Biostatistics for Health Care Researchers: A Short Course

Hypothesis Testing and Confidence Interval Estimation

Presented by:

Susan M. Perkins, Ph.D.

Division of Biostatistics

Indiana University School of Medicine

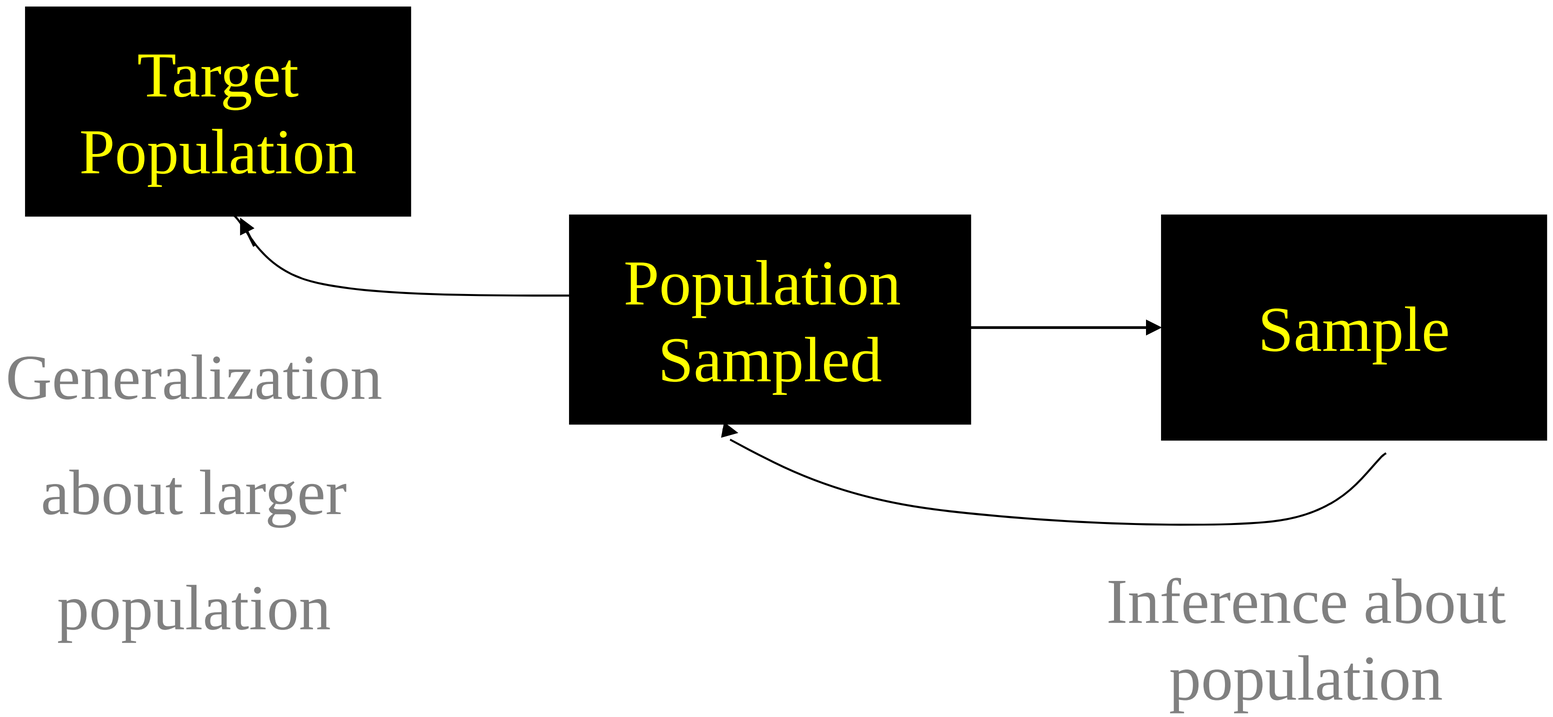
Objectives

- Introduce concepts of hypothesis testing and confidence interval estimation
- Understand how decisions are made based on the sampled data
- Distinguish statistical and clinical significance

Examples

- Do children with new onset seizures have lower levels of neuropsychological functioning than a normative sample?
- Are women of childbearing age drinking more caffeine than recommended?

Statistical Inference



Example 1

- Target Population: All children ages 6-14 with new onset seizures
- Population Sampled: Children in Indiana and in Cincinnati area with new onset seizures
- Sample: Children that participate in study

Example 2

- Target Population: All women age 16-45 in UK
- Population Sampled: women in Manchester area
- Sample: women recruited from business offices

Hypothesis Testing: Basic Concepts

- Null and Alternative (H_0 and H_1)
- One-sided and two-sided tests
- Test statistic and its sampling distribution
- Rejection region
- Type I and Type II errors
- Sample size and power

Example 1

Question: Is processing speed (measured by WISC-III Coding) for children with new onset seizures different the normative score of 10?

Population:

- Normally distributed
- Standard deviation σ=3

Sample summary:

- $n = 282$, $\bar{X} = 9.8$

Fastenau et al. under review

Example 1 (Continued)

- Null Hypothesis $H_0 : \mu = 10$
- Alternative Hypotheses:

  One-sided: $H_1 : \mu < 10$, or $H_1 : \mu > 10$

  Two-sided: $H_1 : \mu \neq 10$
- In this case, we are interested in

  $H_0 : \mu = 10$ versus $H_1 : \mu \neq 10$

Test Statistic

Two characteristics:

- The value of a test statistic depends on a sample
- The sampling distribution of a test statistic under the null hypothesis is known

One Sample Test of the Mean

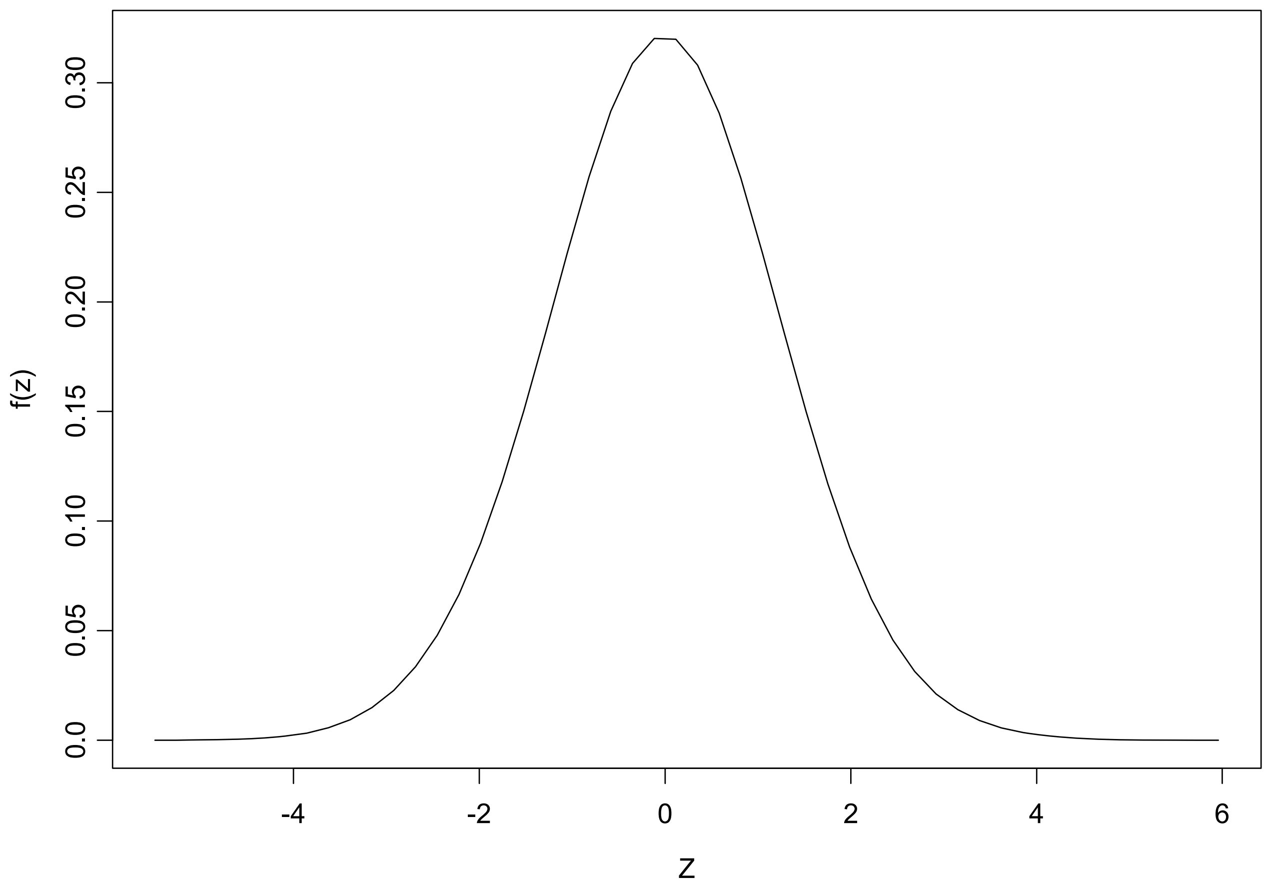
- When the sample is from a normal population (OR when the sample size is large), one can use Z-Test to test the mean of the population
- Test statistic

$$Z = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}}$$

Central Limit Theorem and the Sampling Distribution of Z Under H_0

- CLT: Let $X_1, \ldots X_n$ be a random sample from a population with mean μ and standard deviation σ.
- If H_0: $\mu=\mu_0$ is true, the distribution of Z can be approximated by a N(0,1) distribution.

Sampling Distribution of Z



- Centered at 0
- Most values between ±5
- Area under curve = 1

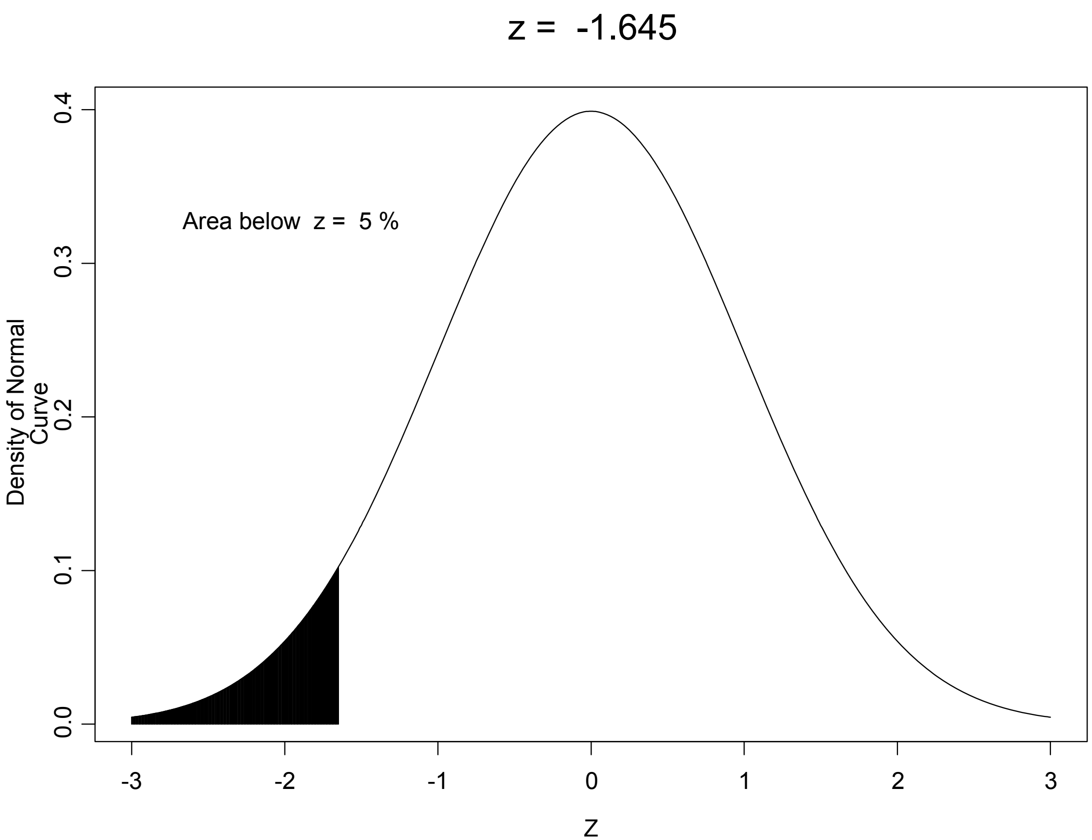
Steps in Hypothesis Testing

- Assume H_0 is true
- Under H_0, Z will have a known sampling distribution (normal distribution)
- If the Z from your sample takes an unlikely (large) value, H_0 is not likely to be true (proof by contraction).

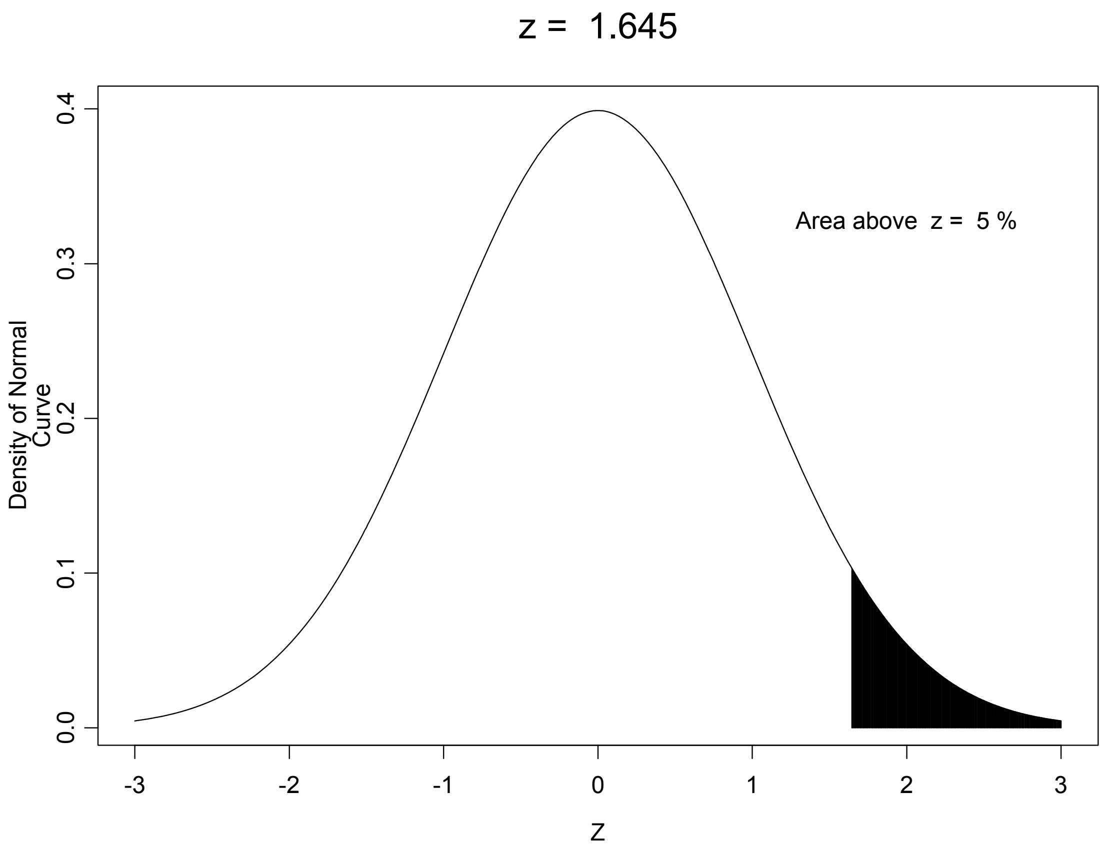
Significance Level and Rejection Region

- We reject the null hypothesis when the Z value is “large”. But how large is “large”?
- Significance level (α): The maximum probability of incorrectly rejecting H_0 you are willing to accept.
- Rejection region: We reject H_0 if the value of Z falls into the region.
- Rejection Region depends on H_1.

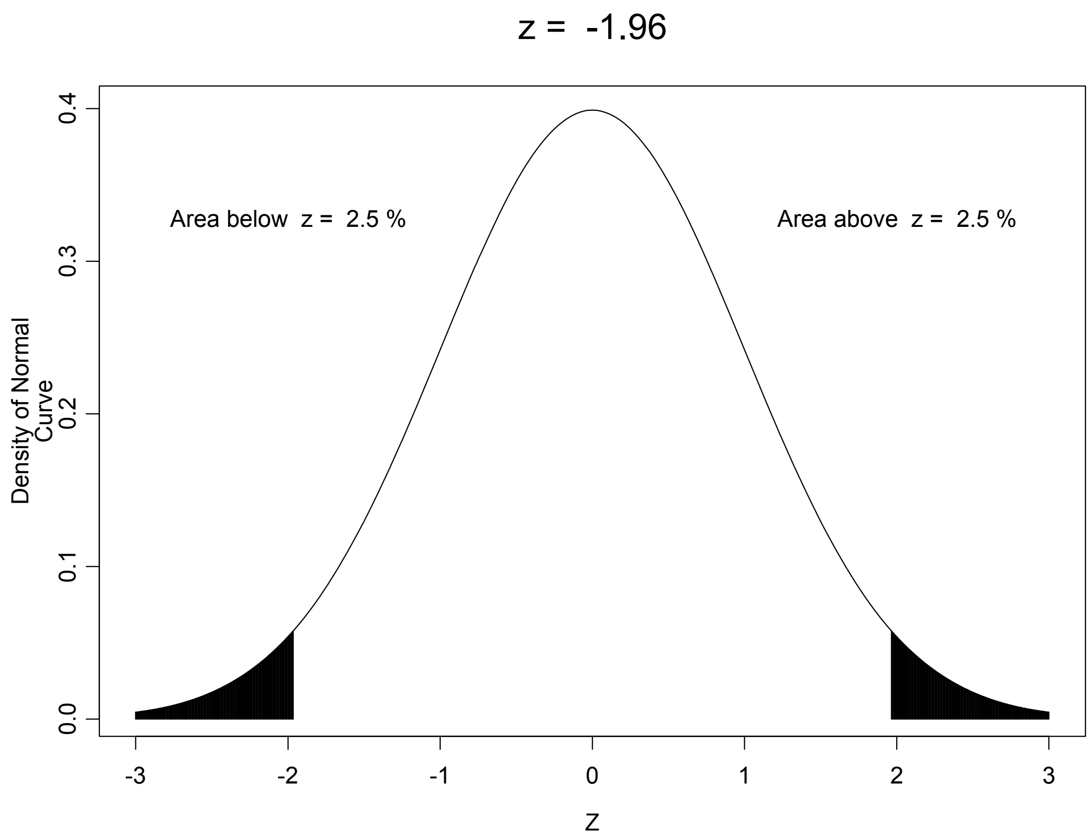
Rejection Region for H_1: $\mu<\mu_0$ (α=0.05)



Rejection Region for H_1: $\mu>\mu_0$ (α=0.05)



Rejection Region for H_1: $\mu \neq \mu_0$ (α=0.05)



Significance Level and p-value

- Significance level is also called α (alpha) – decided up-front
- If the probability of the observed results is below this number, you reject the null hypothesis
- P-value: probability of obtaining a result as or more extreme than the observed data if H_0 is true – calculated from the data
- A small p-value leads to the rejection of the null hypothesis

Example 1 (Continued)

- H_0: μ= 10 versus H_1: $\mu \neq$ 10
- α=0.05

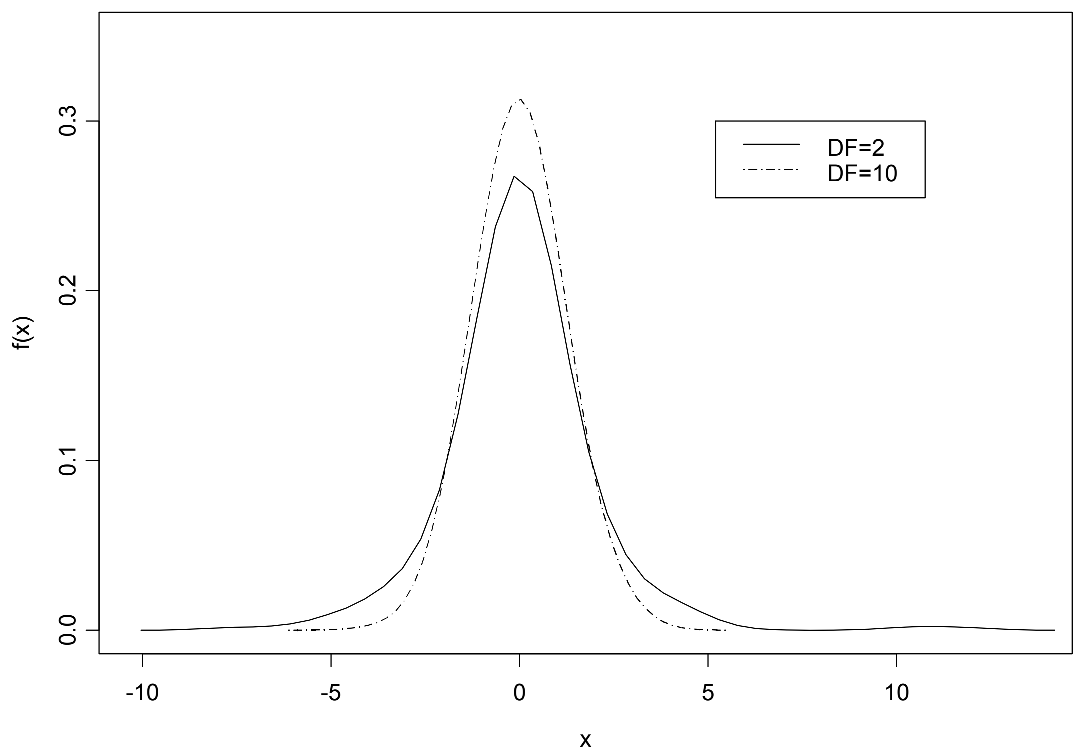
$$Z = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}} = \frac{9.8 - 10}{3 / \sqrt{282}} = 1.1$$

- Rejection Region: Z<-1.96, Z>1.96
- P-value = .26

One Sample t Test

- When the population standard deviation σ is unknown, use sample standard deviation *s* to estimate σ.
- Test statistics $T = \dfrac{\bar{X} - \mu_0}{s / \sqrt{n}}$
- T has t-distribution with *n*-1 degrees of freedom

T Distribution with Various Degrees of Freedom



Example 2

Question: Is the mean intake of caffeine in Women of childbearing age different that the FDA guidelines of 300 mg/day?

Population:

- Normally distributed
- Standard deviation s=128 mg/day

Sample summary:

- $n = 70$, $\bar{X} = 174$ mg/day

Derbyshire & Abdula. Habitual Caffeine Intake in Women of Childbearing Age. J Hum Nutr Diet, 21 (159-164)

Example 2 (Continued)

- H_0: μ= 300 versus H_1: $\mu \neq$ 300
- α=0.05

$$T = \frac{\bar{X} - \mu_0}{s/\sqrt{n}} = \frac{174 - 300}{128/\sqrt{70}} = -8.2$$

- Rejection Region: Z<-1.96, Z>1.96
- P-value < .0001

Type I and Type II Errors

	Null Hypothesis	
Decision	True	False
Reject H_0	Type I (α)	Correct
Not Reject H_0	Correct	Type II (β)

Type I and Type II Errors

- Type I : Pr{rejecting H_0| H_0 true} = α
- Type II: Pr{not rejecting H_0| H_1 true} = β
- Pr{rejecting H_0| H_1 is true} = 1-β = Power
- Typically use: α = 0.05; Power = 0.80 for hypothesis testing and calculating sample sizes

Sample Size and Power

- How does sample size affect the power?
- Example Z statistic $Z = \dfrac{\bar{X} - \mu_0}{\sigma / \sqrt{n}}$
- One can always increase the power of a test by increasing the sample size
- Sample size calculation in the planning stage

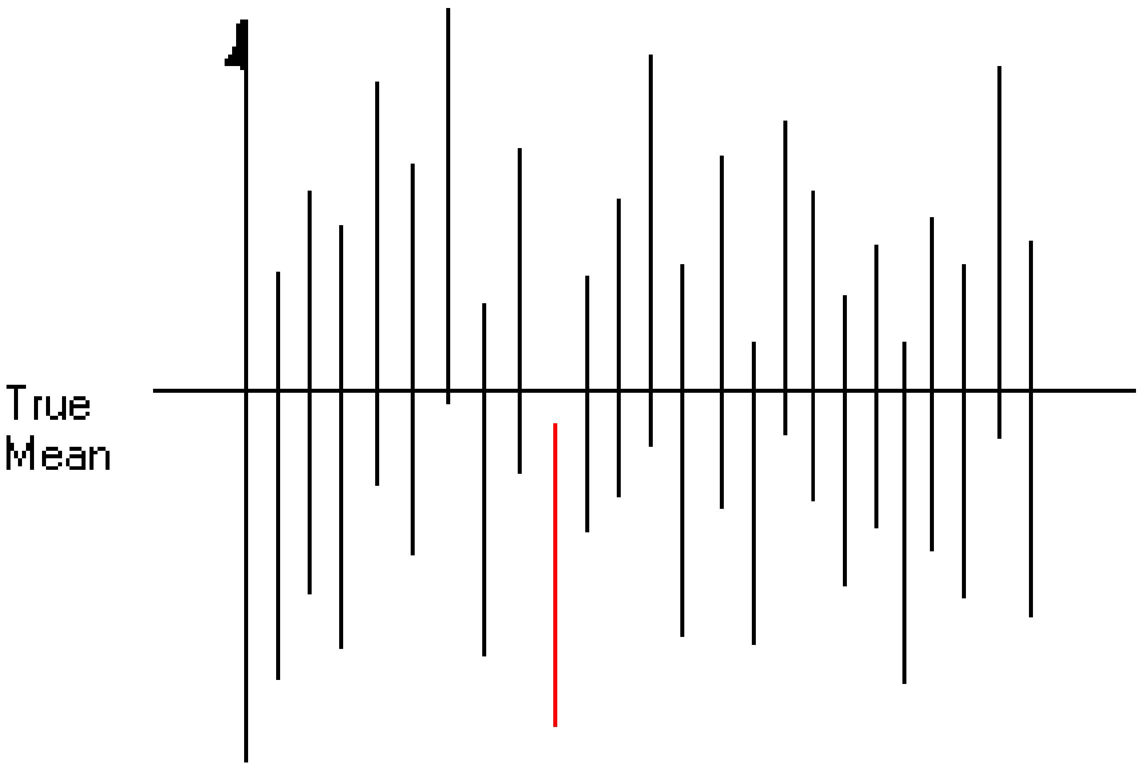
Clinical and Statistical Significance

- The two do not always agree
- Clinically significant but not statistically significant: *n* too small
- Statistically significant but not clinically significant: *n* too large
- Think about what clinical difference is important when designing a study or critiquing someone else's study.

Confidence Intervals

- 100(1-α)% confidence interval for the mean
- σ known: $\bar{X} \pm Z_{\alpha/2}\sigma/\sqrt{n}$
- σ unknown: $\bar{X} \pm t_{\alpha/2(n-1)} s/\sqrt{n}$
- Interpretation: 100(1-α)% confident that this interval will contain the true mean μ

Interpreting Confidence Intervals



Example: Confidence Interval

- 95% CI for mean processing speed

$$9.8 \pm 1.96 \times 3 / \sqrt{282}$$

- (9.4, 10.2): We are 95% confident that the interval between 9.4 and 10.2 captures the true mean processing speed for this group.

Summary

- Hypothesis testing: concepts and procedure
- Errors, power and sample size
- Confidence intervals