

“Experimentos de un solo factor”

Gustavo Ramírez Valverde

DISEÑO COMPLETAMENTE AL AZAR

Modelo de efectos Fijos

- Diseño experimental más simple,
- Es adecuado cuando las unidades experimentales son uniformes
- No se visualiza la presencia de ningún factor de confusión

DISEÑO EXPERIMENTAL COMPLETAMENTE AL AZAR

DISEÑO CON UN CRITERIO DE CLASIFICACION

DISEÑOS EXPERIMENTALES CON UN FACTOR

Ideal cuando :

El material es homogéneo

No presenta claramente factores de confusión

Aleatorización en:

- La asignación de los tratamientos a las unidades experimentales
- El orden en que se realizan las pruebas
- La toma de los datos

Ejemplo :

- Se desea comparar 3 diferentes dietas para engorda de ganado ovino.
- Objetivo: Definir que dieta aumenta el peso de borregas.
- Unidad experimental: 1 borrego
- Numero de repeticiones: 5
- Diseño experimental : Diseño completamente al azar

DISEÑO COMPLETAMENTE AL AZAR



1



2



3



4



5



6



7



8



9



10



11



12



13



14



15

DISEÑO COMPLETAMENTE AL AZAR



1



2



3



4



5



6



7



8



9



10



11



12



13



14



15

$$T1=\{2, 6, 11, 4, 13\}$$

$$T2=\{1, 15, 3, 8, 10\}$$

$$T3=\{5, 7, 9, 12, 14\}$$

DISEÑO COMPLETAMENTE AL AZAR

T2 1	T1 2	T2 3	T1 4	T3 5
T1 6	T3 7	T2 8	T3 9	T2 10
T1 11	T3 12	T1 13	T3 14	T2 15

$$T1=\{2, 6, 11, 4, 13\}$$

$$T2=\{1, 15, 3, 8, 10\}$$

$$T3=\{5, 7, 9, 12, 14\}$$

Planeación DCA en R

- # DISEÑO COMPLETAMENTE AL AZAR
- #####
- #Planeación de un diseño completamente al azar
- #####
- library(agricolae)
- tratamientos = c("Dital1", "Dieta2", "Dieta3")
- repeticion = c(5, 5, 5)
- diseño = design.crd(tratamientos,r=repeticion,)
- diseño1 = diseño\$book
- print(diseño1)

Notación

Tratamiento	Observaciones				Totales	Prom
1	Y_{11}	Y_{12}	Y_{1n}	$Y_{1.}$	$\bar{Y}_{1.}$
2	Y_{21}	Y_{22}	Y_{2n}	$Y_{2.}$	$\bar{Y}_{2.}$
.....
t	Y_{t1}	Y_{t2}	Y_{tn}	$Y_{t.}$	$\bar{Y}_{t.}$
					$Y_{..}$	$\bar{Y}_{..}$

Modelo del diseño completamente al azar

$$y_{ij} = \mu + \tau_i + \varepsilon_{ij}$$

μ es el promedio general si no se hubiese aplicado ningún tratamiento,

τ_i es el efecto del tratamiento i

ε_{ij} es el error experimental cometido en la repetición j del tratamiento i

$$\varepsilon_{ij} \sim NIID(0, \sigma^2)$$

a) Independencia

b) Normalidad

c) Homogeneidad

de varianzas

Análisis de un experimento completamente al azar

$$H_0 : \mu_1 = \mu_2 = \dots = \mu_t \quad \text{Vs.}$$

H_a : Al menos una de las medias de los tratamientos es diferente

Cuadro 2.1.- Tabla de ANDEVA para un Diseño Completamente al azar.

Factor de Variación	Suma de Cuadrados	Grados de libertad	Cuadrados Medios	$F_{\text{calculada}}$	F_{tabulada}
Tratamientos	$SCT = r \sum_{i=1}^t (\bar{y}_{i..} - \bar{y}_{...})^2$	$t - 1$	$CMT = \frac{SCT}{t-1}$	$\frac{CMT}{CME}$	$F_{g T}^{glT}$
Error Experimental	$SCE = \sum_{i=1}^t \sum_{j=1}^r (y_{ij} - \bar{y}_{i..})^2$	$t(r-1)$	$CME = \frac{SCE}{t(r-1)}$		
Total	$SCTotal = \sum_{i=1}^t \sum_{j=1}^r (y_{ij} - \bar{y}_{...})^2$	$rt - 1$			

El modelo general es :

$$y_{ij} = \mu + \tau_i + \varepsilon_{ij}$$

Para cada dato o unidad de observación, se tiene

$$Y_j = \mu + \tau_1 X_{1j} + \tau_2 X_{2j} + \tau_3 X_{3j} + \dots + \tau_p X_{pj} + \varepsilon_j$$

Con $X_{ij} = \begin{cases} 1 & \text{Si la observaciòn es del tratamiento i} \\ 0 & \text{Si la observaciòn no es del tratamiento i} \end{cases}$

El modelo general en forma matricial es :

$$Y = X\beta + \varepsilon$$

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \dots \\ Y_n \end{bmatrix} = \begin{bmatrix} 1 & X_{11} & \dots & X_{p1} \\ 1 & X_{12} & \dots & X_{p2} \\ 1 & \dots & \dots & \dots \\ 1 & X_{1n} & \dots & X_{pn} \end{bmatrix} \begin{bmatrix} \mu \\ \tau_1 \\ \dots \\ \tau_p \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \dots \\ \varepsilon_n \end{bmatrix}$$

Entonces,

$$\hat{\varepsilon}_i = Y_i - \hat{\mu}(X_1, X_2, \dots, X_p)$$

$$\hat{\varepsilon}_i = Y_i - (\hat{\mu} + \hat{\tau}_1 X_{1i} + \hat{\tau}_2 X_{2i} \dots + \hat{\tau}_p X_{pi})$$

$$SCE = \sum_{i=1}^n \hat{\varepsilon}^2 = \sum_{i=1}^n (Y_i - \hat{Y}_i)^2$$

$$SCE = \sum_{i=1}^n \left\{ Y_i - (\hat{\mu} + \hat{\tau}_1 X_{1i} + \hat{\tau}_2 X_{2i} \dots + \hat{\tau}_p X_{pi}) \right\}^2$$

El objetivo es encontrar los valores del parámetro que minimizan SCE

$$Q = \varepsilon' \varepsilon = (Y - X\beta)'(Y - X\beta)$$

$$Q = Y'Y - Y'X\beta - \beta'X'Y + \beta'X'X\beta$$

$$\frac{\partial Q}{\partial \beta} = \frac{\partial Y'Y}{\partial \beta} - \frac{\partial Y'X\beta}{\partial \beta} - \frac{\partial \beta'X'Y}{\partial \beta} + \frac{\partial \beta'X'X\beta}{\partial \beta}$$

$$\frac{\partial Q}{\partial \beta} = -X'Y - X'Y + 2X'X\beta$$

Al igualar a cero el sistema , se obtiene el estimador de mínimos cuadrados de

$$-2X'Y + 2X'X\hat{\beta} = 0 \quad X'X\hat{\beta} = X'Y$$

Ecuaciones Normales

$$y_{ij} = \mu + \tau_i + \varepsilon_{ij}$$

$$X'X\hat{\beta} = X'Y$$

$$\sum_{i=1}^t \tau_i = 0$$

Entonces, la solución al sistema de ecuaciones es:

$$\begin{aligned}\hat{\mu} &= \bar{y}_{..} \\ \hat{\tau}_i &= \bar{y}_{i.} - \bar{y}_{..} \\ \hat{\mu}_i &= \bar{y}_{i.}\end{aligned}$$

Ejemplo 1

- Un ingeniero de desarrollo de productos está interesado en maximizar la resistencia a la tensión de una nueva fibra sintética que se empleará en la manufactura de tela para camisas de hombre.
- El ingeniero sabe por experiencia que la resistencia está influida por el porcentaje de algodón presente en la fibra. Además, sospecha que el contenido de algodón debe estar aproximadamente entre un 10 y 40% para que la tela resultante tenga otras características de calidad que se desean (como la capacidad de recibir un tratamiento de planchado permanente).
- El ingeniero decide probar muestras a cinco niveles de porcentaje de algodón: 15, 20, 25, 30 y 35%.
- Asimismo, decide ensayar cinco muestras a cada nivel de contenido de algodón. Las 25 observaciones se asignan al azar.

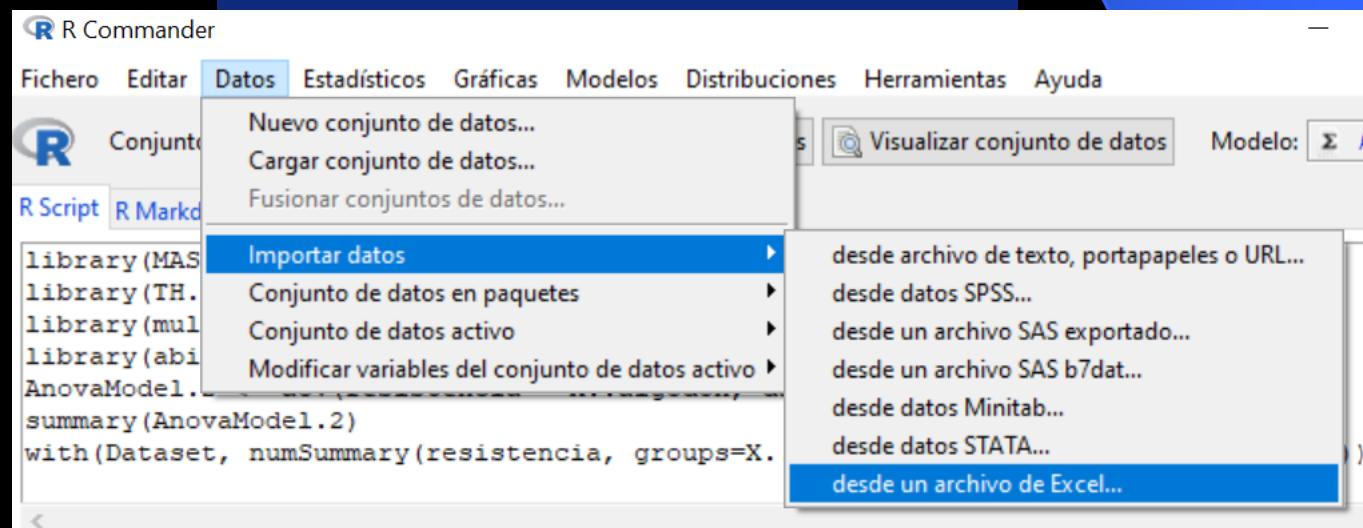
Ejemplo 1

Algodón Porcent.	Observaciones				
15	7	7	15	11	9
20	12	17	12	18	18
25	14	18	18	19	19
30	19	25	22	19	23
35	7	10	11	15	11

Análisis en R

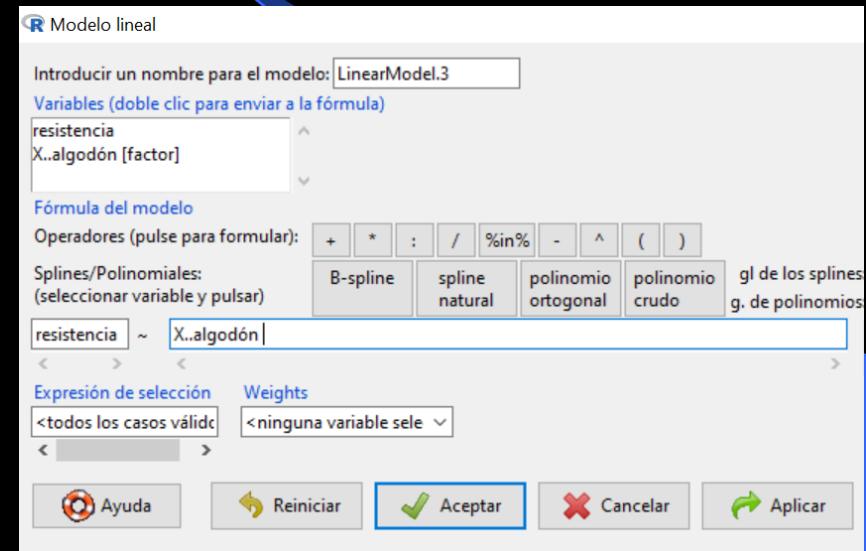
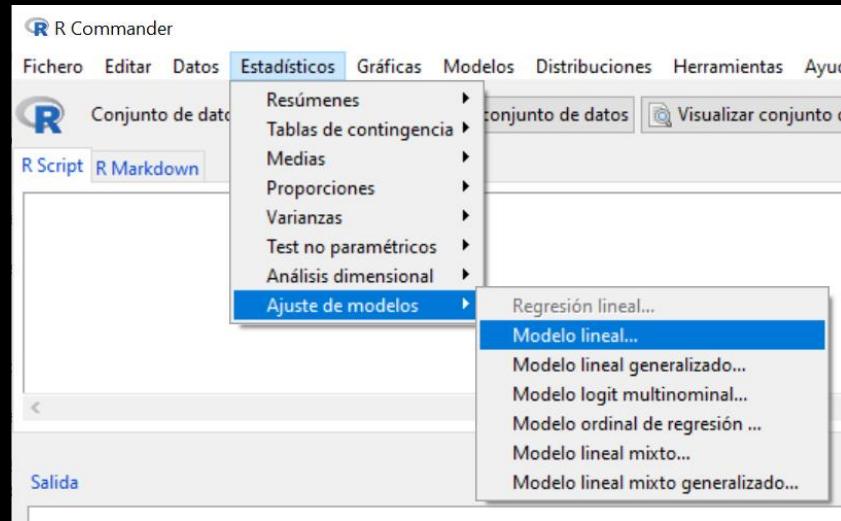
	A	B
1	% algodón	resistencia
2	15	7
3	20	12
4	25	14
5	30	19
6	35	7
7	15	7
8	20	17
9	25	18
10	30	25
11	35	10
12	15	15
13	20	12
14	25	18
15	30	22
16	35	11
17	15	11
18	20	18
19	25	19
20	30	19
21	35	15
22	15	9
23	20	18
24	25	19
25	30	23
26	35	11

- El archivo de datos en Excel
- Dos variables:
 - % algodon
 - Resistencia
- Cargar los datos en R



Análisis en R

• Realizar ANOVA en Rcmdr



```
Salida

> summary(LinearModel.3)

Call:
lm(formula = resistencia ~ X..algodón, data = Dataset)

Residuals:
    Min      1Q  Median      3Q     Max 
 -3.8   -2.6    0.4    1.4    5.2 

Coefficients:
            Estimate Std. Error t value    Pr(>|t|)    
(Intercept)  9.800     1.270   7.719 0.000000202 ***
X..algodón[T.20] 5.600     1.796   3.119  0.005409 ** 
X..algodón[T.25] 7.800     1.796   4.344  0.0000315 ***
X..algodón[T.30] 11.800    1.796   6.572  0.000002108 ***
X..algodón[T.35]  1.000     1.796   0.557  0.583753
```

```
> Anova(LinearModel.3)
Anova Table (Type II tests)

Response: resistencia
          Sum Sq Df F value    Pr(>F)    
X..algodón 475.76  4 14.757 0.000009128 ***
Residuals  161.20 20
---
Signif. codes:  0 '****' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Conclusión

$H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5$

$H_a: Al \ menos \ una \ media \ es \ diferente$

- Nivel de significancia $\alpha = 0.05$
- Estadístico de Prueba $F_c = 14.76$
- Valor de $p < 0.001$
- Decisión rechazo H_0 con nivel de significancia $\alpha = 0.05$ y se concluye que si hay efecto del porcentaje de algodón en la resistencia promedio de las telas

Conclusión

$H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5$

$H_a: Al menos una media es diferente$

- Nivel de significancia $\alpha = 0.05$
- Estadístico de Prueba $F_c = 14.76$
- Valor de $p < 0.001$

Análisis en R

```
#####
```

```
#Análisis de un diseño experimental en un diseño  
completamente al azar
```

```
#####
```

```
resistencia=c(7,7,15,11,9,12,17,12,18,18,14,18,18,19,19,19,  
,25,22,19,23,7,10,11,15,11)
```

```
algodon=c(15,15,15,15,15,20,20,20,20,20,25,25,25,25,25,25,30,  
30, 30,30,35,35,35,35,35)
```

```
Algodon_DCA=data.frame(resistencia,algodon)
```

```
attach(Algodon_DCA)
```

```
algodon=as.factor(algodon)
```

```
modelo<-lm(resistencia~algodon)
```

```
resultado=anova(modelo)
```

```
resultado
```

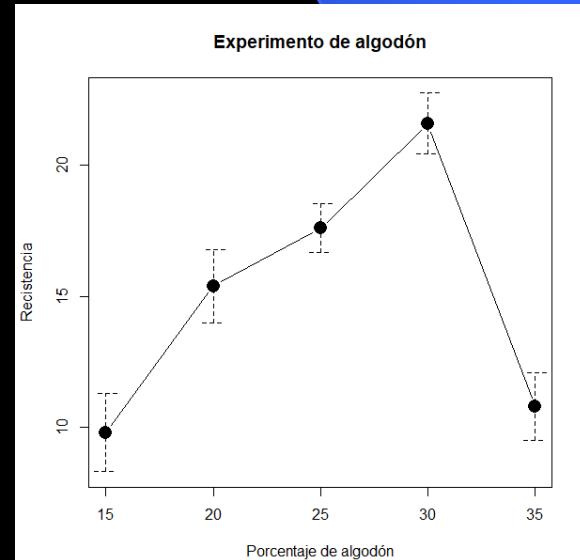
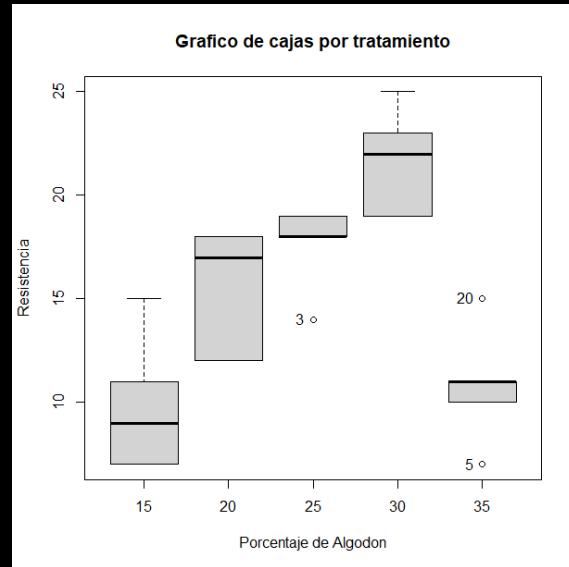
Análisis en R

```
#####
```

```
# Gráfico de boxplot para los cinco tratamientos
```

```
#####
```

```
boxplot(resistencia~algodon, main="Experimento Algodon DCA",  
        xlab="Grafico de cajas por tratamiento", ylab="Resistencia")  
plotMeans(resistencia, algodon, error.bars="se",  
          xlab="Porcentaje de algodón", ylab="Recistencia",  
          main="Experimento de algodón",  
          connect=TRUE))
```



Experimentos de un solo factor

Supuestos

Modelo del diseño completamente al azar

$$Y_{ij} = \mu + \tau_i + \varepsilon_{ij}$$

μ es el promedio general si no se hubiese aplicado ningún

τ_i es el efecto del tratamiento i

ε_{ij} es el error experimental cometido en la repetición j del tratamiento i

$$\varepsilon_{ij} \sim NIID(0, \sigma^2)$$

- a) Independencia
- b) Normalidad
- c) Homogeneidad de varianzas

Independencia

- Las observaciones deben ser independientes
 - Los errores no deben estar correlacionados
 - La magnitud del error en una observación no debe influir en la magnitud del error cometido en otra observación

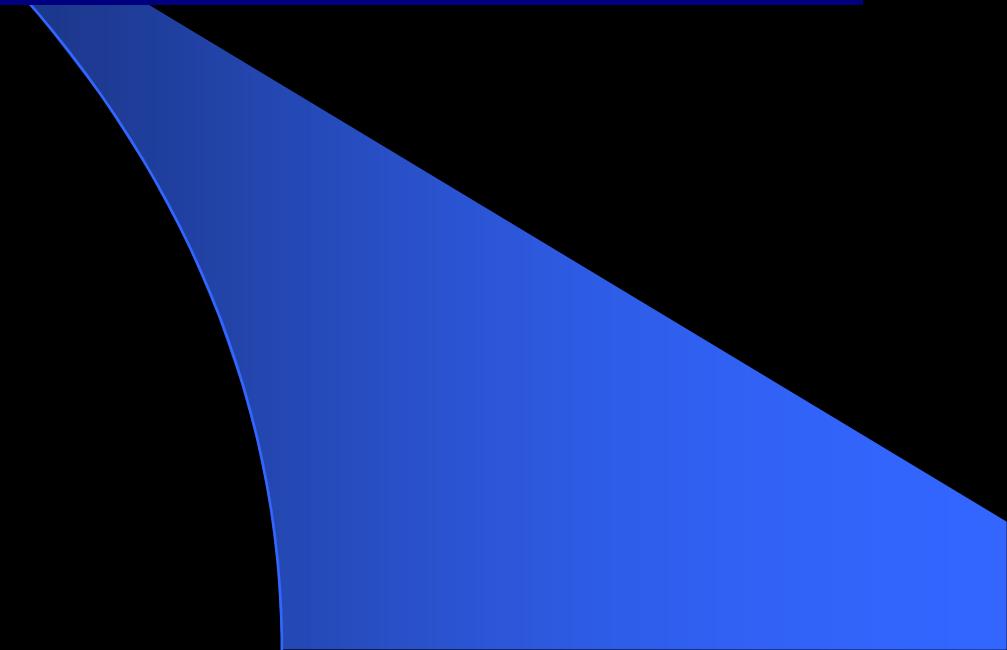
Homogeneidad de Varianzas

- La prueba de F es robusta contra la heterogeneidad de varianzas, sobre todo si el diseño es balanceado y/o de efectos fijos.
- Las pruebas para comparar medias son mucho mas sensibles afectando fuertemente el valor del tamaño de la prueba.

Normalidad

- La prueba de F es muy robusta a no normalidad, especialmente en modelos balanceados y/o de efectos fijos
- Tamaños de muestra grandes aproximarán a la normalidad por el teorema central del límite
- El efecto por falta de normalidad en la prueba de F es una perdida de potencia y una alteración en el nivel de significancia de la prueba (nivel de significancia real diferente a α).

Métodos gráficos para detectar no cumplimiento de los supuestos



Análisis en R

```
#####
#Métodos gráficos para la verificación de supuestos
#####
## QQPlot de residuos
residuos = resid(modelo)
predichos = fitted(modelo)
qqnorm(residuos,main="Grafica Q-Q",
       xlab="Cuantil teórico", ylab="Cuantil muestral")
qqline(residuos)
plot(density(residuals(modelo)))
#####
# Residuos frente a valores ajustados
plot(predichos,residuos,main="Experimento Algodon DCA",
      xlab="Porcentaje de algodón", ylab="Residuos")
abline(h=0,lty=2)
```

Métodos Analíticos para detectar no cumplimiento de los supuestos



Pruebas de Normalidad

- Existe una gran variedad de pruebas para verificar la bondad de ajuste del modelo normal.
- Ejemplos:
 - Especificas para Normalidad
 - Shapiro-Wilk
 - Jarque-Bera
 - Goodness-of-fit tests,
 - Kolmogorov-Smirnov D
 - Cramer-von Mises W^2
 - Chi-Cuadrada

Prueba de Shapiro-Wilk's W

Sean $\epsilon_1, \epsilon_2, \dots$, los residuos ordenados de menor a mayor

H_0 : Los residuos provienen de una distribución normal.

H_A : Los residuos **no provienen** de una distribución normal.

El estadístico

$$W = \frac{1}{d} \left[\sum_{j=1}^k a_j (\varepsilon_{n-j+1} - \bar{\varepsilon}) \right]^2$$

Los coeficientes a_i se obtienen de
Una tabla.

$$d = \sum_{i=1}^n (\varepsilon_i - \bar{\varepsilon})^2$$

Regla.R.

Rechazo H_0 si $W < W_{0.05}$

Los valores de tablas.



$$k = \frac{n}{2}$$

$$k = \frac{(n-1)}{2}$$

Si n es par

si n es impar.

Table A6 Coefficients a_i for the Shapiro-Wilk W Test for Normality

<u>n</u>	2	3	4	5	6	7	8	9	10
1	0.7071	0.7071	0.6872	0.6646	0.6431	0.6233	0.6052	0.5888	0.5739
2	-	0.0000	0.1677	0.2413	0.2806	0.3031	0.3164	0.3244	0.3291
3	-	-	-	0.0000	0.0875	0.1401	0.1743	0.1976	0.2141
4	-	-	-	-	-	0.0000	0.0561	0.0947	0.1224
5	-	-	-	-	-	-	-	0.0000	0.0399

<u>n</u>	11	12	13	14	15	16	17	18	19	20
1	0.5601	0.5475	0.5359	0.5251	0.5150	0.5056	0.4968	0.4886	0.4808	0.4734
2	0.3315	0.3325	0.3325	0.3318	0.3306	0.3290	0.3273	0.3253	0.3232	0.3211
3	0.2260	0.2347	0.2412	0.2460	0.2495	0.2521	0.2540	0.2553	0.2561	0.2565
4	0.1429	0.1586	0.1707	0.1802	0.1878	0.1939	0.1988	0.2027	0.2059	0.2085
5	0.0695	0.0922	0.1099	0.1240	0.1353	0.1447	0.1524	0.1587	0.1641	0.1686
6	0.0000	0.0303	0.0539	0.0727	0.0880	0.1005	0.1109	0.1197	0.1271	0.1334
7	-	-	0.0000	0.0240	0.0433	0.0593	0.0725	0.0837	0.0932	0.1013
8	-	-	-	-	0.0000	0.0196	0.0359	0.0496	0.0612	0.0711
9	-	-	-	-	-	-	0.0000	0.0163	0.0303	0.0422
10	-	-	-	-	-	-	-	-	0.0000	0.0140

<u>n</u>	21	22	23	24	25	26	27	28	29	30
1	0.4643	0.4590	0.4542	0.4493	0.4450	0.4407	0.4366	0.4328	0.4291	0.4254
2	0.3185	0.3156	0.3126	0.3098	0.3069	0.3043	0.3018	0.2992	0.2968	0.2944
3	0.2578	0.2571	0.2563	0.2554	0.2543	0.2533	0.2522	0.2510	0.2499	0.2487
4	0.2119	0.2131	0.2139	0.2145	0.2148	0.2151	0.2152	0.2151	0.2150	0.2148
5	0.1736	0.1764	0.1787	0.1807	0.1822	0.1836	0.1848	0.1857	0.1864	0.1870
6	0.1399	0.1443	0.1480	0.1512	0.1539	0.1563	0.1584	0.1601	0.1616	0.1630
7	0.1092	0.1150	0.1201	0.1245	0.1283	0.1316	0.1346	0.1372	0.1395	0.1415
8	0.0804	0.0878	0.0941	0.0997	0.1046	0.1089	0.1128	0.1162	0.1192	0.1219
9	0.0530	0.0618	0.0696	0.0764	0.0823	0.0876	0.0923	0.0965	0.1002	0.1036
10	0.0263	0.0368	0.0459	0.0539	0.0610	0.0672	0.0728	0.0778	0.0822	0.0862
11	0.0000	0.0122	0.0228	0.0321	0.0403	0.0476	0.0540	0.0598	0.0650	0.0697
12	-	-	0.0000	0.0107	0.0200	0.0284	0.0358	0.0424	0.0483	0.0537
13	-	-	-	-	0.0000	0.0094	0.0178	0.0253	0.0320	0.0381
14	-	-	-	-	-	-	0.0000	0.0084	0.0159	0.0227
15	-	-	-	-	-	-	-	-	0.0000	0.0076

Shapiro-Wilk Coeficientes

Shapiro-Wilk Coeficientes

<u>n</u>	31	32	33	34	35	36	37	38	39	40
1										
2	0.4220	0.4188	0.4156	0.4127	0.4096	0.4068	0.4040	0.4015	0.3989	0.3964
3	0.2921	0.2898	0.2876	0.2854	0.2834	0.2813	0.2794	0.2774	0.2755	0.2737
4	0.2475	0.2462	0.2451	0.2439	0.2427	0.2415	0.2403	0.2391	0.2380	0.2368
5	0.2145	0.2141	0.2137	0.2132	0.2127	0.2121	0.2116	0.2110	0.2104	0.2098
6	0.1874	0.1878	0.1880	0.1882	0.1883	0.1883	0.1883	0.1881	0.1880	0.1878
7	0.1641	0.1651	0.1660	0.1667	0.1673	0.1678	0.1683	0.1686	0.1689	0.1691
8	0.1433	0.1449	0.1463	0.1475	0.1487	0.1496	0.1505	0.1513	0.1520	0.1526
9	0.1243	0.1265	0.1284	0.1301	0.1317	0.1331	0.1344	0.1356	0.1366	0.1376
10	0.1066	0.1093	0.1116	0.1140	0.1160	0.1179	0.1196	0.1211	0.1225	0.1237
11	0.0899	0.0931	0.0961	0.0988	0.1013	0.1036	0.1056	0.1075	0.1092	0.1108
12	0.0739	0.0777	0.0812	0.0844	0.0873	0.0900	0.0924	0.0947	0.0967	0.0986
13	0.0585	0.0629	0.0669	0.0706	0.0739	0.0770	0.0798	0.0824	0.0848	0.0870
14	0.0435	0.0485	0.0530	0.0572	0.0610	0.0645	0.0677	0.0706	0.0733	0.0759
15	0.0289	0.0344	0.0395	0.0441	0.0484	0.0523	0.0559	0.0592	0.0622	0.0651
16	0.0144	0.0206	0.0262	0.0314	0.0361	0.0404	0.0444	0.0481	0.0515	0.0546
17	0.0000	0.0068	0.0131	0.0187	0.0239	0.0287	0.0331	0.0372	0.0409	0.0444
18	-	-	0.0000	0.0062	0.0119	0.0172	0.0220	0.0264	0.0305	0.0343
19	-	-	-	-	0.0000	0.0057	0.0110	0.0158	0.0203	0.0244
20	-	-	-	-	-	-	-	-	0.0000	0.0049

<u>n</u>	41	42	43	44	45	46	47	48	49	50
1										
2	0.3940	0.3917	0.3894	0.3872	0.3850	0.3830	0.3808	0.3789	0.3770	0.3751
3	0.2719	0.2701	0.2684	0.2667	0.2651	0.2635	0.2620	0.2604	0.2589	0.2574
4	0.2357	0.2345	0.2334	0.2323	0.2313	0.2302	0.2291	0.2281	0.2271	0.2260
5	0.2091	0.2085	0.2078	0.2072	0.2065	0.2058	0.2052	0.2045	0.2038	0.2032
6	0.1876	0.1874	0.1871	0.1868	0.1865	0.1862	0.1859	0.1855	0.1851	0.1847
7	0.1693	0.1694	0.1695	0.1695	0.1695	0.1695	0.1695	0.1693	0.1692	0.1691
8	0.1531	0.1535	0.1539	0.1542	0.1545	0.1548	0.1550	0.1551	0.1553	0.1554
9	0.1384	0.1392	0.1398	0.1405	0.1410	0.1415	0.1420	0.1423	0.1427	0.1430
10	0.1249	0.1259	0.1269	0.1278	0.1286	0.1293	0.1300	0.1306	0.1312	0.1317
11	0.1123	0.1136	0.1149	0.1160	0.1170	0.1180	0.1189	0.1197	0.1205	0.1212
12	0.1004	0.1020	0.1035	0.1049	0.1062	0.1073	0.1085	0.1095	0.1105	0.1113
13	0.0891	0.0909	0.0927	0.0943	0.0959	0.0972	0.0986	0.0998	0.1010	0.1020
14	0.0782	0.0804	0.0824	0.0842	0.0860	0.0876	0.0892	0.0906	0.0919	0.0932
15	0.0677	0.0701	0.0724	0.0745	0.0765	0.0783	0.0801	0.0817	0.0832	0.0846
16	0.0575	0.0602	0.0628	0.0651	0.0673	0.0694	0.0713	0.0731	0.0748	0.0764
17	0.0476	0.0506	0.0534	0.0560	0.0584	0.0607	0.0628	0.0648	0.0667	0.0685
18	0.0379	0.0411	0.0442	0.0471	0.0497	0.0522	0.0546	0.0568	0.0588	0.0608
19	0.0283	0.0318	0.0352	0.0383	0.0412	0.0439	0.0465	0.0489	0.0511	0.0532
20	0.0188	0.0227	0.0263	0.0296	0.0328	0.0357	0.0385	0.0411	0.0436	0.0459
21	0.0094	0.0136	0.0175	0.0211	0.0245	0.0277	0.0307	0.0335	0.0361	0.0386
22	0.0000	0.0045	0.0087	0.0126	0.0163	0.0197	0.0229	0.0259	0.0288	0.0314
23	-	-	0.0000	0.0042	0.0081	0.0118	0.0153	0.0185	0.0215	0.0244
24	-	-	-	-	0.0000	0.0039	0.0076	0.0111	0.0143	0.0174
25	-	-	-	-	-	-	0.0000	0.0037	0.0071	0.0104

Table A7 Quantiles of the Shapiro-Wilk W Test for Normality (Values of W Such That $100p\%$ of the Distribution of W Is Less Than W_p)

n	$W_{0.01}$	$W_{0.02}$	$W_{0.05}$	$W_{0.10}$	$W_{0.50}$
3	0.753	0.756	0.767	0.789	0.959
4	0.687	0.707	0.748	0.792	0.935
5	0.686	0.715	0.762	0.806	0.927
6	0.713	0.743	0.788	0.826	0.927
7	0.730	0.760	0.803	0.838	0.928
8	0.749	0.778	0.818	0.851	0.932
9	0.764	0.791	0.829	0.859	0.935
10	0.781	0.806	0.842	0.869	0.938
11	0.792	0.817	0.850	0.876	0.940
12	0.805	0.828	0.859	0.883	0.943
13	0.814	0.837	0.866	0.889	0.945
14	0.825	0.846	0.874	0.895	0.947
15	0.835	0.855	0.881	0.901	0.950
16	0.844	0.863	0.887	0.906	0.952
17	0.851	0.869	0.892	0.910	0.954
18	0.858	0.874	0.897	0.914	0.956
19	0.863	0.879	0.901	0.917	0.957
20	0.868	0.884	0.905	0.920	0.959
21	0.873	0.888	0.908	0.923	0.960
22	0.878	0.892	0.911	0.926	0.961
23	0.881	0.895	0.914	0.928	0.962
24	0.884	0.898	0.916	0.930	0.963
25	0.886	0.901	0.918	0.931	0.964
26	0.891	0.904	0.920	0.933	0.965
27	0.894	0.906	0.923	0.935	0.965
28	0.896	0.908	0.924	0.936	0.966
29	0.898	0.910	0.926	0.937	0.966
30	0.900	0.912	0.927	0.939	0.967
31	0.902	0.914	0.929	0.940	0.967
32	0.904	0.915	0.930	0.941	0.968
33	0.906	0.917	0.931	0.942	0.968
34	0.908	0.919	0.933	0.943	0.969
35	0.910	0.920	0.934	0.944	0.969
36	0.912	0.922	0.935	0.945	0.970
37	0.914	0.924	0.936	0.946	0.970
38	0.916	0.925	0.938	0.947	0.971
39	0.917	0.927	0.939	0.948	0.971
40	0.919	0.928	0.940	0.949	0.972
41	0.920	0.929	0.941	0.950	0.972
42	0.922	0.930	0.942	0.951	0.972
43	0.923	0.932	0.943	0.951	0.973
44	0.924	0.933	0.944	0.952	0.973
45	0.926	0.934	0.945	0.953	0.973
46	0.927	0.935	0.945	0.953	0.974
47	0.928	0.936	0.946	0.954	0.974
48	0.929	0.937	0.947	0.954	0.974
49	0.929	0.937	0.947	0.955	0.974
50	0.930	0.938	0.947	0.955	0.974

Source: After Shapiro and Wilk, 1965.

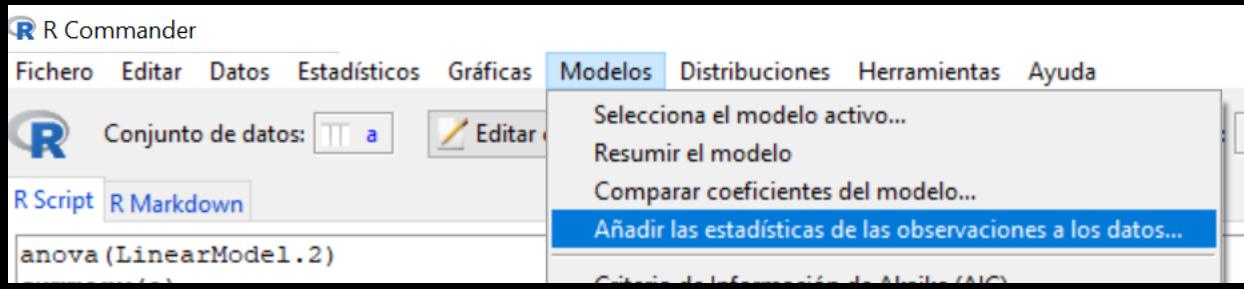
The null hypothesis of a normal distribution is rejected at the α significance level if the calculated W is less than W_α .

Prueba de Shapiro-Wilk's en R

H_0 : Los residuos provienen de una distribución normal.

H_A : Los residuos **no provienen** de una distribución normal.

Primero se guardan los residuos



The screenshot shows the R Commander interface. The 'Modelos' menu is open, displaying options such as 'Selecciona el modelo activo...', 'Resumir el modelo', 'Comparar coeficientes del modelo...', and 'Añadir las estadísticas de las observaciones a los datos...'. The fourth option is highlighted with a blue selection bar. Below the menu, a sub-dialog titled 'Añadir las estadísticas de las observaciones al conjunto de datos' is displayed. It contains several checkboxes: 'Valores ajustados' (unchecked), 'Residuos' (checked), 'Residuos studentizados' (unchecked), 'Valores hat' (unchecked), 'Distancias de Cook' (unchecked), and 'Indices de las observaciones' (unchecked). At the bottom of the dialog are four buttons: 'Ayuda' (Help), 'Reiniciar' (Restart), 'Aceptar' (Accept) with a green checkmark, and 'Cancelar' (Cancel) with a red X.

Prueba de Shapiro-Wilk's en R

H_0 : Los residuos provienen de una distribución normal.

H_A : Los residuos **no provienen** de una distribución normal.

Después:

The screenshot shows the R Commander interface. The 'Estadísticos' menu is open, and the 'Test de normalidad...' option is selected. A sub-dialog box titled 'Test de Normalidad' is displayed. In the 'Variable (elegir una)' section, 'residuals.LinearModel.2' is selected. Under 'Test de Normalidad', the 'Shapiro-Wilk' test is selected. At the bottom, there are buttons for 'Ayuda', 'Reiniciar', 'Aceptar' (with a green checkmark), 'Cancelar' (with a red X), and 'Aplicar'.

R Commander

Fichero Editar Datos Estadísticos Gráficas Modelos Distribuciones Herramientas Ayuda

Conjunto de datos R Script R Markdown

```
algodón <- as.factor(algodón)
LinearModel.2 <- lm(algodón ~ resistencia)
summary(LinearModel.2)
a<- within(a, {
  residuals.LinearModel.2 <- residuals(LinearModel.2)
})
a
```

Resúmenes Tablas de contingencia Medias Proporciones Varianzas Test no paramétricos Análisis dimensional Ajuste de modelos

Conjunto de datos activo Resúmenes numéricos... Distribución de frecuencias... Número de observaciones ausentes Tabla de estadísticas... Matriz de correlaciones... Test de correlación Test de normalidad... Transformación para normalizar...

Test de Normalidad

Variable (elegir una)
residuals.LinearModel.2
resistencia

Test de Normalidad

Shapiro-Wilk Anderson-Darling
 Cramer-von Mises Lilliefors (Kolmogorov-Smirnov)
 Shapiro-Francia Chi cuadrado de Pearson

El número de divisiones para chi cuadrado de Pearson

Ayuda Reiniciar Aceptar Cancelar Aplicar

Prueba de Shapiro-Wilk's en R

H_0 : Los residuos provienen de una distribución normal.

H_A : Los residuos **no provienen** de una distribución normal.

Después:

```
> normalityTest(~residuals.LinearModel.2, test="shapiro.test", data=a)

Shapiro-Wilk normality test

data: residuals.LinearModel.2
W = 0.94387, p-value = 0.1818
```

$\alpha = 0.05$, $valor - p = 0.1818$ Conclusión: **No se** rechaza H_0 con $\alpha = 0.05$ y se concluye que se satisface el supuesto de normalidad satisfactoriamente.

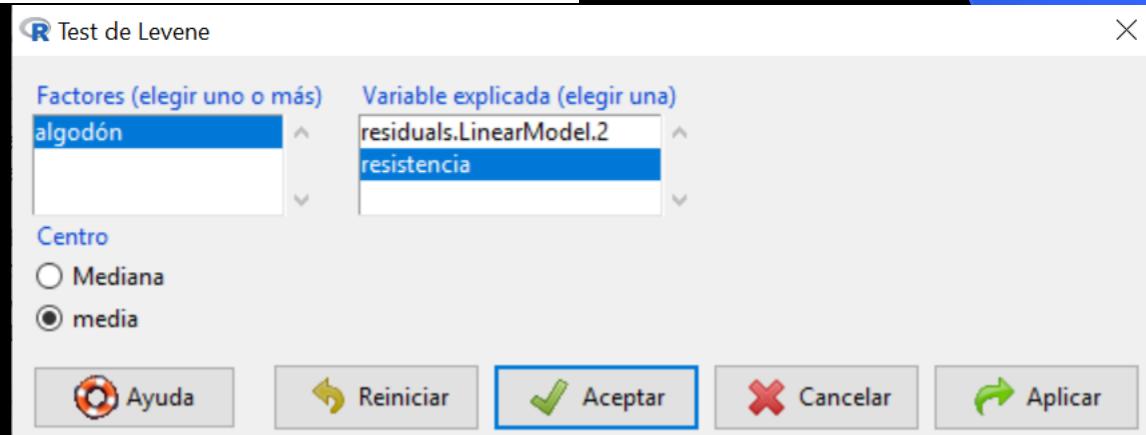
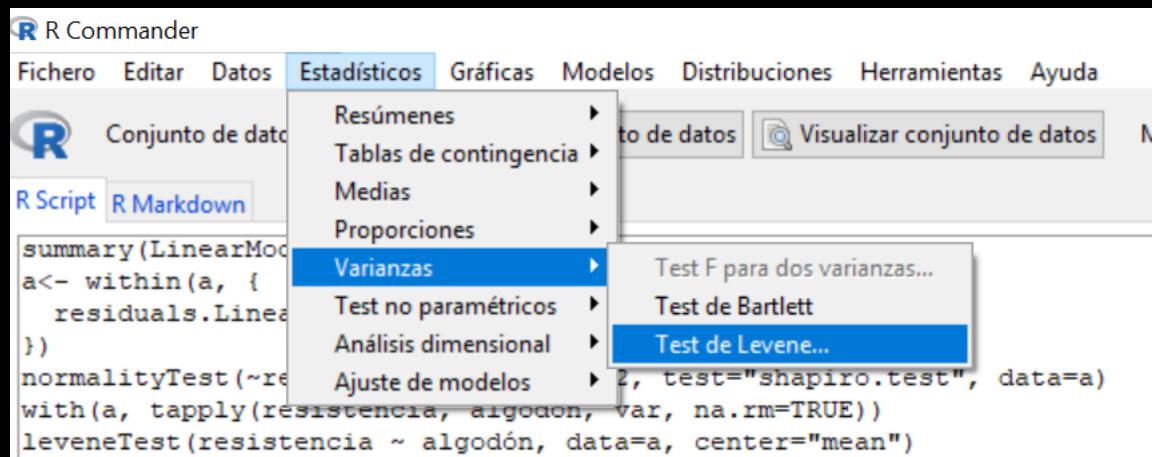
Pruebas para homogeneidad de Varianzas

- Prueba de Levene
 - Calcula un ANOVA en los valores absolutos (o los cuadrados) de los residuales, $|y_{ij} - \hat{y}_i|$ con $t-1, N-t$ grados de libertad.
 - Es robusta a no normalidad, pero es demasiado conservadora
- Prueba Brown-Forsythe
 - Una pequeña modificación a la prueba de Levene, donde la mediana es substituida por la media (Kuehl (2000) la llama Levene (med) Test)
- La prueba de F_{max}
 - El cociente entre la varianza mas grande entre la mas chica es comparada a un valor de tablas
 - Tabachnik and Fidell (2001) usa el cociente F_{max} mas como regla del dedo que para ser comparada con un valor de tablas.
 - F_{max} no es mas grande que 10
 - Los tamaños de muestra son aproximadamente iguales (el cociente entre el mayor y el menor son menores de 4)
 - Supone normalidad

Prueba de Levene

$$H_0: \sigma_{15}^2 = \sigma_{20}^2 = \sigma_{25}^2 = \sigma_{30}^2 = \sigma_{35}^2$$

H_a : Al menos la varianza de un tratamiento *es* diferente



Prueba de Levene

$$H_0: \sigma_{15}^2 = \sigma_{20}^2 = \sigma_{25}^2 = \sigma_{30}^2 = \sigma_{35}^2$$

H_a : Al menos la varianza de un tratamiento *es* diferente

```
> leveneTest(resistencia ~ algodón, data=a, center="mean")
Levene's Test for Homogeneity of Variance (center = "mean")
    Df F value Pr(>F)
group  4  0.6443 0.6372
      20
```

$\alpha = 0.05$, valor - $p = 0.6372$ Conclusión: **No se** rechaza H_0 con $\alpha = 0.05$ y se concluye que se satisface el supuesto de normalidad homogeneidad de varianzas.

Pruebas Analíticas en R

- Las pruebas de homogeneidad de varianzas de Levene se realiza usando un ANOVA con el valor absoluto de los residuos

Pruebas Analíticas en R

- #####
- #Métodos analíticos para la verificación de supuestos
- #####
- ##### Prueba de Normalidad de Shapiro Wilks
- shapiro.test(residuos)
- ##### Prueba de homogeneidad de varianzas
- bartlett.test(resistencia~algodon,data=Algodon_DCA)
- library(car) #requiere tener instalado el paquete car
- leveneTest(resistencia~algodon,data=Algodon_DCA, center="mean")