

CORRELACIONES MULTIPLES Y CORRELACIONES CANONICAS EN LA PREDICCIÓN DE RENDIMIENTO ACADÉMICO¹

Hermelinda M. Fogliatto
Universidad Nacional de Córdoba
Córdoba, Argentina

A sample of Psychology students answered the Kuder preference scales, the verbal reasoning, numeric abilities, abstract reasoning, Space relationships, spelling and phrases sub-tests of the DAT and the Maudsley Personality Inventory. When they finished their first year of school their grades were correlates and a regression equation was calculated. The seven variables which best predicted achievement were: the areas of scientific interest, persuasive interest, musical interest and social service; as well as verbal reasoning, abstract reasoning and spelling. Four years later the total GPA and the areas GPA was calculated for those who did graduate. Canonical correlations were calculated for the seven best predictors and the subject areas as well as a multiple correlation with the general GPA. From the results one can conclude that although the first two canonical correlations were greater than any of the correlation coefficients between the original variables, these values are not significantly different from zero; the multiple correlation coefficient was significant. Two of the predictor variables (musical interest and abstract reasoning) did not correlate significantly with any of the criteria and the regression coefficients are practically zero. The best predictors were persuasive interest, social service and the spelling sub-test.

Se administró el Registro de Preferencias Kuder, los subtests Razonamiento Verbal, Habilidad Numérica, Razonamiento Abstracto, Relaciones Espaciales, Ortografía y Oraciones del Test de Aptitudes Diferenciales (D.A.T.) y el Inventario de Personalidad de Maudsley a una muestra de alumnas en el momento de ingresar a la carrera de Psicología. Al concluir el primer año de estudios universitarios se obtuvo un índice de rendimiento académico (promedio de las materias rendidas) y se calcularon las correlaciones múltiples y la correspondiente ecuación de regresión incluyendo una variable por vez (la mejor predictora). Las siete variables mejores predictoras resultaron ser: las áreas de intereses Científico, Persuasivo, Musical y Servicio Social; y las aptitudes Razonamiento Verbal, Razonamiento Abstracto y Ortografía. Cuatro años más tarde se seleccionaron aquellas alumnas que habían egresado y se calculó el promedio general obtenido en la carrera y el promedio por áreas de materias. Tomando los siete mejores predictores obtenidos a nivel de primer año se calcularon las correlaciones canónicas entre éstos y los promedios de las cinco áreas de mate-

¹ Investigación subvencionada por el Consejo Nacional de Investigaciones Científicas y Técnicas.

rias y la correlación múltiple con el promedio general. De los resultados obtenidos se concluye que si bien las dos primeras correlaciones canónicas son mayores que cualquiera de los coeficientes de correlaciones entre las variables originales, estos valores no son significativamente distintos de cero; por el contrario el coeficiente de correlación múltiple es significativamente distinto de cero a nivel .01. Dos de las variables predictoras (interés Musical y Razonamiento Abstracto) no presentan correlaciones significativas con ninguno de los criterios y los coeficientes de regresión son prácticamente cero. Los mejores predictores son las áreas de intereses Persuasivo, Servicio Social y el subtest Ortografía.

La finalidad del presente trabajo fue, por un lado, estudiar la posibilidad de predecir rendimiento académico de un grupo de alumnas universitarias en base a una batería de tests administrada al ingresar a la Carrera de Psicología. Así se seleccionarían de entre una serie de variables aquellas que fuesen las mejores predictoras del rendimiento del primer año de estudios universitarios; y comprobar al finalizar la carrera si por medio de estas mismas variables (las seleccionadas como mejores predictoras a nivel del primer año) era posible predecir el rendimiento general de la carrera. Si los resultados fuesen altamente satisfactorios se podría contar con instrumentos que permitieran predecir el buen desempeño de las alumnas en una carrera universitaria. Lamentablemente, el único criterio con que se cuenta sobre rendimiento académico son las notas obtenidas por los sujetos en las materias cursadas en el ámbito universitario.

En segundo lugar, se trata de comparar los resultados obtenidos por dos métodos de análisis multivariado. Contando con datos de una sola población y dos series de los mismos se los analizará en dos momentos de un estudio longitudinal de predicción de rendimiento académico (al finalizar el primer año de estudios y al egresar) por medio de correlaciones múltiples; y por medio de correlaciones canónicas al egresar.

La correlación múltiple analiza la relación existente entre una serie de variables predictoras y una variable criterio. El coeficiente de correlación canónica da la máxima relación existente entre las funciones lineales de dos series de variables. Si bien Bartlett (1941) provee un test de significación para los valores de las correlaciones canónicas, suelen con frecuencia considerarse como "no triviales" los coeficientes de correlación canónica mayores de .30 cuando se cuenta con muestras grandes (Cooley y Lohnes, 1971).

Método

A 162 alumnas que ingresaron a la Carrera de Psicología se les administró una batería de pruebas compuesta por el Registro de Preferencias Kuder (forma CH); los subtests Razonamiento Verbal, Habilidad Numérica, Razonamiento Abstracto, Relaciones Espaciales, Oraciones y Ortografía del Test de Aptitudes Diferenciales (D.A.T.) de Bennett, Seashore y Wesman; y el Inventario de Personalidad de Maudsley (M.P.I.). Al concluir el primer año de estudios universitarios se obtuvo un índice de rendimiento académico (promedio de las materias rendidas). Se eliminaron las alumnas que no habían rendido por lo menos una materia; la muestra quedó constituida por 146 sujetos. Se obtuvieron los puntajes para cada alumna en las 19 variables: A_0 variable dependiente: índice de rendimiento académico; A_1 a A_{18} variables independientes. De estas últimas A_1 a A_{10} corresponden a las áreas medidas por el Registro de Preferencias Kuder (Aire Libre, Mecánico, Numérico, Científico, Persuasivo, Artístico, Literario, Musical, Servicio Social y Oficina), A_{11} a A_{16} a los

subtests del D.A.T. (Razonamiento Verbal, Habilidad Numérica, Razonamiento Abstracto, Relaciones Espaciales, Ortografía y Oraciones) y A_{17} a A_{18} a las escalas del M.P.I. (Neuroticismo y Extraversión). Una vez obtenidos estos puntajes se calculó la matriz de correlación de orden 19×19 .

A partir de esta matriz de correlaciones simples se calcularon las correlaciones múltiples y sus correspondientes ecuaciones de regresión ².

Las variables independientes A_1, A_2, \dots, A_{18} fueron entradas en la ecuación de regresión si reducían en .100 la varianza residual muestral, no eran entradas o se las extraía si al hacerlo aumentaba en .050 la varianza residual muestral. Estos valores (.100 y .050) fueron elegidos arbitrariamente ya que no se tenía ninguna información sobre los posibles valores a usar.

Se procesó una variable por vez, eligiendo siempre la variable mejor predictora.

Al ser procesada cada variable se la comparó con los criterios preestablecidos para determinar si se la incluía o no en la ecuación de regresión. Si la variable era entrada se calculaban los coeficientes para la ecuación que contenía dicha variable del número total de variables en el análisis. Esto era impreso y usado en el próximo paso. De este modo se obtuvieron las correlaciones múltiples, los tests F y las ecuaciones de regresiones, llegando hasta incluir 16 variables independientes. Sin embargo, como afirma McNemar (1962) seguir agregando predictores después de cinco o seis variables incrementa poco el coeficiente de correlación múltiple, a no ser que la nueva variable se correlacione en mayor grado con el criterio que las variables utilizadas anteriormente, o que esta variable no se correlacione con los otros predictores (un tercer caso sería que la nueva variable no se correlacione con la variable dependiente, pero si de un modo particular con los otros predictores).

Cuatro años más tarde se seleccionaron aquellas alumnas que habían egresado y se calculó el promedio general obtenido en la carrera y el promedio por áreas de materias (clínica, educacional, experimental, social-laboral y generales, entendiéndose por estas últimas materias tales como Psicología General, Historia de la Psicología, Antropología Cultural, Antropología Filosófica e Introducción a la Filosofía). La muestra quedó reducida a 61 sujetos. Tomando las siete variables mejores predictoras obtenidas a nivel de primer año se calcularon las correlaciones canónicas con éstas y los promedios de las cinco áreas de materias mencionadas. Además se calculó la correlación múltiple con los siete predictores y el promedio general obtenido en la carrera ³.

Resultados

La tabla 1 que presenta la matriz de intercorrelaciones entre las 18 variables independientes entre sí y las de éstas con la variable dependiente, a nivel del primer año, muestra correlaciones significativamente distintas de cero para el índice de rendimiento académico con todas las áreas medidas por el Registro de Preferencias Kuder (con excepción de Interés Mecánico y Numérico), con todos los subtests del D.A.T. y con la escala Extraversión del M.P.I.

La tabla 2 presenta la correlación múltiple, el test F y la ecuación de regresión. Se resolvió tomar los siete mejores predictores ya que a partir de este punto la co-

² El autor agradece la colaboración prestada por el personal del Centro de Cómputos de la Universidad de Puerto Rico.

³ El autor agradece la colaboración prestada por el Dr. J.B. Allende del Centro de Computación y Procesamiento de datos de la Facultad de Ciencias Económicas de la Universidad Nacional de Córdoba.

TABLA I

MATRIZ DE INTERCORRELACIONES (A_0 = VARIABLE DEPENDIENTE: PROMEDIO DE PRIMER AÑO). N=146

	A_0	A_1	A_2	A_3	A_4	A_5	A_6	A_7	A_8	A_9	A_{10}	A_{11}	A_{12}	A_{13}	A_{14}	A_{15}	A_{16}	A_{17}	A_{18}
A_0																			
A_1	25 ^{xx}																		
A_2	07	37 ^{xx}																	
A_3	03	-13	05																
A_4	22 ^{xx}	32 ^{xx}	29 ^{xx}	41 ^{xx}															
A_5	31 ^{xx}	-19 ^x	-07	-12	-19 ^x														
A_6	20 ^x	03	23 ^{xx}	-06	-12	29 ^{xx}													
A_7	17 ^x	13	-06	04	-04	27 ^{xx}	14												
A_8	17 ^x	18 ^x	00	-12	-07	04	09	16 ^x											
A_9	39 ^{xx}	46 ^{xx}	19 ^x	08	28 ^{xx}	04	-01	12	08										
A_{10}	16 ^x	-39 ^{xx}	-15	45 ^{xx}	12	32 ^{xx}	08	06	-06	-08									
A_{11}	43 ^{xx}	11	17 ^x	09	23 ^{xx}	21 ^{xx}	21 ^{xx}	19 ^x	02	12	01								
A_{12}	39 ^{xx}	08	02	12	12	10	-12	10	-03	22 ^{xx}	09	44 ^{xx}							
A_{13}	40 ^{xx}	05	11	14	11	24 ^{xx}	24 ^{xx}	15	-08	15	06	55 ^{xx}	46 ^{xx}						
A_{14}	26 ^{xx}	11	18 ^x	16 ^x	34 ^{xx}	13	10	-03	-04	23 ^{xx}	03	51 ^{xx}	30 ^{xx}	56 ^{xx}					
A_{15}	43 ^{xx}	08	19 ^x	-04	10	26 ^{xx}	15	25 ^{xx}	13	16 ^x	12	47 ^{xx}	33 ^{xx}	34 ^{xx}	31 ^{xx}				
A_{16}	33 ^{xx}	08	11	-05	16 ^x	18 ^x	05	21 ^{xx}	04	18 ^x	09	32 ^{xx}	39 ^{xx}	32 ^{xx}	27 ^{xx}	62 ^{xx}			
A_{17}	13	07	12	-03	11	20 ^x	19 ^x	14	17 ^x	-05	07	18 ^x	10	12	14	13	15		
A_{18}	25 ^{xx}	03	16 ^x	20 ^x	19 ^x	24 ^{xx}	12	16 ^x	-07	10	30 ^{xx}	23 ^{xx}	21 ^{xx}	26 ^{xx}	15	21 ^{xx}	24 ^{xx}	22 ^{xx}	

xp <.05

xxxp <.01

Se omitieron los puntos decimales.

TABLA 2

CORRELACION MULTIPLE, TEST F Y ECUACION DE REGRESION (SIETE MEJORES PREDICTORES).

$$R_{m\acute{u}lt} = .65^{***} \quad F = 14.39^{***} \quad N = 146$$

$$P' = .0189 X_4 + .0296 X_5 + .0365 X_8 + .0309 X_9 + .0259 X_{11} \\ + .0250 X_{13} + .0151 X_{15} + .6950$$

X_4	=	Interés Científico
X_5	=	Interés Persuasivo
X_8	=	Interés Musical
X_9	=	Interés Servicio Social
X_{11}	=	Subtest Verbal (D.A.T.)
X_{13}	=	Subtest Abstracto (D.A.T.)
X_{15}	=	Subtest Ortografía (D.A.T.)

***p. < .001

TABLA 3

INTERCORRELACIONES ENTRE LAS SIETE VARIABLES PREDICTORAS Y LOS PROMEDIOS DE LAS CINCO AREAS DE MATERIAS; MEDIAS Y DESVIACIONES STANDARDS (N = 61)

	Clínica	Educativa	Experimental	Social Laboral	Generales
Interés Científico	.028	.115	.016	-.095	.071
Interés Persuasivo	.281*	.156	.142	.308*	.182
Interés Musical	-.156	-.080	-.081	-.165	.142
Interés Serv.Social	.130	.178	.138	.151	.060
Subtest Verbal	.183	.328**	.305*	.268*	.305*
Subtest Ortografía	.174	.249	.307*	.158	.314*
M	7.02	6.77	6.31	7.59	6.94
D.S.	1.09	1.17	1.21	.99	1.31

*p < .05

**p < .001

rrelación múltiple incrementaba muy poco con cada variable predictor agregada a la ecuación de regresión. Estos resultaron ser las áreas de intereses Científico, Persuasivo, Musical y Servicio Social; y las aptitudes Razonamiento Verbal, Razonamiento Abstracto y Ortografía. Se observa una correlación múltiple entre estos predictores y el criterio (promedio obtenido en las materias rendidas durante el primer año de estudios universitarios) de .65 significativamente distinta de cero a nivel .001.

La tabla 3 presenta la matriz de correlaciones entre las siete variables seleccionadas a nivel de primer año como las mejores predictoras (variables independientes) y los promedios obtenidos por las egresadas en las cinco áreas de materias mencionadas anteriormente (variables dependientes); como así las medias y desviaciones standards de las variables dependientes. Se observan correlaciones significativamente distintas de cero entre Interés Persuasivo y los promedios de las materias del área clínica y Social-laboral; entre el subtest Verbal y los promedios de materias de las áreas educacional, experimental, social-laboral y generales; y entre Ortografía y los promedios de las áreas de materias experimental y generales.

La Tabla 4 que presenta las raíces características y las correspondientes correlaciones canónicas muestra que si bien las dos primeras correlaciones canónicas son mayores que cualquiera de los coeficientes de correlaciones entre las variables originales, estos valores ($R_{c.1} = .58$ y $R_{c.2} = .45$) no son significativamente distintos de cero. Considerando que estos valores son mayores de .30 se presenta en la tabla 5 los vectores canónicos normalizados asociados a dichas correlaciones. El primer vector canónico asociado a las variables independientes refleja las áreas de intereses Persuasivo, Servicio Social y Científico y las aptitudes Ortografía y Verbal y el primer vector canónico asociado a las variables dependientes refleja las áreas de materias Educacional, Social-laboral y en cierto grado el área Clínica. El segundo vector canónico asociado a las variables predictoras refleja las áreas de intereses Musical y Científico y la aptitud Ortografía y el vector canónico asociado a las variables criterio es bipolar reflejando en el extremo positivo el área de materias generales y en el extremo negativo las áreas Social-laboral y Educacional. Esto podría sugerir que existen dos factores comunes a ambas series de variables, es decir que hay dos formas por medio de las cuales se correlacionan entre sí las variables dependientes y las variables independientes.

Haciendo un análisis detallado de la cuarta columna de la tabla 6 donde se presentan las correlaciones de las variables independientes con el promedio general obtenido en la carrera por las egresadas se observan sólo correlaciones significativamente distintas de cero con los subtests Verbal y Ortografía. Sin embargo el coeficiente de correlación múltiple utilizando como predictoras las siete variables y como criterio el promedio general da un valor altamente significativo ($R_{múlt} = .57$, significativamente distinto de cero a nivel .01). La última columna de la tabla 6 que presenta los valores "t" de los coeficientes de regresión muestra valores significativamente distintos de cero para las variables: interés Persuasivo, Servicio Social y el Subtest Ortografía.

Discusión

Teniendo presente los objetivos planteados en la Introducción se concluye que es posible predecir el 42,425% de la varianza del criterio (rendimiento académico del primer año de estudios universitarios) en base a cuatro áreas de intereses y a tres áreas de aptitudes; y luego de cinco años de estudios universitarios es posible predecir en base a las mismas variables el 33% de la varianza del segundo criterio utilizado

TABLA 4

CORRELACIONES CANONICAS ENTRE LAS SIETE VARIABLES PREDICTORAS Y LOS PROMEDIOS DE LAS CINCO AREAS DE MATERIAS

Número de raíz.caract. extraídas	Raíces Características	Correlac. canónicas	lambda	Chi Cuadrado	Grados de libertad
0	.355	.579	.444	44.22	35
1	.200	.418	.668	21.98	24
2	.091	.301	.836	9.78	15
3	.049	.221	.919	4.59	8
4	.034	.184	.966	1.87	3

TABLA 5

VECTORES CANONICOS NORMALIZADOS ASOCIADOS CON LAS DOS PRIMERAS CORRELACIONES CANONICAS

Predictores	Vector I	Vector II	Criterios	Vector I	Vector II
Científico	.374	.306	Clínica	.314	-.219
Persuasivo	.684	-.021	Educacional	.735	-.305
Musical	-.088	.803	Experimental	.097	.175
Serv.Social	.402	-.085	Soc.Laboral	.589	-.391
Verbal	.329	.179	Generales	.066	.822
Abstracto	.011	-.212			
Ortografía	.337	.405			

TABLA 6

MEDIAS, DESVIACIONES STANDARDS, CORRELACIONES CON EL PROMEDIO GENERAL, COEFICIENTES DE REGRESION, ERROR STANDARD DE LOS COEFICIENTES DE REGRESION, VALORES "t"; CORRELACION MULTIPLE Y TEST F. (N= 61).

Variable	M	D.S.	r_{xy}	Coefic. Regres.	Error st. coef.regr.	Valor "t" calculado
Int.Cientif.	37.54	8.76	.043	.030	.017	1.78
Int.Persuas.	37.26	9.31	.241	.047	.016	2.94**
Int.Musical	12.33	5.34	-.071	-.002	.022	-.11
Int.Ser.Soc.	61.66	9.55	.160	.028	.013	2.17*
Subt.Verbal	26.33	7.82	.352*	.031	.016	1.91
Subt.Abstrac.	27.95	9.30	.172	-.003	.014	-.25
Subt.Ortogr.	56.00	15.80	.309*	.018	.008	2.32*
Variable dep.	6.83	.99				
$R_{m\acute{u}lt.} = .573^{**}$ $F = 3.70^{**}$						
ord. origen = .5808						

*p <.05

**p <.01

(promedio general de la carrera). Resulta promisor observar que entre las áreas de intereses destacadas como mejores predictoras se encuentran los intereses Científico, Persuasivo y Servicio Social que concuerdan con hallazgos de previas investigaciones las cuales consideran de importancia dichos intereses para seguir estudios y ejercer la profesión de psicólogo⁴. Así por ejemplo se considera importante obtener puntajes altos en las áreas de intereses Científicos y Servicio Social para los profesionales de Psicología clínica y Psicología infantil, de intereses Persuasivos y Servicio Social para los profesionales en Orientación Vocacional, en interés Científico para los Psicólogos e investigadores.

El área de interés Musical que aparece como buena predictora a nivel de primer año no presenta ninguna significación como predictora del promedio general de la carrera. Esto puede ser debido a que dicho interés fuese considerado como hobby por las alumnas ingresantes o que la formación que luego recibieron no tuvo en cuenta dicha área, ya que no existe, por ejemplo, ninguna especialidad de la carrera relacionada con la musicoterapia.

No es raro observar que las aptitudes Verbal y Ortografía se mantienen como buenas variables predictoras en ambos niveles, ya que la primera tiene como fin evaluar la capacidad de abstracción y generalización del estudiante lo cual está íntimamente ligado con las actividades que debe desarrollar un estudiante universitario, y por supuesto la variable Ortografía está muy relacionada con la variable verbal. El subtest Razonamiento Abstracto resulta ser buen predictor a nivel de primer año, pero no con el promedio general de la carrera, ni tiene pesaje en las funciones lineales analizadas por medio de las correlaciones canónicas. Pareciera ser que de los tres aspectos considerados como medida de inteligencia general (Verbal, Numérica y Abstracta) la variable Verbal es la más importante para seguir estudios de Psicología.

Haciendo un estudio comparativo de la predicción de rendimiento académico de las egresadas tomando como variables independientes las siete variables mencionadas anteriormente y como criterio por un lado, el promedio general de la carrera y por otro el promedio en las cinco áreas de materias, se observa que los valores del coeficiente de correlación múltiple en el primer caso y la primera correlación canónica en el segundo son prácticamente iguales. La ventaja de las correlaciones canónicas al permitir trabajar con una serie de variables dependientes no produce resultados significativos. Esto puede ser debido a que los promedios por áreas de materias no se diferencian entre sí. Es decir que a pesar de ser una Facultad en la cual a la carrera de psicología se le da gran importancia al área clínica, el rendimiento de los alumnos es aparentemente igual en esta área que en las otras; sin dejar de reconocer, por supuesto, la poca confiabilidad de las notas como índice de rendimiento.

En resumen, tres áreas de intereses (científico, persuasivo y servicio social) y la aptitud Verbal mantienen su poder predictivo a lo largo de los cinco años de estudios universitarios, llegando a predecir un tercio de la varianza del criterio. Es posible suponer que los dos tercios de dicha varianza son debidos a variables tales como motivación, personalidad, hábitos de estudios, etc. En segundo lugar, el coeficiente de correlación múltiple puede considerarse un buen estimador de la predicción del criterio cuando se cuenta sólo con las notas como índice de rendimiento académico.

Referencias

Bartlett, M.S. The statistical significance of canonical correlations. *Biometrika*, 1947, 9.

⁴ Clasificación de las ocupaciones según los intereses principales, Manual del Registro de Preferencias Kuder.

- Cooley, W.W., & Lohnes, R.P. *Multivariate data analysis*. London: 1971.
Dubois, P.H. *Multivariate correlational analysis*. N.Y.: Harper, 1957.
Horst, P. Relations among sets of measures. *Psychometrika*, 1961, 26.
Hotelling, H. Relations between two sets of variates. *Biometrika*, 1936, 28.
McNemar, Q. *Psychological statistics*. N. Y.: Wiley, 1962.

PRESENTACION DE LA PRIMERA VERSION: 27 de febrero de 1975

PRESENTACION DE LA SEGUNDA VERSION: 23 de mayo de 1975