\sqrt{n}} > z_\alpha$

$$\Rightarrow \quad \Phi\left(z_\alpha - \frac{(\mu-\mu_0)}{\sigma/\sqrt{n}}\right) > \Phi(z_\alpha)$$

$$\Rightarrow \quad \beta(\mu) = 1 - \Phi\left(z_\alpha - \frac{(\mu-\mu_0)}{\sigma/\sqrt{n}}\right) < \beta(\mu_0) = 1 - \Phi(z_\alpha)$$

Rejecting H_0

$$H_0 : \theta \in \Theta_0 \qquad H_A : \theta \in \Theta_1$$

> A hypothesis test rejects H_0 if $(X_1, \ldots, X_n)$ is in the rejection region R. Are there any issues if we only use a rejection region to test hypotheses?

- picking one rejection region only allows us to look at one α
- reject / fail to reject only provides binary information ?,

Question: what other values of α do we reject for?

p-values

$$H_0 : \theta \in \Theta_0 \qquad H_A : \theta \in \Theta_1$$

Given α, we construct a rejection region R and reject H_0 when $(X_1, \ldots, X_n) \in R$. Let $(x_1, \ldots, x_n)$ be an observed set of data.

Definition: The **p-value** for the observed data $(x_1, \ldots, x_n)$ is the smallest α for which we reject H_0.

Intuition: reject H_0 when p-value $< \alpha$

$\Rightarrow \inf\{\alpha : \text{reject } H_0\} = \inf\{\alpha : \text{p-value} < \alpha\}$

$= \text{p-value}$

Suppose we have a test which rejects H_0 when $T(X_1, \dots, X_n) > c_\alpha$,

where c_α is chosen such that $\sup_{\theta \in \Theta_0} \beta(\theta) = \sup_{\theta \in \Theta_0} P_\theta(T(X_1, \dots, X_n) > c_\alpha) = \alpha$

Let $(x_1, \dots, x_n)$ be a set of observed data.

Theorem: The p-value for $(x_1, \dots, x_n)$ is

$$p = \sup_{\theta \in \Theta_0} P_\theta\big(T(X_1, \dots, X_n) > \underbrace{T(x_1, \dots, x_n)}_{\text{observed test statistic}}\big)$$

("probability of our test statistic or more extreme")

Proof: $p = \inf\{\alpha : \text{reject } H_0\} = \inf\{\alpha : T(x_1, \dots, x_n) > c_\alpha\}$

As $\alpha \downarrow$, $c_\alpha \uparrow$ $\Rightarrow$ $c_p = \sup\{c_\alpha : T(x_1, \dots, x_n) > c_\alpha\} = T(x_1, \dots, x_n)$

$$\sup_{\theta \in \Theta_0} P_\theta(T(X_1, \dots, X_n) > c_p) = p$$

$$\Rightarrow \sup_{\theta \in \Theta_0} P_\theta(T(X_1, \dots, X_n) > T(x_1, \dots, x_n)) = p$$

//

Example

$X_1, \ldots, X_n$ iid from a population with mean μ and variance σ^2.

$$H_0 : \mu = \mu_0 \quad H_A : \mu > \mu_0$$

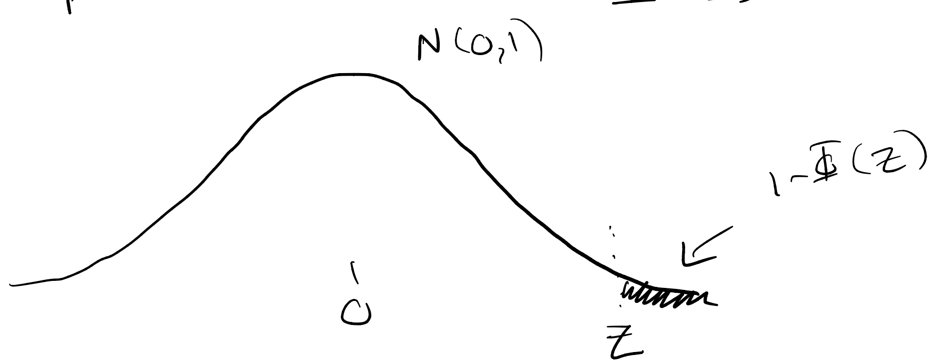
reject H_0 when $Z_n = \dfrac{\bar{X}_n - \mu_0}{\sigma/\sqrt{n}} > z_\alpha$

Let z be observed test statistic

$$\text{p-value} = P_{\mu_0}(Z_n > z)$$

If $\mu = \mu_0$, $Z_n \approx N(0,1)$

$\Rightarrow$ p-value $= 1 - \Phi(z)$



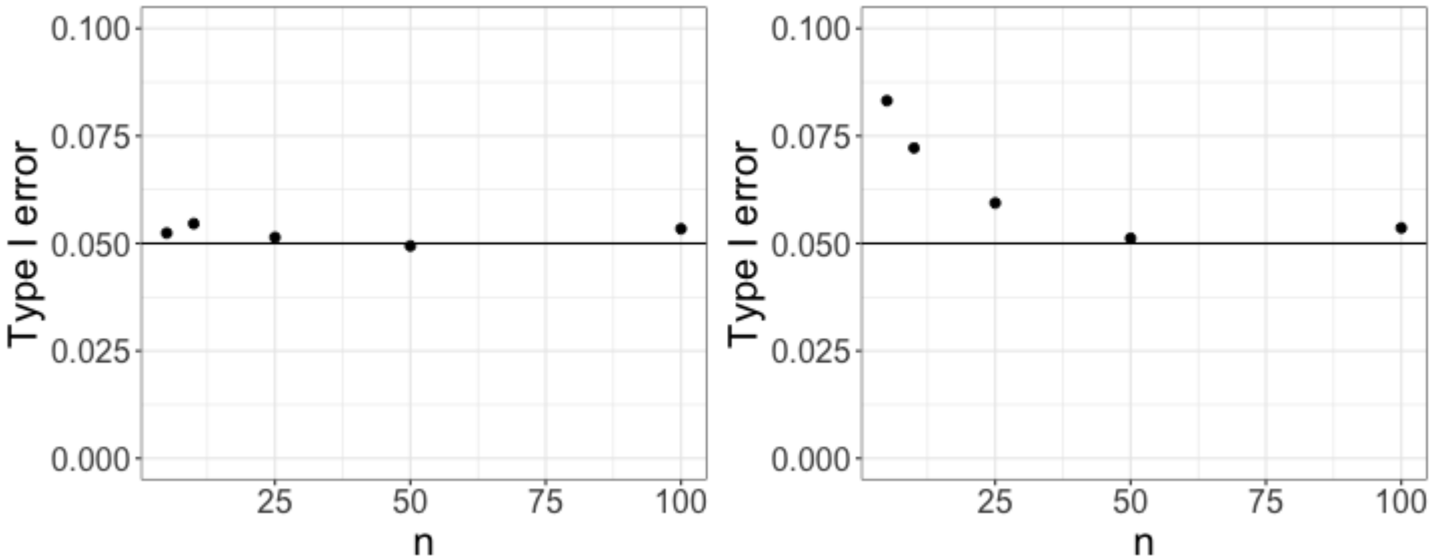
Next steps

> So far, we have discussed the Wald test in detail. What other hypothesis tests have you seen in statistics courses?

Class activity

https://sta711-s23.github.io/class_activities/ca_lecture_20.html

Class activity



If we reject $H_0 : \mu = \mu_0$ when $\frac{\sqrt{n}(\overline{X}_n - \mu_0)}{s} > z_\alpha$, why does type I error increase as n decreases?