### Sobre Normalidad Multivariada

Evaluación, Generación, Gráficos, etc.

Para: Estudiantes de IAM

2021 - 04 - 06

# Prueba de Normalidad Univariada del Coeficiente de Correlación, Ejemplo de notas de Clase

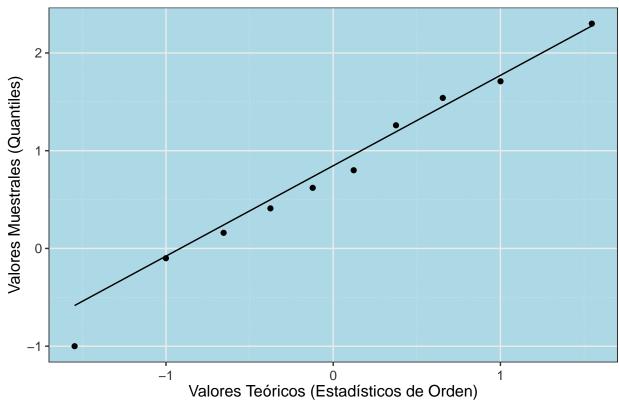
Prueba de Normalidad Univariada mediante la Prueba del Coeficiente de Correlación.

En la tabla 1, están los datos de un ejemplo visto en clase.

Tabla 1: Datos

-1 -0.1 0.16 0.41 0.62 0.8 1.26 1.54 1.71 2.3

### Normal Probaility Plot o QQ-plot



Resultados de la prueba de normalidad univariada vía la PH del coeficiente de Correlación:

#### Probability Plot Correlation Coefficient Test

data: ejemplo ppcc = 0.99387, n = 10, p-value = 0.9902 alternative hypothesis: ejemplo differs from a Normal distribution

#### Ejemplo de Notas de Clases (Manual)

Prueba del coeficiente de correlación realizada de manera manual.

El Coeficiente de Correlación es r=:0.9943596

El Valor tabulado en la tabla es:  $T_{tabla}=R(n,\alpha)=R(10,0.05)=0.9351$  luego, Como r>T-tabulado, luego no se rechaza  $H_0$ : Los datos son Normales Univariados

### TRANSFORMACIONES DE POTENCIA INDIVIDUALES (Box-Cox)

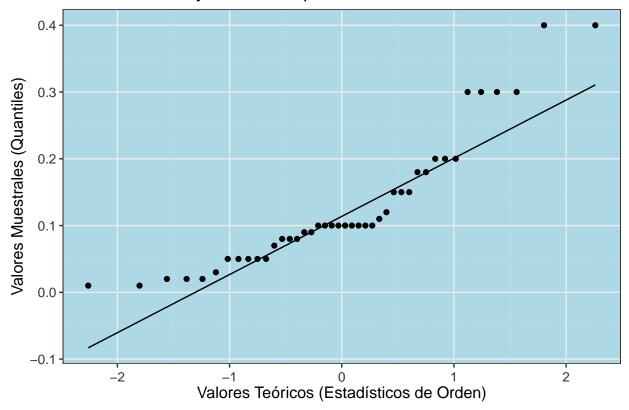
Ahora se realizan transformaciones de Box-Cox para acercar los datos a la normalidad univariada.

Primero se tiene la prueba de Shapiro Wilk para Normalidad Univariada para los datos crudos.

En la tabla 2, están los datos de un segundo ejemplo visto en clase.

Tabla 2: Datos

## Normal Probaility Plot o QQ-plot

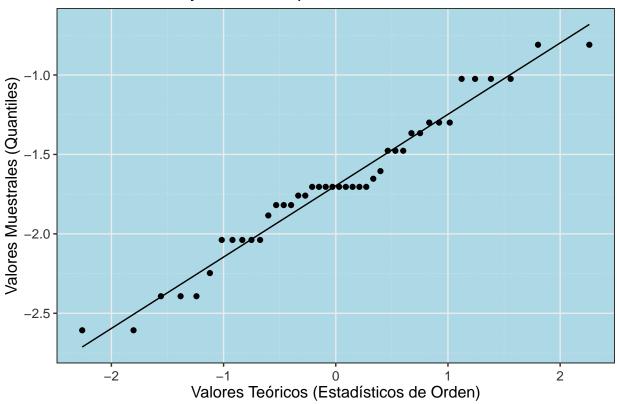


Shapiro-Wilk normality test

data: x[, 1] W = 0.85793, p-value = 9.902e-05

Ahora se tiene la prueba de Shapiro Wilk para Normalidad Univariada para los datos transformados.

## Normal Probaility Plot o QQ-plot



Shapiro-Wilk normality test

data: xt[, 1] W = 0.96519, p-value = 0.2257

Cálclulo del coeficiente de asimetría de Fisher y el coeficiente de Kurtosis para un conjunto de datos-univariado

Tabla 3: Datos Normales Univariados Generados

0.0836325	0.5487834	0.6423876	0.8294203	-0.8811785
0.4410065	1.2056960	-0.6697220	0.6689003	0.1120762
0.0803371	0.4363650	0.0319504	-1.2480960	-0.6586225
-0.2911459	0.9449975	-0.1133443	-0.1426866	0.0276856
0.1207408	-0.1470738	1.2949744	0.7510650	1.1413309
0.7947157	-1.6066513	-0.1839536	-1.1944635	-0.2230719
0.1391854	-0.7772863	-1.2232062	-0.9499444	1.2376793
1.5396700	1.8069851	-1.3582475	-0.5349300	-1.0276010
0.5320363	-0.7646701	-0.6739721	1.1738375	0.0249950
-0.9801376	1.2863604	-0.3219088	1.0494364	-0.7755493
-1.4999641	0.9882556	-1.5025817	0.0951810	1.1223061
-0.8431066	-1.7044359	0.4634107	0.3464936	-0.4544363
-1.1921148	0.5279541	0.4412060	0.3245776	1.2644989
0.3347165	0.4808874	0.1069668	2.0026733	-1.0867275
0.3527966	-0.2740565	-0.3783730	0.2218035	-1.1735483
2.0335552	-0.4755359	-0.6961383	-0.0586670	0.8524958
0.3350160	-0.4755399		0.2857717	0.8324938
		-2.1477180		
0.2445010	1.3779389	0.1344326	-0.1211018	1.6545608
2.1934106	2.6339265	0.4425881	-1.1826515	1.2226567
-0.0175913	0.5887241	-0.7600120	-2.0140733	0.2237705

El coeficiente de Asimetría para el conjunto de datos Univariado Generados es: [1] 0.08161626

El coeficiente de Kurtosis para el conjunto de datos Univariado Generados es: [1] 2.699508

#### Test Para Normalidad Univariada basado en el Coeficiente de Asimetría.

Ahora se tiene una PH de Normalidad Univariada basada en el Coeficiente de asimetría, cuyo estadístico de prueba es:

$$Z = \sqrt{n/6} * A \sim N(0,1)$$
, donde:  $A$  – Coeficiente de Asimetría.

El estadístico de prueba es: [1] 0.3332

El Valor-p de la prueba de Asimetria Normal Univariada es: [1] 0.739

#### Test Para Normalidad Univariada basado en el Coeficiente de Kurtosis.

Ahora se tiene una PH de Normalidad Univariada basada en el Coeficiente de Kurtosis, cuyo estadístico de prueba es:

$$Z = \sqrt{n/24}*(K-3) \sim N(0,1), \; \; \text{donde:} \; \; K - \text{Coeficiente de Kurtosis.}$$

El estadístico de prueba es: [1] -0.6134

El Valor-p de la prueba de Kurtosis Normal Univariada es: [1] 0.5396

### Pruebas de Normalidad Multivariada

Se utiliza la función mvn del paquete  $\mathbf{MVN}$  para realizar pruebas de Normalidad Multivariada y Univariada con distintas opciones de visualización.

En la tabla 14, se encuentra el ancabezado del conjunto de datos seleccionado.

Tabla 4: Encabezado de Datos

	V1	V2	V3	V4	V5	V6	V7
1	8	98	7	2	12	8	2
2	7	107	4	3	9	5	3
3	7	103	4	3	5	6	3
4	10	88	5	2	8	15	4
5	6	91	4	2	8	10	3
6	8	90	5	2	12	12	4
7	9	84	7	4	12	15	5
8	5	72	6	4	21	14	4
9	7	82	5	1	11	11	3
10	8	64	5	2	13	9	4
1.1	C	71	-	4	10	0	
11	6	71	5	4	10	3	3
12	6 7	91	4 7	2 4	12	7	3
13 14	10	72 70	4	2	18 11	10 7	3
15	10	70	4	1	8	10	3
10	10	12	4	1	0	10	
16	9	77	4	1	9	10	3
17	8	76	4	1	7	7	3
18	8	71	5	3	16	4	4
19	9	67	4	2	13	2	3
20	9	69	3	3	9	5	3
21	10	62	5	3	14	4	4
22	9	88	4	2	7	6	3
23	8	80	4	2	13	11	4
24	5	30	3	3	5	2	3
25	6	83	5	1	10	23	4
26	8	84	3	2	7	6	3
27	6	78	4	2	11	11	3
28	8	79	2	1	7	10	3
29	6	62	4	3	9	8	3
30	10	37	3	1	7	2	3
31	8	71	4	1	10	7	3
32	7	52	4	1	12	8	4
33	5	48	6	5	8	4	3
34	6	75	4	1	10	24	3
35	10	35	4	1	6	9	2
36	8	85	4	1	9	10	2
37	5	86	3	1	6	12	2
38	5	86	7	2	13	18	2
39	7	79	7	4	9	25	3
40	7	79	5	2	8	6	2
41	6	68	6	2	11	14	3
42	8	40	4	3	6	5	2

En la tabla 5, se tienen las salidas básicas de la función mvn. La prueba multivariada usada es la de Mardia. Favor ver la ayuda del R para esta función y explorar los distintos argumentos de la misma.

Tabla 5: Salidas Básicas de PH-NM

		Test	Ç	Statistic		p va	alue		Result	
		Mardia	a Skewness 2	20.8643968	397274	0.00	00336886	45173655	54 NO	
		Mardia	a Kurtosis 1	1.51113770	487766	0.13	30753369	968473	YES	
		MVN	1	NA		NA			NO	
			Test	Variable	Stat	tistic	p value	Norma	ality	
			Shapiro-Wilk	V2	0.93	388	0.026	NO		
		_	Shapiro-Wilk	V5	0.93	300	0.013	NO		
	n	Mean	Std.Dev	Median	Min	Max	25th	$75 \mathrm{th}$	Skew	Kurtosis
V2	42	73.85714	17.335388	76.5	30	107	68.25	84.75	-0.7332377	0.3090125
V5	42	10.04762	3.370984	9.5	5	21	8.00	12.00	0.9854394	1.2219732

En la tabla 6, se tienen las salidas de la función mvn para la prueba normal multivariada usada, en este caso la prueba de Mardia. Favor ver la ayuda del R para esta función y explorar los distintos argumentos de la misma.

Tabla 6: Prueba de Normalidad Multivariada

	Prueba	Valor Estadística	Valor-p	Resultado
1	Mardia Skewness	20.8643968397274	0.000336886451736554	NO
2	Mardia Kurtosis	1.51113770487766	0.130753369968473	YES
3	MVN	NA	NA	NO

En la tabla 7, se tienen las salidas de la función mvn para las pruebas normal univariada usada, en este caso la prueba de Shapiro-Wilk. Favor ver la ayuda del R para esta función y explorar los distintos argumentos de la misma.

Tabla 7: Prueba de Normalidad Univariada

	Prueba	Variables	Valor Estadística	Valor-p	Resultado
1	Shapiro-Wilk	V2	0.9388	0.026	NO
2	Shapiro-Wilk	V5	0.9300	0.013	NO

En la tabla 8, se tiene un resumen descriptivo de las variables del conjunto de datos p-variado usado, en este caso el conjunto de datos consta de dos variables. Favor ver la ayuda del R para esta función y explorar los distintos argumentos de la misma.

Tabla 8: Resumen Descriptivo

	n	Media	DesvEst	Mediana	Min	Máx	Per.25	Per.75	Asimetría	Kurtosis
V2	42	73.857	17.335	76.5	30	107	68.25	84.75	-0.733	0.309
V5	42	10.048	3.371	9.5	5	21	8.00	12.00	0.985	1.222

#### Coeficientes de Asimetría y Kurtosis Multivariado-Distancia de Mahalanobis

Los coeficientes de Asimetría y Kurtosis **UNIVARIADOS** son respectivamente: -0.7332377 y 0.3090125 para la variable  $V_2$  y 0.9854394 y 1.2219732 para la variable  $V_5$ .

En la tabla 9, se tienen las distracias de Mahalanobis (al cuadrado) de cada Observación al vector de medias de los datos.

Tabla 9: Distancias de Mahalanobis a la media

2.1683	0.0279	2.1058	0.0279	0.2243
4.0383	1.2260	1.7170	2.1744	4.8939
5.8674	5.8510	0.8522	2.4754	0.2243
1.1933	0.1494	8.0632	0.0049	4.8939
1.5428	0.3794	0.2908	6.0668	0.2243
1.1190	0.1479	1.3312	0.5773	4.8939
0.6223	0.8914	0.1259	2.2085	0.2243
11.0559	3.3355	1.0045	1.1584	4.8939
0.2801	1.0422	0.5349	0.2138	0.2243
1.2518	0.1609	5.0802	0.5178	4.8939

luego, en este caso el Coeficiente de Asimetría Multivariado es  $A_m=2.9806281$ .

#### El Coeficiente de Asimetría Multivariado esta dado por:

$$A_m = \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n r_{ij}^3$$

donde: 
$$r_{ij} = (\underline{X}_i - \overline{X})^T \mathbf{S}^{-1} (\underline{X}_j - \overline{X}).$$

Para i = j se tienen las distancias de Mahalanobis de cada observación al vector de medias, es decir,

$$d_M^2 \left( \ \underline{X}_i \ , \ \overline{X} \ \right) = r_{ii} = (\underline{X}_i - \overline{X})^T \mathbf{S}^{-1} (\underline{X}_i - \overline{X}).$$

luego, en este caso el Coeficiente de Asimetría Multivariado es  $A_m = 2.9806281$ .

#### El Coeficiente de Kurtosis Multivariado esta dado por:

$$K_m = \frac{1}{n} \sum_{i=1}^n r_{ii}^2 = \frac{1}{n} \sum_{i=1}^n d_M^4 \left( \underline{X}_i , \overline{X} \right)$$

luego, en este caso el Coeficiente de Kurtosis Multivariado es  $K_m$ =9.8653889=9.8653889

En la tabla 10, están las Salidas básicas de la función mvn usando la prueba multivariada de Royston. Favor ver la ayuda del R para esta función y explorar los distintos argumentos de la misma.

Tabla 10: Salidas Básicas PH NM- Royston

Test		Н	p value	MVN	Test		Variable	Statist	ic p value	Normality
Roysto	on	11.66612	0.0029292	NO	Shapiro Shapiro		V2 V5	0.9388 0.9300	0.026 0.013	NO NO
	n	Mean	Std.Dev	Media	n Min	Max	25th	75th	Skew	Kurtosis
V2 V5	42 42	73.85714 10.04762	11.000000		.5 30 .5 5	107 21	68.25 8.00		0.7332377 0.9854394	$0.3090125 \\ 1.2219732$

Tabla 11: Prueba de Normalidad Multivariada

	Prueba	Valor Estadística	Valor-p	Resultado
1	Royston	11.66612	0.00293	NO

Tabla 12: Prueba de Normalidad Univariada

	Prueba	Variables	Valor Estadística	Valor-p	Resultado
1	Shapiro-Wilk	V2	0.9388	0.026	NO
2	Shapiro-Wilk	V5	0.9300	0.013	NO

Tabla 13: Resumen Descriptivo

	n	Media	DesvEst	Mediana	Min	Máx	Per.25	Per.75	Asimetría	Kurtosis
V2	42	73.857	17.335	76.5	30	107	68.25	84.75	-0.733	0.309
V5	42	10.048	3.371	9.5	5	21	8.00	12.00	0.985	1.222

Salidas básicas de la función mvn. La prueba multivariada usada es la de Mardia. Favor ver la ayuda del R para esta función y explorar los distintos argumentos de la misma.

En la tabla 14, se encuentra el ancabezado del conjunto de datos seleccionado.

Tabla 14: Encabezado de Datos

	V1	V2	V3	V4	V5	V6	V7
1	8	98	7	2	12	8	2
2	7	107	4	3	9	5	3
3	7	103	4	3	5	6	3
4	10	88	5	2	8	15	4
5	6	91	4	2	8	10	3
6	8	90	5	2	12	12	4
7	9	84	7	4	12	15	5
8	5	72	6	4	21	14	4
9	7	82	5	1	11	11	3
10	8	64	5	2	13	9	4

En la tabla 19, se encuentran los resultados de la prueba de normalidad multivariada realizada según la prueba seleccionada de las disponibles en el argumento mvnTest de la función mvn.

Tabla 15: Prueba de Normalidad Multivariada

	Prueba	Valor Estadística	Valor-p	Resultado
1	Mardia Skewness	20.8643968397274	0.000336886451736554	NO
2	Mardia Kurtosis	1.51113770487766	0.130753369968473	YES
3	MVN	NA	NA	NO

En la tabla 20, se encuentran los resultados de las pruebas de normalidad univariadas para cada una de las variables del conjunto de datos p-variados.

Tabla 16: Prueba de Normalidad Univariada

	Prueba	Variables	Valor Estadística	Valor-p	Resultado
1	Shapiro-Wilk	V2	0.9388	0.026	NO
2	Shapiro-Wilk	V5	0.9300	0.013	NO

En la tabla 21 aparece un resumen descriptivo de las variables del conjunto de datos p-variado.

Tabla 17: Resumen Descriptivo

	n	Media	DesvEst	Mediana	Min	Máx	Per.25	Per.75	Asimetría	Kurtosis
V2	42	73.857	17.335	76.5	30	107	68.25	84.75	-0.733	0.309
V5	42	10.048	3.371	9.5	5	21	8.00	12.00	0.985	1.222

Salidas básicas de la función mvn. La prueba multivariada usada es la de Mardia. Favor ver la ayuda del R para esta función y explorar los distintos argumentos de la misma.

En la tabla 18, se encuentra el ancabezado del conjunto de datos Normal Multivariado generados a partir del R.

Tabla 18: Conjunto de Datos NM Generados

	X1	X2	Х3	X4	X5
1	-1.02154	-0.95486	-2.58488	0.91944	-0.25447
2	-1.13166	-0.32030	-0.46909	-2.83509	1.84775
3	-1.29405	0.87176	0.20089	4.10162	-3.86428
4	0.21078	0.88121	0.98641	-3.36501	4.24784
5	-1.33392	-0.31332	-2.08717	3.56056	-3.67253
6	-1.15591	-1.06091	0.30703	1.28932	0.12197
7	0.14518	-1.12889	0.97631	-0.12695	1.45841
8	1.52517	-0.50440	2.26913	1.06475	-0.11732
9	3.76283	0.68598	2.60191	-1.41957	-0.01121
10	3.30394	1.35101	0.03987	-4.07021	1.53133
11	-7.29025	-1.10183	0.33070	6.48908	-2.52212
12	1.32084	-0.25233	1.41249	-3.04809	6.51867
13	1.74325	2.99098	1.59822	-2.20685	3.80124
14	4.42005	1.47914	0.71246	-1.78364	-2.40192
15	2.70106	1.06246	1.83892	-2.17507	4.57495
16	3.98925	-0.76703	2.71813	-4.14902	2.56521
17	1.53866	-0.66269	1.51498	-0.89699	-2.43129
18	3.11135	0.74419	0.89964	-1.09409	1.21614
19	-2.23444	0.03788	0.03304	-1.54158	0.93263
20	1.42637	1.70911	1.42321	2.33529	-1.47213
21	0.23868	0.76795	-0.37300	1.76390	-1.07336
22	-0.62009	0.47643	-0.42191	1.89293	-1.30188
23	0.15392	0.11263	0.29961	-2.22380	2.52758
24	2.50588	3.81394	2.55068	-1.52580	2.06287
25	-2.81040	-2.81039	-1.55920	1.19751	-0.40857
26	7.24521	2.97247	3.03046	-1.44950	-5.16393
27	1.86221	0.23827	0.72347	-2.33896	0.40029
28	4.85007	-0.44272	1.89565	-1.71617	-1.96932
29	0.68351	0.18905	2.83752	-3.82942	-1.01478
30	-4.11431	-1.94044	-1.85048	5.20943	-1.81481
31	1.93861	-0.11652	0.98225	-2.77996	4.84061
32	3.94124	0.23650	-1.21577	-1.64965	1.07104
33	-1.86371	-0.39749	1.31339	1.45344	1.41066
34	-7.41821	-0.80707	-1.99599	4.22120	-0.86521
35	0.20430	-0.81617	2.19419	-3.80162	0.62253
36	-1.92423	-1.28060	0.62917	2.78448	-1.32278
37	-0.26518	2.92913	0.58968	2.18401	0.66489
38	2.31949	1.03405	-0.21198	0.77126	0.76629
39	-4.21663	0.16215	-2.37171	1.60848	-0.38357
40	-5.61054	-0.29921	-3.96069	4.38575	0.67226
41	1.66345	-0.41656	1.45259	1.27480	-2.82159
42	-2.85202	-1.50603	-2.32171	4.09524	-1.46580
43	-1.06118	-0.16111	-2.01110	4.59373	-0.54219
44	4.17062	-1.88230	3.99086	-3.85255	-1.26062
45	-1.35428	-0.50082	-2.13318	0.57066	3.62532
46	-5.52704	-1.11257	-1.75735	0.50975	2.11102
47	-0.21790	-0.42447	2.32306	1.14839	-0.87815
48	-4.30035	-0.53403	-1.80654	6.08585	-3.45625
49	-4.19950	0.52671	-1.83512	1.67652	-1.05329
-10	-4.13300	0.02011	-1.00012	1.01002	-1.00023

continúa en la siguiente página

Tabla 18: Conjunto de Datos NM Generados continuación

	X1	X2	X3	X4	X5
50	-0.80714	0.53535	-2.80091	2.07808	-0.67127
51	-0.74393	1.03399	1.66456	-1.79601	1.10103
52	0.10329	0.00989	0.82779	-1.73393	-0.52545
53	-4.92261	-1.77899	-4.31957	2.02500	1.09175
54	0.27296	0.17369	-0.75574	1.83803	-0.88368
55	2.33660	1.39564	1.24118	-1.03203	5.87136
56	2.01165	-0.54538	0.06815	1.82692	-4.00201
57	0.14803	3.04789	2.48776	-1.86963	-2.52940
58	-0.90899	-3.05355	-0.99737	1.82891	1.20895
59	-2.93558	0.94612	-3.48703	1.03129	0.10265
60	2.70924	-0.83158	-1.39859	-0.77266	-1.26348
61	2.41247	0.66265	-2.49441	0.09199	2.04732
62	-1.37785	-1.07894	0.78969	-1.93465	1.29888
63	4.39536	-0.58142	-1.08146	-9.62967	6.49816
64	-1.14704	-2.12408	1.38290	-0.27353	-1.39912
65	0.59054	0.30183	-0.24062	2.55165	-2.13210
66	-1.68030	2.04551	-1.61834	3.27335	0.56734
67	-5.22804	-0.76430	-1.90802	2.79183	-1.32397
68	-2.59045	-0.21452	-1.53819	0.46878	-2.01951
69	-0.56617	-0.96598	-1.09064	-0.84423	1.64677
70	-0.14117	-1.42261	0.65305	-1.03501	3.79784
71	1.15917	2.05592	-0.02840	1.14404	-3.47058
72	2.57954	-0.47218	2.11127	1.47729	-2.57592
73	1.07730	0.52812	1.21573	-1.13897	2.29351
74	-1.34087	-0.93208	1.17278	-0.64352	0.03861
75	1.14114	-0.98710	0.56826	4.34932	-5.52051
76	2.08781	-1.95101	0.68525	-1.16967	0.01281
77	-2.53567	-0.10189	4.02983	-0.39781	0.31120
78	-0.81161	-0.22715	-2.80660	-0.02615	-2.07158
79	2.96201	1.80089	-2.01712	-2.56236	1.57248
80	0.69088	-1.16016	-0.90998	-1.21926	2.25672
81	-5.23845	-1.92635	-2.32693	6.56575	-1.62383
82	-2.59621	-0.67814	0.28293	2.20755	0.37916
83	0.10326	1.29792	1.74642	-2.39373	0.10025
84	4.34946	0.67539	1.97643	-3.00793	-1.44987
85	-6.47526	-0.06404	-3.54983	2.45862	1.95258
86	0.93178	-0.88638	1.14672	-1.14625	3.02507
87	-0.82441	-0.09047	1.64685	1.48646	-1.70413
88	3.45273	1.01059	-1.62313	-1.04363	-4.87207
89	1.20201	0.10104	-0.34841	3.25711	-3.70088
90	-5.14719	0.90696	-2.01743	1.53629	-2.59760
91	0.02087	-0.46520	0.67317	3.44808	-3.53086
92	-1.65554	1.65405	-2.26081	2.53969	-1.40206
93	0.63819	0.60806	1.48027	-2.88751	0.64075
94	-2.00110	0.10395	4.13533	-2.33523	-0.02002
95	-2.59511	-1.08845	1.37599	2.37458	-3.25269
96	-2.58800	-1.04592	-0.04995	-4.48449	3.34674
97	0.18053	-1.21273	0.60488	1.59338	-4.42216
98	-1.04478	2.05068	1.11769	0.49881	-2.68938
99	2.30935	0.28834	1.45076	1.52281	-0.45415
100	0.75530	-0.64297	-1.81115	-1.24801	-0.44714

En la tabla 19, se encuentran los resultados de la prueba de normalidad multivariada realizada según la prueba seleccionada de las disponibles en el argumento mvnTest de la función mvn.

Tabla 19: Prueba de Normalidad Multivariada

	Prueba	Valor Estadística	Valor-p	Resultado
1	Mardia Skewness	46.7418263358847	0.0886296085370614	YES
2	Mardia Kurtosis	0.185753341764549	0.852638195507249	YES
3	MVN	NA	NA	YES

En la tabla 20, se encuentran los resultados de las pruebas de normalidad univariadas para cada una de las variables del conjunto de datos p-variados.

Tabla 20: Prueba de Normalidad Univariada

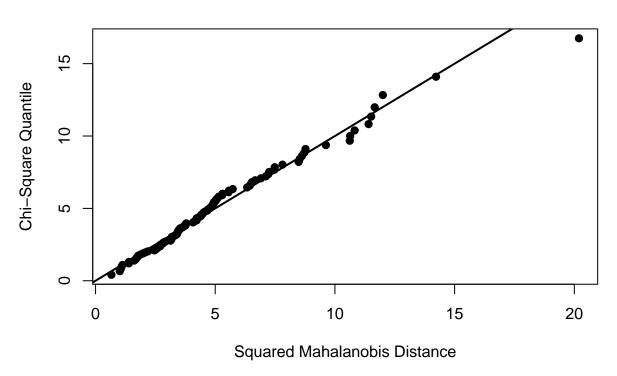
	Prueba	Variables	Valor Estadística	Valor-p	Resultado
1	Shapiro-Wilk	V1	0.9862	0.3860	YES
2	Shapiro-Wilk	V2	0.9751	0.0548	YES
3	Shapiro-Wilk	V3	0.9784	0.0993	YES
4	Shapiro-Wilk	V4	0.9784	0.1000	YES
5	Shapiro-Wilk	V5	0.9864	0.4008	YES

En la tabla 21 aparece un resumen descriptivo de las variables del conjunto de datos p-variado.

Tabla 21: Resumen Descriptivo

	n	Media	DesvEst	Mediana	Min	Máx	Per.25	Per.75	Asimetría	Kurtosis
V1	100	-0.201	2.862	0.103	-7.418	7.245	-1.726	1.773	-0.285	-0.071
V2	100	-0.013	1.272	-0.188	-3.054	3.814	-0.845	0.701	0.515	0.466
V3	100	0.068	1.860	0.319	-4.320	4.135	-1.620	1.415	-0.148	-0.652
V4	100	0.229	2.734	0.484	-9.630	6.566	-1.721	1.926	-0.164	0.568
V5	100	-0.153	2.501	-0.319	-5.521	6.519	-1.732	1.327	0.361	0.040

Resultados con Gráfico QQ-Plot para la PH de Normalidad Multivariada utilizando la función **mvn**.



Prueba de Shapiro\_Multivariada para normalidad con la función *mshapiro.test* del paquete: **RVAideMemoire**.

Multivariate Shapiro-Wilk normality test

data: (V1,V2,V3,V4,V5) W = 0.95386, p-value = 0.001504

#### Prueba de Norm. multivariada basada en Asimetría

A continuación se tienen los resultados de la PH para NM basada en Asimetría, utilizando la función munorm.skew.test del paquete ICS. Favor exlorar este paquete y sus funciones disponibles.

Multivariate Normality Test Based on Skewness

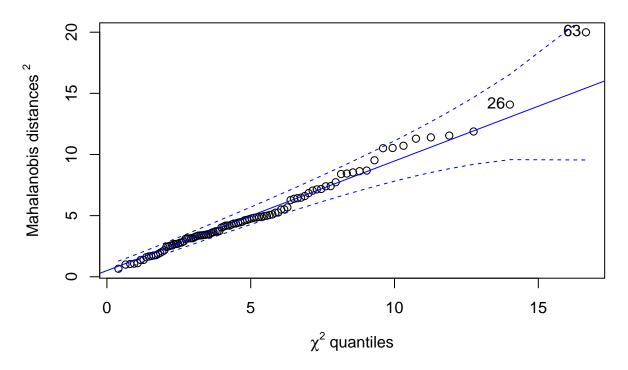
data: datos\_gen U = 7.1974, df = 5, p-value = 0.2064

Multivariate Normality Test Based on Skewness

data: datos[, c(2, 5, 6)] U = 5.0067, df = 3, p-value = 0.1713

Gráfico qq-plot para Normalidad Multivariada con la función mqqnorm.

## Gráfico QQ-Plot para Normalidad Multivariada



[1] 63 26

Resultados con Gráfico Box-Plot para la PH de Normalidad Multivariada utilizando la función **mvn**.

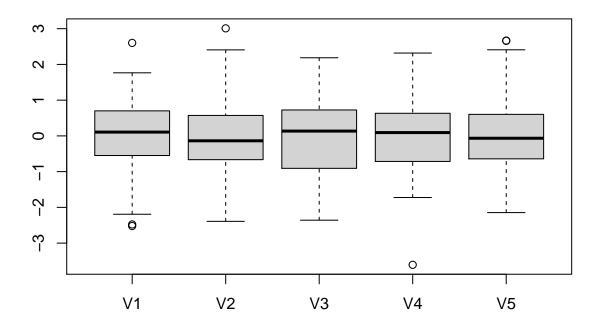


Tabla 22: Salidas sin Resumen Descriptivo

Test	Statistic	p value	Result
Mardia Skewness	46.7418263358847	0.0886296085370614	YES
Mardia Kurtosis	0.185753341764549	0.852638195507249	YES
MVN	NA	NA	YES

Test	Variable	Statistic	p value	Normality
Shapiro-Wilk	V1	0.9862	0.3860	YES
Shapiro-Wilk	V2	0.9751	0.0548	YES
Shapiro-Wilk	V3	0.9784	0.0993	YES
Shapiro-Wilk	V4	0.9784	0.1000	YES
Shapiro-Wilk	V5	0.9864	0.4008	YES

En la tabla 23, se tienen los resultados de PH con la función mvn junto con Observaciones atípicas y transformaciones de Box-Cox.

### Chi-Square Q-Q Plot

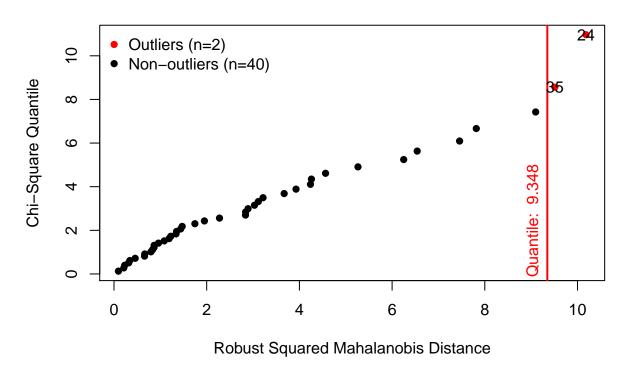


Tabla 23: Salidas Obs. Atípicas y Tranf. de Box-Cox

Statistic

Test

24

35

24

35

Observation

				dia Skewness dia Kurtosis N		17.0256138296024 0.0737991289800862 0.373894975940776 0.708482454435366 NA NA		YES YES YES						
				Test		Variable	Statistic	р	value	Normalit	у			
				Shapiro-W		V2	0.9694	-	.3155	YES				
				Shapiro-W Shapiro-W		V5 V6	0.9815 $0.9729$		.7200 .4109	YES YES				
				Shapho-w	IIK	VO	0.9129	0.	.4109	1 E3				
	n	Mea	n	Std.Dev		Median		Min		Max	25th	75th	Skew	Kur
V2	42	1205.106759	8 42	4.7391971	12	40.5067401	266.6349	9930	2152	.3514225	1028.5059285	1467.7200408	-0.2782877	-0.035
V5	42	0.682780	5	0.0375328		0.6824644	0.5962	2895	0	7608499	0.6557382	0.7024814	0.0573767	-0.404
V6	42	1.370977	4	0.1262033		1.3798332	1.1099	9916	1	.6235217	1.3096332	1.4347636	-0.1617803	-0.258

Outlier

TRUE

TRUE

V2

V5

1.6423277

-0.1698227

p value

Result

V6 0.1505487

Mahalanobis Distance

En la tabla 24, se encuentran los resultados de la prueba de normalidad multivariada realizada según la prueba seleccionada de las disponibles en el argumento mvnTest de la función mvn.

10.182

9.511

Tabla 24: Prueba de Normalidad Multivariada

	Prueba	Valor Estadística	Valor-p	Resultado
1	Mardia Skewness	17.0256138296024	0.0737991289800862	YES
2	Mardia Kurtosis	0.373894975940776	0.708482454435366	YES
3	MVN	NA	NA	YES

En la tabla 25, se encuentran los resultados de las pruebas de normalidad univariadas para cada una de las variables del conjunto de datos p-variados.

Tabla 25: Prueba de Normalidad Univariada

	Prueba	Variables	Valor Estadística	Valor-p	Resultado
1	Shapiro-Wilk	V2	0.9694	0.3155	YES
2	Shapiro-Wilk	V5	0.9815	0.7200	YES
3	Shapiro-Wilk	V6	0.9729	0.4109	YES

En la tabla 26 aparece un resumen descriptivo de las variables del conjunto de datos p-variado.

Tabla 26: Resumen Descriptivo

	n	Media	DesvEst	Mediana	Min	Máx	Per.25	Per.75	Asimetría	Kurtosis
V2	42	1205.107	424.739	1240.507	266.635	2152.351	1028.506	1467.720	-0.278	-0.036
V5	42	0.683	0.038	0.682	0.596	0.761	0.656	0.702	0.057	-0.404
V6	42	1.371	0.126	1.380	1.110	1.624	1.310	1.435	-0.162	-0.259

En la tabla 27, se muestran las observaciones atípicas en los datos.

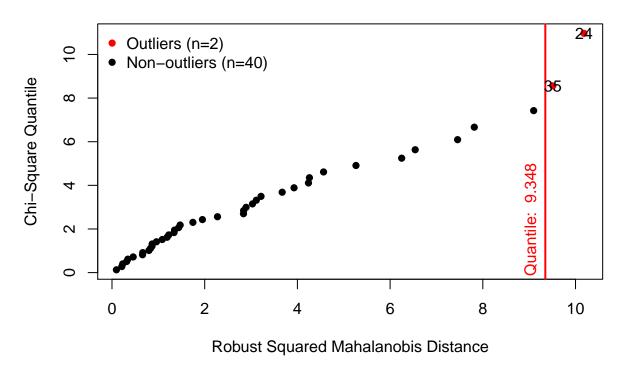


Tabla 27: Observaciones Atípicas

	Observation	Mahalanobis Distance	Outlier
24	24	10.182	TRUE
35	35	9.511	TRUE

En la tabla 28, se muestran los valores de Lambda para las transformaciones de Box-Cox.

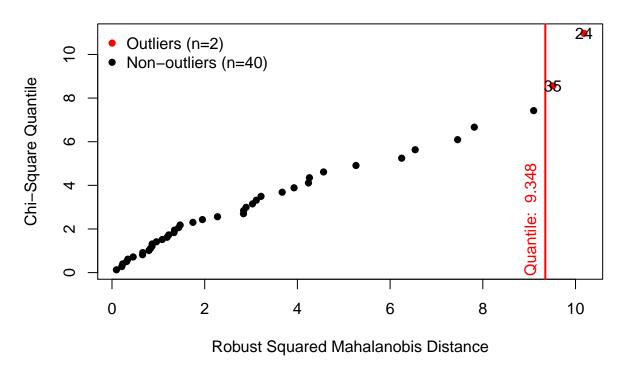


Tabla 28: Valores de lambda para Box-Cox

	x
V2	1.6423277
V5	-0.1698227
V6	0.1505487

En la tabla 29, se muestran los nuevos datos transfromados con Box-Cox.

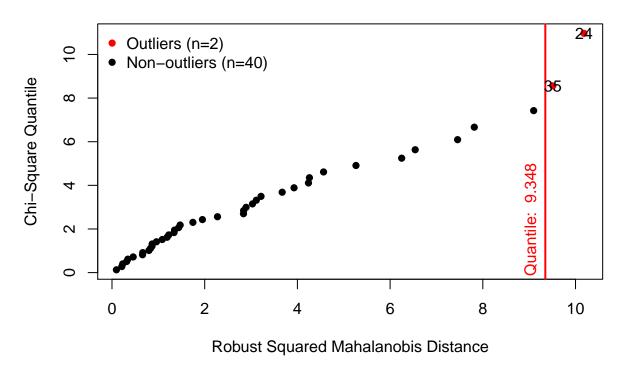


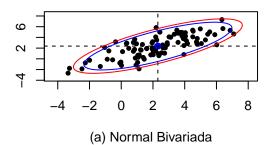
Tabla 29: Datos transformados con Box-Cox

	V2	V5	V6
1	1863.1410	0.6557382	1.367600
10	925.4211	0.6468851	1.392067
11	1097.4200	0.6763590	1.179859
12	1649.6353	0.6557382	1.340381
13	1122.9194	0.6121054	1.414323
14	1072.1502	0.6654997	1.340381
15	1122.9194	0.7024814	1.414323
16	1253.8222	0.6885698	1.414323
17	1227.1913	0.7185933	1.340381
18	1097.4200	0.6244722	1.232081
19	997.7304	0.6468851	1.109992
2	2152.3514	0.6885698	1.274175
20	1047.1112	0.6885698	1.274175
21	878.4044	0.6387949	1.232081
22	1561.2690	0.7185933	1.309633
23	1335.0496	0.6468851	1.434764
25	1418.2575	0.6763590	1.603269
26	1446.4291	0.7185933	1.309633
27	1280.6763	0.6654997	1.434764
28	1307.7524	0.7185933	1.414323
29	878.4044	0.6885698	1.367600
3	2021.8006	0.7608499	1.309633
30	376.2713	0.7185933	1.109992
31	1097.4200	0.6763590	1.340381
32	658.0211	0.6557382	1.367600
33	576.9644	0.7024814	1.232081
34	1200.7844	0.6763590	1.613575
36	1474.8170	0.6885698	1.414323
37	1503.4202	0.7376532	1.453682
38	1503.4202	0.6468851	1.545182
39	1307.7524	0.6885698	1.623522
4	1561.2690	0.7024814	1.503346
40	1307.7524	0.7024814	1.309633
41	1022.3042	0.6654997	1.487812
42	427.6687	0.7376532	1.274175
5	1649.6353	0.7024814	1.414323
6	1619.9686	0.6557382	1.453682
7	1446.4291	0.6557382	1.503346
8	1122.9194	0.5962895	1.487812
9	1390.3031	0.6654997	1.434764

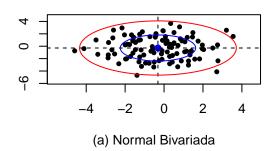
### Función R que representa datos bivariados, junto con elipses de confianza

Con esta función se realzia un gráfico de dispersión de dos variables junto con dos elipses de confianza del  $(1 - \alpha_1)100\%$  y  $(1 - \alpha_2)100\%$ . El centro de la elipse está representado por un punto de color azul.

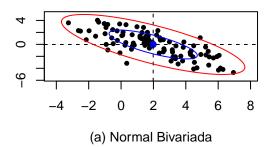
Con  $\mu$  y  $\Sigma$  dados,  $(1-\alpha_1)$ ,  $(1-\alpha_2)100\%$ 



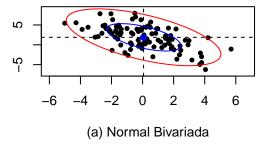
Con  $\mu$  y  $\Sigma$  dados,  $(1-\alpha_1)$ ,  $(1-\alpha_2)100\%$ 



Con  $\mu$  y  $\Sigma$  dados,  $(1-\alpha_1)$ ,  $(1-\alpha_2)100\%$ 



Con  $\mu$  y  $\Sigma$  dados,  $(1-\alpha_1)$ ,  $(1-\alpha_2)100\%$ 



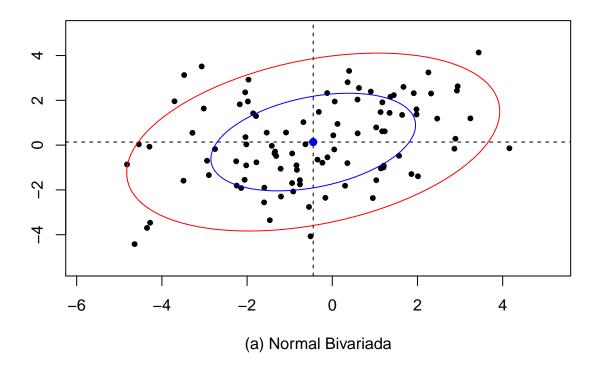
Se utiliza la función genPositiveDefMat del paquete **clusterGeneration** para generar Matrices Simétricas Definida-Positiva que se usan como Matrices de Var-Cov  $\Sigma$  de los datos Normales Multivariados generados posteriormente.

La matriz de Var-Cov Utilizada es:

Tabla 30: Matrix Var-Cov

5.783935	1.690361
1.690361	3.868633

Con  $\underline{\mu}$  y  $\Sigma$  dados,  $(1-\alpha_1)$ ,  $(1-\alpha_2)100\%$ 



Ahora para un conjunto de **datos bi-variados dado**, se utiliza la función anterior para representar los datos gráficamente junto a elipses de confianza específicas.

Tabla 31: Encabezado de Datos

12 8 9 5 5 6 8 15 8 10	V5	V6
5 6 8 15	12	8
8 15	9	5
	5	6
8 10	8	15
	8	10

El Vector de medias y La matriz de Var-Cov Utilizadas son:

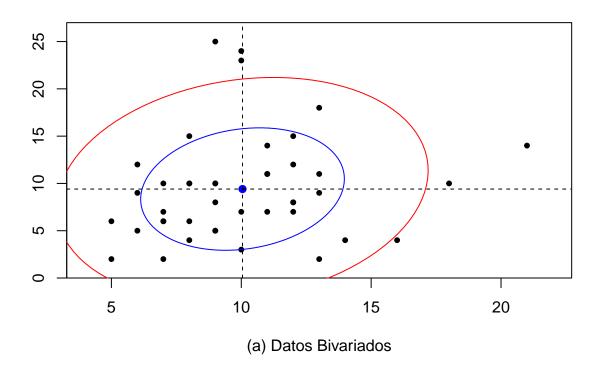
Tabla 32: Vector de Medias

	X
V5	10.047619
V6	9.404762

Tabla 33: Matriz de var-Cov

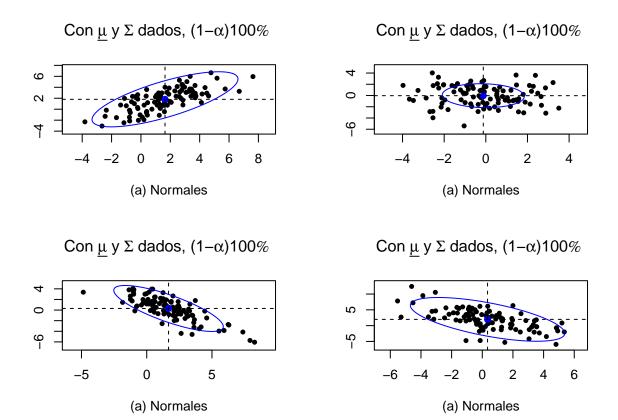
	V5	V6
V5	11.363531	3.126597
V6	3.126597	30.978513

Con  $\underline{\mu}$  y  $\Sigma$  dados, (1– $\alpha_1$ ), (1– $\alpha_2$ )100%



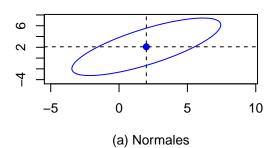
### Función R que representa datos bivariados, junto con el vector de medias

Con esta función se realzia un gráfico de dispersión de dos variables junto con una elipse de confianza del  $(1-\alpha)100\%$ . El centro de la elipse está representado por un punto de color azul.

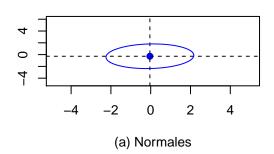


Con esta función se realzia una elipse de confianza del  $(1-\alpha)100\%$ . El centro de la elipse está representado por un punto de color azul.

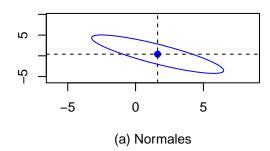
Con  $\mu$  y  $\Sigma$  dados,  $(1-\alpha)100\%$ 



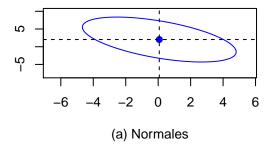
Con  $\mu$  y  $\Sigma$  dados,  $(1-\alpha)100\%$ 



Con  $\mu$  y  $\Sigma$  dados,  $(1-\alpha)100\%$ 



Con  $\underline{\mu}$  y  $\Sigma$  dados,  $(1-\alpha)100\%$ 



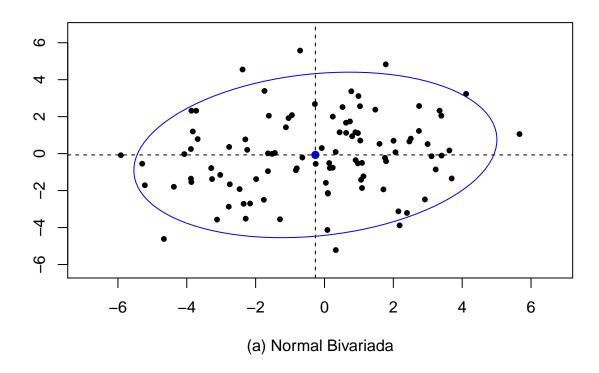
Se utiliza la función genPositiveDefMat del paquete **clusterGeneration** para generar Matrices Simétricas Definida-Positiva que se usan como Matrices de Var-Cov  $\Sigma$  de los datos Normales Multivariados generados posteriormente.

La matriz de Var-Cov Utilizada es:

Tabla 34: Matriz de var-Cov

6.829572	1.697678
1.697678	5.921997

Con  $\mu$  y  $\Sigma$  dados,  $(1-\alpha)100\%$ 



Con esta función se realzia un gráfico de dispersión de dos variables junto con una elipse de confianza del  $(1-\alpha)100\%$ . El centro de la elipse está representado por un punto de color azul.

Tabla 35: Encabezado de Datos

V5	V6
12	8
9	5
5	6
8	15
8	10

El Vector de medias y La matriz de Var-Cov Utilizadas son:

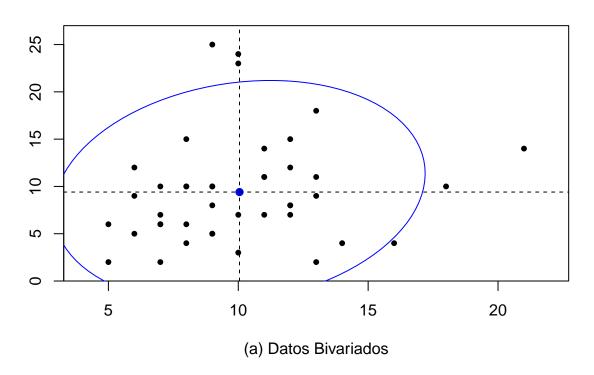
Tabla 36: Vector de Medias

	X
V5	10.047619
V6	9.404762

Tabla 37: Matriz de var-Cov

	V5	V6
V5	11.363531	3.126597
V6	3.126597	30.978513

Con  $\mu$  y  $\Sigma$  dados,  $(1-\alpha)100\%$ 



Ahora para un conjunto de **datos bi-variados dado**, se utiliza la función anterior para representar una elipse de confianza específica.

Tabla 38: Encabezado de Datos

V5	V6
12	8
9	5
5	6
8	15
8	10

El Vector de medias y La matriz de Var-Cov Utilizadas son:

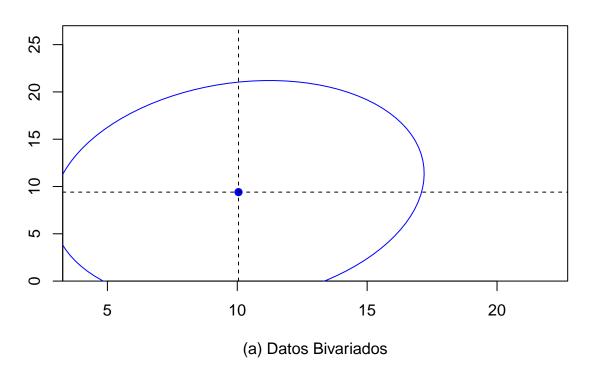
Tabla 39: Vector de Medias

	X
V5	10.047619
V6	9.404762

Tabla 40: Matriz de var-Cov

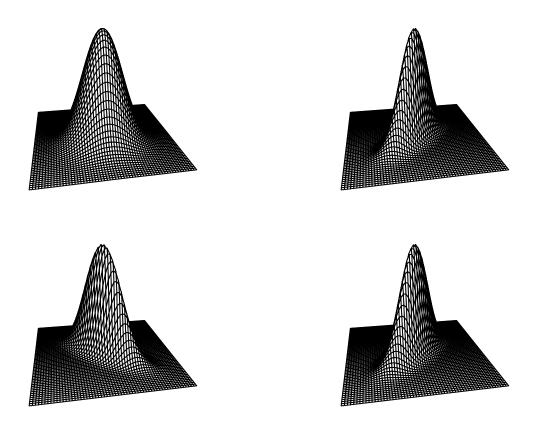
	V5	V6
V5	11.363531	3.126597
V6	3.126597	30.978513

Con  $\underline{\mu}$  y  $\Sigma$  dados,  $(1-\alpha)100\%$ 

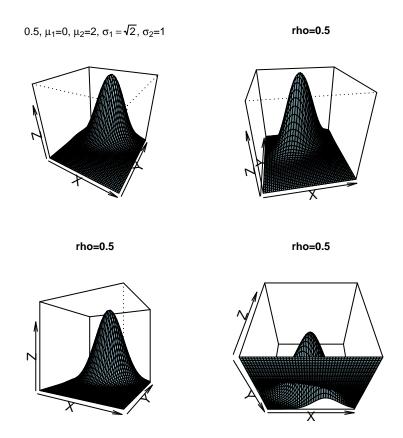


### Grafica de Superficies Normaales Bivariadas

Ahora se grafican las superficies de varias f.d.p. normal multivariada con disptintas perspectivas de vistas.



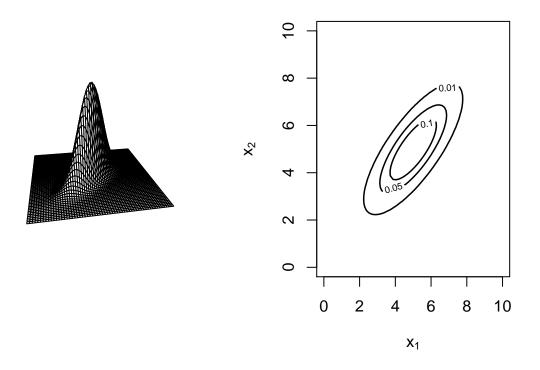
Más ejemplos de superficies de f.d.p normal multivariadas en varias perspectivas de vistas:



### Grafica de Contornos de Superficies de Normaales Bivariadas

Ahora se tienen gráficos de contornos de f.f.p normal multivariada.

### Contornos de verosimilitud del 90%,95% y 99



## Contornos de verosimilitud del 90%,95% y 99

