

3. Agrupación de las medias según métodos LSD, Tukey, Duncan

datos del ejemplo 3.2.7						
Nivel (Lb N/acre)	Lechugas/parcela				$\bar{y}_{i.}$	S_i^2
0	104	114	90	140	112.0	445.33
50	134	130	144	174	145.5	395.67
100	146	142	152	156	149.0	38.67
150	147	160	160	163	157.5	51.00
200	131	148	154	163	149.0	182.00
$MSE = S_p^2 = \frac{\sum_{i=1}^a (n_i - 1) S_i^2}{N - a} = 222.533334$					$\bar{y}_{..} = 142.6$	$\bar{S}^2 = 222.533334$

Mínima diferencia significativa: $LSD_{ij} = t_{0.025, N-a} \times \sqrt{MSE \left(\frac{1}{n_i} + \frac{1}{n_j} \right)} = 2.13145 \times \sqrt{222.533334 \left(\frac{1}{4} + \frac{1}{4} \right)} = 22.483177$

Diferencia “Honestamente” significativa: $HSD_{ij} = \frac{q_{0.05}(a, N-a)}{\sqrt{2}} \times \sqrt{MSE \left(\frac{1}{n_i} + \frac{1}{n_j} \right)} = \frac{4.366985}{\sqrt{2}} \times \sqrt{222.533334 \left(\frac{1}{4} + \frac{1}{4} \right)} = 32.57236$

			Método LSD				Método Tukey			
Comparaciones			LSD _{ij}	Intervalo de Confianza LSD		Decisión	HSD _{ij}	Intervalo de confianza Tukey		Decisión
i vs. j	$\bar{y}_{i.} - \bar{y}_{j.}$	$ \bar{y}_{i.} - \bar{y}_{j.} $		$(\bar{y}_{i.} - \bar{y}_{j.}) - LSD_{ij}$	$(\bar{y}_{i.} - \bar{y}_{j.}) + LSD_{ij}$			$(\bar{y}_{i.} - \bar{y}_{j.}) - HSD_{ij}$	$(\bar{y}_{i.} - \bar{y}_{j.}) + HSD_{ij}$	
1 vs. 2	-33.5	33.5	22.4832	-55.9832	-11.0168	diferentes	32.572	-66.0724	-0.9276	diferentes
1 vs. 3	-37.0	37.0	22.4832	-59.4832	-14.5168	diferentes	32.572	-69.5724	-4.4276	diferentes
1 vs. 4	-45.5	45.5	22.4832	-67.9814	-23.0168	diferentes	32.572	-78.0724	-12.9276	diferentes
1 vs. 5	-37.0	37.0	22.4832	-59.4832	-14.5168	diferentes	32.572	-69.5724	-4.4276	diferentes
2 vs. 3	-3.5	3.5	22.4832	-25.9832	18.9832	iguales	32.572	-36.0724	29.0724	iguales
2 vs. 4	-12.0	12.0	22.4832	-34.4832	10.4832	iguales	32.572	-44.5724	20.5724	iguales
2 vs. 5	-3.5	3.5	22.4832	-25.9832	18.9832	iguales	32.572	-36.0724	29.0724	iguales
3 vs. 4	-8.5	8.5	22.4832	-30.9832	13.9832	iguales	32.572	-41.0724	24.0724	iguales
3 vs. 5	0.0	0.0	22.4832	-22.4832	22.4832	iguales	32.572	-32.5724	32.5724	iguales
4 vs. 5	8.5	8,5	22.4832	-13.9832	30.9832	iguales	32.572	-24.0724	41.0724	iguales

Grupos de medias iguales según:	
LSD	Tukey
Grupo 1: $\mu_2 = \mu_3 = \mu_4 = \mu_5$	Grupo 1: $\mu_2 = \mu_3 = \mu_4 = \mu_5$
Grupo 2: μ_1	Grupo 2: μ_1

Resultados para comparaciones de Duncan

	p=2	3	4	5
Valor Crítico rangos significativos: $r_{0.05}(p, N - a)$	3,014327	3,159828	3,25025	3,31185
Rangos de Mínima significancia: $R_p = r_{0.05}(p, N - a) \sqrt{\frac{MSE}{n}}$	22,48317182	23,5684303	24,2428672	24,7023274

	Medias ordenadas (mayor a menor)				
tratamiento	i=4	5	3	2	1
Media $\bar{y}_{i.}$	157,5	149	149	145,5	112

Comparaciones según rangos	$D_{ij} = \bar{y}_{i.} - \bar{y}_{j.} $	p	R_p	Decisión Duncan	Observación
4 vs. 1	45.5	5	24,70232744	diferentes	
4 vs. 2	12.0	4	24,24286721	iguales	Implica la igualdad no solo de medias tratamientos 4 y 2 sino entre todas las que están en el rango entre 4 y 2: $\mu_4 = \mu_2 = \mu_3 = \mu_5$, por tanto, no es necesario comparar 4 vs. 3, 4 vs. 5, 5 vs. 2, 5 vs. 3 ni 3 vs. 2.
5 vs. 1	37.0	4	24,24286721	diferentes	Con estas tres pruebas y la primera, se concluye que μ_1 no es igual a las demás medias, y junto con la segunda prueba se tiene que los grupos de medias son: Grupo 1: μ_1 , Grupo 2: $\mu_2 = \mu_3 = \mu_4 = \mu_5$
3 vs. 1	37.0	3	23,56843031	diferentes	
2 vs. 1	33.5	2	22,48317182	diferentes	

4. Comparaciones de Dunnett tomando como control al nivel 1: 0 Lbs N/acre

Para las comparaciones de Dunnett tomando como control al nivel 1: 0 Lbs N/acre y un nivel de significancia de 0.05, tenemos para el caso bilateral el valor crítico $D_{0.05}(a-1, N-a) = D_{0.05}(4,15) = 2.7286$ (valor arrojado por función **glht**, el cual es obtenido por simulación y puede variar en diferentes corridas), por tanto, la mínima diferencia significativa de Dunnett es

$$MSD_{Dunnett} = D_{0.05}(a-1, N-a) \times \sqrt{MSE\left(\frac{1}{n_i} + \frac{1}{n_j}\right)} = 2.7286 \times \sqrt{222.533334\left(\frac{1}{4} + \frac{1}{4}\right)} = 28.7821$$

Comparaciones			MSD _{Dunnett}	Intervalo de Confianza Dunnett bilateral		Decisión
i vs. 1	$\bar{y}_i - \bar{y}_1$	$ \bar{y}_i - \bar{y}_1 $		$(\bar{y}_i - \bar{y}_1) - MSD_{Dunnett}$	$(\bar{y}_i - \bar{y}_1) + MSD_{Dunnett}$	
2 vs. 1	33.5	33.5	28.7821	4.7179	62.2821	diferentes
3 vs. 1	37.0	37.0	28.7821	8.2179	65.7821	diferentes
4 vs. 1	45.5	45.5	28.7821	16.7179	74.2821	diferentes
5 vs. 1	37.0	37.0	28.7821	8.2179	65.7821	diferentes

Lo anterior indica que todos los tratamientos donde se aplica un nivel de lbs N/acre diferente de cero, da una respuesta media diferente a este nivel de control. Pero si el interés es determinar cuáles de los tratamientos dan una respuesta media mayor que el control, entonces, el interés es probar para $i = 2, 3, 4, 5$, que

$$H_0: \mu_i - \mu_1 \leq 0 \text{ vs. } H_1: \mu_i > \mu_1$$

A un nivel de significancia de 0.05, tenemos para el caso unilateral el valor crítico $D_{1,0.05}(a-1, N-a) = D_{1,0.05}(4,15) = 2.3565$ (valor arrojado por función **glht**, el cual es obtenido por simulación y puede variar en diferentes corridas), por tanto, la mínima diferencia significativa de Dunnett es

$$MSD_{Dunnett} = D_{1,0.05}(a-1, N-a) \times \sqrt{MSE\left(\frac{1}{n_i} + \frac{1}{n_j}\right)} = 2.3565 \times \sqrt{222.533334\left(\frac{1}{4} + \frac{1}{4}\right)} = 24.8571$$

Se rechaza H_0 al 5% de significancia, si $\bar{y}_i - \bar{y}_1 > 24.8459$

i vs. 1	$\bar{y}_i - \bar{y}_1$	MSD _{Dunnett}	I.C. unilateral de Dunnett de cota inferior $\mu_i - \mu_1 \geq (\bar{y}_i - \bar{y}_1) - MSD_{Dunnett}$	Decisión
2 vs. 1	33.5	24.8571	$\mu_i - \mu_1 \geq 8.6429$	Se rechaza H_0 y por tanto $\mu_2 > \mu_1$
3 vs. 1	37.0	24.8571	$\mu_i - \mu_1 \geq 12.1429$	Se rechaza H_0 y por tanto $\mu_3 > \mu_1$
4 vs. 1	45.5	24.8571	$\mu_i - \mu_1 \geq 20.6429$	Se rechaza H_0 y por tanto $\mu_4 > \mu_1$
5 vs. 1	37.0	24.8571	$\mu_i - \mu_1 \geq 12.1429$	Se rechaza H_0 y por tanto $\mu_5 > \mu_1$

Apéndice: LSD considerando varianzas desiguales, según lo visto en los residuos del ANOVA en este ejemplo (página 81 de notas de clase):

Mínima diferencia significativa: $LSD_{ij} = t_{0.025, v_{ij}} \times \sqrt{\frac{S_i^2}{n_i} + \frac{S_j^2}{n_j}}$, con

$$v_{ij} = \frac{\left(\frac{S_i^2}{n_i} + \frac{S_j^2}{n_j}\right)^2}{\frac{\left(\frac{S_i^2}{n_i}\right)^2}{n_i - 1} + \frac{\left(\frac{S_j^2}{n_j}\right)^2}{n_j - 1}}$$

i vs. j	$D_{ij} = \bar{y}_i - \bar{y}_j $	$Std_{ij} = \sqrt{\frac{S_i^2}{n_i} + \frac{S_j^2}{n_j}}$	v_{ij}	LSD_{ij}	Conclusión
1 vs 2	33.5	14.5000	5.9791	35.510239	medias iguales
1 vs 3	37.0	11.0000	3.5171	32.268071	medias diferentes
1 vs 4	45.5	11.1393	3.6782	32.024362	medias diferentes
1 vs 5	37.0	12.5233	5.1012	32.001131	medias diferentes
2 vs 3	3.5	10.4203	3.5808	30.317443	medias iguales
2 vs 4	12.0	10.5672	3.7607	30.090389	medias iguales
2 vs 5	3.5	12.0173	5.2779	30.40866	medias iguales
3 vs 4	8.5	4.7346	5.8886	11.638546	medias iguales
3 vs 5	0.0	7.4274	4.2197	20.205275	medias iguales
4 vs 5	8.5	7.6322	4.5589	20.204041	medias iguales

O bien construyendo intervalos de confianza LSD para $\mu_i - \mu_j$, como $(\bar{y}_i - \bar{y}_j) \pm LSD_{ij}$, de nuevo con $LSD_{ij} = t_{0.025, v_{ij}} \sqrt{\frac{S_i^2}{n_i} + \frac{S_j^2}{n_j}}$

			Intervalos de confianza LSD			
i-j	$\bar{y}_i - \bar{y}_j$	v_{ij}	$Std_{ij} = \sqrt{\frac{S_i^2}{n_i} + \frac{S_j^2}{n_j}}$	$(\bar{y}_i - \bar{y}_j) - LSD_{ij}$	$(\bar{y}_i - \bar{y}_j) + LSD_{ij}$	Conclusión
1-2	-33.5	5.9791	14.5000	-69.0102	2.0102	medias iguales
1-3	-37.0	3.5171	11.0000	-69.2681	-4.7319	medias diferentes
1-4	-45.5	3.6782	11.1393	-77.5244	-13.4756	medias diferentes
1-5	-37.0	5.1012	12.5233	-69.0011	-4.9989	medias diferentes
2-3	-3.5	3.5808	10.4203	-33.8174	26.8174	medias iguales
2-4	-12.0	3.7607	10.5672	-42.0904	18.0904	medias iguales
2-5	-3.5	5.2779	12.0173	-33.9087	26.9087	medias iguales
3-4	-8.5	5.8886	4.7346	-20.1385	3.1385	medias iguales
3-5	0.0	4.2197	7.4274	-20.2053	20.2053	medias iguales
4-5	8.5	4.5589	7.6322	-11.7040	28.7040	medias iguales

Los grupos de medias son: Grupo 1: $\mu_2 = \mu_3 = \mu_4 = \mu_5$ y Grupo 2: $\mu_1 = \mu_2$ (hay traslape entre los dos grupos)

Efectos de tratamientos estimados, sus errores estándar e I.C del 95% bajo varianza no constante, usando aproximación Satterwaite: Aquí es necesario expresar cada efecto como un contraste de las medias, así: $\alpha_i = \mu_i - \mu = \mu_i - \frac{1}{a} \sum_{j=1}^a \mu_j = \frac{a-1}{a} \mu_i - \frac{1}{a} \sum_{j \neq i}^a \mu_j$, entonces aplicando fórmula (A.12), con $c_i = \frac{a-1}{a}$ y $c_j = -\frac{1}{a}$ para $j \neq i$, tenemos que un intervalo de confianza del 95% para α_i es

$$\hat{\alpha}_i \pm t_{0.025, v} \times \sqrt{\left(\frac{a-1}{a}\right)^2 \frac{S_i^2}{n_i} + \left(\frac{1}{a}\right)^2 \sum_{j \neq i}^a \frac{S_j^2}{n_j}}$$

Con

$$v = \frac{\left[\left(\frac{a-1}{a}\right)^2 \frac{S_i^2}{n_i} + \sum_{j \neq i}^a \left(\frac{1}{a}\right)^2 \frac{S_j^2}{n_j}\right]}{\frac{\left[\left(\frac{a-1}{a}\right)^2 \frac{S_i^2}{n_i}\right]}{n_i - 1} + \sum_{j \neq i}^a \frac{\left[\left(\frac{1}{a}\right)^2 \frac{S_j^2}{n_j}\right]}{n_j - 1}}$$

Intervalos de confianza para α_i bajo varianza no constante					
i	Estimación	v	Std	LIC	LSC
1	-30.6	3.5746	8.8276	-56.3036	-4.8964
2	2.9	3.6963	8.3950	-21.1834	26.9834
3	6.4	11.1145	4.1142	-2.6439	15.4439
4	14.9	10.0222	4.3332	5.2479	24.5521
5	6.4	5.0118	6.1989	-9.5235	22.3235

Donde $Std = \sqrt{Var(\hat{\alpha}_i)} = \sqrt{\left(\frac{a-1}{a}\right)^2 \frac{S_i^2}{n_i} + \left(\frac{1}{a}\right)^2 \sum_{j \neq i}^a \frac{S_j^2}{n_j}}$