
ANALISIS DISCRIMINANTE

Dr. Porfirio Gutiérrez González

El análisis discriminante es una técnica multivariante cuya finalidad es analizar si existen diferencias significativas entre grupos de objetos respecto a un conjunto de variables medidas sobre los mismos para, en el caso de que existan, explicar en que sentido se dan y proporcionar procedimientos de clasificación sistemática de nuevas observaciones de origen desconocido en uno de los grupos analizados.

En el análisis discriminante , la información de las variables clasificadoras se sintetiza en unas funciones, denominadas funciones *discriminantes*, que son las que finalmente se utilizan en los procesos de clasificación.

Los objetivos del análisis discriminante se pueden sintetizar en dos:

- 1) Para fines *explicativos*, se trata de determinar la contribución de cada función clasificadora en la clasificación correcta de cada uno de los individuos.
- 2) Una aplicación *predictiva*, se trata de determinar el grupo al que pertenece un individuo para el que se conocen los valores que toman las funciones clasificadoras.

Sea un conjunto de n objetos divididos en q grupos $\{G_i; i = 1, \dots, q\}$ de tamaños $\{n_g; g = 1, \dots, q\}$ que constituyen una información de la población de la que dichos objetos proceden.

Sea $\mathbf{Y} = (Y_1, \dots, Y_p)'$ un conjunto de variables numéricas observadas sobre dichos objetos con el fin de utilizar dicha información para discriminar entre los q grupos anteriores.

VARIABLES

<i>INDIVIDUOS</i>	Y_1	Y_2	\dots	Y_P	
1	Y_{111}	Y_{112}	\dots	Y_{11P}	Grupo 1
2	Y_{211}	Y_{212}	\dots	Y_{21P}	
\vdots	\vdots	\vdots	\vdots	\vdots	
n_1	Y_{n_111}	Y_{n_112}	\dots	Y_{n_11P}	
1	Y_{121}	Y_{122}	\dots	Y_{12P}	Grupo 2
2	Y_{221}	Y_{222}	\dots	Y_{22P}	
\vdots	\vdots	\vdots	\vdots	\vdots	
n_2	Y_{n_221}	Y_{n_222}	\dots	Y_{n_22P}	
\vdots	\vdots	\vdots	\vdots	\vdots	
1	Y_{1q1}	Y_{1q2}	\dots	Y_{1qp}	Grupo q
2	Y_{2q1}	Y_{2q2}	\dots	Y_{2qp}	
\vdots	\vdots	\vdots	\vdots	\vdots	
n_g	Y_{n_gq1}	Y_{n_gq2}	\dots	Y_{n_gqp}	

CALCULO DE LAS FUNCIONES DISCRIMINANTES

La discriminación entre los q grupos se realiza mediante el cálculo de unas funciones matemáticas denominadas *funciones discriminantes*. Considerando la existencia de q grupos, en el caso general, el número máximo de funciones discriminantes que se pueden obtener viene dado por

$$\min(q - 1, p)$$

Procedimiento discriminante de Fisher

El estadístico Fisher se resolvió analíticamente en 1936 para el caso general de p variables. El procedimiento de Fisher toma como funciones discriminantes, combinaciones lineales de las variables clasificadoras de la forma

$$D_{igk} = u_{kg1}Y_1 + u_{kg2}Y_2 + \dots + u_{kgp}Y_p$$

<i>Parámetro</i>	<i>Clave</i>	<i>Valor de referencia</i>
<i>Cuenta leucocitaria</i>	WBC	5.0 - 10.0 mil
<i>Cuenta eritrocitaria</i>	RBC	4.5 - 5.5 10e6/ μ L
<i>Hemoglobina</i>	Hgb	11.8 - 14.5 gr/dL
<i>Hematocrito</i>	Hct	36.0 - 47.0 %
<i>Volumen corpuscular medio</i>	MCV	90 \pm 10 fL
<i>Concentración de hemoglobina corpuscular</i>	MCH	29.5 \pm 2.5 pg
<i>Concentración media de hemoglobina corpuscular</i>	MCHC	30.0 - 35.0 %
<i>Plaquetas</i>	Plt	130 - 450 mil
<i>Linfocitos</i>	LNF	20 - 50 %
<i>Monocitos</i>	MNC	1 - 10 %
<i>Neutrófilos</i>	NTF	45 - 75 %
<i>Eosinófilos</i>	ESF	1 - 6 %
<i>Basófilos</i>	BSF	0 - 1 %
<i>Segmentados</i>	SE	50 - 70%
<i>Bandas</i>	BN	0 - 5 %

Santa cruz del valle

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

Las pintas de arriba

MUESTRA	WBC	RBC	Hgb	Hct	MCV	MCH	MCHC	Plt	LNf	MNC	NTF	ESF	BSF	SE	BN
B1	5,3	5,06	14,1	43,3	85,6	27,9	32,6	3,81	44	5	50	1	0	49	1
B2	6,6	4,91	14,6	43,7	89	29,8	33,5	4,75	53	4	42	2	0	42	0
B3	5,6	5,21	16,7	51,2	98,2	32,1	32,7	37	26	2	61	2	0	6	1
B4	11,5	4,73	14,4	43,7	92,4	30,4	32,9	14,5	34	2	64	0	0	63	1
B5	6,7	4,81	14	42,8	89	29,1	32,6	409	51	4	44	1	0	42	2
B6	4,7	5,33	14,9	44,8	84	28	33,3	40	24	5	70	1	0	69	1
B7	8,6	4,89	13,5	41	83,7	27,6	33	432	35	1	6	4	0	58	2
B8	8,1	5,2	14,5	45,5	87,6	27,9	31,8	298	36	2	61	1	0	58	3
B9	3,4	4,98	13,6	41,8	84	29,4	32,6	58	56	7	36	1	0	34	2
B10	8,5	4,9	13,3	42	85,6	27,1	31,7	331	27	5	64	3	1	61	3
B11	7,8	5,54	15,3	46,7	84,2	27,6	32,7	302	54	5	38	2	1	36	2
B12	6,7	4,63	12,9	40,2	86,9	27,9	32,1	397	66	1	32	1	0	31	1
B13	5,3	5,06	14,1	43,3	85,6	27,9	32,6	3,81	44	5	50	1	0	49	1
B14	6,6	4,91	14,6	43,7	89	29,8	33,5	4,75	52	4	42	2	0	42	0
B15	5,6	5,21	16,7	51,2	98,2	32,1	32,7	37	30	4	64	2	0	63	1
B16	5,9	4,87	13,7	40,6	82,9	27,9	33,7	183	32	3	60	5	0	57	3
B17	12,6	5,51	16	47,4	86	29,1	33,8	121	38	4	48	10	0	46	2
B18	8,7	5,38	15,8	45,9	85,3	29,3	34,4	341	32	5	61	2	0	60	1
B19	5,5	4,72	14,7	41,4	87,6	31,2	35,6	353	42	2	48	7	1	46	2
B20	5,5	4,96	13,1	40,3	81,2	30,4	37,4	230	45	5	48	2	0	45	3

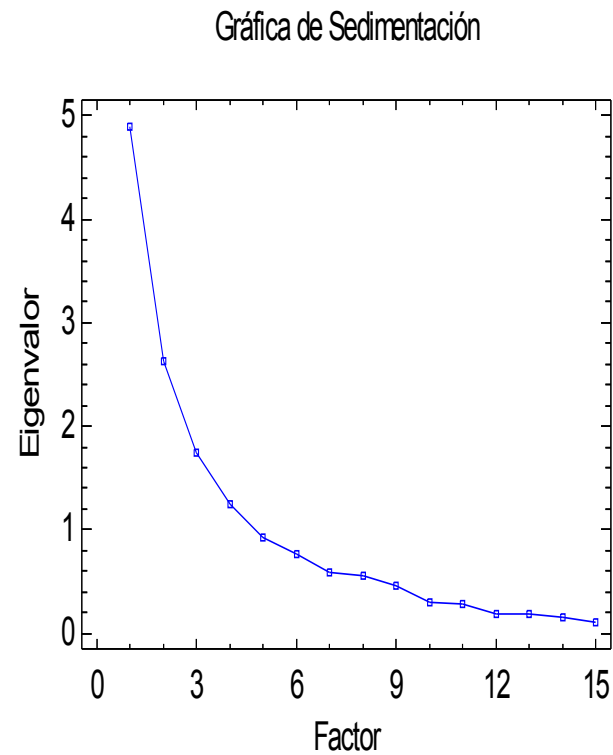
José del Castillo

MUESTRA	WBC	RBC	Hgb	Hct	MCV	MCH	MCHC	Plt	LNF	MN C	NTF	ESF	BSF	SE	BN
C1	8	4,87	14,9	43,3	91,3	30,6	32,9	326	22	3	73	2	0	71	2
C2	9,3	5,64	15,8	49,3	87,4	28	32	221	28	3	67	1	1	65	2
C3	9,1	4,85	14,9	45,7	94,3	30,7	32,5	262	11	3	84	2	0	78	6
C4	8,7	4,61	13,6	41,6	90,2	29,5	32,8	205	36	2	56	6	0	56	0
C5	5,7	4,31	13,3	40,1	93,1	30,8	33,1	198	24	4	67	5	0	65	7
C6	10,6	5,03	14	43,6	86,8	27,9	32,1	320	56	7	34	3	0	30	4
C7	7	5,26	15,2	46,5	88,3	28,8	32,6	239	36	5	54	5	0	53	1
C8	8,7	5,43	15	46,1	84,1	27,7	32,7	216	52	3	38	6	1	37	1
C9	6,3	4,95	14,6	44,9	90,2	29,2	32,4	248	44	3	49	4	0	47	2
C10	5,5	4,73	13,9	41,6	88	29,4	33,4	202	30	6	40	24	0	40	0
C11	2,8	5,26	14,9	46,6	88,6	28,4	32	138	20	9	71	0	0	54	17
C12	8,8	4,67	14	42,6	91,1	30	32,9	185	35	2	58	5	0	57	1
C13	6	4,03	12,3	37,7	93,6	30,3	32,4	163	27	7	64	2	0	59	5
C14	6,3	4,88	14	49,7	87,3	28,7	34,9	210	21	5	74	0	0	54	15
C15	5,5	5,02	14,5	44	87,7	28,9	32,9	188	35	10	55	0	0	42	13
C16	5,9	5,14	14,1	43,6	83,3	27,4	32,2	225	35	6	50	9	0	49	1
C17	6,9	5,42	15,6	47,5	87,6	28,3	32,3	240	29	5	62	4	0	62	0
C18	7,1	4,52	12,6	59,7	87,8	27,8	31,7	176	31	4	61	3	1	58	3
C19	6,2	5,23	15,4	49,9	91,5	29,5	32,2	189	41	4	45	10	0	45	0
C20	8,6	4,89	13,5	41	83,7	27,6	33	432	35	1	60	4	0	58	2

Grupo control

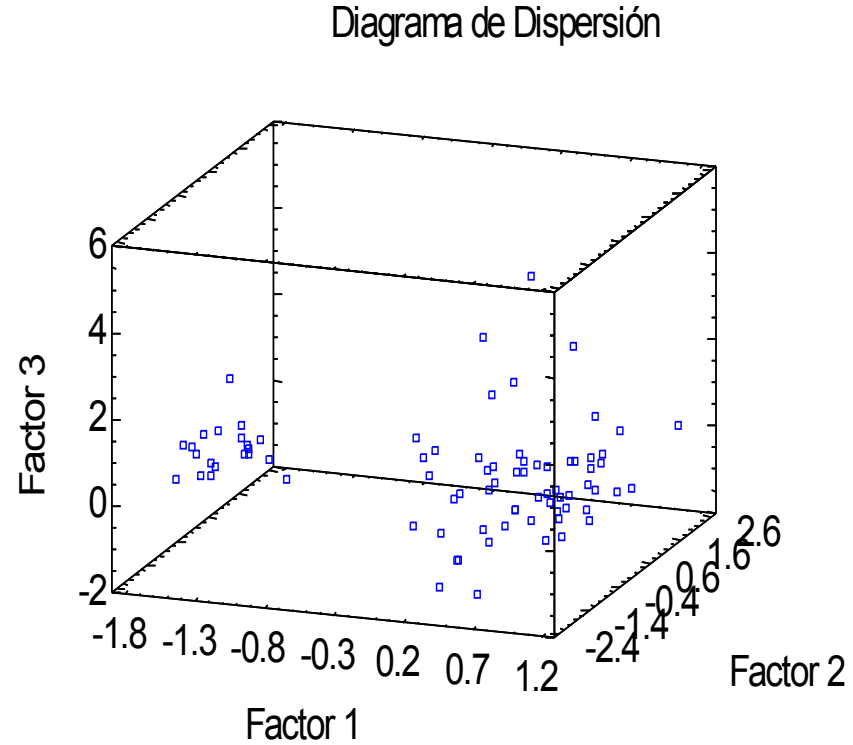
MUESTRA	WBC	RBC	Hgb	Hct	MCV	MCH	MCHC	Plt	LNF	MNC	NT F	ESF	BSF	SE	BN
CT1	6,7	4,45	12,8	39,5	88,8	28,6	32,3	333	33	3	59	5	0	59	0
CT2	5	4,3	13	40,9	95,2	30,3	31,8	213	40	3	52	4	1	51	1
CT3	7,4	4,58	13,7	41,5	90,6	29,9	33	368	44	4	43	9	0	42	1
CT4	6,1	5,29	16,2	48,7	92,1	30,7	33,3	325	39	3	53	4	1	52	1
CT5	8,6	4,51	13,4	41,6	92,3	29,8	32,3	227	42	4	51	3	0	49	2
CT6	7,5	4,79	14,5	44	91,8	30,2	32,9	382	35	1	64	0	0	64	0
CT7	7,6	5,31	13,8	45	81	26,1	32,2	253	21	4	71	3	1	69	2
CT8	9,7	4,44	13	40,5	91	29,2	32,1	298	29	6	63	2	0	62	1
CT9	6	4,2	13,1	40,6	96,6	31,2	32,3	242	48	6	46	0	0	44	2
CT10	8,1	4,79	14,2	43,4	90,6	29,7	32,7	247	32	4	63	1	0	63	0
CT11	8,5	4,61	13,5	41,6	90,1	29,2	32,4	308	25	3	69	2	1	68	1
CT12	6,7	4,89	13,7	42,8	87,4	28	32	219	40	6	52	2	0	50	2
CT13	8,1	4,85	13	40,9	84,5	26,8	31,7	216	2	1	74	3	0	74	0
CT14	7,8	5,83	16	50	85,8	27,5	32	183	39	1	58	2	0	57	1
CT15	9,1	4,68	14,6	44,2	94,3	31,1	33	353	35	3	60	1	1	59	1
CT16	10,1	4,52	13,1	40,2	89	28,9	32,5	380	26	4	68	2	0	67	1
CT17	10,3	5,17	13,2	42,7	82,7	25,5	30,9	327	38	4	57	1	0	56	1
CT18	6,5	4,74	13,3	42,1	88,8	28	31,6	349	39	4	57	0	0	55	2
CT19	6,9	4,7	12,8	40,3	85,9	27,3	31,8	262	58	3	36	3	0	35	1
CT20	10,1	5,05	14,4	48,5	96	28,5	29,7	355	27	3	67	3	0	65	2

Factor	Eigenvalor	Porcentaje de Varianza	Porcentaje Acumulado
1	4.88842	32.589	32.589
2	2.62263	17.484	50.074
3	1.73903	11.594	61.667
4	1.24659	8.311	69.978
5	0.924236	6.162	76.139
6	0.761316	5.075	81.215
7	0.580762	3.872	85.087
8	0.558133	3.721	88.808
9	0.457712	3.051	91.859
10	0.297706	1.985	93.844
11	0.286455	1.910	95.753
12	0.192882	1.286	97.039
13	0.182237	1.215	98.254
14	0.159877	1.066	99.320
15	0.102005	0.680	100.000



	<i>Factor</i>	<i>Factor</i>	<i>Factor</i>	<i>Factor</i>	<i>Factor</i>
	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>5</i>
WBC	0.767959	0.0354975	-0.258968	0.033567	-0.0495602
RBC	0.875517	0.011845	0.0345457	-0.00297758	-0.116168
Hgb	0.782399	-0.000474736	0.0107607	0.106731	-0.0437317
Hct	0.84148	-0.0533366	0.130653	-0.116292	-0.15223
MCV	0.770044	0.0241018	0.058268	0.202422	0.164831
MCH	0.869707	-0.10901	0.0241044	0.0543458	-0.0225022
MCHC	0.806556	-0.0490486	0.0720906	-0.103554	-0.20088
Plt	-0.111312	-0.194182	-0.645589	0.439975	0.120916
LNF	0.0342576	-0.926661	-0.107198	0.150422	-0.0941627
MNC	-0.0596618	-0.172736	0.793543	0.0788196	-0.200549
NTF	-0.00218834	0.917889	0.130801	0.105022	-0.0256782
ESF	-0.100939	-0.10688	-0.0892397	-0.839439	0.229666
BSF	-0.230687	-0.0114807	-0.0984909	-0.215812	0.848762
SE	-0.0740668	0.871887	-0.225842	0.151303	-0.059042
BN	0.0644759	0.172008	0.725941	0.360893	0.362159

		<i>Varianza</i>
<i>Variable</i>	<i>Comunalidad</i>	<i>Específica</i>
WBC	0.661669	0.338331
RBC	0.781368	0.218632
Hgb	0.625568	0.374432
Hct	0.764701	0.235299
MCV	0.665087	0.334913
MCH	0.772314	0.227686
MCHC	0.709212	0.290788
Plt	0.675081	0.324919
LNF	0.90286	0.0971402
MNC	0.709541	0.290459
NTF	0.871323	0.128677
ESF	0.78698	0.21302
BSF	0.83002	0.16998
SE	0.843056	0.156944
BN	0.822136	0.177864



<i>Función</i>	<i>Eigenvalor</i>	<i>Porcentaje</i>	<i>Correlación</i>
<i>Discriminante</i>		<i>Relativo</i>	<i>Canónica</i>
1	15.7462	93.55	0.96968
2	0.792869	4.71	0.66501
3	0.293657	1.74	0.47644

<i>Funciones</i>	<i>Lambda</i>			
<i>Derivadas</i>	<i>de Wilks</i>	<i>Chi-Cuadrada</i>	<i>GL</i>	<i>Valor-P</i>
1	0.0257463	254.3327	45	0.0000
2	0.431154	58.4697	28	0.0006
3	0.773002	17.8944	13	0.1616

<i>Actual</i>	<i>Tamaño</i>	<i>Predicho</i>	<i>LOCALIDAD</i>		
<i>LOCALIDAD</i>	<i>de Grupo</i>	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>
1	20	20	0	0	0
		(100.00%)	(0.00%)	(0.00%)	(0.00%)
2	20	0	15	2	3
		(0.00%)	(75.00%)	(10.00%)	(15.00%)
3	20	0	1	17	2
		(0.00%)	(5.00%)	(85.00%)	(10.00%)
4	20	0	4	1	15
		(0.00%)	(20.00%)	(5.00%)	(75.00%)

Porcentaje de casos correctamente clasificados: 83.75%

A	1	2	3
WBC	0.402175	-0.285192	0.34663
RBC	0.540263	0.326481	-0.133977
Hgb	0.322019	-0.372278	-0.364374
Hct	0.301709	-0.0127484	0.0174842
MCV	0.417257	0.642467	0.4442
MCH	0.37392	-0.192789	0.342375
MCHC	0.475158	-0.139397	-0.44431
Plt	-0.111629	0.0352436	0.269736
LNF	0.0151479	0.60496	0.149099
MNC	0.0045892	0.315659	0.517749
NTF	-0.472559	0.675146	0.750854
ESF	-0.0245404	1.02919	-0.0374246
BSF	-0.0896152	-0.213624	0.239778
SE	0.296109	0.418694	0.219829
BN	0.286368	0.52682	-0.7546

	1	2	3
WBC	0.0176288	-0.012501	0.015194
RBC	0.00488116	0.00294969	-0.00121046
Hgb	0.00756802	-0.0087492	-0.00856344
Hct	0.00259187	-0.000109517	0.000150201
MCV	0.00156676	0.0024124	0.00166793
MCH	0.00573551	-0.00295717	0.00525166
MCHC	0.00544961	-0.00159876	-0.00509582
Plt	-0.00121271	0.000382879	0.00293036
LNF	0.00142653	0.0569712	0.0140411
MNC	0.00259686	0.17862	0.292976
NTF	-0.0398341	0.0569111	0.0632928
ESF	-0.006414	0.268994	-0.00978148
BSF	-0.175572	-0.418528	0.469769
SE	0.0258861	0.0366027	0.0192177
BN	0.105107	0.193362	-0.276966
CONSTANTE	-6.51609	-9.12171	-7.04186

	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>
WBC	-0.0396235	0.127196	0.10516	0.126408
RBC	0.0263848	0.0683437	0.0754701	0.0646135
Hgb	-0.0744509	0.00831357	-0.0147959	-0.0159041
Hct	-0.0184545	0.00524109	0.00533587	0.0032848
MCV	0.0210255	0.031889	0.0384252	0.0346137
MCH	0.0291659	0.0820888	0.0776727	0.0830113
MCHC	-0.0455863	0.00840223	0.00333657	-0.00437977
Plt	0.0169496	0.00401355	0.00587297	0.00945191
LNF	3.88913	3.83365	3.97448	3.89045
MNC	3.52942	3.21465	3.7485	3.75127
NTF	2.70358	2.24718	2.4014	2.40798
ESF	5.10922	4.76507	5.40021	4.93484
BSF	-0.00810875	-1.39175	-2.23243	-0.853813
SE	1.52917	1.71632	1.81347	1.74722
BN	1.51787	2.40583	2.77473	2.05194
CONSTANTE	-208.129	-235.699	-261.277	-244.656

Gráfica de Funciones Discriminantes

