Lecture 3

Statistical background review (II)

MCB 416A/516A
Statistical Bioinformatics and Genomic Analysis

Prof. Lingling An
Univ of Arizona

Last time we reviewed:

- Population – parameter
- Sample – statistic – estimate
- Mean, standard deviation
- Normal distribution
- Central limit theorem
- T-distribution
- Confidence interval
    - One sample case
        - for population mean with *known* standard deviation
        - for population mean with *unknown* standard deviation
    - Two sample case
        - for difference of two population means with *known* s.d.
        - for difference of two population means with *unknown* s.d.
        - for difference of paired samples

Hypotheses tests

A **test of statistical significance** tests a specific hypothesis (an assumption, or a theory about the characteristics of one or more variables in one or more populations) using sample data to decide on the validity of the hypothesis.

Example*:* a researcher implements protocols for performing intubation on pediatric patients in the prehospital setting. To evaluate whether these protocols were successful in improving intubation rates, he measures the intubation rate over time in one group randomly assigned to training in the new protocols, and compare this to the intubation rate over time in another control group that did not receive training in the new protocols.

Hypothesis tests - procedure

Follow these steps:

1) State the null and alternative hypotheses

2) Calculate the test statistic

3) Get the corresponding p-value

(the chance of obtaining a test statistic **at least** as extreme as the one that was actually observed result, assume the null hypothesis is true).

4) Draw a conclusion

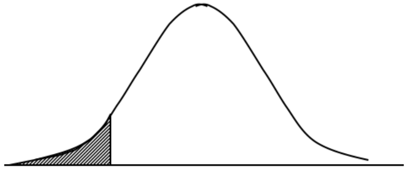
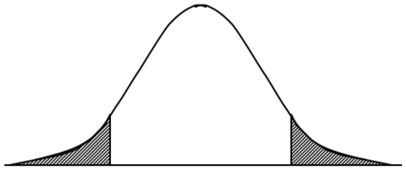
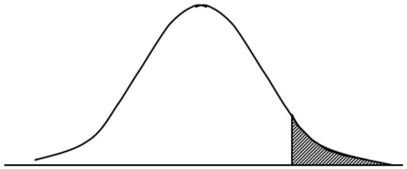
- ❖ If the *P*-value is equal to or less than *α* (***p* ≤ *α***), then we **reject *H*₀**.
- ❖ If the *P*-value is greater than *α* (***p* > *α***), then we **fail to reject *H*₀**.

The significance level, *α*, is the largest *P*-value tolerated for rejecting a true null hypothesis

One-sided and two-sided tests

- Two-sided test
  - No *a priori* reason 1 group should have stronger effect
  - Used for most tests
- One-sided test
  - Specific interest in only one direction

Example: one sample case

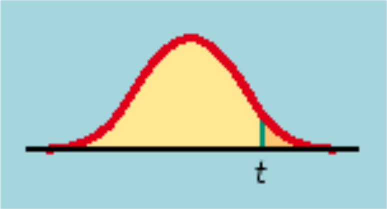
One-sided Test (left tail)	Two-sided Test (both tails)	One-sided Test (right tail)
$H_0 : \mu = \mu_0$ $H_1 : \mu < \mu_0$	$H_0 : \mu = \mu_0$ $H_1 : \mu \neq \mu_0$	$H_0 : \mu = \mu_0$ $H_1 : \mu > \mu_0$
		

P-value in one-sided and two-sided tests

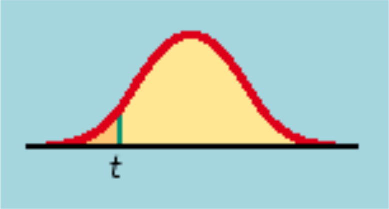
use one sample with unknown population standard deviation as an example:

One-sided (one-tailed)

$H_a: \mu > \mu_0 \Longrightarrow P(T \geq t)$



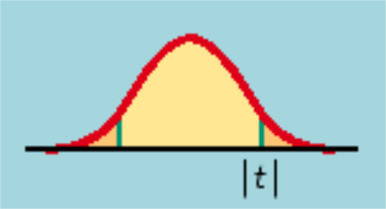
$H_a: \mu < \mu_0 \Longrightarrow P(T \leq t)$



$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

Two-sided (two-tailed)

$H_a: \mu \neq \mu_0 \Longrightarrow 2P(T \geq |t|)$

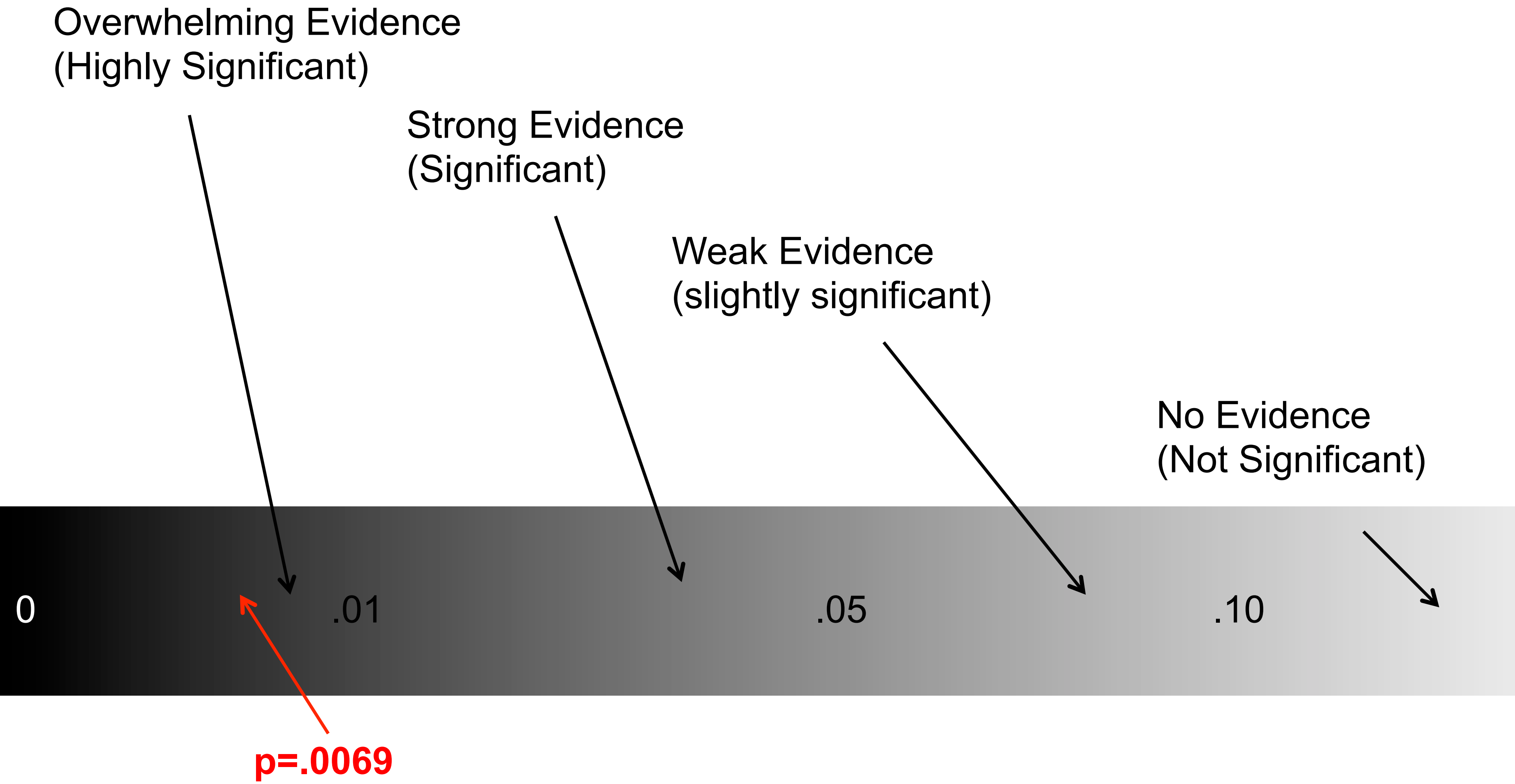


when the standard deviation is known, everything is same except $t \longrightarrow z$ (and use z-table).

T-table …

- To find the p-value for t-statistic in t-table, df=n-1 (one population/sample case)
- What happens if your degrees of freedom isn’t on the table, for example df = 79? Always round DOWN to the next lowest degrees of freedom to be conservative.
- No worries if use calculator or computer

Interpreting the p-value… if no significance level is pre-specified/given



Overwhelming Evidence
(Highly Significant)

Strong Evidence
(Significant)

Weak Evidence
(slightly significant)

No Evidence
(Not Significant)

0

.01

.05

.10

p=.0069

Sample size

- For the confidence interval:

  estimate ± margin of error

  =estimate ± critical value * Std. Error(estimate)

- Margin of error (*E*): maximum difference between the true parameter and its estimate from the sample.
- If specify a margin of error, then the sample size *n* can be solved through:

  critical value * std. error (estimate)=*E*

Eg. the sample mean from a population with known std. deviation:

$$Z_{\alpha/2}*\sigma/\sqrt{n}=E$$

i.e., $n= \{Z_{\alpha/2}*\sigma/E\}$^2

Then any number larger than $\{Z_{\alpha/2}*\sigma/E\}$^2 can result in smaller *E,* but ----

- Cautions about P-Values

Sample size directly impacts the p-value. Large sample sizes produce small p-values even when differences between groups are not meaningful. You should always verify the practical relevance of your results.

On the other hand, a sample size that is too small can result in a failure to identify a difference when one truly exists.

$$t = \frac{\bar{x} - \mu}{s / \sqrt{n}} = \sqrt{n}\,\frac{(\bar{x} - \mu)}{s}$$

$$z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} = \sqrt{n}\,\frac{(\bar{x} - \mu)}{\sigma}$$

=> Same size justification, (t-test, z-test, type II error, etc.)

Hypothesis Testing: Comparing two populations

- We have **two independent SRSs** (simple random samples) coming from two distinct populations (like men vs. women) with (μ_1,s_1) and (μ_2,s_2) unknown. should be σ
- Both populations should be Normally distributed. We can use some tests to check whether the normality assumption is satisfied.
- In practice, it is enough that the two distributions have similar shapes and that the sample data contain no strong outliers.

Two-sample *t*-test

The null hypothesis is that both population means μ_1 and μ_2 are equal, thus their difference is equal to zero.

$$H_0\text{: } \mu_1 = \mu_2 \Leftrightarrow \mu_1 - \mu_2 = 0$$

with either a one-sided or a two-sided alternative hypothesis.

We find how many standard errors (SE) away from ($\mu_1 - \mu_2$) is ($\bar{x}_1 - \bar{x}_2$) by standardizing with t:

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{SE}$$

If two standard deviations are not equal, the *t*-statistic is:

$$\boxed{t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}}$$

with **df = min(*n*$_1$ − 1,*n*$_2$ − 1)** approximately.

If two standard deviations are equal, the t-statistic is:

$$\boxed{t = \frac{\bar{x}_1 - \bar{x}_2}{S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}}$$

with **df = (*n*$_1$+*n*$_2$ − 2)**

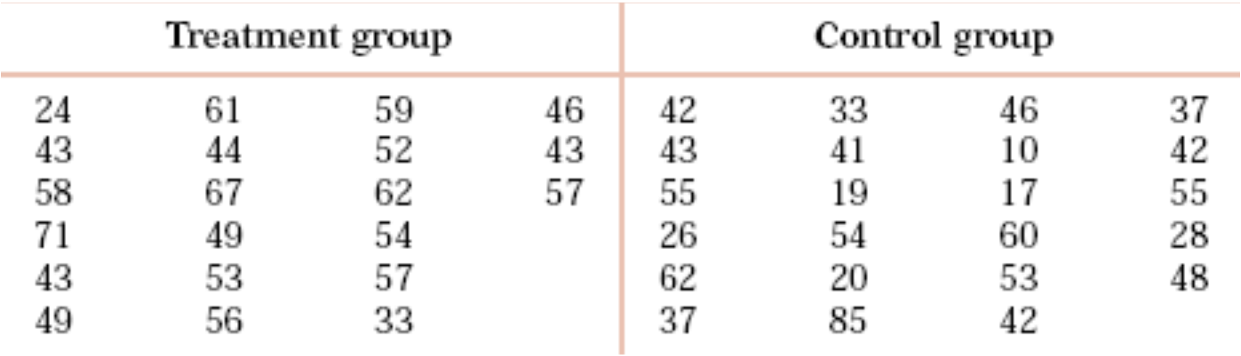
$$S_p = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$$

Relationship between HT and CI

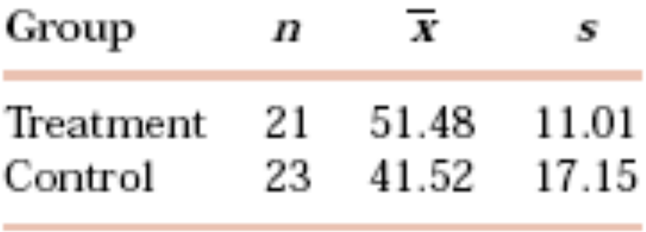
- There is an extremely close relationship between confidence intervals and hypothesis testing.
- When a 95% confidence interval is constructed, all values in the interval are considered plausible values for the parameter being estimated. Values outside the interval are rejected (at significant level = 0.05) as relatively implausible.
  - If the value of the parameter specified by the null hypothesis is contained in the 95% interval then the null hypothesis cannot be rejected at the 0.05 level.
  - If the value specified by the null hypothesis is not in the interval then the null hypothesis can be rejected at the 0.05 level.
- If a 99% confidence interval is constructed, then values outside the interval are rejected at the 0.01 level.

Example: Can directed reading activities in the classroom help improve reading ability?

- A class of 21 third-graders participates in these activities for 8 weeks while a control classroom of 23 third-graders follows the same curriculum without the activities. After the 8 weeks, all children take a reading test (scores in table).



Treatment group				Control group			
24	61	59	46	42	33	46	37
43	44	52	43	43	41	10	42
58	67	62	57	55	19	17	55
71	49	54		26	54	60	28
43	53	57		62	20	53	48
49	56	33		37	85	42	



Group	n	$\overline{x}$	s
Treatment	21	51.48	11.01
Control	23	41.52	17.15

Relationship between HT and CI -2

Confidence Interval approach:

- 1. Construct the (1-α)*100% confidence interval on the difference between means.
- 2. check if zero, the value specified by the null hypothesis, is in the interval.
- 3. draw a conclusion.

Hypothesis testing approach:

- 1. state H_0 and H_a hypothesis
- 2. calculate statistic
- 3. get p-value
- 4. draw a conclusion, compared with significance level α.

Relationship between HT and CI -3

Although the relationship between confidence intervals and hypothesis testing is very close, the objectives of the two methods are different:

- Hypothesis testing relates to a single conclusion of statistical significance vs. no statistical significance.
- Confidence intervals provide a range of plausible values for the population.

Relationship between HT and CI - 4

Which one?

- Use hypothesis testing when you want to do a strict comparison with a pre-specified hypothesis and significance level.

- Use confidence intervals to describe the magnitude of an effect (e.g., mean difference, odds ratio, etc.) or when you want to describe a single sample.

Unknown Truth and the Data

Truth Data	H_0 Correct (no disease)	H_a Correct (disease)
Decide H_0 "fail to reject H_0" (test shows no disease)	True Negative	False Negative
Decide H_a "reject H_0" (test shows disease)	False Positive	True Positive

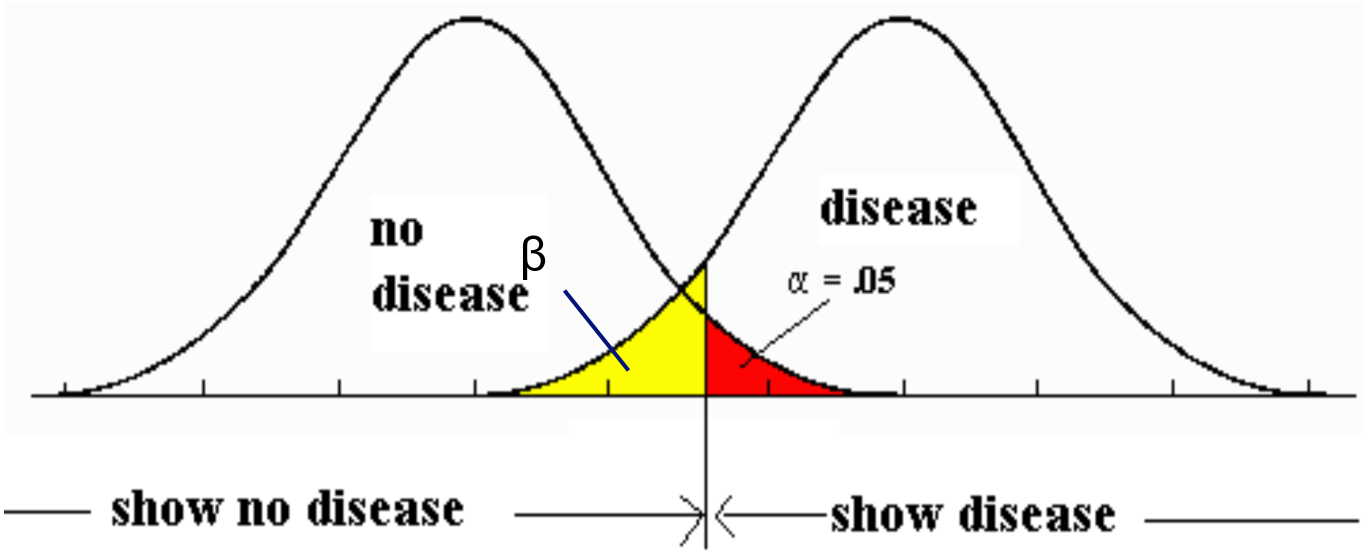
Two types of error

Truth Data	H_0 Correct (no disease)	H_a Correct (disease)
Decide H_0 "fail to reject H_0" (test shows no disease)	1- α True Negative	β False Negative
Decide H_a "reject H_0" (test shows disease)	α False Positive	1- β True Positive

- Type I error : False positive rate
- α = P(reject H_0 | H_0 true)
  — Probability reject the true null hypothesis
- α is significance level

- Type II error: False negative rate
- β = P(do not reject H_0 | H_a true)
  — Probability not reject a false null hypothesis
- **Power** = 1-β = P(reject H_0 | H_a true)

Illustration of Types of error



α decreases -> β increases