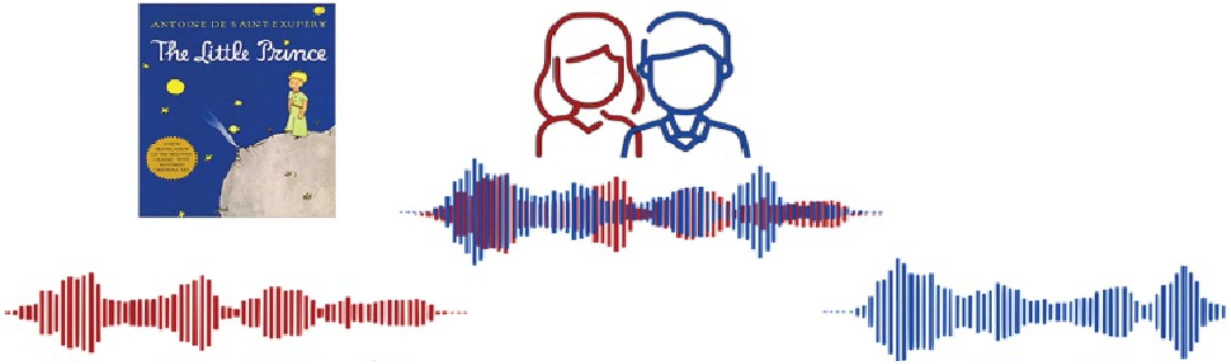
Department of Linguistics and Translation
香港城市大學
City University of Hong Kong

Fundamentals of Statistics for Language Sciences LT2206

Jixing Li

Lecture 7: ANOVA II

Example: Cocktail party experiment



After listening, rate how clear the speech was from 1-5

1: not clear at all

5: very clear

Participants:

hearing-impaired adults (N=45): High-frequency hearing loss (> 8kz)

normal hearing adults (N=49)

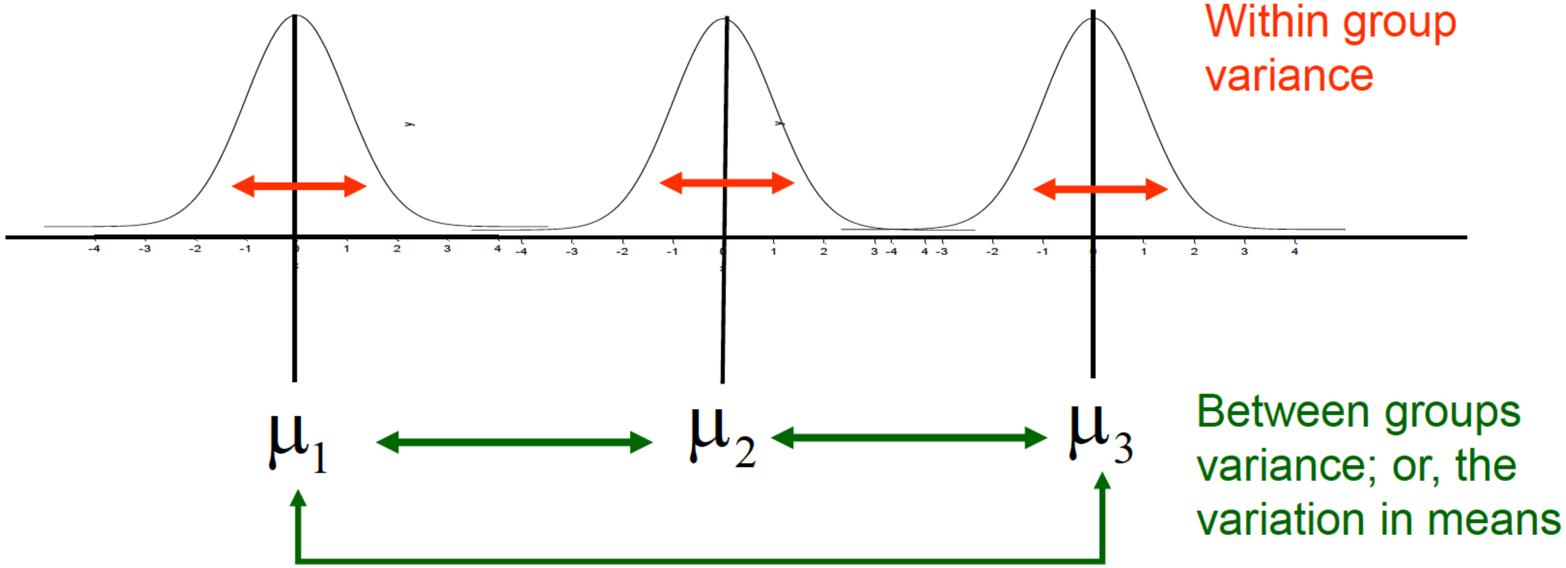
children (N=47)

One-way ANOVA

1. **Compute sum of squares between group (SSB):**

$$\sum_{1}^{k} n_k(\overline{X_k} - \overline{X})^2$$

2. **Compute sum of squares within group / error (SSE):**
   sum of the squared differences between each individual observation and the group mean of that observation.
3. **Compute sum of squares total (SST): SSB+SSE**



The ANOVA table

Source of variation	Degrees of freedom	Sum of squares	Mean squares	*F* statistic
Between	k-1	SSB	MSB =SSB/(k-1)	MSB/MSE
Within (Error)	n-k	SSE	MSE =SSE/(n-k)	
Total	n-1	SST		

Under H_0, *F* should tend to be close to 1. Under H_a, *F* should exceed 1, by an amount depending on both *n* and *k*.

k**:** number of groups
n**:** total number of samples

Example: The cocktail party experiment

group	mixed	mean	grand mean
hearing-impaired	2,2,3	2.33	2.78
normal	2,4,4	3.33	
children	2,3,3	2.67	

SSB = $3*(2.33-2.78)^2 + 3*(3.33-2.78)^2 + 3*(2.67-2.78)^2 = 2.77$

SSE = $(2-2.33)^2 + (2-2.33)^2 + (3-2.33)^2$
$+ (2-3.33)^2 + (4-3.33)^2 + (4-3.33)^2$
$+ (2-2.67)^2 + (3-2.67)^2 + (3-2.67)^2 = 4$

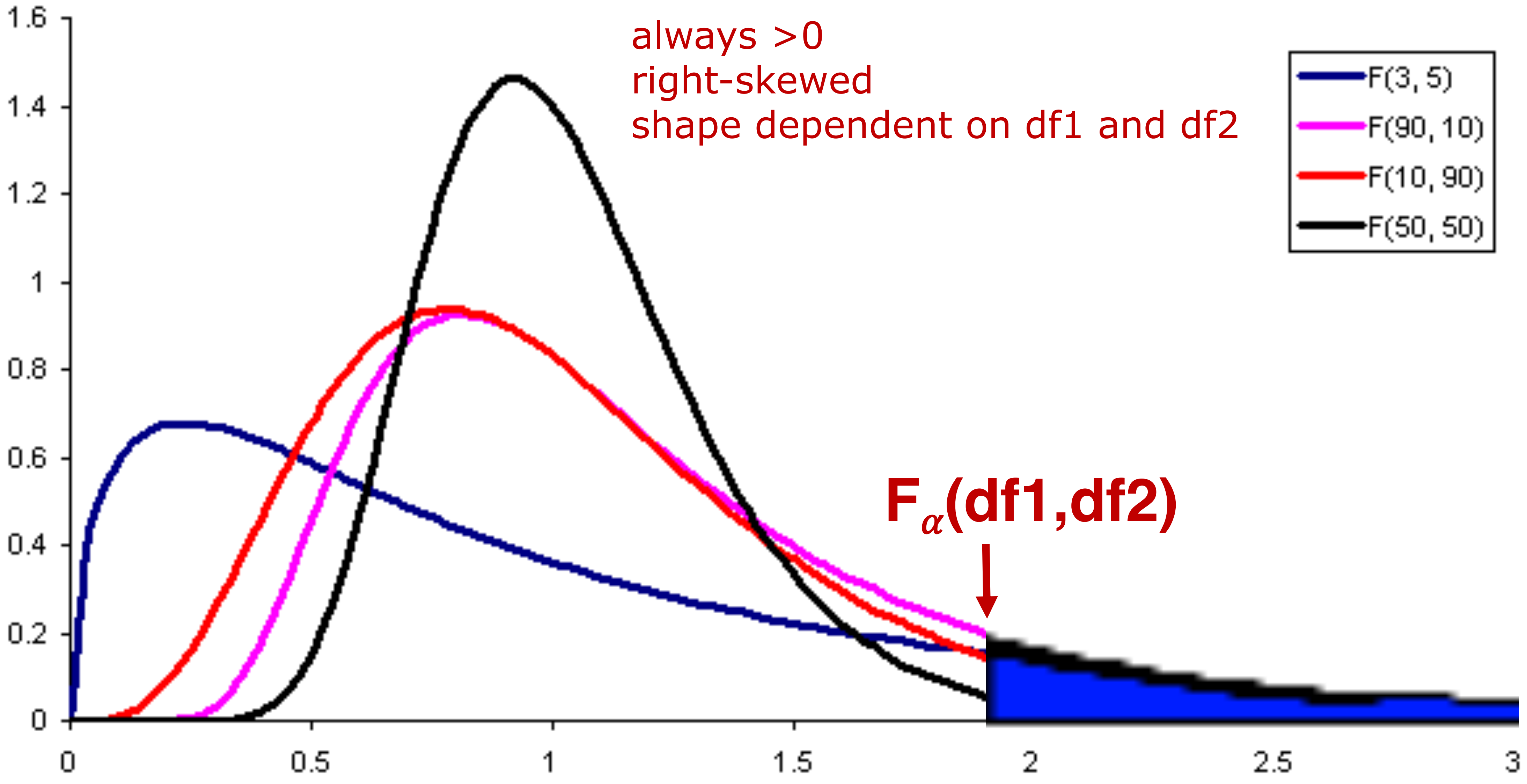
k = 3, **n** = 9

MSB= SSB / (k-1) = 2.77/2 = 1.385

MSE = SSE / (n-k) = 0.67

F = MSB/MSE = 2.07

F-distribution



Two-way ANOVA (factorial design)

Use two-way ANOVA when you have more than one **categorical** variables

2 factors

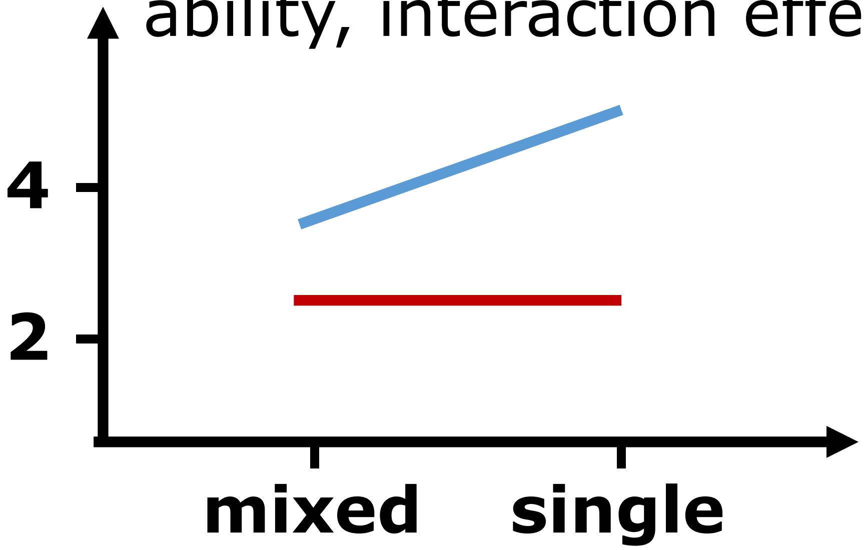
hearing ability	**speech type**	
	mixed	single
hearing-impaired	2,2,3	3,2,2
normal	2,4,4	4,5,4

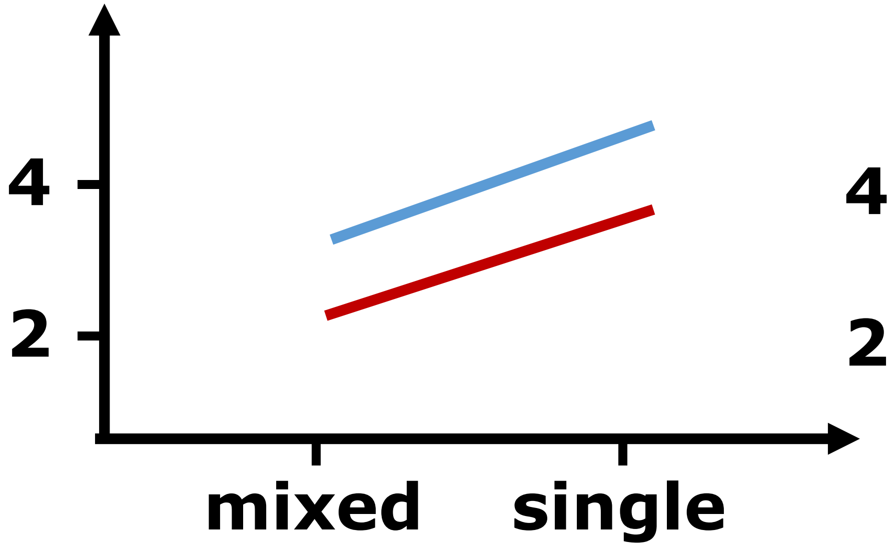
Whether speech type or hearing ability or both affect the intelligibility score?

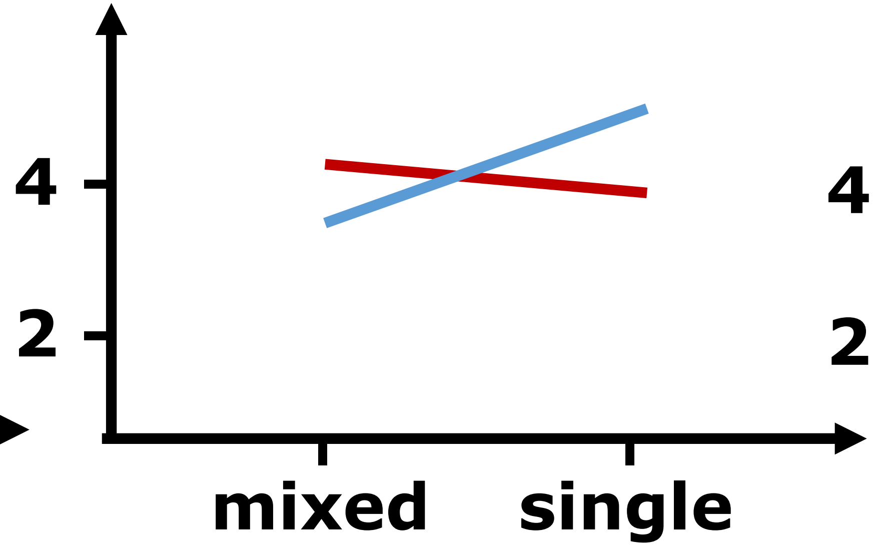
Main and interaction effects

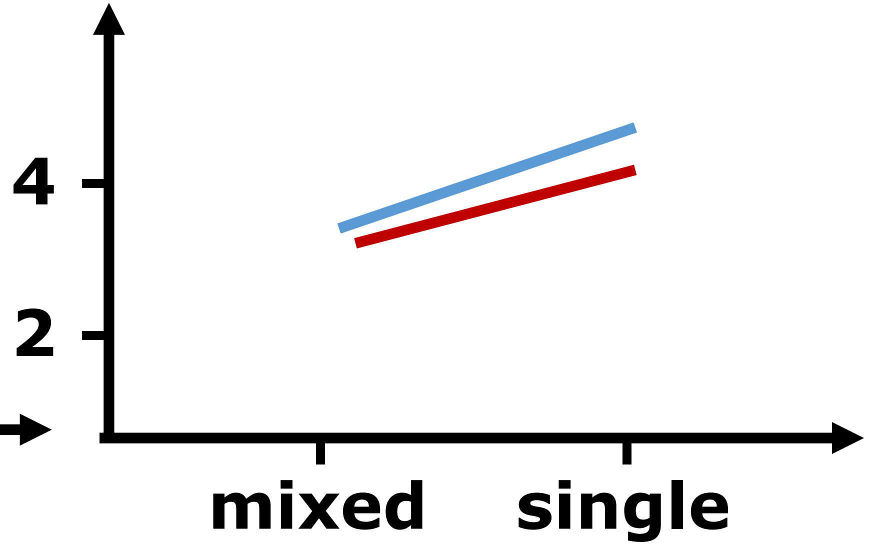
hearing ability	speech type	
	mixed	single
hearing-impaired	2,2,3	3,2,2
normal	2,4,4	5,4,5

main effect of hearing ability, interaction effects









Perform a two-way ANOVA

1. Compute **sum of squares 1st factor**
(group mean – grand mean)^2 * n

hearing ability: $(2.3-3.2)^2*6 + (4-3.2)^2*6 = 8.7$

mean table

hearing ability	**speech type**		**mean**
	mixed	single	
hearing-impaired	2,2,3	3,2,2	2.3
normal	2,4,4	5,4,5	4
mean	2.8	3.5	3.2

grand mean

Perform a two-way ANOVA

2. Compute **sum of squares 2nd factor**
(group mean – grand mean)^2 * n

speech type: $(2.8-3.2)^2*6 + (3.5-3.2)^2*6 = 1.5$

mean table

hearing ability	speech type		mean
	mixed	single	
hearing-impaired	2,2,3	3,2,2	2.3
normal	2,4,4	5,4,5	4
mean	2.8	3.5	3.2

→ **grand mean**

Perform a two-way ANOVA

3. Compute **sum of squares within (error)**

$(2-2.33)^2 + (2-2.33)^2 + (3-2.33)^2 + (3-2.33)^2 + (2-2.33)^2 + (2-2.33)^2$
$+ (2-3.33)^2 + (4-3.33)^2 + (4-2.33)^2 + (5-4.67)^2 + (4-4.67)^2 + (5-4.67)^2 = 4.67$

hearing ability	**speech type**		**mean**
	mixed	single	
hearing-impaired	2,2,3 (M=2.33)	3,2,2 (M=2.33)	2.3
normal	2,4,4 (M=3.33)	5,4,5 (M=4.67)	4
mean	2.8	3.5	3.2

Perform a two-way ANOVA

4. Compute **sum of squares total**

$$\sum_{i=1}^{n}(x_i - \bar{x})^2 \quad = 15.67$$

mean table

hearing ability	speech type		mean
	mixed	single	
hearing-impaired	2,2,3	3,2,2	2.3
normal	2,4,4	5,4,5	4
mean	2.8	3.5	3.2

Perform a two-way ANOVA

5. Compute **sum of squares both factors**

SSTotal = SS1Factor + SS2Factor + SSBothFactor + SSError

SSBothFactor = SSTotal–SS1Factor-SS2Factor-SSError
= 15.67 – 8.7 – 1.5 – 4.67
= 0.8

Calculate *F* score

Source of variation	Degrees of freedom	Sum of squares	Mean squares	*F* statistic
1st Factor	k_1-1 = 1	8.7	8.7/1	8.7/0.58
2nd Factor	k_2-1 = 1	1.5	1.5/1	1.5/0.58
Both Factor	$k_1 * k_2$ = 1	0.8	0.8/1	0.8/0.58
Within	n-k_1-k_2 = 8	4.67	4.67/8=0.58	
Total	n-1 = 11	15.67		

Example: Semantic composition vs. association