FACULTAD DE CIENCIAS AGRONÓMICAS
UNIVERSIDAD DE CHILE

Pruebas de comparaciones múltiples de medias

Erika Kania Kuhl
Ing. Agr. Dr.

1

Prueba de comparaciones múltiples

El análisis de varianza es un procedimiento poderoso para probar la homogeneidad de un conjunto de medias poblacionales (o grupos, o tratamientos).

Sin embargo, si se rechaza la hipótesis nula, de igualdad de medias:

¿Cuál o cuáles de las medias poblacionales en estudio son las diferentes?

2

Prueba de comparaciones múltiples

Generalmente el investigador desea ir más allá de la simple conclusión de que hay diferencias entre las medias del conjunto, pues le interesa saber entre que poblaciones (o grupos) están las diferencias poblacionales en términos de sus medias.

Si se ha rechazado Ho, es obvio que la máxima diferencia observada entre los promedios de dos tratamientos debe corresponder a una diferencia verdadera entre las μ, pero ¿Qué puede decirse de las diferencias entre los promedios intermedios?

Para responder esta interrogante se han planteado diversos procedimientos denominados pruebas de Comparaciones Múltiples.

3

Prueba de comparaciones múltiples

Básicamente las pruebas consisten en calcular cada una de las diferencias (($\bar{Y}1.$ - $\bar{Y}2.$)por ejemplo) y compararlas con un **valor crítico**

4

Prueba de comparaciones múltiples

Existen varias alternativas para llevar adelante este tipo de pruebas:

- Tukey
- SNK
- Duncan
- LSD Fisher
- DGC
- Entre otras

5

Ejemplo

Supongamos un ensayo en donde se aplican 4 tratamientos:

Para 4 tratamientos existen 6 pares posibles combinaciones:

$\mu 1 - \mu 2$

$\mu 1 - \mu 3$

$\mu 1 - \mu 4$

$\mu 2 - \mu 3$

$\mu 2 - \mu 4$

$\mu 3 - \mu 4$

$$H_0: u_1 = u_2 = u_3 = u_4 = u$$

$$H_A: \text{existe } ui \neq u$$

6

Procedimiento

1) Establecer un cuadro con los promedios de los tratamientos

		MENOR			MAYOR
Tratamiento		**Ȳ 4**	**Ȳ 3**	**Ȳ 2**	**Ȳ 1**
MAYOR	**Ȳ 1**				
	Ȳ 2				
	Ȳ 3	**I (Ȳ fila - Ȳ columna) I = d**			
MENOR	**Ȳ 4**				

Media muestreal = Ȳ

Media poblacional = μ

Procedimiento

1) Establecer un cuadro con los promedios de los tratamientos

Ȳ 1 = 18

Ȳ 2 = 12

Ȳ 3 = 20

Ȳ 4 = 15

Tratamiento		**T2**	**T4**	**T1**	**T3**
	Ȳi.				
T3					
T1					
T4					
T2					

Procedimiento

1) Establecer un cuadro con los promedios de los tratamientos

$\bar{Y}\ 1 = 18$

$\bar{Y}\ 2 = 12$

$\bar{Y}\ 3 = 20$

$\bar{Y}\ 4 = 15$

Tratamiento		T2	T4	T1	T3
	$\bar{Y}$i.	12	15	18	20
T3	20				
T1	18				
T4	15				
T2	12				

9

Procedimiento

1) Establecer un cuadro con los promedios de los tratamientos

$\bar{Y}\ 1 = 18$

$\bar{Y}\ 2 = 12$

$\bar{Y}\ 3 = 20$

$\bar{Y}\ 4 = 15$

Tratamiento		T2	T4	T1	T3
	$\bar{Y}$i.	12	15	18	20
T3	20	8	5	2	-
T1	18	6	3	-	
T4	15	3	-		
T2	12	-			

10

Procedimiento

2) Se establece un estadígrafo basado en la variabilidad de los datos (medida en términos de la variabilidad del CME) y de un valor de tabla que depende del tipo de prueba a utilizar

Estadígrafo → **VALOR CRÍTICO**

(VC)

Que es función de (CME, valor de tabla)

11

Procedimiento

3) Se compara cada diferencia obtenida con un Valor Crítico

Si resulta que:

d > VC → d es una diferencia significativa (promedios unidos por letras diferentes)

d ≤ VC → d es una diferencia no significativa (promedios unidos por letras iguales)

$H_0 : ui = ui$

$H_A : ui \neq ui$

Ns: no significativo

* : Significativo

12

Diferencias significativas

Es una diferencia observada muy improbablemente atribuida al mero azar. La diferencia significativa entre tratamientos indica evidencia estadística de una diferencia real entre parámetros, es decir, diferencias significativas entre medias muestreales implica diferencias significativas entre medias poblacionales.

d > VC → d es una diferencia significativa (promedios unidos por letras diferentes)

13

Diferencias no significativas

Es una diferencia observada atribuible al puro azar, es decir, diferencias no significativas entre medias muestreales implica igualdad de medias poblacionales

Tratamientos que comparten una misma letra no se pueden declarar como estadísticamente diferentes, es decir las medias muestrales observadas pueden haberse dado por el azar y por lo tanto no ser repetibles

$d \leq VC$ → d es una diferencia no significativa (promedios unidos por letras iguales)

14

Procedimiento

3) Se compara cada diferencia obtenida con un Valor Crítico

Supongamos un VC = 4

Tratamiento		T2	T4	T1	T3
	$\bar{Y}i.$	12	15	18	20
T3	20	8	5	2	-
T1	18	6	3	-	
T4	15	3	-		
T2	12	-			

15

Procedimiento

3) Se compara cada diferencia obtenida con un Valor Crítico

Supongamos un VC = 4

Tratamiento		T2	T4	T1	T3
	$\bar{Y}i.$	12	15	18	20
T3	20	8*	5*	2^{ns}	-
T1	18	6*	3^{ns}	-	
T4	15	3^{ns}	-		
T2	12	-			

16

Procedimiento

3) Se compara cada diferencia obtenida con un Valor Crítico

$\mu3 \neq \mu2$

$\mu3 \neq \mu4$

$\mu1 \neq \mu2$

$\mu3 = \mu1$

$\mu1 = \mu4$

$\mu4 = \mu2$

Tratamiento		T2	T4	T1	T3
	$\bar{Y}i.$	12	15	18	20
T3	20	8*	5*	2^{ns}	-
T1	18	6*	3^{ns}	-	
T4	15	3^{ns}	-		
T2	12	-			

17

Procedimiento

4) Se resume el resultado de las comparaciones múltiples en una tabla con promedios unidos por letras (Tabla con letras)

a=MejorTratamiento (+ grande en este caso)

Tratamiento	Nombre variable (unidad)
T3	20
T1	18
T4	15
T2	12

18

Procedimiento

4) Se resume el resultado de las comparaciones múltiples en una tabla con promedios unidos por letras (Tabla con letras)

a=MejorTratamiento (+ grande en este caso)

Tratamiento	Nombre variable (unidad)
T3	20
T1	18
T4	15
T2	12

Leer de derecha a izquierda

-Partir con una "a" para todas las ns en la primera fila

-Partir con una "b" para todas las ns en la segunda fila….etc.

19

		T2	T4	T1	T3
	$\bar{Y}i.$	**12**	**15**	**18**	**20**
T3	**20**	8*	5*	2^{ns}	-
T1	**18**	6*	3^{ns}	-	
T4	**15**	3^{ns}	-		
T2	**12**	-			

Trat	Nombre variable (unidad)	
T3	20	a
T1	18	a
T4	15	
T2	12	

20

		T2	T4	T1	T3
	$\bar{Y}i.$	12	15	18	20
T3	20	8*	5*	2^{ns}	-
T1	18	6*	3^{ns}	-	
T4	15	3^{ns}	-		
T2	12	-			

Trat	Nombre variable (unidad)		
T3	20	a	
T1	18	a	b
T4	15		b
T2	12		

21

		T2	T4	T1	T3
	$\bar{Y}i.$	12	15	18	20
T3	20	8*	5*	2^{ns}	-
T1	18	6*	3^{ns}	-	
T4	15	3^{ns}	-		
T2	12	-			

Trat	Nombre variable (unidad)			
T3	20	a		
T1	18	a	b	
T4	15		b	c
T2	12			c

22

Procedimiento

4) Se resume el resultado de las comparaciones múltiples en una tabla con promedios unidos por letras (Tabla con letras)

Tratamiento	Nombre variable (unidad)			
T3	20	a		
T1	18	a	b	
T4	15		b	c
T2	12			c

a=MejorTratamiento (+ grande en este caso)

Promedios unidos por letras iguales en sentido vertical indican diferencias estadísticamente no significativas entre los tratamientos, según la prueba de _________________(p-value > 0,05)

Procedimiento

4) Se resume el resultado de las comparaciones múltiples en una tabla con promedios unidos por letras (Tabla con letras)

Tratamiento	Nombre variable (unidad)			
T3	20	a		
T1	18	a	b	
T4	15		b	c
T2	12			c

En la PCM no se aplican las relaciones de equivalencia, es <u>incorrecto decir</u>:

<u>Si μ 1 = μ 3 y μ 1 = μ 4, entonces μ 3 = μ 4</u>

Procedimiento

4) Se resume el resultado de las comparaciones múltiples en una tabla con promedios unidos por letras (Tabla con letras)

Tratamiento	Nombre variable (unidad)			
T3	20	a		
T1	18	a	b	
T4	15		b	c
T2	12			c

En la PCM no se aplican las relaciones de equivalencia, es incorrecto decir:

Si $\mu_1 = \mu_3$ y $\mu_1 = \mu_4$, entonces $\mu_3 = \mu_4$

25

Procedimiento

4) Se resume el resultado de las comparaciones múltiples en una tabla con promedios unidos por letras (Tabla con letras)

Tratamiento	Nombre variable (unidad)			
T3	20	a		
T1	18	a	b	
T4	15		b	c
T2	12			c

Es incorrecto hablar de tratamientos "similares", solamente:

-Existen diferencias estadísticamente significativas o

-No existen diferencias estadísticamente significativas

26

Ejemplo 2

Tratamiento		T6	T5	T2	T4	T1	T3
	$\bar{Y}_{i.}$						
T3		*	*	*	*	ns	
T1		*	*	*	ns		
T4		*	*	ns			
T2		*	*				
T5		*					
T6							

27

Ejemplo 2

Tratamiento	Promedio
T3	
T1	
T4	
T2	
T5	
T6	

28

Ejemplo 2

Tratamiento	Promedio
T3	a
T1	a b
T4	b c
T2	c
T5	
T6	

Trat		T6	T5	T2	T4	T1	T3
	$\bar{Y}_{i.}$						
T3		*	*	*	*	ns	
T1		*	*	*	ns		
T4		*	*	ns			
T2		*	*				
T5		*					
T6							

29

Ejemplo 2

Tratamiento	Promedio
T3	a
T1	a b
T4	b c
T2	c d
T5	e
T6	

Trat		T6	T5	T2	T4	T1	T3
	$\bar{Y}_{i.}$						
T3		*	*	*	*	ns	
T1		*	*	*	ns		
T4		*	*	ns			
T2		*	*				
T5		*					
T6							

30

Ejemplo 2

Tratamiento	Promedio
T3	a
T1	a b
T4	b c
T2	c $\boxed{d}$
T5	e
T6	f

31

Ejemplo 2

Tratamiento	Promedio	
T3	a	a
T1	a b	a b
T4	b c	b c
T2	c $\boxed{d}$	c
T5	e	d
T6	f	e

32

Ejemplo 3

Tratamiento		T6	T5	T2	T4	T1	T3
	Ȳi.						
T3		*	*	*	*	ns	
T1		*	ns	ns	ns		
T4		*	ns	ns			
T2		*	ns				
T5		*					
T6							

33

Ejemplo 3

Tratamiento	Promedio
T3	
T1	
T4	
T2	
T5	
T6	

34

Ejemplo 3

Tratamiento	Promedio
T3	a
T1	a b
T4	b c
T2	b c d
T5	b c d e
T6	f

35

Ejemplo 3

Tratamiento	Promedio
T3	a
T1	a b
T4	b c
T2	b c d
T5	b c d e
T6	f

36

Ejemplo 3

Tratamiento	Promedio	
T3	a	a
T1	a b	a b
T4	b c	b
T2	b c d	b
T5	b c d e	b
T6	f	c

37

¿Cómo leer o interpretar los Cuadros?

Hay que expresar todas las diferencias significativas

Ejemplo 4

Tratamiento	Promedio
T4	a
T3	a b
T5	a b
T1	a b
T2	b

38

¿Cómo leer o interpretar los Cuadros?

Hay que expresar todas las diferencias significativas

Ejemplo 4

Tratamiento	Promedio
T4	a
T3	a b
T5	a b
T1	a b
T2	b

¿Cuál diferencia es estadísticamente significativa?

39

¿Cómo leer o interpretar los Cuadros?

Hay que expresar todas las diferencias significativas

Ejemplo 4

Tratamiento	Promedio
T4	a
T3	a b
T5	a b
T1	a b
T2	b

$\mu 4 \neq \mu 2$

(Promedio poblacional 4 presenta diferencias estadísticamente significativas con el promedio poblacional 2)

40

¿Cómo leer o interpretar los Cuadros?

Hay que expresar todas las diferencias significativas

Ejemplo 4

Tratamiento	Promedio	
T4	a	Mayor
T3	a b	
T5	a b	
T1	a b	
T2	b	Menor

Supongamos que la variable respuesta es:

Rendimiento

-Encuentre el o los mejores tratamientos

-Encuentre el o los peores tratamientos

-Justifique

41

¿Cómo leer o interpretar los Cuadros?

Hay que expresar todas las diferencias significativas

Ejemplo 4

Tratamiento		Promedio
T4	Mayor	a
T3		a b
T5		a b
T1		a b
T2	Menor	b

Los mejores tratamientos : T4, T3, T5, T1 (por tener los mayores rendimientos y estar unidos por la letra "a")

Los peores tratamientos : T3, T5, T1, T2 (por tener los menores rendimientos y estar unidos por la letra "b")

42

¿Cómo leer o interpretar los Cuadros?

Hay que expresar todas las diferencias significativas

Ejemplo 5

Tratamiento	Promedio
T4	a
T3	a
T5	a b
T1	b
T6	b c
T2	c

43

¿Cómo leer o interpretar los Cuadros?

Hay que expresar todas las diferencias significativas

Ejemplo 5

Tratamiento	Promedio
T4	a
T3	a
T5	a b
T1	b
T6	b c
T2	c

$\mu 4 \neq \mu 1$

$\mu 4 \neq \mu 6$

$\mu 4 \neq \mu 2$

$\mu 3 \neq \mu 1$

$\mu 3 \neq \mu 6$

$\mu 3 \neq \mu 2$

$\mu 5 \neq \mu 2$

$\mu 1 \neq \mu 2$

44

Prueba de Tukey

45

Valor Crítico

$d > VC$ → d es una diferencia significativa (*)

$d \leq VC$ → d es una diferencia no significativa

$$VC = q \times \sqrt{(CME/r)}$$

<u>Donde</u>:

q: es un **valor de tabla**

CME: cuadrado medio del error

r: número de repeticiones por tratamiento (observaciones por grupo)

46

Ejemplo Prueba de Tukey

En un ensayo con 5 tratamientos y 5 repeticiones por tratamiento se obtuvieron los siguientes resultados:

CME: 108

$\bar{Y}1 = 52$

$\bar{Y}2 = 69$

$\bar{Y}3 = 82$

$\bar{Y}4 = 85$

$\bar{Y}5 = 94$

Variable respuesta: altura de planta (cm)

47

Procedimiento

1) Establecer un cuadro con los promedios de los tratamientos

<table>
<tr><td></td><td></td><td>MENOR</td><td></td><td></td><td>MAYOR</td></tr>
<tr><td>Tratamiento</td><td></td><td>$\bar{Y}$ 4</td><td>$\bar{Y}$ 3</td><td>$\bar{Y}$ 2</td><td>$\bar{Y}$ 1</td></tr>
<tr><td>MAYOR</td><td>$\bar{Y}$ 1</td><td colspan="4" rowspan="4">I ($\bar{Y}$ fila - $\bar{Y}$ columna) I = d</td></tr>
<tr><td></td><td>$\bar{Y}$ 2</td></tr>
<tr><td></td><td>$\bar{Y}$ 3</td></tr>
<tr><td>MENOR</td><td>$\bar{Y}$ 4</td></tr>
</table>

48

Procedimiento

1) Establecer un cuadro con los promedios de los tratamientos

$\bar{Y}1= 52$
$\bar{Y}2= 69$
$\bar{Y}3= 82$
$\bar{Y}4= 85$
$\bar{Y}5= 94$

Tratamiento		T1	T2	T3	T4	T5
	$\bar{Y}i.$	52	69	82	85	94
T5	94					
T4	85					
T3	82					
T2	69					
T1	52					

49

Procedimiento

1) Establecer un cuadro con los promedios de los tratamientos

$\bar{Y}1= 52$
$\bar{Y}2= 69$
$\bar{Y}3= 82$
$\bar{Y}4= 85$
$\bar{Y}5= 94$

Tratamiento		T1	T2	T3	T4	T5
	$\bar{Y}i.$	52	69	82	85	94
T5	94	42	25	12	9	-
T4	85	33	16	3	-	-
T3	82	30	13	-	-	-
T2	69	17	-	-	-	-
T1	52	-	-	-	-	-

50

Procedimiento

2) Se establece un estadígrafo basado en la variabilidad de los datos (medida en términos de la variabilidad del CME) y de un valor de tabla que depende del tipo de prueba a utilizar

Estadígrafo → **VALOR CRÍTICO**

(VC)

VC (Tukey) = q x √(CME/r)

51

Prueba de Tukey

2) Se establece un estadígrafo basado en la variabilidad de los datos (medida en términos de la variabilidad del CME) y de un valor de tabla que depende del tipo de prueba a utilizar

VC (Tukey) = q x √(CME/r)

VC (Tukey) = 4,232 x √(108/5)

VC (Tukey) = 19,7 (cm)

52

Procedimiento

3) Se compara cada diferencia obtenida con un Valor Crítico

Si resulta que:

d > VC (19,7) → d es una diferencia significativa

d ≤ VC (19,7) → d es una diferencia no significativa

53

Procedimiento

1) Establecer un cuadro con los promedios de los tratamientos

$\bar{Y}1= 52$
$\bar{Y}2= 69$
$\bar{Y}3= 82$
$\bar{Y}4= 85$
$\bar{Y}5= 94$

VC: 19,7

Tratamiento		T1	T2	T3	T4	T5
	$\bar{Y}i.$	52	69	82	85	94
T5	94	42	25	12	9	-
T4	85	33	16	3	-	-
T3	82	30	13	-	-	-
T2	69	17	-	-	-	-
T1	52	-	-	-	-	-

54

Procedimiento

1) Establecer un cuadro con los promedios de los tratamientos

$\bar{Y}1= 52$
$\bar{Y}2= 69$
$\bar{Y}3= 82$
$\bar{Y}4= 85$
$\bar{Y}5= 94$

VC: 19,7

Tratamiento		T1	T2	T3	T4	T5
	$\bar{Y}i.$	52	69	82	85	94
T5	94	42 *	25 *	12	9	-
T4	85	33 *	16	3	-	-
T3	82	30 *	13	-	-	-
T2	69	17	-	-	-	-
T1	52	-	-	-	-	-

Procedimiento

4) Se resume el resultado de las comparaciones múltiples en una tabla con promedios unidos por letras (Tabla con letras)

Tratamiento	Altura planta (cm)
T5	94 a
T4	85 b
T3	82 c
T2	69 d
T1	52 e

Procedimiento

4) Se resume el resultado de las comparaciones múltiples en una tabla con promedios unidos por letras (Tabla con letras)

Tratamiento	Altura planta (cm)
T5	94 a
T4	85 a b
T3	82 a b c
T2	69 b c d
T1	52 d e

57

Procedimiento

4) Se resume el resultado de las comparaciones múltiples en una tabla con promedios unidos por letras (Tabla con letras)

Tratamiento	Altura planta (cm)	
T5	94 a	a
T4	85 a b	ab
T3	82 a b c	ab
T2	69 b c d	bc
T1	52 d e	c

58

Presentación de resultados

Cuadro XX. Título..

Tratamiento	Altura planta (cm)	
T5	94	a
T4	85	ab
T3	82	ab
T2	69	bc
T1	52	c

Promedios unidos por letras iguales en sentido vertical indican diferencias estadísticamente no significativas entre los tratamientos, según la prueba de Tukey (p-value > 0,05)

59

Tratamiento	Promedio	Tukey	SNK	Duncan
T5	94	a	a	a
T4	85	ab	ab	a
T3	82	ab	ab	ab
T2	69	bc	b	b
T1	52	c	c	c

60

Expresando todas las diferencias significativas

Tukey	SNK	Duncan
μ5 ≠ μ2	μ5 ≠ μ2	μ5 ≠ μ2

61

Expresando todas las diferencias significativas

Tukey	SNK	Duncan
μ5 ≠ μ2	μ5 ≠ μ2	μ5 ≠ μ2
μ5 ≠ μ1	μ5 ≠ μ1	μ5 ≠ μ1
μ4 ≠ μ1	μ4 ≠ μ1	μ4 ≠ μ2
μ3 ≠ μ1	μ3 ≠ μ1	μ4 ≠ μ1
	μ2 ≠ μ1	μ3 ≠ μ1
		μ2 ≠ μ1

62

Tukey	•Es + exigente •Muy conservadora •Encuentra menos diferencias
SNK	•Es + sensible que Tukey •Encuentra + fácil las diferencias que Tukey
Duncan	•Es menos estricta •Es + sensible que Tukey y SNK •Encuentra + fácil las diferencias que Tukey y SNK

63

FACULTAD DE CIENCIAS AGRONÓMICAS
UNIVERSIDAD DE CHILE

Pruebas de comparaciones múltiples de medias

Erika Kania Kuhl
Ing. Agr. Dr.

64