Department of Linguistics and Translation
香港城市大學
City University of Hong Kong

Fundamentals of Statistics for Language Sciences LT2206

Jixing Li

Lecture 5: t-tests

Slides adapted from Cecilia Earls

Inference for μ

Scenario: Often we would like to say SOMETHING about the mean of a population we are sampling from.

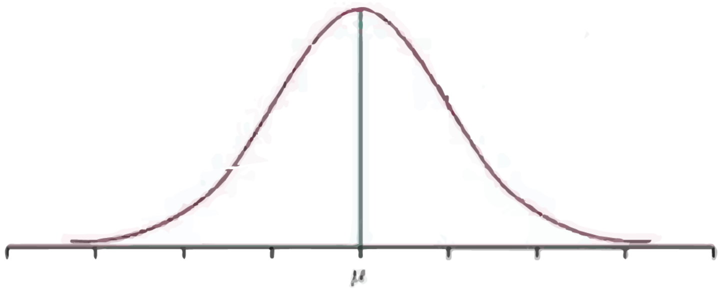
What do we have to work with?

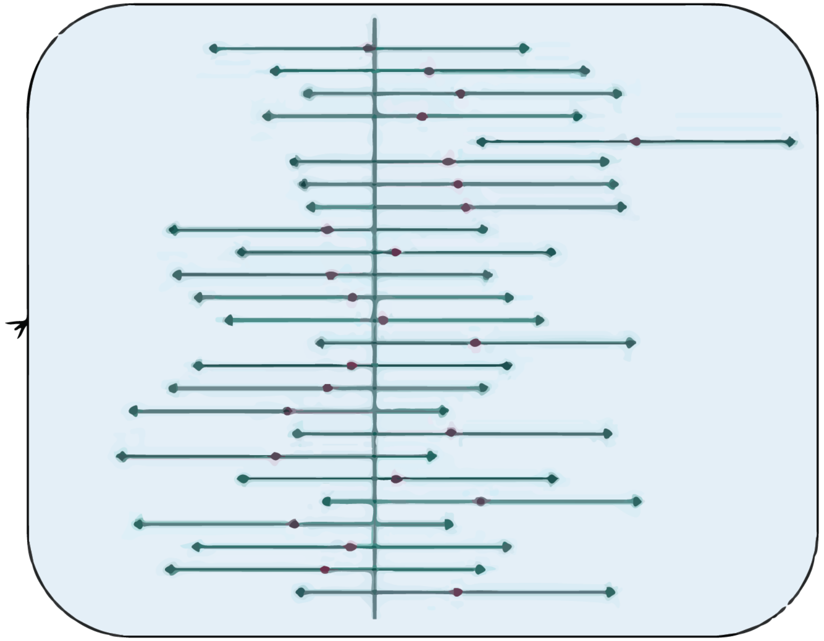
- Take one SRS from this population
  - $X_i, i = 1, \dots, n$
  - $E(X_i) = \mu, Var(X_i) = \sigma^2$

  Basic Assumptions (might have more)

- From our sample, we get an outcome: $\bar{x}_n$ (e.g., 17)

Often know the standardized sampling distribution of $\bar{X}_n$(at least approximately)





Hypothesis Testing

- **First steps:**
  - Set type I error rate, α= P(Reject H_0| H_0 is true)
  - H_a ($\neq$ or $>$ or $<$) & H_0 ($=$ or $\leq$ or $\geq$)
  - Determine sample size to control β= P(Fail to reject H_0 | H_a is true)
  - Collect SRS
- **Determine:**
  - Distribution of standardized test statistic under H_0
  - Using this distribution make a decision:
    - Reject H_0? or Fail to reject H_0?
    - Decision can be based on rejection region
    - Decision can be based on a *p*-value

Suppose under H_0:

$$\frac{\bar{X}_n - \mu_0}{\sigma/\sqrt{n}} \dot{\sim} N(0,1)$$

standardized test statistic — standard normal distribution

Critical value: z_α

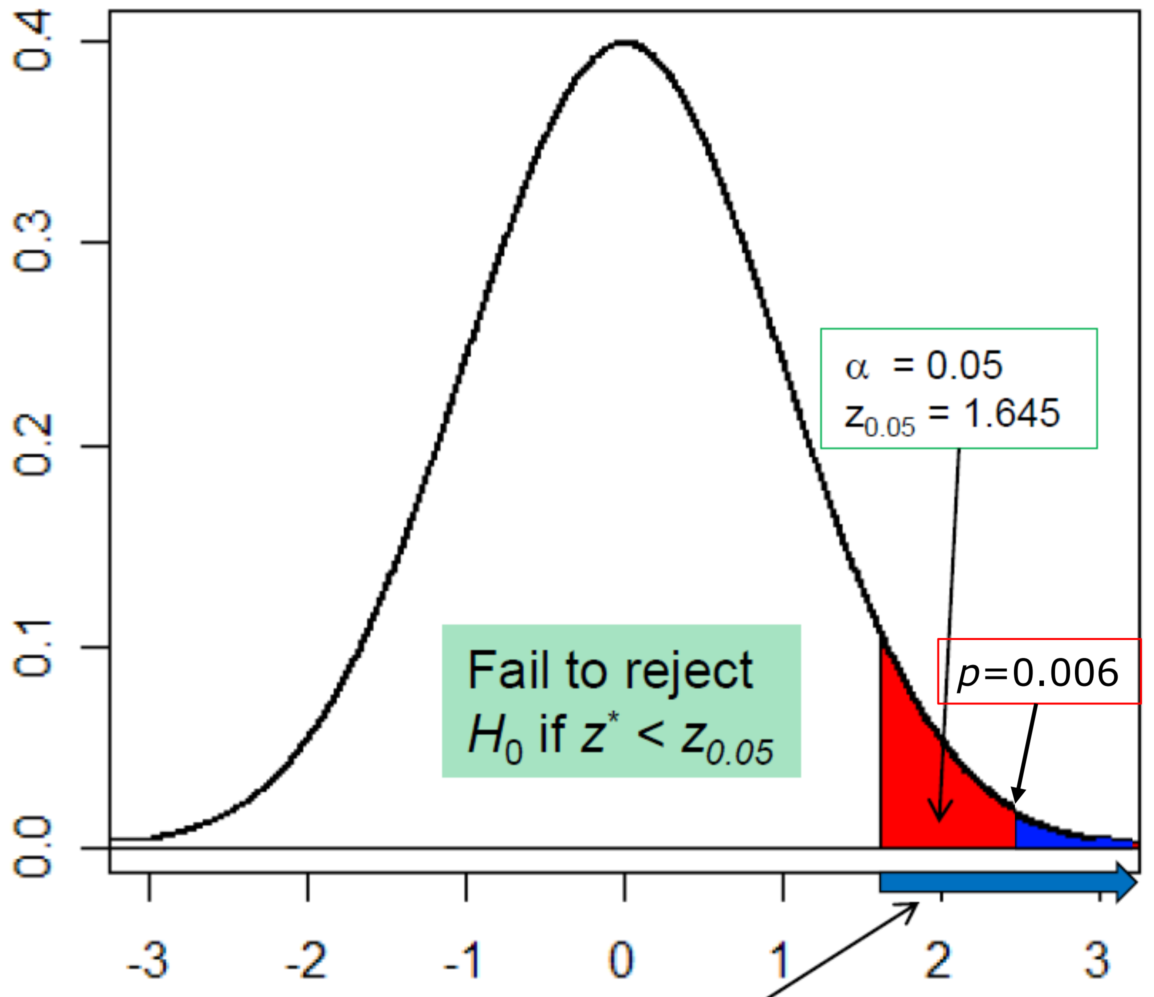
$(1-\alpha)$100th percentile of N(0,1)

In R: $z_{0.05}$= qnorm(1–0.05)=1.645

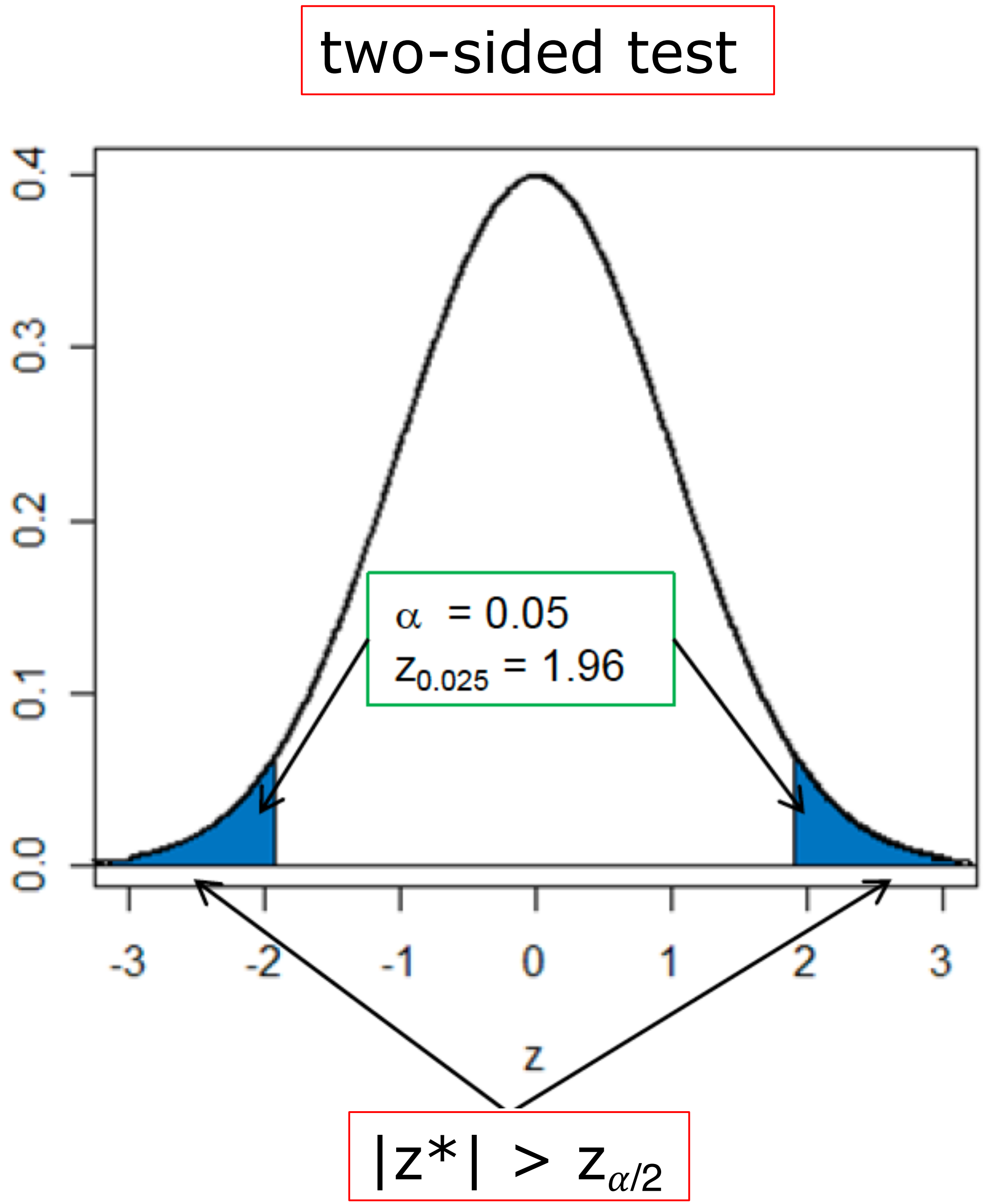
p-value: $P(z_\alpha > z^*)$

In R: p = 1-pnorm(z*)

α = 0.05, one-sided



Reject H_0 if $z^* > z_{0.05}$ or $p < \alpha$



Example

Reading comprehension in L1 and L2

Mean accuracy of L1: 0.85

	L2
N	46
Mean	0.8
SD	0.2

Are L2 speakers' accuracy lower than L1?

Are L2 speakers' accuracy different than L1?

t-distribution

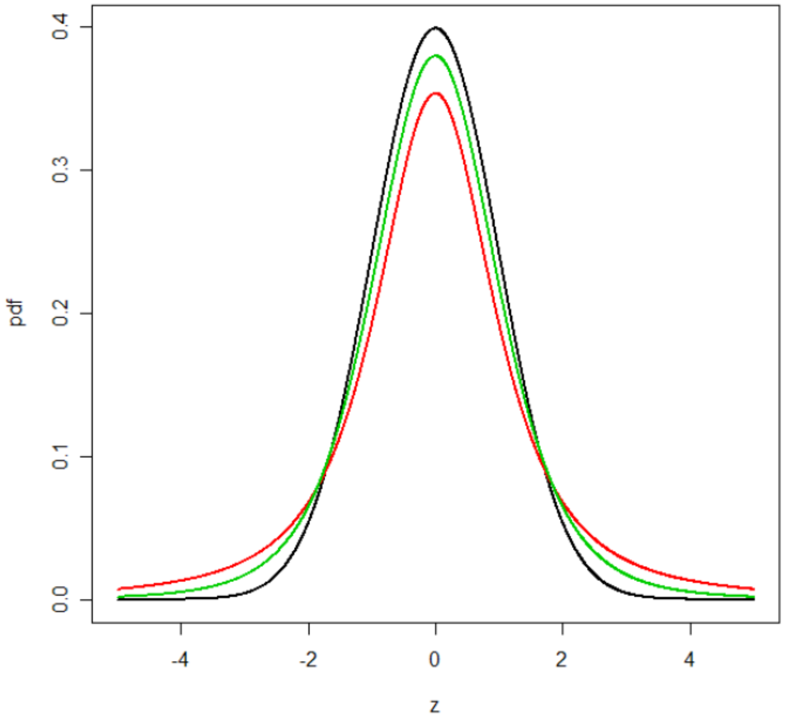
$$\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \sim N(0,1) \text{ does not } \Longrightarrow \frac{\bar{X}_n - \mu}{S/\sqrt{n}} \sim N(0,1)$$

Reason: There is extra variability introduced when we replace a fixed σ by a varying, sample-dependent s.

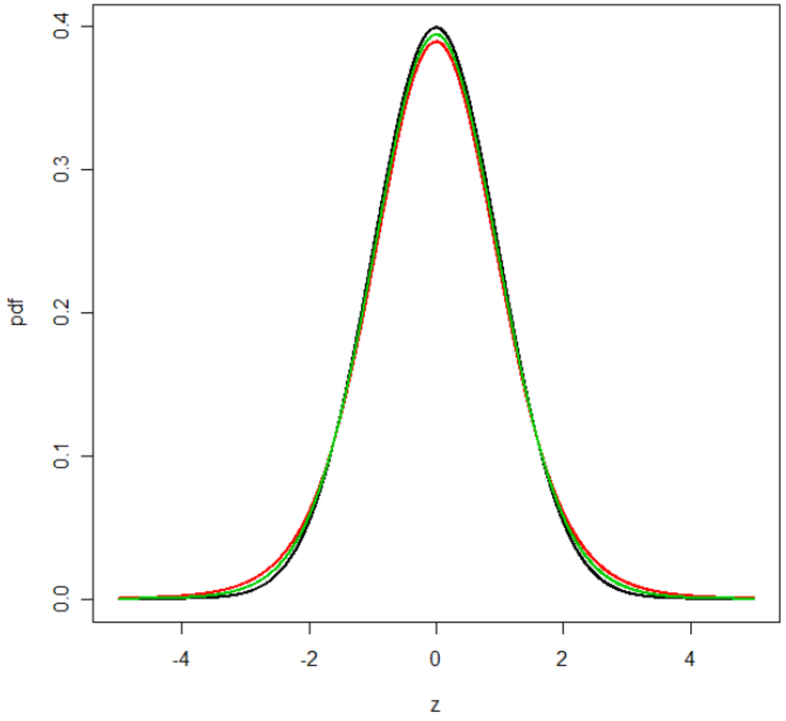
Sampling distribution of $$\frac{\bar{X}_n - \mu_0}{S/\sqrt{n}} \sim t_{n-1}$$

degrees of freedom

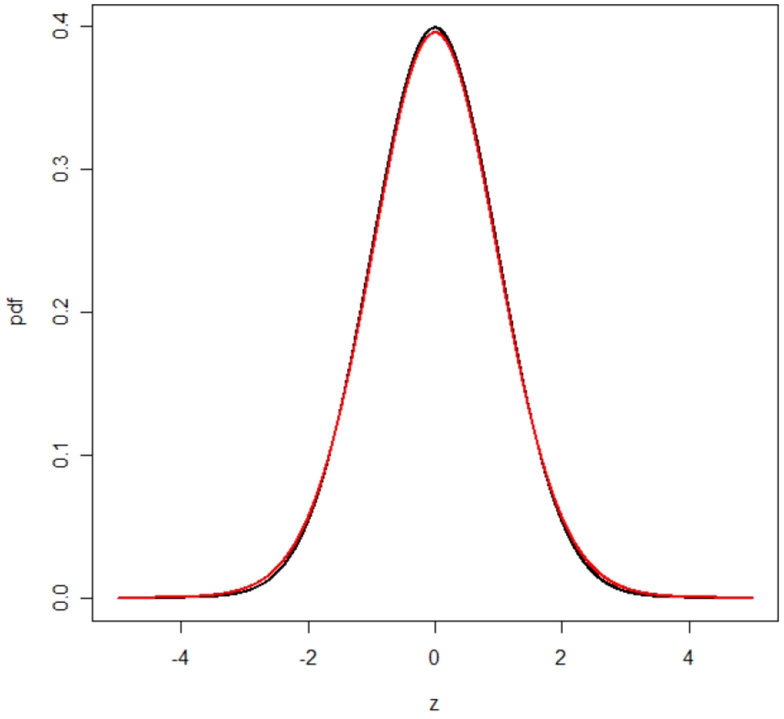
How different is a t_{df} distribution from a N(0,1) distribution?



Black: N(0,1)
Green: t_5
Red: t_2



Black: N(0,1)
Green: t_{20}
Red: t_{10}



Black: N(0,1)

Red: t_{30}

Hypothesis Testing: Normal Samples, σ Unknown

Example: $H_0: \mu \leq \mu_0$
$H_a: \mu > \mu_0$

T.S. $$t^* = \frac{\bar{X}_n - \mu_0}{s/\sqrt{n}}$$

R.R. $t^* > t_{n-1,\alpha}$

p-value $p = P(t_{n-1} > t^*)$

For the two other forms of H_a, the relevant *p*-values are:
$H_a: \mu < \mu_0 \leftrightarrow p = P(t_{n-1} < t^*)$
$H_a: \mu \neq \mu_0 \leftrightarrow p = 2\,P(t_{n-1} > |t^*|)$

Example: Students quiz score

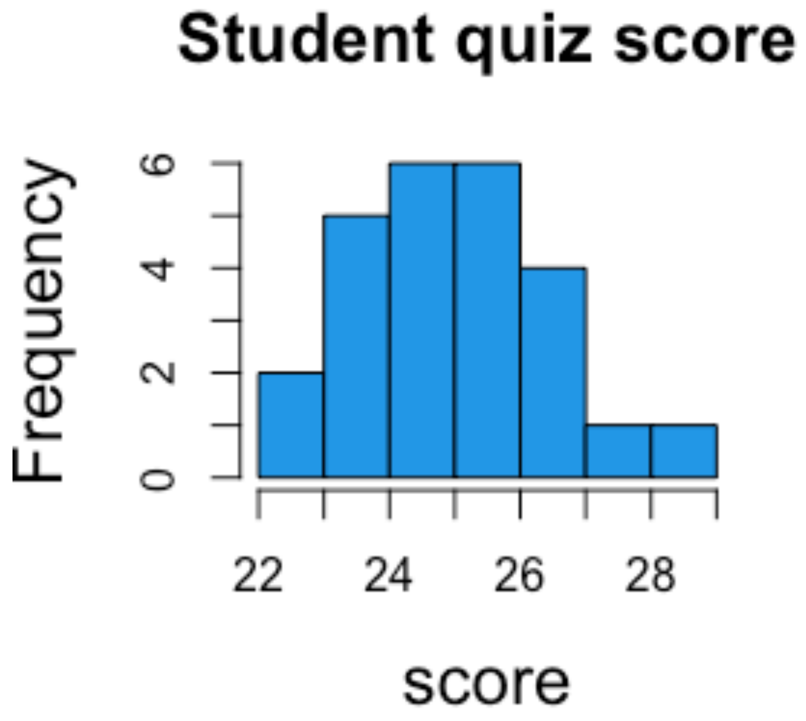
- Average quiz score of 25 students:

  25.8, 24.6, 26.1, 22.9, 25.1, 27.3, 24.0, 24.5, 23.9, 26.2, 24.3, 24.6, 23.3, 25.5, 28.1, 24.8, 23.5, 26.3, 25.4, 25.5, 23.9, 27.0, 24.8, 22.9, 25.4.

- Questions: Based on this SRS, is the current mean different from that in 2022 (24.5)? (Testing)

Look at the data first…

```
> quizscore<-c(25.8, 24.6, 26.1, 22.9, 25.1, 27.3, 24.0, 24.5, 23.9, 26.2,
24.3, 24.6, 23.3, 25.5, 28.1, 24.8, 23.5, 26.3, 25.4, 25.5, 23.9, 27.0, 2
4.8, 22.9, 25.4)
> hist(quizscore,col=4,main="Student quiz score",xlab="score",cex.lab=1.5,
cex.main=1.5)
```



Do these data look like they were sampled from a normal distribution?

25 might not be "large enough" to approximate the sampling distribution of $\frac{\bar{X}_n - \mu}{S/\sqrt{n}}$ by N(0,1)

Test: $H_0: \mu = 24.5$ vs. $H_a: \mu \neq 24.5$

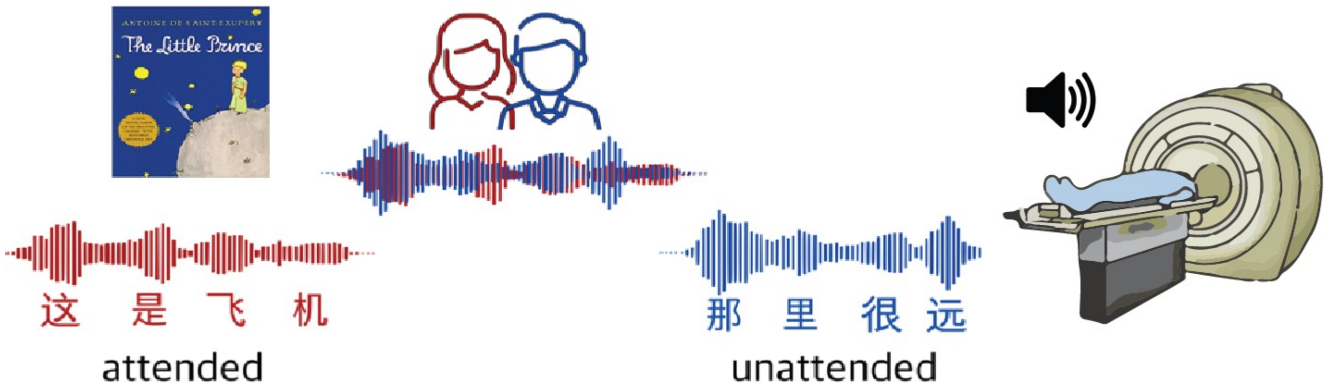
t.test() reports all output available for the test specified by the user

```
> t.test(quizscore,alternative="two.sided",mu=24.5)

        One Sample t-test

data:  quizscore
t = 1.9675, df = 24, p-value = 0.06079
alternative hypothesis: true mean is not equal to 24.5
95 percent confidence interval:
 24.47413 25.58187
sample estimates:
mean of x
   25.028
```

```
> 2*(1-pt(1.9675,24))
[1] 0.06078647
```

Inference for two samples

Difference in analysis: Form of the test statistic depends on whether the samples are independent or paired



two participant groups:
hearing-impaired
normal hearing

- Paired → inference for the mean difference is identical to inference for μ
- Independent → can we assume $\sigma_1 = \sigma_2$?
  - Equal variance assumption
  - Unequal variance assumption

Example: Cocktail party experiment

- Both hearing-impaired (N=49) and normal hearing participants (N=45) listened to mixed speech and single-talker speech in the fMRI scanner.
- Participants reported self-rated intelligibility score (on a 1-7 Likert scale) for each speech after the experiment

- **Question 1:** Is mixed speech more difficult to understand than single-talker speech for hearing-impaired and normal hearing listeners?
- **Question 2:** Are hearing-impaired listeners have more difficulty understanding mixed and single-talker speech than normal hearing listeners?

Sampling distributions for sums and differences of independent normal means

- Let $\bar{X}$ be the sample mean from a SRS of size n_x from a $N(\mu_x, \sigma_x^2)$ population.
- Let $\bar{Y}$ be the sample mean from a SRS of size n_y from a $N(\mu_y, \sigma_y^2)$ population.
- Assume that the two populations and the samples from them are independent. Then:

$$\bar{X} \pm \bar{Y} \sim N(\mu_x \pm \mu_y, \frac{\sigma_x^2}{n_x} + \frac{\sigma_y^2}{n_y})$$

H_o: $\mu_1 - \mu_2 \leq D_0$ vs. H_a: $\mu_1 - \mu_2 > D_0$, reject H_o if $t^* > t_{\alpha, df}$

H_o: $\mu_1 - \mu_2 \geq D_0$ vs. H_a: $\mu_1 - \mu_2 < D_0$, reject H_o if $t^* < -t_{\alpha, df}$

H_o: $\mu_1 - \mu_2 = D_0$ vs. H_a: $\mu_1 - \mu_2 \neq D_0$, reject H_o if $|t^*| > t_{\alpha/2, df}$

$$\text{where: } T^* = \frac{(\bar{X}_1 - \bar{X}_2) - D_0}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

$$df = \frac{(n_1 - 1)(n_2 - 1)}{(1-c)^2(n_1 - 1) + c^2(n_2 - 1)} \text{ for } c = \frac{s_1^2/n_1}{s_1^2/n_1 + s_2^2/n_2}$$