Lecture 22: t-tests

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 $\frac{\sqrt{n}(\overline{X}_n - \mu)}{s}$?

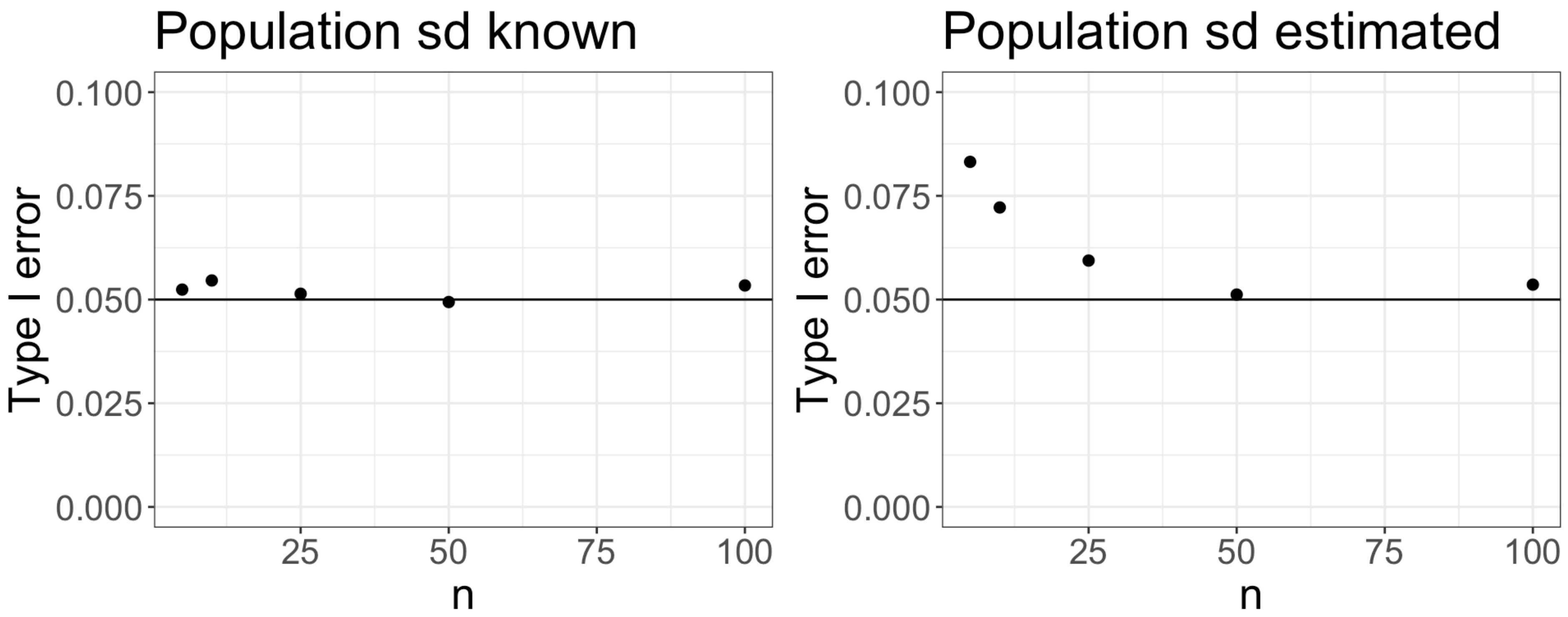
t-tests

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

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

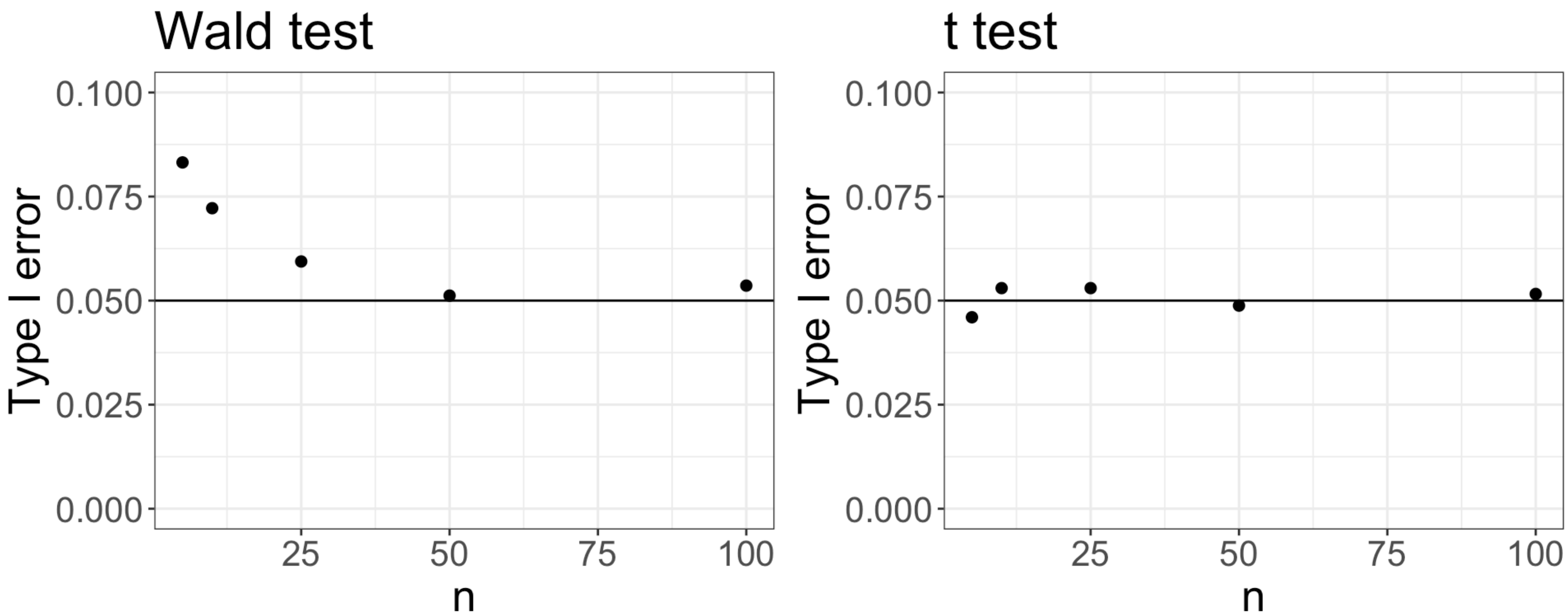
Class activity

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?

t distribution

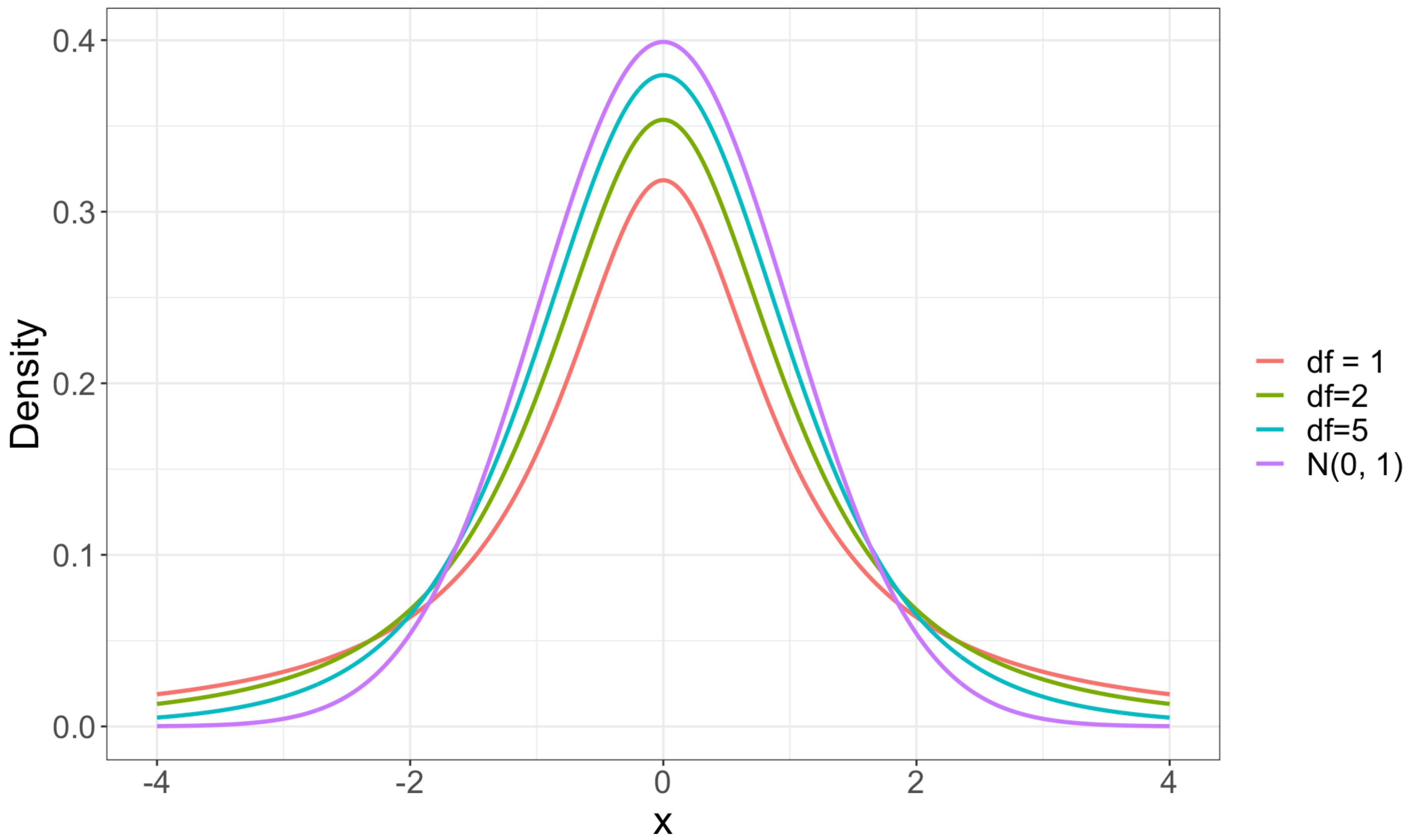
If $X_1, \ldots, X_n \overset{iid}{\sim} N(\mu, \sigma^2)$, then

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

Definition: Let $Z \sim N(0, 1)$ and $V \sim \chi^2_d$ be independent. Then

$$T = \frac{Z}{\sqrt{V/d}} \sim t_d$$

t-distribution



Cochran's theorem

Let $Z_1, \ldots, Z_n \overset{iid}{\sim} N(0,1)$, and let $Z = [Z_1, \ldots, Z_n]^T$. Let $A_1, \ldots, A_k \in \mathbb{R}^{n \times n}$ be symmetric matrices such that $Z^T Z = \sum_{i=1}^{k} Z^T A_i Z$, and let $r_i = rank(A_i)$. Then the following are equivalent:

- $r_1 + \cdots + r_k = n$
- The $Z^T A_i Z$ are independent
- Each $Z^T A_i Z \sim \chi^2_{r_i}$

Application to t-tests