

INTRODUCCION A LOS METODOS MULTIVARIADOS

Dr. Porfirio Gutiérrez González

JJMV/PGG

EJEMPLO: DATOS DE INDICE DE DESARROLLO HUMANO EN ALGUNOS PAISES

	V ₁	V ₂	V ₃	V ₄	V ₅	V ₆	V ₇	V ₈	V ₉	V ₁₀	V ₁₁	V ₁₂	V ₁₃	V ₁₄	V ₁₅	V ₁₆
URU	72.2	7.8	5916	0.881	2620	3.3	101	21	96	14	510	163	2.4	47	41	67
CHI	71.8	7.5	5099	0.864	1950	2.9	102	17	93	14	1230	68	2.7	74	26	52
ARG	71.0	8.7	4295	0.832	2380	1.5	131	30	95	14	370	115	2.8	62	34	53
VEN	70.0	6.3	6169	0.824	2560	4.1	99	34	90	9	700	92	3.2	71	21	56
COL	68.8	7.1	4237	0.770	1260	2.9	106	38	86	30	1230	81	2.7	45	39	68
BRA	65.6	3.9	4718	0.730	2680	3.9	114	59	80	25	1080	96	2.9	25	21	47
EQU	66.0	5.6	3074	0.646	960	2.7	105	59	84	44	820	36	3.8	121	33	46
PAR	67.1	4.9	2790	0.641	1090	1.0	116	48	88	52	1460	27	4.4	41	11	31
PER	63.0	6.4	2622	0.592	1100	3.5	87	80	79	30	1040	31	3.7	59	11	53
BOL	54.5	4.0	1572	0.398	630	2.4	84	89	71	49	1530	29	4.7	101	40	34

V₁ : Esperanza de vida al nacimiento, 1990

V₂ : Tiempo medio de escolarización (años), 1990

V₃ : Producto Interior Bruto, per capita (US\$), 1990

(*) V₄ : Índice de Desarrollo Humano

(-) V₅ : Producto Nacional Bruto per capita, (US\$), 1990

V₆ : Gasto Público en la Instrucción (en % del PNB), período 1988-90

(-) V₇ : Aporte Calórico Cotidiano (en % de las necesidades normalizadas), 1988-90

V₈ : Taza de mortalidad de menores de 15 años (por 1000 nacidos vivos), 1991

**V₉ : Taza de alfabetización de las mujeres
(en % de la pop. femenina de edad sup. a 15 años), 1990**

V₁₀ : Población Rural (en % de la Pop. Total), 1991

V₁₁ : Cantidad de habitantes por médico, período 1984-89

V₁₂ : Cantidad de teléfonos por 1000 hab., período 1986-88

V₁₃ : Taza de fertilidad, 1991

V₁₄ : Monto de la Deuda Total (en % del PNB), 1990

**V₁₅ : Importancia de la carga de la deuda
(en % de las exportaciones de bienes y servicios), 1990**

V₁₆ : Empleados del Sector Servicios (en % de la Pop. Activa), 1989-91

EJEMPLO: Resultados de biometrías hemáticas

<i>MUESTRA</i>	<i>WBC</i>	<i>RBC</i>	<i>Hgb</i>	<i>Hct</i>	<i>MCV</i>	<i>MCH</i>	<i>MCHC</i>	<i>Plt</i>	<i>LNF</i>	<i>MNC</i>	<i>NTF</i>	<i>ESF</i>	<i>BSF</i>	<i>SE</i>	<i>BN</i>
<i>A1</i>	7.6	4.28	13	38.7	90.5	30.5	33.7	265	32	6	55	6	1	54	1
<i>A2</i>	7.2	5.43	15.8	46.7	85.9	29.2	34	367	22	3	71	4	0	68	3
<i>A3</i>	6.6	5.33	15.3	46.1	86.5	28.8	53.3	260	44	4	50	2	0	48	2
<i>A4</i>	7.6	5.58	7	50.6	90.7	30.5	33.6	281	39	3	53	5	0	51	2
<i>A5</i>	6.7	5.11	14.9	43.7	85.6	29.1	34	239	39	2	53	4	2	48	5
<i>A6</i>	8.7	4.81	14.1	41.2	85.6	39.2	34.1	308	50	5	41	2	1	40	1
<i>A7</i>	7.7	4.95	15.9	46.4	93.8	32.2	34.3	316	31	3	64	1	1	62	2
<i>A8</i>	7.6	4.3	13	38.3	89.1	30.3	34	287	40	3	48	5	0	47	1
<i>A9</i>	11.1	4.35	13.7	41.1	84.8	28.2	53.3	318	19	2	55	22	2	54	1
<i>A10</i>	8.2	4.56	14.1	41.8	91.7	30.9	33.7	329	38	4	51	6	1	49	2
<i>A11</i>	9	4.49	14.2	41.6	92.6	41.6	34.1	242	43	6	48	3	0	48	0
<i>A12</i>	8.5	4.45	13.5	40	89.8	30.2	33.6	187	32	4	60	3	1	58	2
<i>A13</i>	8.5	4.93	14	41.5	84.2	28.3	33.6	291	39	3	54	4	0	54	0
<i>A14</i>	9.5	4.83	14.4	42.4	87.8	29.9	34	274	34	3	52	9	2	48	5
<i>A15</i>	11.3	4.44	18.7	41.6	93.7	31	33	294	28	4	66	2	0	65	1
<i>A16</i>	8.9	4.96	15	44.8	90.3	30.3	33.5	222	25	7	65	3	0	62	3
<i>A17</i>	7	4.51	14.5	42.3	93.6	32.1	34.3	253	42	5	51	2	0	50	1
<i>A18</i>	5.7	4.52	14.1	41.9	92.7	51.2	33.2	286	35	6	54	5	0	53	1
<i>A19</i>	8.5	4.75	14.1	40.9	86.3	29.8	34.5	386	21	2	76	1	0	75	1
<i>A20</i>	8.4	4.69	15.5	44.4	94.8	33	34.8	275	28	3	66	3	0	65	1

EJEMPLO: Resultados de medidas de marcas de tequila

Category	Núm de mue	pH	conductividad	viscosidad	Densidad	Velocidad de sonido	Indice Refraccion
Aged	1	3.72	48.9	0.0025	0.94786	1596.5	1.35297
Aged	2	3.44	162.87	0.00221	0.98448	1550.21	1.3376
Aged	3	4.11	41.6	0.00238	0.95481	1612.52	1.35156
Aged	4	3.82	29.28	0.00231	0.9641	1620.17	1.35063
Aged	5	3.98	25.68	0.00252	0.94738	1595.07	1.35327
Aged	6	4.22	34.1	0.00247	0.95261	1608.45	1.3516
Silver	7	4.05	18.42	0.00248	0.94851	1597.98	1.35303
Gold	8	4.87	41.6	0.00234	0.95409	1611.64	1.3515
Aged	9	3.99	30.3	0.00249	0.94831	1596.51	1.3534
Aged	10	4.07	39.8	0.00249	0.94552	1589.53	1.3535
Silver	11	3.62	103.47	0.00194	0.96814	1613.03	1.3449
Silver	12	3.95	50.03	0.0025	0.94802	1596.7	1.353
Aged	13	4.14	35.17	0.00218	0.95232	1606.35	1.346
Silver	14	4.76	13.09	0.0025	0.94612	1591.84	1.35353
Gold	15	3.9	44.5	0.00243	0.95196	1606.42	1.3522
Gold	16	3.97	55.8	0.00246	0.94927	1599.58	1.35267
Gold	17	5.25	222.03	0.00267	1.01763	1631.45	1.3617
Aged	18	4.23	19.43	0.00243	0.95353	1607.88	1.35213
Aged	19	4.05	28.18	0.00247	0.9505	1603.46	1.35243
Silver	20	4.06	25.55	0.00245	0.94787	1596.21	1.3529
Aged	21	4.72	56.77	0.00239	0.95392	1611.18	1.3513
Aged	22	3.73	38.13	0.00237	0.95315	1609.13	1.35126
Aged	23	3.8	31.67	0.00241	0.9535	1607.76	1.3515
Aged	24	3.63	40.67	0.00248	0.9481	1596.28	1.35143
Aged	25	3.64	48.43	0.00247	0.94774	1595.07	1.35313
Aged	26	4.07	35	0.00248	0.9479	1595.59	1.35303
Aged	27	4.2	37.73	0.00243	0.95142	1604.65	1.3516
Aged	28	4.61	20.04	0.00236	0.95572	1611.06	1.35126
Aged	29	3.98	36.47	0.00245	0.95489	1609.68	1.3523
Aged	30	4.55	29.13	0.00251	0.95106	1601.03	1.3525
Extra aged	31	3.95	31.33	0.00251	0.95167	1600.14	1.3535
Extra aged	32	3.97	24.59	0.00247	0.94933	1596.28	1.35293
Silver	33	5.27	13.46	0.00247	0.95005	1602.26	1.35233
Extra aged	34	4.06	34.43	0.00249	0.9497	1599	1.35363
Aged	35	4.08	17.67	0.00248	0.94911	1597.41	1.35303
Silver	36	4.56	38.87	0.00249	0.9489	1598.81	1.35297
Gold	37	4.52	40.83	0.00244	0.94922	1599.07	1.3532
Aged	38	4.56	49.5	0.00246	0.95025	1602.03	1.35263
Aged	39	4.86	20.19	0.00247	0.95143	1602.65	1.35306
Extra aged	40	4.7	28.52	0.00244	0.9487	1598.43	1.3531
Aged	41	4.02	20.12	0.00247	0.95046	1598.69	1.35316
Silver	42	4.28	19.99	0.00249	0.94872	1597.97	1.35313
Aged	43	4.58	27.26	0.00235	0.95348	1608.99	1.35087
Extra aged	44	4.42	40.3	0.00253	0.94612	1590.15	1.3541
Silver	45	3.97	18.13	0.00232	0.95471	1610.59	1.3509
Aged	46	4.15	27.5	0.00239	0.95343	1607.91	1.35076
Extra aged	47	4.12	43.67	0.00248	0.95042	1601.57	1.35166
Aged	48	4.3	11.72	0.00248	0.95272	1607.36	1.352
Aged	49	3.61	32.1	0.00248	0.94476	1586.2	1.35406
Extra aged	50	3.75	33.4	0.00252	0.94637	1585.99	1.3543
Silver	51	4.19	12.18	0.0026	0.94476	1586.95	1.3535
Aged	52	4.14	14.72	0.00258	0.9455	1587.57	1.3541
Extra aged	53	3.97	23.26	0.00254	0.94484	1586.91	1.35417

ESTADISTICA MULTIVARIADA

Organización de los datos

		VARIABLES				
		X_{11}	X_{12}	X_{13}	X_{1p}
		X_{21}	X_{22}	X_{23}	X_{2p}
		X_{31}	X_{32}	X_{33}	X_{3p}
		X_{41}	X_{42}	X_{43}	X_{4p}
	
	
	
	
		X_{n1}	X_{n2}	X_{n3}	X_{np}
INDIVIDUOS	$\mathbf{X} =$					

MEDIDAS DE TENDENCIA CENTRAL Y VARIACION DE MARCAS DE TEQUILA

	Recuento	Promedio	Varianza	Desviación Estándar
pH	53	4.17283	0.161632	0.402035
conductividad	53	39.0109	1200.96	34.6549
viscosidad	53	0.0024438	1.19E-08	0.000109267
Densidad	53	0.952549	0.000122716	0.0110777
Velocidad de sonido	53	1600.49	129.061	11.3605
Índice refracción	53	1.35221	8.65159E-06	0.00294136
	VARIANZA TOTAL		1330.182763	

Desviación estándar

$$S = \sqrt{S^2}$$

Varianza

$$S^2 = \left(\frac{1}{n-1} \right) \sum_{i=1}^n (X_i - \bar{X})^2$$

Media

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$$

Estandarización de una matriz

Una razón para estandarizar una matriz de datos, se debe a que las diferentes unidades que se utilizan para medir diferentes variables pueden afectar los resultados de los análisis.

De esta forma se evitan los problemas de escala que pudieran existir entre las variables.

Si se estandarizan las variables se logra que todas estén en igualdad de circunstancias. Para estandarizar las variables, lo que se hace es restar a cada variable su media y el resultado se divide por su desviación estándar; con lo que se logra que la media de los datos estandarizados es cero y su desviación estándar es uno.

La matriz de datos estandarizados se denota por

$$Z_{ij} = \frac{X_{ij} - \bar{X}_i}{S_i}$$
$$\mathbf{Z} = \begin{bmatrix} z_{11} & z_{12} & z_{13} & \cdots & z_{1p} \\ z_{21} & z_{22} & z_{23} & \cdots & z_{2p} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ z_{n1} & z_{n2} & z_{n3} & \cdots & z_{np} \end{bmatrix}$$

Ejemplo

$$\mathbf{X} = \begin{array}{c|cccc} & X_1 & X_2 & X_3 & X_4 \\ \hline & 10 & 15 & 21 & 12 \\ & 15 & 16 & 29 & 16 \\ & 20 & 18 & 16 & 25 \\ & 30 & 25 & 19 & 30 \end{array}$$

Promedio de las variables

$$\overline{X}_1 = 18.75, \overline{X}_2 = 18.5, \overline{X}_3 = 21.25, \overline{X}_4 = 20.75$$

Desviación estándar de las variables

$$Sx_1 = 8.539125638, Sx_2 = 4.509249753, Sx_3 = 5.560275773, Sx_4 = 8.220908303$$

Matriz estandarizada

$$Z = \begin{vmatrix} -1.024695077 & -0.776182335 & -0.044961799 & -1.064359275 \\ -0.439155033 & -0.554415953 & 1.39381573 & -0.577795035 \\ 0.146385011 & -0.110883191 & -0.944197773 & 0.516974505 \\ 1.317465098 & 1.441481478 & -0.404656488 & 1.125179805 \end{vmatrix}$$

El promedio de cada columna es cero y la desviación estándar de cada columna es uno.

MATRIZ ESTANDARIZADA MARCAS DE TEQUILA

pH	conductividad	viscosidad	Densidad	Velocidad de sonido	Indice Refraccion
-1.12634472	0.285359358	0.51459687	-0.42328281	-0.351216936	0.258383877
-1.8228015	3.574071776	-2.13937952	2.88245755	-4.425861538	-4.967090054
-0.15627993	0.074710936	-0.58360026	0.20410374	1.058932265	-0.22098621
-0.87761016	-0.280794347	-1.22421525	1.04272548	1.73231812	-0.537166481
-0.47963486	-0.3846576	0.69762972	-0.46661311	-0.477091677	0.360377512
0.117328093	-0.141708676	0.24004759	0.00550656	0.700673386	-0.207387059
-0.30552066	-0.594169944	0.33156402	-0.36460637	-0.22094098	0.278782604
1.734102752	0.074710936	-0.94966597	0.1391083	0.981470886	-0.241384938
-0.4547614	-0.251361279	0.42308044	-0.38266066	-0.350336693	0.404574755
-0.25577375	0.022770229	0.42308044	-0.634518	-0.96474627	0.438572633
-1.37507928	1.860028452	-4.61032305	1.40742212	1.103824656	-2.485244921
-0.55425523	0.317966579	0.51459687	-0.40883938	-0.333612077	0.26858324
-0.08165956	-0.110832812	-2.4139288	-0.02067216	0.515822367	-2.111268257
1.460494733	-0.747972148	0.51459687	-0.58035513	-0.761410149	0.448771997
-0.67862251	0.158393185	-0.12601812	0.05316988	0.521984068	-0.003399788
-0.50450831	0.4844654	0.14853116	-0.29600007	-0.080102108	0.156390241
2.679294091	5.281189673	2.07037613	5.87495599	2.725232164	3.226398673
0.14220155	-0.565025437	-0.12601812	0.08855629	0.650499538	-0.027198303
-0.30552066	-0.31253589	0.24004759	-0.18496619	0.261432155	0.074795333
-0.28064721	-0.388427033	0.05701473	-0.4223801	-0.376743981	0.234585362
1.361000908	0.51245567	-0.49208383	0.12376215	0.94097971	-0.309380695
-1.10147126	-0.025419205	-0.67511668	0.05425314	0.760529906	-0.322979846
-0.92735707	-0.21182863	-0.30905097	0.08584815	0.639936623	-0.241384938
-1.35020583	0.047874904	0.33156402	-0.40161766	-0.370582281	-0.265183453
-1.32533237	0.271797062	0.24004759	-0.43411538	-0.477091677	0.312780482
-0.25577375	-0.115738323	0.33156402	-0.41967195	-0.431319044	0.278782604
0.067581181	-0.036961584	-0.12601812	-0.10191646	0.366181066	-0.207387059
1.087392889	-0.547423308	-0.76663311	0.28625076	0.930416795	-0.322979846
-0.47963486	-0.073320079	0.05701473	0.21132546	0.808943268	0.030598091
0.938152151	-0.285122739	0.6061133	-0.13441418	0.047533119	0.098593848
-0.55425523	-0.221639653	0.6061133	-0.0793486	-0.030808503	0.438572633
-0.50450831	-0.416128744	0.24004759	-0.29058379	-0.370582281	0.244784725
2.729041004	-0.737295447	0.24004759	-0.22558834	0.155803002	0.040797454
-0.28064721	-0.132186213	0.42308044	-0.25718335	-0.131156199	0.482769875
-0.23090029	-0.615811905	0.33156402	-0.3104435	-0.271114828	0.278782604
0.963025607	-0.004065803	0.42308044	-0.32940051	-0.147880815	0.258383877
0.863531782	0.052491855	-0.03450169	-0.30051364	-0.124994498	0.336578997
0.963025607	0.302672926	0.14853116	-0.20753405	0.135557414	0.14279109
1.709229296	-0.543094916	0.24004759	-0.10101375	0.190132477	0.288981968
1.311253995	-0.302724867	-0.03450169	-0.3474548	-0.181330047	0.302581119
-0.38014103	-0.545114832	0.24004759	-0.18857705	-0.15844373	0.322979846
0.266568831	-0.548866106	0.42308044	-0.34564937	-0.221821223	0.312780482
1.01277252	-0.339083362	-0.85814954	0.08404272	0.748206505	-0.455571572
0.614797219	0.037198203	0.78914615	-0.58035513	-0.910171207	0.642559904
-0.50450831	-0.602538169	-1.13269882	0.1950766	0.889045377	-0.445372209
-0.0567861	-0.332157934	-0.49208383	0.07952914	0.653140267	-0.492969239
-0.13140647	0.134442748	0.33156402	-0.19218791	0.095066238	-0.186988332
0.316315744	-0.787504797	0.33156402	0.01543642	0.604726905	-0.071395545
-1.39995274	-0.199420573	0.33156402	-0.70312429	-1.257867171	0.628960753
-1.05172435	-0.16190784	0.69762972	-0.55778727	-1.276352273	0.710555661
0.042707724	-0.774231061	1.42976114	-0.70312429	-1.19184895	0.438572633
-0.08165956	-0.700936953	1.24672829	-0.63632342	-1.137273888	0.642559904
-0.50450831	-0.454507155	0.88066258	-0.69590258	-1.195369922	0.666358419

$$Z_{ij} = \frac{X_{ij} - \bar{X}_i}{S_i}$$

Cada Variable tiene una media igual a cero y una desviación estándar igual 1

MEDIDAS DE TENDENCIA CENTRAL Y VARIACION DE MARCAS DE TEQUILA

	Recuento	Promedio	Varianza	Desviación Estándar
pH	53	0	1	1
conductividad	53	0	1	1
viscosidad	53	0	1	1
Densidad	53	0	1	1
Velocidad de sonido	53	0	1	1
Índice refracción	53	0	1	1
	VARIANZA TOTAL		6	

Desviación estándar

$$S = \sqrt{S^2}$$

Varianza

$$S^2 = \left(\frac{1}{n-1} \right) \sum_{i=1}^n (X_i - \bar{X})^2$$

Media

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$$

MEDIDAS DE TENDENCIA CENTRAL SIN ESTANDARIZAR

	Recuento	Promedio	Varianza	Desviación Estándar
pH	53	4.17283	0.161632	0.402035
conductividad	53	39.0109	1200.96	34.6549
viscosidad	53	0.0024438	1.19E-08	0.000109267
Densidad	53	0.952549	0.000122716	0.0110777
Velocidad de sonido	53	1600.49	129.061	11.3605
Índice refracción	53	1.35221	8.65159E-06	0.00294136
		VARIANZA TOTAL	1330.182763	

MEDIDAS DE TENDENCIA CENTRAL CON ESTANDARIZAR

	Recuento	Promedio	Varianza	Desviación Estándar
pH	53	0	1	1
conductividad	53	0	1	1
viscosidad	53	0	1	1
Densidad	53	0	1	1
Velocidad de sonido	53	0	1	1
Índice refracción	53	0	1	1
		VARIANZA TOTAL	6	

Matriz de Varianzas y Covarianzas

$$\mathbf{S} = \begin{vmatrix} S_{11} & S_{12} & S_{13} & \cdots & S_{1p} \\ S_{21} & S_{22} & S_{23} & \cdots & S_{2p} \\ S_{31} & S_{32} & S_{33} & \cdots & S_{3p} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ S_{p1} & S_{p2} & S_{p3} & \cdots & S_{pp} \end{vmatrix}$$

$$S_{jk} = \left(\frac{1}{n-1} \right) \sum_{i=1}^n (\mathbf{X}_{ij} - \bar{\mathbf{X}}_j)(\mathbf{X}_{ik} - \bar{\mathbf{X}}_k) \quad \mathbf{j,k=1,2,\dots,p}$$

Covarianza

	pH	conductivida d	viscosidad	Densidad	Velocidad de sonido	RI
pH	0.161632	0.480767	1.1233E-05	0.00085408	1.75714	0.000451423
conductividad	0.480767	1200.96	-0.00058535	0.342538	14.3726	-0.0140171
viscosidad	1.12333E-05	-0.000585352	1.19E-08	-2.12E-07	-0.000168604	2.61E-07
Densidad	0.000854079	0.342538	-2.12E-07	0.00012272	0.0368873	-0.0000028
Velocidad de sonido	1.75714	14.3726	-0.0001686	0.0368873	129.061	0.0108572
Indice refraccion	0.000451423	-0.0140171	2.61E-07	-0.0000028	0.0108572	8.65159E-06

 covarianza

 varianza

$$S_{jk} = \frac{1}{(n-1)} \sum_{i=1}^n (X_{ij} - \bar{X}_j)(X_{ik} - \bar{X}_k)$$

	Varianza
pH	0.161632
conductividad	1200.96
viscosidad	1.19E-08
Densidad	0.000122716
Velocidad de sonido	129.061
Índice refracción	8.65159E-06

Coeficiente de correlación

$$r_{jk} = \frac{S_{jk}}{S_j S_k}$$

S_j desviación estandar de la variable j

S_k desviación estandar de la variable k

S_{jk} = covarianza de la variable X_j y variable X_k

Matriz de Correlaciones

$$\mathbf{R} = \begin{pmatrix} 1 & R_{12} & R_{13} & \cdots & R_{1p} \\ R_{21} & 1 & R_{23} & \cdots & R_{2p} \\ R_{31} & R_{32} & 1 & \cdots & R_{3p} \\ \vdots & \vdots & \vdots & \ddots & \\ R_{p1} & R_{p2} & R_{p3} & & 1 \end{pmatrix}$$

$$R_{jk} = \frac{\left(\frac{1}{n-1}\right) \sum_{i=1}^n (X_{ij} - \bar{X}_j)(X_{ik} - \bar{X}_k)}{S_j S_k}$$

$$j, k = 1, 2, \dots, p$$

Matriz de correlaciones

	pH	conductividad	viscosidad	Densidad	Velocidad de sonido	Índice refracción
pH		0.0345	0.2557	0.1918	0.3847	0.3817
conductividad	0.0345		-0.1546	0.8923	0.0365	-0.1375
viscosidad	0.2557	-0.1546		-0.1749	-0.1358	0.8124
Densidad	0.1918	0.8923	-0.1749		0.2931	-0.0859
Velocidad de sonido	0.3847	0.0365	-0.1358	0.2931		0.3249
Indice refraccion	0.3817	-0.1375	0.8124	-0.0859	0.3249	