Revista de Psicología, 1977-1978, *22-23*, 79-96.

MODELO PRACTICO PARA EL DESARROLLO DE ECUACIONES DE PREDICCION

Antonio Ospina Cubillos, M. S. (ICFES).

Fernando Angel Reyes (U. N.).

Los autores presentan este modelo desarrollado en el Servicio Nacional de Pruebas del Instituto Colombiano para el Fomento de la Educación Superior, ICFES, cuyo objetivo fue el de elaborar un manual sobre predicción del éxito académico en la universidad, que ofreciera a los responsables de la admisión universitaria una técnica de trabajo fácil de utilizar por personas no especialistas, para estudio, evaluación y mejor utilización de los instrumentos de admisión.

I. *Introducción.*

Los textos de estadística educativa ofrecen fórmulas para el cálculo de coeficientes de múltiple regresión para el caso de dos o tres variables, pero refieren a métodos como el de Doolittle o a métodos de álgebra de matrices, los cálculos para más de tres variables.

En un artículo titulado "Some Simple Computational Formulas" [1], el profesor Aiken presenta algunas fórmulas para el cálculo directo de los pesos beta, el error estándar de los pesos beta y los coeficientes de correlación múltiple para el caso de tres variables independientes y una variable dependiente. Aun cuando estas fórmulas no implican álgebra de matrices, su desarrollo e interpretación requiere formación matemática y psicológica que en general no posee el jefe de admisiones de las universidades colombianas.

El objetivo del presente trabajo, consistió básicamente en transformar en la forma más simple posible el desarrollo de estas fórmulas que presenta Aiken y en proporcionar la información, igualmente en forma simple, de tal manera que pudiera ser utilizada por personas no expertas en los campos de la psicología y las matemáticas.

---

[1] Aiken, Some Computational Formulas for Multiple Regression. Educational and Psychological Measurement (Vol. 34), 1974.

II. *La ecuación de predicción.*

La ecuación de predicción es una fórmula matemática por medio de la cual se estima o "predice" el puntaje de un estudiante (por ejemplo, la calificación semestral que obtendrá un aspirante a determinado programa universitario), a partir de otros puntajes que ya se conocen de él (por ejemplo, sus puntajes en tres pruebas de admisión).

Al puntaje o calificación que se va a estimar de cada estudiante, se le denomina *criterio,* y a las pruebas o instrumentos de admisión se les denomina *predictores.*

Matemáticamente esta ecuación sería de la siguiente forma:

$$Y = P_1 X_1 + P_2 X_2 + P_3 X_3 + C.$$

Donde:

Y sería el puntaje que se va a estimar o a "predecir" del estudiante. Es decir, el *criterio.* Por ejemplo, su calificación final en el primer semestre de estudios universitarios.

X_1, X_2, X_3, serían los puntajes en cada una de las tres pruebas de admisión. Es decir, los predictores. Por ejemplo, sus puntajes en aptitud verbal (X_1), biología (X_2), y química (X_3).

P_1, P_2, P_3, serían los pesos o ponderaciones de cada prueba. P_1 sería el peso, ponderación o coeficiente de X_1, P_2 de X_2 y P_3 de X_3.

C, es una constante. Es un número, que es el mismo para las ecuaciones de cada uno de los estudiantes. Esta constante tiene como función hacer que el puntaje estimado para cada estudiante resulte en su propia escala, por ejemplo, de 0 a 5, en el caso de las calificaciones.

Ejemplo:

Para ingreso a la escuela de química de una universidad se encontró con base en una muestra de estudiantes que concluyeron su primer semestre, la siguiente ecuación de predicción:

$$Y = .20Q + .10R + .08E + 18.$$

Donde:

Y es el promedio que se va a estimar para cada estudiante. El criterio.

.20 es el peso o ponderación para la primera prueba o predictor, en este caso Q.

Q es química, en este caso el predictor uno.

.10 es el peso o ponderación para el predictor dos, R.

R es razonamiento abstracto, en este caso el predictor dos.

.08 es el peso o ponderación para el predictor tres, E.

E es español, en este caso el predictor tres.

18 es la constante C.

Para mayor comodidad en el cálculo de esta ecuación la escala de calificación de 1 a 5 se tomó sin decimales, es decir, se tomó de 10 a 50.

Los siguientes son los resultados de 3 estudiantes A, B, C, en las pruebas de admisión.

	Español	Razonamiento	Química
A	45	40	60
B	50	30	35
C	55	60	65

El cálculo para la estimación de los promedios que probablemente obtendrían estos tres estudiantes en su primer semestre universitario, sería:

$$\text{A:}\quad Y = .20(60) + .10(40) + .08(45) + 18$$
$$Y = 12 + 4 + 3.6 + 18$$
$$\boxed{Y = 37.6}$$

$$\text{B:}\quad Y = .20(35) + .10(30) + .08(50) + 18$$
$$Y = 7 + 3 + 4 + 18$$
$$\boxed{Y = 32.0}$$

$$\text{C:}\quad Y = .20(65) + .10(60) + .08(55) + 18$$
$$Y = 13 + 6 + 4.4 + 18$$
$$\boxed{Y = 41.4}$$

De acuerdo con esta ecuación los estudiantes A, B, C, obtendrían con mayor probabilidad promedios de 3.76, 3.20 y 4.14 en su primer semestre universitario.

En este caso el estudiante C sería el que tendría el mayor rendimiento en su primer semestre universitario.

En este caso el estudiante B sería el que tendría el menor rendimiento en su primer semestre universitario.

Muestra de estudiantes para la obtención de la ecuación de predicción.

La ecuación de predicción se obtiene con un grupo de estudiantes que han tomando las pruebas de admisión (predictores) y que han completado su primer semestre de estudios universitarios y por lo tanto se tiene de ellos su calificación semestral en la universidad (criterio). Por ejemplo los estudiantes que han finalizado su primer semestre de química.

Muestra de estudiantes para la utilización de la ecuación de predicción.

La ecuación de predicción, una vez obtenida, se puede utilizar con el grupo de aspirantes a la facultad o programa para el que se obtuvo la ecuación. Esta ecuación se puede utilizar para estimar o predecir, el criterio (calificación semestral), de los aspirantes de los cuales se tienen únicamente los puntajes de las pruebas de admisión (predictores).

Como se puede ver, la ecuación se obtiene con un grupo de universitarios y se aplica a un grupo de aspirantes del mismo programa.

El dato de la predicción del éxito criterio, puede ser utilizado para la selección o admisión de estudiantes a determinado programa.

III. *Modelo práctico para el desarrollo de la ecuación de predicción.*

El modelo práctico para el desarrollo de la ecuación de predicción, en el que ha venido trabajando el SNP, consiste en un esquema de trabajo, que consta de tres pasos:

1. Cálculo de las estadísticas utilizadas.
2. Desarrollo del modelo.
3. Cálculo de los pesos o ponderaciones de los predictores.

El diseño de este modelo, que se presenta en este documento, es utilizable para obtener una ecuación de predicción, en una calculadora manual, empleando los datos o puntajes de 3 predictores y 1 criterio, es decir, 4 variables en total.

La práctica ha demostrado que más de 3 predictores, no son necesarios o no aportan a la predicción del criterio.

El primer paso para el desarrollo de este modelo de predicción, es el cálculo de tres estadísticas: la correlación cuyo símbolo usual es r, la desviación estándar DE y la media M. Antes de entrar al cálculo de estas estadísticas se hará una breve explicación de cada una de ellas.

1. *Definición y cálculo de las estadísticas utilizadas.*

— *Correlación.*

La correlación es el grado de relación o de asociación existente entre dos variables, es decir, la relación que existe entre la variación de dos puntajes. Esta relación puede ser de tres tipos diferentes: primero, cuando un puntaje alto o un grado alto en una característica puede asociarse con un puntaje alto o un grado alto en otra característica; segundo, cuando no hay una correspondencia definida entre los puntajes obtenidos en las dos variables y tercero, cuando un alto grado de una característica puede asociarse con un bajo grado de otra. En el primer caso se tiene una correlación positiva, en el segundo una falta o ausencia de correlación y en el tercero, una correlación negativa.

La correlación es el sustento de la predicción. Para que un puntaje pueda servir como base para estimar otro, es decir, como predictor, es necesario que tenga correlación con el criterio o puntaje que se va a estimar, que permita suponer que existe algún patrón uniforme de variación entre los puntajes del predictor o predictores y los puntajes de criterio. El grado de correlación entre dos variables se expresa por medio del coeficiente de correlación.

El coeficiente de correlación es un índice estadístico que mide la relación existente entre dos variables o sea la tendencia de una variable a variar concomitantemente con otra. Una relación perfecta se expresa por un coeficiente de 1.00 y la falta completa de relación, por un coeficiente de 0.00. Un coeficiente de correlación que esté entre 0.00 y 1.00 siempre implica algún grado de asociación positiva, y el grado de correspondencia depende de la magnitud del coeficiente. La relación también puede ser negativa, lo que significa que un alto grado de una característica puede asociarse con un bajo grado de otra. Cuando la relación negativa o inversa es perfecta el coeficiente es —1.00. Los coeficientes negativos pueden extenderse de 0.00 a —1.00. El rango de variación de un coeficiente de correlación está entre —1 y 1.

El hecho de que un coeficiente de correlación encontrado entre dos variables sea significativo, es decir, exprese una verdadera relación entre ellas, depende del número de casos —de parejas— entre las cuales se hizo el cálculo del coeficiente. En otras palabras, si se establece un coeficiente de correlación entre las variables peso y estatura en 5 personas, este coeficiente tendría que ser igual o mayor a 0.878 para que tuviera significado, mientras si este mismo coeficiente se calcula en 127 personas, el coeficiente sería significativo si fuera igual o mayor a 0.174 (consúltese tabla número 1). La tabla siguiente número 1, señala los valores mínimos requeridos para que un coeficiente sea significativo, de acuerdo con el número de casos o parejas de datos sobre el que es calculado (N).

TABLA NUMERO 1

Significación del coeficiente de correlación según número de casos.

N	r	N	r
3	0.997	26	0.388
4	0.950	27	0.381
5	0.878	28	0.374
6	0.811	29	0.367
7	0.754	30	0.361
8	0.707	31	0.355
9	0.666	32	0.349
10	0.632	37	0.325
11	0.602	42	0.304
12	0.576	47	0.288
13	0.553	52	0.273
14	0.532	62	0.250
15	0.514	72	0.232
16	0.497	82	0.217
17	0.482	92	0.205
18	0.468	102	0.195
19	0.456	127	0.174
20	0.444	152	0.159
21	0.433	202	0.138
22	0.423	302	0.113
23	0.413	402	0.098
24	0.404	502	0.088
25	0.396	1.002	0.062

N = número de sujetos.

r = correlación.

—Media, M.

La media es una medida que indica el puntaje promedio en una variable y cuyo resultado se obtiene al sumar todos los datos y dividir esta suma por el número de ellos.

—Desviación estándar, DE.

La desviación estándar es una medida de la variabilidad, dispersión o alejamiento de los datos o puntuaciones con respecto a su media.

—Cálculo de las estadísticas.

Con el objeto de hacer didáctico el desarrollo de la ecuación de predicción, se explicará detalladamente con un ejemplo, paso a paso, cada uno de los pasos, hasta llegar a la obtención de la ecuación.

Partiendo del registro de los datos, es decir de los puntajes de los sujetos en el criterio (su calificación universitaria en el primer semestre), y de los predictores (sus puntajes en tres pruebas de admisión), se desarrollará el siguiente ejemplo.

Aunque no es aconsejable desarrollar ecuaciones de predicción con poco número de casos, se toma este ejemplo con datos ficticios de 10 estudiantes en cuatro variables, así: su promedio académico en el primer semestre universitario (criterio), y sus puntajes en tres pruebas de admisión (predictores).

Para facilitar las operaciones, los datos del criterio o promedio en la universidad, se toman sin la coma decimal.

REGISTRO DE LOS DATOS

TABLA 1

No. del sujeto	VARIABLES			
	Promedio o criterio	PREDICTORES		
	0	1	2	3
1	34	37	71	48
2	32	46	40	37
3	37	63	56	45
4	40	66	70	45
5	35	44	38	50
6	28	36	47	32
7	40	70	65	45
8	39	60	70	65
9	30	42	44	40
10	37	36	53	34
T	352	500	554	441

— *Cálculo de la correlación.*

Dado que hay cuatro variables, el promedio o criterio y cada uno de los tres predictores, se hace necesario calcular las siguientes seis correlaciones:

Criterio con predictor uno.

Criterio con predictor dos.

Criterio con predictor tres.

Predictor uno con predictor dos.

Predictor uno con predictor tres.

Predictor dos con predictor tres.

Para efectos del ejemplo se calculará, paso a paso, una de estas seis correlaciones. Todas las seis correlaciones se calculan en la misma forma.

Los pasos a seguir en el cálculo de la correlación entre dos variables, son los siguientes:

Paso 1.

Diseñar una hoja de trabajo como la tabla número 2.

Paso 2.

Numerar los sujetos y escribir los puntajes de las 2 variables de cada uno de ellos en las columnas X e Y, respectivamente.

Paso 3.

Elevar al cuadrado cada valor de la columna X, cada valor de la columna Y, anotar los cuadrados en las columnas X^2, Y^2, respectivamente.

Paso 4.

Multiplicar cada valor de X por el correspondiente valor de Y en la misma fila, anotar su producto en la columna XY.

Paso 5.

Sumar cada una de las columnas y colocar su resultado en la parte correspondiente al total.

Paso 6.

Elevar al cuadrado la suma o total de la columna X, elevar al cuadrado el total de la columna Y, colocar los resultados en la fila 1 debajo del total de la columna correspondiente.

Paso 7.

Multiplicar el total de las columnas X^2, Y^2, XY por el número de sujetos y colocar su producto en la fila 1 debajo del total de la columna correspondiente.

Paso 8.

Preparar el quebrado final:

$$r = \frac{\overset{\text{Paso 9}}{\boxed{\qquad\qquad}}}{\underset{\text{Paso 10}}{\boxed{\qquad\qquad}} \quad \text{por} \quad \underset{\text{Paso 11}}{\boxed{\qquad\qquad}}}$$

TABLA NUMERO 2

	Promedio	Predictor uno			
Sujeto	**X**	**Y**	**X^2**	**Y^2**	**XY**
Número de sujetos	**Total**	**Total**	**Total**	**Total**	**Total**
Fila 1	a	b	c	d	e

Paso 9

$$r = \frac{\boxed{}_{\text{Paso 9}}}{\boxed{}_{\text{Paso 10}} \text{ por } \boxed{}_{\text{Paso 11}}}$$

Paso 10 **Paso 11**

Cumplidos estos 8 pasos, se tomaría la siguiente tabla:

TABLA NUMERO 2A

Correlación entre el critero y el predictor uno.

	Promedio	Predic. uno			
Sujeto	X	Y	X^2	Y^2	XY
1	34	37	1.156	1.369	1.258
2	32	46	1.024	2.116	1.472
3	37	63	1.369	3.969	2.331
4	40	66	1.600	4.356	2.640
5	35	44	1.225	1.936	1.540
6	28	36	784	1.296	1.008
7	40	70	1.600	4.900	2.800
8	39	60	1.521	3.600	2.340
9	30	42	900	1.764	1.260
10	37	36	1.369	1.296	1.332
Número de sujetos	Total	Total	Total	Total	Total
10	352	500	12.548	26.602	17.981
Fila 1	123.904	250.000	125.480	266.020	179.810
	a	b	c	d	e

$$r_{01} = \frac{\overset{\text{Paso 9}}{\boxed{3.810}}}{\underset{\text{Paso 10}}{\boxed{39{,}70}} \text{ por } \underset{\text{Paso 11}}{\boxed{126{,}57}}} = 0{,}76$$

Paso 9.

Reste al valor (e) el producto de la multiplicación del total de la columna X por el total de la columna Y. Coloque este resultado en el numerador del quebrado.

Paso 10.

Reste al valor (c) el valor (a). Obtenga la raíz cuadrada a este resultado y colóquelo en la primera casilla (paso 10) del denominador del quebrado.

Paso 11.

Reste al valor (d) el valor (b). Obtenga la raíz cuadrada a este resultado y colóquelo en la segunda casilla (paso 11) del denominador del quebrado.

Paso 12.

Resuelva el quebrado.

Todos los seis coeficientes de correlación se calculan en la misma forma. Para efectos de este ejemplo, el proceso para el cálculo de los otros cinco coeficientes no se incluye, pero sí sus resultados.

Una vez obtenidos los seis coeficientes de correlación, en esta misma forma, se registran en una tabla de doble entrada, así:

TABLA NUMERO 3

		0	1	2	3
		CRITERIO O PROMEDIO	PREDICTOR	PREDICTOR	PREDICTOR
0	CRITERIO O PROMEDIO		.76	.66	.55
1	PREDICTOR			.50	.47
2	PREDICTOR				.52
3	PREDICTOR				

NOTA: Los datos o coeficientes de las casillas arriba de la diagonal son los mismos de las casillas debajo de la diagonal. Para efectos prácticos sólo se registran en las casillas superiores.

— *Cálculo de la media M, y de la desviación estándar DE, de cada variable.*

Para cada una de las cuatro variables, debe calcularse la media M y la desviación estándar DE. Para ello deben seguirse las siguientes instrucciones:

— *Media, M.*

El valor de la media de una variable se obtiene dividiendo el valor del total de la suma de la columna del criterio y de cada uno de los predictores de la tabla número 1, por el número de sujetos, que para el caso de este ejemplo son 10. Las medias de este ejemplo, son:

$$M_0 = \frac{352}{10} = 32{,}2 \quad \text{Media del criterio.}$$

$M_1 \quad = \quad 50{,}0$ Media predictor uno.
$M_2 \quad = \quad 55{,}4$ Media predictor dos.
$M_3 \quad = \quad 44{,}1$ Media predictor tres.

— *Desviación estándar, DE.*

El valor de la DE del criterio, se obtiene dividiendo el dato de la primera casilla del denominador del quebrado, con el que se obtiene la correlación r_{01} entre el criterio y el predictor uno. (Paso 10, tabla número 2A), por el número de sujetos.

El valor de la DE del predictor uno se obtiene haciendo esta misma operación con el quebrado en el cual se obtiene la correlación r_{12} entre el predictor uno y el predictor dos 12.

El valor de la DE del predictor dos, se obtiene haciendo esta misma operación con el quebrado en el cual se obtiene la correlación r_{23} entre el predictor dos y el predictor tres 23.

El valor de la DE del predictor tres, se obtiene dividiendo la *segunda* casilla del denominador del quebrado con el que se obtiene la correlación anterior entre el predictor dos y tres, por el número de sujetos.

El valor de la DE del criterio para este ejemplo, es:

$$DE_0 = \frac{39{,}70}{10} = 3{,}97 \quad \text{Desviación estándar del criterio.}$$

Los valores de las desviaciones de los predictores serían:

$DE_1 = \quad 12{,}657$ Desviación estándar del predictor uno.
$DE_2 = \quad 12{,}282$ Desviación estándar del predictor dos
$DE_3 = \quad 8{,}972$ Desviación estándar del predictor tres.

Una vez obtenidas estas estadísticas, se registran en una tabla de la siguiente manera:

TABLA NUMERO 4

		MEDIA	DESVIACION ESTANDAR
O	CRITERIO O PROMEDIO	35.2	3,970
I	PREDICTOR	50.0	12,657
2	PREDICTOR	55.4	12,282
3	PREDICTOR	44.I	8,972

2. *Desarrollo del modelo.*

El desarrollo del modelo se hace a partir de las correlaciones encontradas y de sus valores elevados al cuadrado, en la siguiente forma:

Símbolo usual			Denominación en el modelo	
$r_{01} =$	.76	A	.58	A^2
$r_{02} =$	.66	B	.44	B^2
$r_{03} =$	.55	C	.30	C^2
$r_{12} =$		D		D^2
$r_{13} =$		E		E^2
$r_{23} =$		F		F^2

A la correlación r_{01} entre el criterio y el predictor uno, se le denominará A de ahora en adelante y a su cuadro se le denominará A^2.

Una vez encontradas las correlaciones el modelo consiste en el desarrollo de las operaciones que se indican al margen izquierdo y que a su vez reciben una nueva denominación, margen derecho.

Esta denominación va de las letras G a la R .

Explicaciones.

El dato G , por ejemplo, resulta de restar F^2 de 1, es decir 1-0.27 = 0.73.

El dato M , por ejemplo, resulta de multiplicar los datos B pör H o sea $0.66 \times 0{,}27 = 0.18$.

El dato S , por ejemplo, de restar a B , el producto de C por F

Control de operaciones.

Con el objeto de verificar si las operaciones en el modelo han sido correctamente realizadas, el valor W que es igual al valor de R elevado al cua-

MODELO

$1 - F^2$ —— G

$D - E\ F$ —— H

$E - D\ F$ —— I

$1 - D^2 - E^2 - F^2$ —— J

$2\ D\ E\ F$ —— K

$A\ G$ —— L

$B\ H$ —— M

$C\ I$ —— N

$L - M - N$ —— O

$J + K$ —— P

$\beta_1 \longleftarrow O \div P$ —— Q

$B - C\ F$ —— S

$S - Q\ H$ —— T

$\beta_2 \longleftarrow T \div G$ —— U

$\beta_3 \longleftarrow C - E\ Q - F\ U$ —— V

$R^2 \longleftarrow Q\ A + U\ B + V\ C$ —— W

$R \longleftarrow \sqrt{W}$ —— R

El desarrollo del modelo, utilizando los datos del ejemplo, sería el siguiente:

TABLA NUMERO 5

r_{01}	0.76	A	0.58	A^2
r_{02}	0.66	B	0.44	B^2
r_{03}	0.55	C	0.30	C^2
r_{12}	0.51	D	0.26	D^2
r_{13}	0.47	E	0.22	E^2
r_{23}	0.52	F	0.27	F^2

$1 - F^2$	0.73	G
$D - E\ F$	0.27	H
$E - D\ F$	0.20	I
$1 - D^2 - E^2 - F^2$	0.25	J
$2\ D\ E\ F$	0.25	K
$A\ G$	0.55	L
$B\ H$	0.18	M
$C\ I$	0.11	N
$L - M - N$	0.26	O
$J + K$	0.50	P
$\beta_1 \longleftarrow O \div P$	0.52	Q
$B - C\ F$	0.37	S
$S - Q\ H$	0.23	T
$\beta_2 \longleftarrow T \div G$	0.32	U
$\beta_3 \longleftarrow C - E\ Q - F\ U$	0.14	V
$A\ Q + B\ U + C\ V$	0.68	W
$R^2 \longleftarrow 0.39 + 0.21 + 0.08$		
$R \longleftarrow \sqrt{W}$	0.83	R

drado, debe calcularse nuevamente por otro procedimiento, utilizando el siguiente pequeño modelo:

TABLA NUMERO 6

$$1 - E^{2} \quad \underline{\qquad} \quad X$$

$$1 - D^{2} \quad \underline{\qquad} \quad Y$$

$$F \quad - D \quad E \quad \underline{\qquad} \quad Z$$

Para el caso del ejemplo los cálculos son los siguientes:

TABLA NUMERO 6A

$$1 - E^{2} \quad \underline{\quad .78 \quad} \quad X$$

$$1 - D^{2} \quad \underline{\quad .74 \quad} \quad Y$$

$$F \quad - \quad D \quad E \quad \underline{\quad .28 \quad} \quad Z$$

Una vez obtenidos los cálculos de este pequeño modelo, se obtiene el valor de R^2 con la siguiente fórmula:

$$R^2 = \frac{2\,A\,O + B^{2}\,X + C^{2}\,Y - 2\,B\,C\,Z - A\,L}{P} = W$$

En caso de obtener un valor para W igual al obtenido en el modelo, es porque las operaciones han sido realizadas correctamente.

Para el caso del ejemplo, las operaciones serían las siguientes:

$$R^2 = \frac{.40 + .34 + .22 - .20 - .42}{50} = \frac{.34}{.50} = .68 = W$$

Este valor de $R^2 = .68 = W$, confirma que las operaciones han sido bien calculadas.

Por medio del modelo anterior se han calculado tres índices estadísticos que aparecen señalados en la parte izquierda del modelo, éstos con el R, el R^2 y los β (betas) de los cuales se hará una breve explicación.

β, ponderaciones beta. Indican la importancia relativa de cada predictor en la predicción del criterio. Estos β dan origen a los pesos o ponderaciones que deberán darse a cada uno de los predictores.

R, coeficiente de correlación múltiple. Se define como la correlación existente entre los puntajes efectivamente logrados en el criterio y los puntajes predichos para la misma muestra de estudiantes. Así un $R = 0.72$ significa que los puntajes en el criterio (promedio), pronosticados en función de la ecuación de predicción, tienen una correlación de 0.72 con los puntajes reales que obtuvieron los mismos sujetos en el criterio.

R^2: Numéricamente corresponde al cuadrado del coeficiente de correlación múltiple. Indica la proporción de la variación de la medida del criterio que es atribuible a la acción conjunta de los tres predictores. Un $R^2 = (0.72)^2 = 0.52$ indica que un 52% de lo que causa la variación en el promedio en el primer semestre puede atribuirse a los tres predictores considerados en la ecuación.

Mediante la fórmula $R^2 = Ⓦ = Ⓠ Ⓐ + Ⓤ Ⓑ + Ⓥ Ⓒ$, la contribución total del 52% puede separarse en la contribución independiente de cada uno de los predictores así:

Ⓠ Ⓐ = Contribución del predictor uno.

Ⓤ Ⓑ = Contribución del predictor dos.

Ⓥ Ⓒ = Contribución del predictor tres.

El 48% restante de la variación del criterio debe atribuirse a factores no medidos o no considerados en la ecuación.

El hecho de poder considerar el aporte de cada predictor a la predicción del criterio, es de gran importancia, pues permite reemplazar aquellos predictores que aportan muy poco o nada a la predicción.

3. *Cálculo de las ponderaciones o pesos.*

Como el objetivo final es el cálculo de las ponderaciones o de los pesos que se deban dar a las pruebas de admisión o predictores para lograr las mejores decisiones en la admisión universitaria, se procede al cálculo de éstos a partir de los datos ya encontrados.

Para el cálculo de ellos así como para el cálculo de la constante de la ecuación, se utilizará la siguiente nomenclatura:

P_1 = Peso o ponderación del predictor uno.
P_2 = Peso o ponderación del predictor dos.
P_3 = Peso o ponderación del predictor tres.
C = Constante.

El valor del peso de los predictores se obtiene mediante la aplicación de las siguientes fórmulas (consúltese la tabla número 4):

$$P_1 = \frac{\text{Desviación estándar del criterio}}{\text{Desviación estándar del predictor uno}} \quad Q \quad = \frac{DE_0}{DE_1} \quad Q$$

$$P_2 = \frac{\text{Desviación estándar del criterio}}{\text{Desviación estándar del predictor dos}} \quad U \quad = \frac{DE_0}{DE_2} \quad U$$

$$P_3 = \frac{\text{Desviación estándar del criterio}}{\text{Desviación estándar del predictor tres}} \quad V \quad = \frac{DE_0}{DE_3} \quad V$$

TABLA NUMERO 6B

$$P_1 = \frac{3.97}{12.657} (.52) = .163$$

$$P_2 = \frac{3.97}{12.282} (.32) = .103$$

$$P_3 = \frac{3.97}{8.972} (.14) = .062$$

Una vez obtenidos los pesos, el valor de la constante C se obtiene con la siguiente fórmula:

$$C = M_0 - P_1M_1 - P_2M_2 - P_3M_3$$

Donde:

M_0 = Media del criterio.
M_1 = Media del predictor uno.
M_2 = Media del predictor dos.
M_3 = Media del predictor tres.
P_1 = Peso del predictor uno.
P_2 = Peso del predictor dos.
P_3 = Peso del predictor tres.

Operaciones.

(Consúltese tablas números 4 y 6).

$$C = 35.2 - .163(50) - .103(55.4) - .062(44.1) = 18.61.$$

Ecuación final:

$$X_0 = P_1X_1 + P_2X_2 + P_3X_3 + C.$$
$$X_0 = .163X_1 + .103X_2 + .062X_3 + 18.61.$$

IV. *Ejemplo en el manejo de la ecuación en programas de admisión.*

Una vez calculada la ecuación de predicción, se puede calcular el promedio esperado de un grupo de sujetos diferentes a aquél que sirvió como base para calcular la ecuación. Para ello se requiere obtener los puntajes alcanzados por cada uno de los sujetos en cada uno de los tres predictores. Sustituyendo las incógnitas por el puntaje respectivo y resolviendo la ecuación se obtiene el promedio estimado para cada sujeto.

Así, por ejemplo, para la ecuación de predicción obtenida:

$$\text{Promedio} = 0{,}163\times + 0{,}103\times_2 + 0{,}062\times_3 + 18{,}61.$$

El promedio de Juan Pérez, quien obtuvo respectivamente en los tres predictores 50, 40 y 30, será:

$$\begin{aligned}\text{Promedio} &= 0{,}163(50) + 0{,}103(40) + 0{,}062(30) + 18{,}61.\\ &= 8{,}15 + 4{,}12 + 1{,}86 + 18{,}61.\\ &= 32{,}74.\end{aligned}$$

El promedio estimado para Juan Pérez en el primer semestre es de 33 en una escala de calificación de 0 a 50. Es decir, obtendría como nota más probable 3.3 al finalizar su primer semestre universitario.

BIBLIOGRAFIA

AIKEN, L. R. Some computational formulas for multiple regression. Educational and Psychological Measurement, 1974, *34*, 767-770.

DARLINGTON, R. B. Multiple regression in psychological research and practice. *Psychological Bulletin,* 1968, *69,* 161-182.

HURST, P. *Factor analysis of data matrices.* New York: Holt, Rinehart and Winston, 1965.

KERLINGER, F. N. *Foundations of behavioral research.* New York: Hol, Rinehart and Winston, 1964.

KERLINGER, F. N. *Multiple regression in behavioral research.* New York: Holt, Rinehart and Winston, 1973.

MC NEMAR, Q. *Psychological Statistic.* (3rd ed.) New York: Wiley, 1962.

THURSTONE, L. L. *Multiple-factor analysis.* Chicago: University of Chicago Press, 1947.