Lecture 21: More hypothesis testing

Rejecting H_0

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

Question: 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?

- reject / fail to reject: only get binary information
- to choose a rejection region, need to specify α first – don't know what would have happened had I used a different α

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 so that $\sup\limits_{\theta\in\Theta_0} \beta(\theta) = \sup\limits_{\theta\in\Theta_0} P_\theta(T(X_1,\dots,X_n) > c_\alpha)$

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

$= \alpha$

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

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

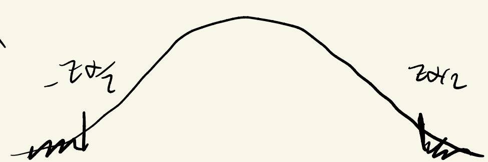
$$p = \sup_{\theta\in\Theta_0} P_\theta\Big(T(X_1,\dots,X_n) > \underbrace{T(x_1,\dots,x_n)}_{\text{observed test statistic}}\Big)$$

("probability of observed test stat or 'more extreme' under H_0")

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

$$T(X_1,\dots,X_n) = \left|\frac{\bar{X}_n-\mu_0}{S/\sqrt{n}}\right| \quad N(0,1) \qquad \text{reject when } \left|\frac{\bar{X}_n-\mu_0}{S/\sqrt{n}}\right| > z_{\frac{\alpha}{2}}$$

rejection region



reject if observed test stat is in tails

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

p-value : $\displaystyle \sup_{\mu=\mu_0} P_{\mu}\left(\left| \frac{\bar{X}_n - \mu_0}{S/\sqrt{n}} \right| > \left| \frac{\bar{x}_n - \mu_0}{s/\sqrt{n}} \right| \right)$

$$= P_{\mu_0}\left(\left| \frac{\bar{X}_n - \mu_0}{S/\sqrt{n}} \right| > \underbrace{\left| \frac{\bar{x}_n - \mu_0}{s/\sqrt{n}} \right|}_{\text{observed test stat}} \right)$$

$=$

$N(0,1) \leftarrow \approx$ sampling distribution of $\left| \frac{\bar{X}_n - \mu_0}{S/\sqrt{n}} \right|$ under H_0

p-value $\approx 2\Phi\left(-\left| \frac{\bar{x}_n - \mu_0}{s/\sqrt{n}} \right|\right)$

$2\Phi(-z_{\frac{\alpha}{2}}) = \alpha$ (by def.)

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

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

("probability of observed test stat or 'more extreme' under H_0")

Proof:

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

As $\alpha \downarrow$, $c_\alpha \uparrow$

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

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

$T(x_1, \dots, x_n)$

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

//

Issue: Wald tests with small n

The Wald test for a population mean μ relies on

$$Z_n = \frac{\sqrt{n}(\overline{X}_n - \mu)}{s} \approx N(0, 1)$$

- $Z_n \xrightarrow{d} N(0, 1)$ as $n \to \infty$
- But for small n, Z_n is not normal, even if $X_1, \ldots, X_n \overset{iid}{\sim} N(\mu, \sigma^2)$

What is the exact distribution of $\frac{\sqrt{n}(\overline{X}_n - \mu)}{s}$?

t-tests

If $X_1, \ldots, X_n \overset{iid}{\sim} N(\mu, \sigma^2)$, then

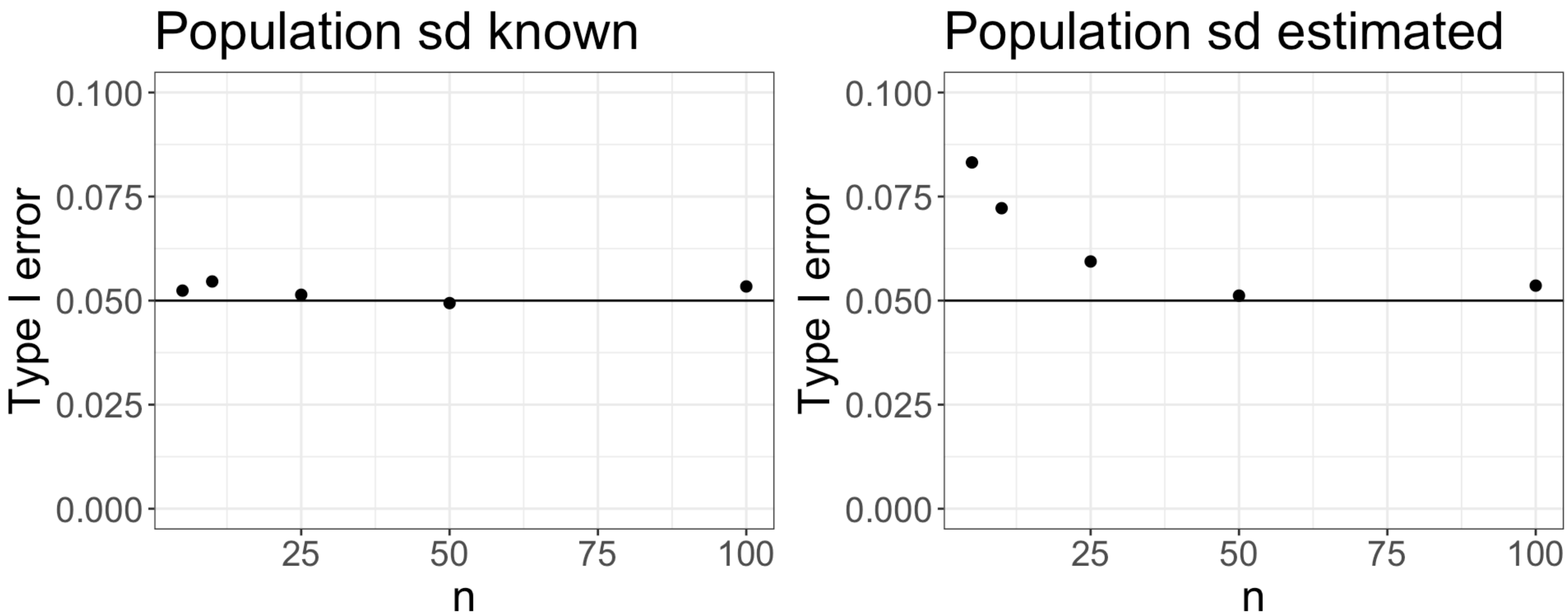
$$\frac{\sqrt{n}(\overline{X}_n - \mu)}{s} \sim t_{n-1}$$

Class activity

https://sta711-s24.github.io/class_activities/ca_lecture_21.html

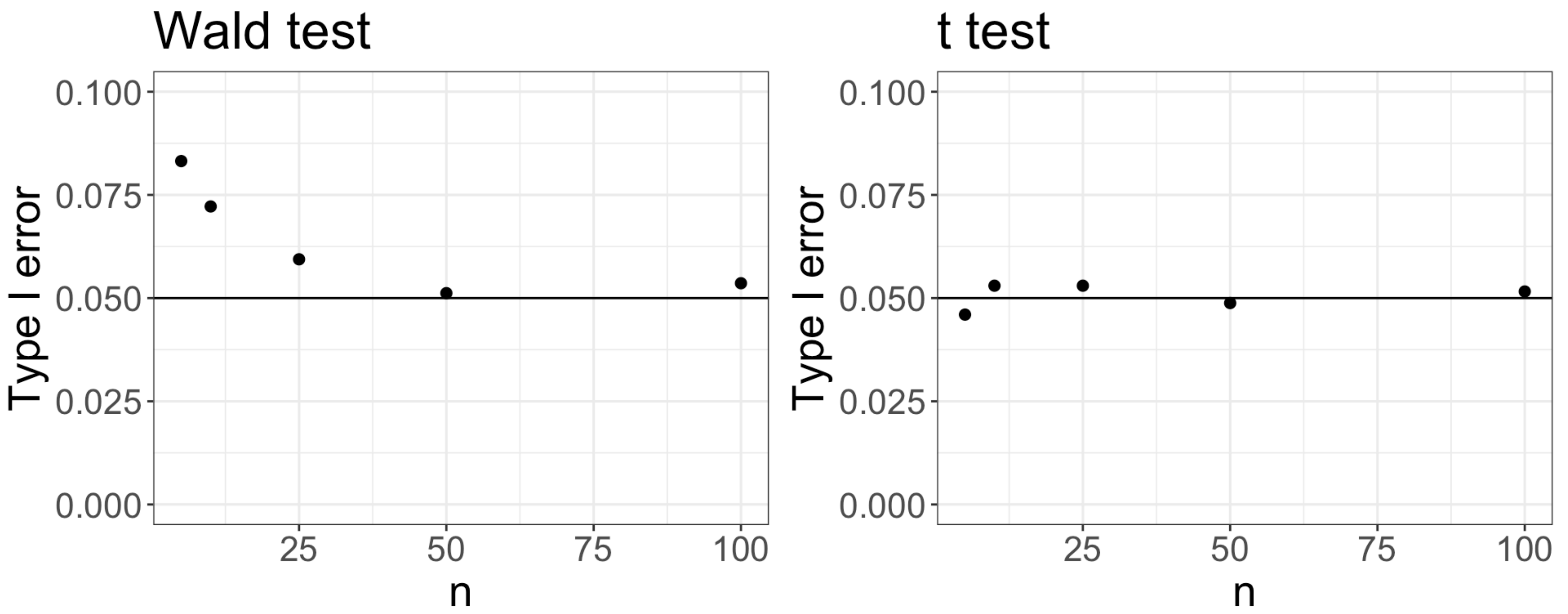
Class activity

Type I error rate with Normal distribution:



Class activity

Wald test vs. t-test:



Philosophical question

- **Position 1:** We should always use a Wald test to test hypotheses about a population mean
- **Position 2:** We should always use a t-test to test hypotheses about a population mean

With which position do you agree?