Reunión anual de la Unión Matemática Argentina

LXIV Reunión Anual de Comunicaciones Científicas

"ANALISIS DE ASOCIACIONES ENTRE VARIABLES CATEGORIZADAS EN UN ESTUDIO SOBRE GRADUADOS UNIVERSITARIOS"

CARRIZO, Jorgelina- RUIZ, Ana María- MALLEA, Adriana

Universidad Nacional de San Juan -UNSJ

Jorgelinav_carrizo@yahoo.com

-BAHIA BLANCA 2016-

Objetivo General

- Indagar asociación entre variables relacionadas con el ingreso y egreso universitario que permitan describir aspectos sobre el transito del graduado en la universidad

Objetivos específicos

- Describir relaciones entre variables categorizadas con información presente en tablas de contingencia

- Modelar interacciones entre variables categorizadas

- Evaluar la significancia de asociaciones entre las distintas variables.

- Interpretar parámetros en contexto del problema de estudio que permita describir aspectos relacionados al ingreso y egreso universitario

Contexto

Implementación Sistema SIU Kolla en la Facultad de Filosofía Humanidades y Artes de la UNSJ

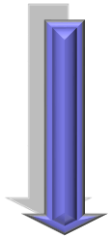
SIU-Kolla es una encuesta que permite a las universidades conocer la realidad de los estudiantes al momento de su graduación y realizar un seguimiento de sus graduados

Se trata de una encuesta para analizar, diagnosticar y conocer los perfiles de los egresados, y constituye una parte de un proyecto integral que aspira a crear una base de conocimiento sobre la población estudiantil.

Diseño experimental: población y muestra

Población:

Carreras de la FFHA, en su mayoría carreras de formación docente pertenecen a las Áreas de Filosofía y Humanidades, Ciencias Básicas y Tecnología, Música y Artes



Muestra :

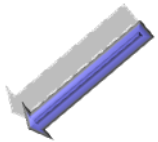
Alumnos de carreras de la FFHA (84) que gestionaron su titulo en el año 2014

Cuadro 1: Variables utilizadas en el estudio

Variables	*Categorías*
Características de los graduados	
▪ **A:** realiza carrera a termino	**SI:** Ingresa después del 2007 **NO**: Ingresa antes del 2007
▪ **B:** motivos de elección de carrera	**VOC:** Por vocación **OTROS:** Otros motivos
▪ **C**: nivel estudios de los padres	**BAS:** Primarios y/o secundarios **SUP:** Superiores y/o universitarios

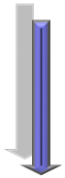
¿Cómo resolver el problema?

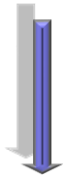
Metodología estadística para datos categorizados



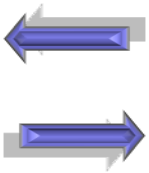
Análisis descriptivo y exploratorio: gráficos y tablas de contingencia

Análisis confirmatorio: modelado estocástico





Análisis de correspondencias (ACS y ACM)



Modelos Log-lineales (MLL)

Todas las técnicas utilizadas se encuentran implementadas en el programa estadístico InfoStat y SAS versión 9.1

Tablas de contingencia bidimensionales

Independencia

$$H_0 : \pi_{ij} = \pi_{i\cdot} \times \pi_{\cdot j}$$

Test estadístico Chi-cuadrado de Pearson

$$\chi^2 = N \sum_{i=1}^{a} \sum_{j=1}^{b} \frac{(p_{ij} - \hat{\pi}_{ij})^2}{\hat{\pi}_{ij}}$$

$$\chi^2 \approx \chi^2_{(a-1)(b-1)}$$

Test de la razón de verosimilitud

$$G^2 = 2N \sum_{i=1}^{a} \sum_{j=1}^{b} p_{ij} \log_e \left(\frac{p_{ij}}{\hat{\pi}_{ij}} \right)$$

$$G^2 \approx \chi_{1-\alpha(a-1)(b-1)}$$

Tablas de contingencia 2x2
Modelos probabilisticos

- **Modelo de homogeneidad de proporciones** (totales filas fijos: muestreo binomial)

$$H_0 : \pi_1 = \pi_2 \qquad \text{vs.} \qquad H_a : \pi_1 \neq \pi_2$$

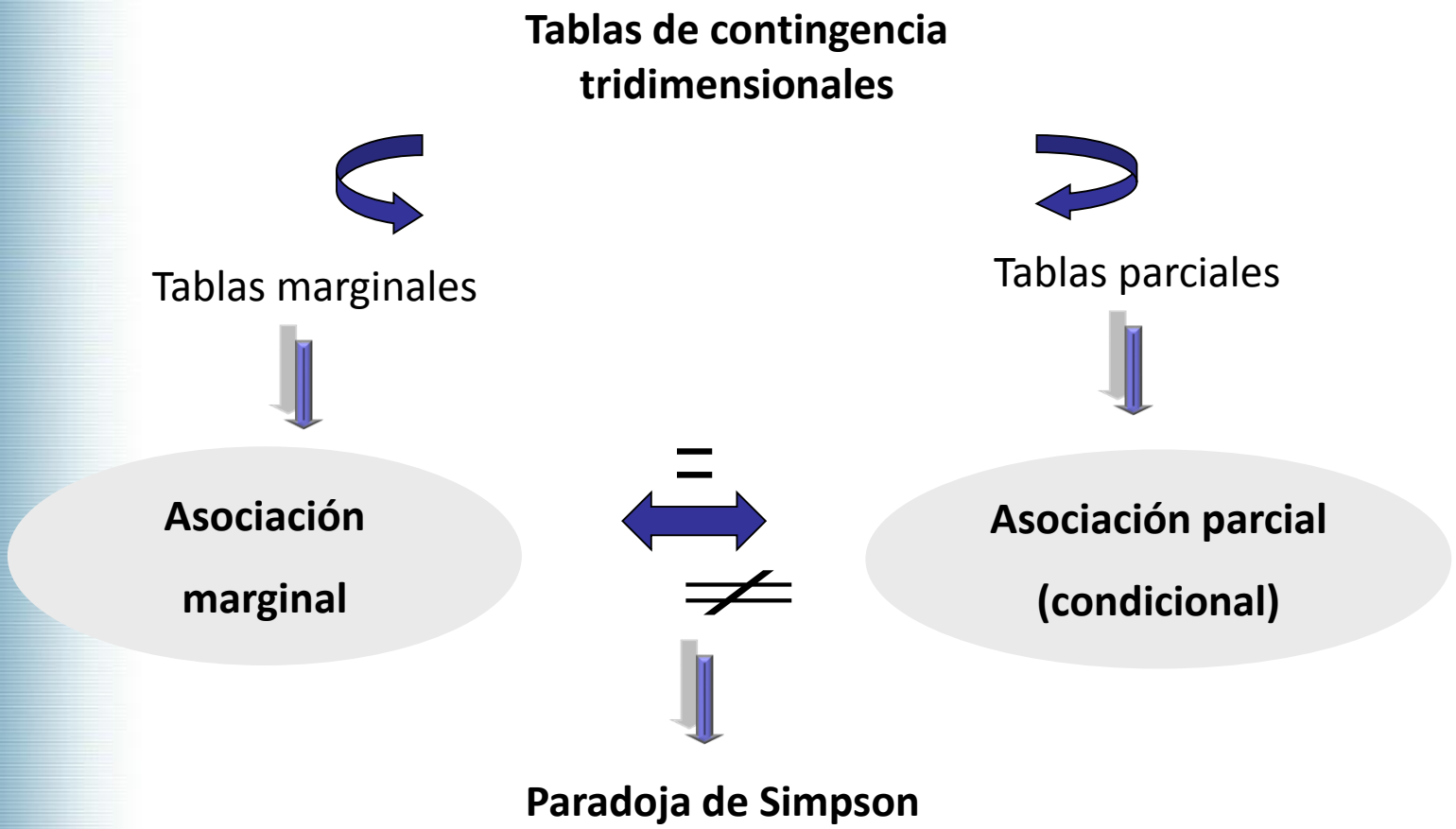
  - Prueba de diferencia de probabilidades
  - Prueba Chi-cuadrado de homogeneidad
  - Prueba de razón de productos cruzados (odds ratio)

- **Modelo de independencia** (total general fijo: muestreo binomial; aleatorio: Poisson)

$$H_0 : \pi_{ij} = \pi_{i.}\pi_{.j}$$

  - Prueba de test de la razón de verosimilitud
  - Prueba Chi cuadrado de Pearson
  - Prueba de la razón de productos cruzados
  - Prueba del coeficiente de correlación

Tablas de contingencia tridimensionales

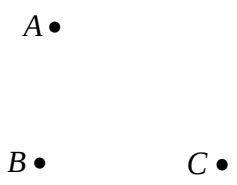


Modelos de independencia con tres variables
Representación gráfica

Modelo de independencia completa

$$M^{(0)} : p_{ijk} = p_{i..}\, p_{.j.}\, p_{..k}$$

Representación gráfica (grafo)



Estimaciones y tests estadísticos

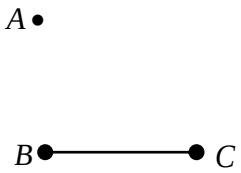
$$\hat{p}_{ijk} = \hat{p}_{i..}\,\hat{p}_{.j.}\,\hat{p}_{..k} = \frac{n_{i..}}{n_{...}} \times \frac{n_{.j.}}{n_{...}} \times \frac{n_{..k}}{n_{...}} \qquad \hat{m}_{ijk} = n_{...} \times \hat{p}^{(0)}_{ijk} = \frac{n_{i..} \times n_{.j.} \times n_{..k}}{n^2_{...}}$$

$$\chi^2 = \sum_i \sum_j \sum_k \frac{(n_{ijk} - \hat{m}^{(0)}_{ijk})^2}{\hat{m}^{(0)}_{ijk}} \qquad G^2 = 2 \sum_i \sum_j \sum_k n_{ijk} \times \log_e \frac{n_{ijk}}{\hat{m}^{(0)}_{ijk}}$$

Modelo con una variable independiente de las dos restantes

$$M^{(1)} : p_{ijk} = p_{i..} \times p_{.jk} \qquad M^{(2)} : p_{ijk} = p_{.j.} \times p_{i.k} \qquad M^{(3)} : p_{ijk} = p_{..k} \times p_{ij.}$$

Representación gráfica de $M^{(1)}$ (grafo)



Estimaciones y tests estadísticos

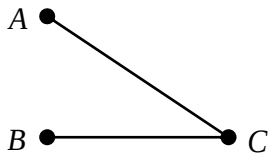
$$\hat{p}^{(1)}_{ijk} = \hat{p}_{i..} \times \hat{p}_{.jk} = \frac{n_{i..}}{n_{...}} \times \frac{n_{.jk}}{n_{...}} \qquad \hat{m}^{(1)}_{ijk} = n_{...} \times \hat{p}^{(1)}_{ijk} = n_{...} \times \frac{n_{i..}}{n_{...}} \times \frac{n_{.jk}}{n_{...}} = \frac{n_{i..} \times n_{.jk}}{n_{...}}$$

$$\chi^2 = \sum_i \sum_j \sum_k \frac{(n_{ijk} - \hat{m}^{(1)}_{ijk})^2}{\hat{m}^{(1)}_{ijk}} \qquad G^2 = 2 \sum_i \sum_j \sum_k n_{ijk} \times \log_e \frac{n_{ijk}}{\hat{m}^{(1)}_{ijk}}$$

Modelo de independencia condicionada

$$M^{(4)}: p_{ijk} = \frac{p_{i\cdot k}\, p_{\cdot jk}}{p_{\cdot\cdot k}} \qquad M^{(5)}: p_{ijk} = \frac{p_{ij\cdot}\, p_{\cdot jk}}{p_{\cdot j\cdot}} \qquad M^{(6)}: p_{ijk} = \frac{p_{ij\cdot}\, p_{i\cdot k}}{p_{i\cdot\cdot}}$$

Representación gráfica de $M^{(4)}$ (grafo)



Estimaciones y tests estadísticos

$$\hat{p}_i^{(4)} = \frac{\hat{p}_{i\cdot k} \times \hat{p}_{\cdot jk}}{\hat{p}_{\cdot\cdot k}} = \frac{\frac{n_{i\cdot k}}{n_{\cdots}} \times \frac{n_{\cdot jk}}{n_{\cdots}}}{\frac{n_{\cdot\cdot k}}{n_{\cdots}}} = \frac{n_{i\cdot k} \times n_{\cdot jk}}{n_{\cdot\cdot k} \times n_{\cdots}} \qquad \hat{m}_{ijk}^{(4)} = n_{\cdots} \times \hat{p}_{ijk}^{(4)} = n_{\cdots} \times \frac{n_{i\cdot k} \times n_{\cdot jk}}{n_{\cdot\cdot k} \times n_{\cdots}} = \frac{n_{i\cdot k} \times n_{\cdot jk}}{n_{\cdot\cdot k}}$$

$$\chi^2 = \sum_i \sum_j \sum_k \frac{(n_{ijk} - \hat{m}_{ijk}^{(4)})^2}{\hat{m}_{ijk}^{(4)}} \qquad G^2 = 2 \sum_i \sum_j \sum_k n_{ijk} \times \log_e \frac{n_{ijk}}{\hat{m}_{ijk}^{(4)}}$$

Tablas de contingencia multidimensionales

¿Cómo analizar información en tablas de contingencia muldimensionales?

Análisis exploratorio

Análisis de Correspondencias Múltiples (ACM)

(explora gráficamente relaciones entre variables categóricas con el objeto de determinar las razones específicas de las desviaciones de la hipótesis de independencia entre estas variables. No es necesario asumir una distribución de la cual provienen los datos).

Modelado estocástico

Modelos Log-lineales

(permite estudiar las relaciones entre variables categóricas y así mismo orienta a encontrar el modelo más adecuado para describirlas. Es necesario asumir una distribución de la cual provienen los datos y el modelo hipotético que estos siguen).

Técnicas de análisis complementarias

El modelo Log-lineal como modelo lineal generalizado (MLG)

Definido como un MLG el MLL puede considerarse como un modelo *de error de Poisson* con las siguientes características:

- No se establece distinción alguna entre variables explicativas y variables de respuesta (son de naturaleza simétricas).
- El componente sistemático solamente admite variables categorizadas (nominales u ordinales).
- El componente aleatorio se distribuye según Poisson: $\varepsilon_i \approx P(\mu_i)$
- La función de enlace canónico es la *transformación logarítmica*, definida por: $\eta_i = \log(\mu_i)$

MLL para dos dimensiones

Modelo de independencia

$$\ln(m_{ij}) = \mu + \lambda_i^A + \lambda_j^B$$

$$\sum_i \lambda_i^x = \sum_j \lambda_j^y = 0$$

Nº de parámetros: 1+(i-1)+(j-1)=i+j-1

Modelo saturado o completo

$$\ln(m_{ij}) = \mu + \lambda_i^A + \lambda_j^B + \lambda_{ij}^{AB}$$

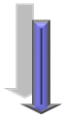
$$\sum_i \lambda_i^x = \sum_j \lambda_j^y = \sum_i \lambda_{ij}^{xy} = \sum_j \lambda_{ij}^{xy} = 0,$$

Nº de parámetros:1+(i-1)+(j-1)+((i-1)(ju-1)=ij

(tantos parámetros como celdas)

Si en una tabla de contingencia bidimensional se verifica que $p_{ij} = p_{i\cdot} p_{\cdot j}$ para todo *i* y para todo *j*,

las relaciones entre las dos variables categóricas puede modelizarse a través del modelo de independencia



$$p_{ij} = p_{i\cdot} p_{\cdot j} \Leftrightarrow \ln(m_{ij}) = \mu + \lambda_i^A + \lambda_j^B$$

Cuadro 2.: Principales modelos jerárquicos con tres variables

MODELO	NOTACION ABREVIADA
Modelo de independencia completa (IC)	
$\ln(m_{ij}) = \mu + \lambda_i^A + \lambda_j^B + \lambda_k^C$	$[A, B, C]$
Modelos de independencia parcial (IP)	
$\ln(m_{ij}) = \mu + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AB}$	$[AB, C]$
$\ln(m_{ij}) = \mu + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ik}^{AC}$	$[AC, B]$
$\ln(m_{ij}) = \mu + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{jk}^{BC}$	$[BC, A]$
Modelos de independencia condicionada (ICO)	
$\ln(m_{ijk}) = \mu + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AB} + \lambda_{ik}^{AC}$	$[AB, AC]$
$\ln(m_{ijk}) = \mu + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ik}^{AC} + \lambda_{jk}^{BC}$	$[AC, BC]$
$\ln(m_{ijk}) = \mu + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{jk}^{BC} + \lambda_{ij}^{AB}$	$[BC, AB]$
Modelo de asociación parcial (AP)	
$\ln(m_{ijk}) = \mu + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AB} + \lambda_{ik}^{AC} + \lambda_{jk}^{BC}$	$[AB, AC, BC]$
Modelo saturado o completo (MS)	
$\ln(m_{ijk}) = \mu + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AB} + \lambda_{ik}^{AC} + \lambda_{jk}^{BC} + \lambda_{ijk}^{ABC}$	$[ABC]$

Modelo saturado o completo

$$\ln(m_{ijk}) = \mu + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AB} + \lambda_{ik}^{AC} + \lambda_{jk}^{BC} + \lambda_{ijk}^{ABC}$$

Cuadro 3: Modelos Log-lineales y sus hipótesis de independencia

Modelo IC	$p_{ijk} = p_{i..} \times p_{.j.} \times p_{..k}$	$\lambda_{ijk}^{ABC} = \lambda_{ij}^{AB} = \lambda_{ik}^{AC} = \lambda_{jk}^{BC} = 0$	$\forall i, j, k$
Modelo IP1	$p_{ijk} = p_{i..} \times p_{.jk}$	$\lambda_{ijk}^{ABC} = \lambda_{ij}^{AB} = \lambda_{ik}^{AC} = 0$	$\forall i, j, k$
Modelo IP2	$p_{ijk} = p_{.j.} \times p_{i\cdot k}$	$\lambda_{ijk}^{ABC} = \lambda_{ij}^{AB} = \lambda_{jk}^{BC} = 0$	$\forall i, j, k$
Modelo IP3	$p_{ijk} = p_{..k} \times p_{ij.}$	$\lambda_{ijk}^{ABC} = \lambda_{ik}^{AC} = \lambda_{jk}^{BC} = 0$	$\forall i, j, k$
Modelo ICO1	$p_{ijk} = \frac{p_{i\cdot k} \times p_{.jk}}{p_{..k}}$	$\lambda_{ijk}^{ABC} = \lambda_{ij}^{AB} = 0$	$\forall i, j, k$
Modelo ICO2	$p_{ijk} = \frac{p_{ij.} \times p_{i\cdot k}}{p_{i..}}$	$\lambda_{ijk}^{ABC} = \lambda_{jk}^{BC} = 0$	$\forall i, j, k$
Modelo ICO3	$p_{ijk} = \frac{p_{ij.} \times p_{.jk}}{p_{.j.}}$	$\lambda_{ijk}^{ABC} = \lambda_{ik}^{AC} = 0$	$\forall i, j, k$
Modelo AP	-----	$\lambda_{ijk}^{ABC} = 0$	$\forall i, j, k$

Estimadores de los parámetros del modelo

$$\ln(m_{ijk}) = \mu + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AB} + \lambda_{ik}^{AC} + \lambda_{jk}^{BC} + \lambda_{ijk}^{ABC}$$

$$u = \log(m_{111}) \qquad \lambda_i^X = \log\frac{m_{i11}}{m_{111}} \qquad \lambda_j^Y = \log\frac{m_{1j1}}{m_{111}} \qquad \lambda_k^Z = \log\frac{m_{11k}}{m_{111}}$$

$$\lambda_{ij}^{XY} = \log\left(\frac{m_{ij1}.m_{111}}{m_{i11}.m_{ij1}}\right) = \log\vartheta_{ij} \qquad \lambda_{ik}^{XZ} = \log\left(\frac{m_{i1k}.m_{111}}{m_{i11}.m_{11k}}\right) = \log\vartheta_{ik}$$

$$\lambda_{jk}^{YZ} = \log\left(\frac{m_{1jk}.m_{111}}{m_{1j1}.m_{11k}}\right) = \log\varphi_{jk} \qquad \lambda_{ijk}^{XYZ} = \log\left(\frac{m_{ijk}.m_{11k}/m_{i1k}.m_{1jk}}{m_{ij1}.m_{111}/m_{i11}.m_{1j1}}\right)$$

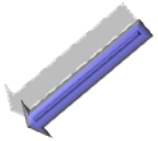
Ajuste de un MLL

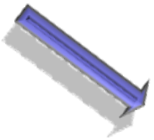
Se utiliza la *deviance D (G^2),* como medida de la discrepancia o bondad de ajuste existente entre valores empíricos y ajustados. Esta medida se basa en la *prueba de la razón de verosimilitud (likelihood ratio test).*

Modelo parsimonioso

(modelo de trabajo)

$M^{[T]}$





Estrategia de ajuste global

(modelo saturado)

$M^{[S]}$

si $\dfrac{D\left[M^{[T]}\right]}{N-P} \leq 1 \Rightarrow$ $M^{[T]}$ presenta buen ajuste

$$\lambda = \frac{L_0^{M^{[T]}}}{L_1^{M[S]}} \Rightarrow D[M^T] = -2\ln\lambda \approx \chi^2_{N-p}$$

Estrategia de ajuste condicional

(modelo restringido)

$M^{[R]}$

$$\Delta D = D\left[M^{[R]}\right] - D\left[M^{[T]}\right] \approx \chi^2_{p-q}$$

si ΔD es estadísticamente significativa se opta por el $M^{[T]}$

Cuadro 4: Tabla de contingencia entre las variables A, B y C

Termina carrera a termino (A)	Motivo elección carrera (B)	Nivel estudio de los padres (C)	Total
NO	OTROS	BAS	20
		SUP	12
	VOC	BAS	13
		SUP	6
SI	OTROS	BAS	7
		SUP	5
	VOC	BAS	7
		SUP	14
		TOTAL	**84**

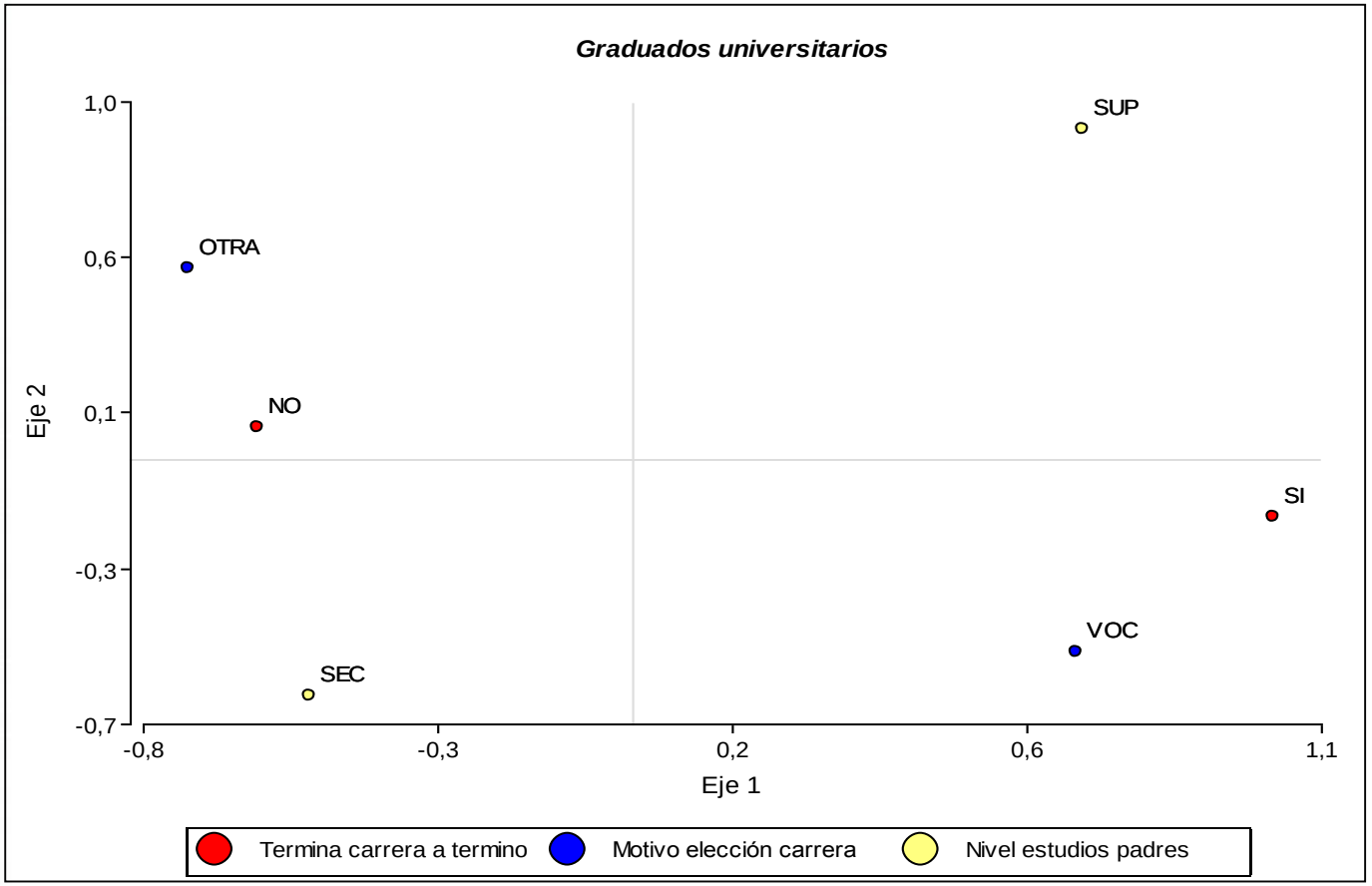


Figura 1 *Biplot del Análisis de Correspondencia Múltiple para visualizar asociaciones entre las variables: T çermina a tiempo carrera (A), Motivo elección carrera (B) y Nivel de estudios padres (C).*

Cuadro 5: Asociaciones marginales de las variables: A-B, A-C y B-C

Variables A : Termina carrera a termino B:Motivo elección carrera C: Nivel estudios padres			
	Termina carrera a termino – Motivo elección carrera (A-B)	Termina carrera a termino – Nivel estudios padres (A-C)	Motivo elección carrera – Nivel estudios padres (B-C)
Test Chi cuadrad de Pearson Grados de libertad Valor de p	5,59 1 0,0181	4,04 1 0,0445	0,65 1 0,419

Cuadro 6*:*Asociaciones parciales

Variables A y B condicionada por C		
	AB/C=BAS	AB/C)=SUP
Test Chi cuadrado de Pearson Grados de libertad Valor de p	5045 1 0,50	6,06 1 0,0138

asociación significativa (P<0,05)

Selección del mejor modelo (Método de ajuste global)

Cuadro 7 : Modelos con tres variables

Modelo	Efectos eliminados del modelo saturado	Diferencia de verosimilitud entre cada modelo y el modelo saturado			
		gl	G^2	G^2 / gl	p
$[AB, AC, BC]$	ABC	1	1,8232	1,8232	0,17
$[AB, AC]$	BC, ABC	2	2,1374	1,06	0,34
$[AB, BC]$	AC, ABC	2	5,0829	2,5414	0,07
$[AC, BC]$	AB, ABC	2	6,6839	3,34	0.03
$[AB]$	AC, BC, ABC	3	6,1823	2,068	0,10
$[AC]$	AB, BC, ABC	3	7,7833	2,5944	0,05
$[BC]$	AB, AC, ABC	3	10,7288	3,5763	0,01

asociación significativa (P<0,05

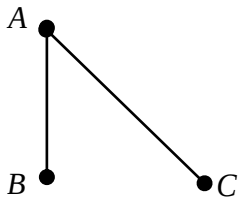
Modelo seleccionado y sus características

- Expresión funcional:

$$\ln(m_{ijkl}) = \mu + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AB} + \lambda_{ik}^{AC}$$

- Notación abreviada: $[AB, AC]$

- Representación Gráfica



- Test de ajuste global

$$\begin{cases} H_0 : \lambda_{jk}^{BC} = \lambda_{ijk}^{ABC} = 0 & G^2 = 2{,}1374 \quad 2gl \\ H_1 : \lambda_{jk}^{BC} \neq 0 \; y \; \lambda_{ijk}^{ABC} \neq 0 & (p = 0{,}34) \quad (P \geq \alpha) \end{cases}$$

Modelo seleccionado : interpretación

$$\ln(m_{ijkl}) = \mu + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AB} + \lambda_{ik}^{AC}$$

En este modelo de doble asociación los parámetros λ_{ij}^{AB} λ_{ik}^{AC}

captan la asociación parcial existente entre las variables A y B y A y C respectivamente, pero las variables B y C son condicionalmente independientes para cada nivel de la variable A.

Este modelo presupone que el terminar a tiempo la carrera (A) depende del motivo que la llevo a elegirla (B) y también depende del nivel de estudios de los padres. En consecuencia, la tabla marginal A-B y A-C exhibirán las misma asociaciones que sus tablas parciales, pero la asociación marginal B-C puede ser diferente de las asociaciones parciales B-C para cada nivel de A

Modelo seleccionado y sus características

Cuadro 8 : Valores predichos por el Modelo (AB,AC)

OBSERVACION	CONTEO	PREDICHO
1	13	12.294118
2	6	6.7058824
3	20	20.705882
4	12	11.294118
5	7	8.9090918
5	14	12.090909
7	7	5.0909133
8	5	6.9090957

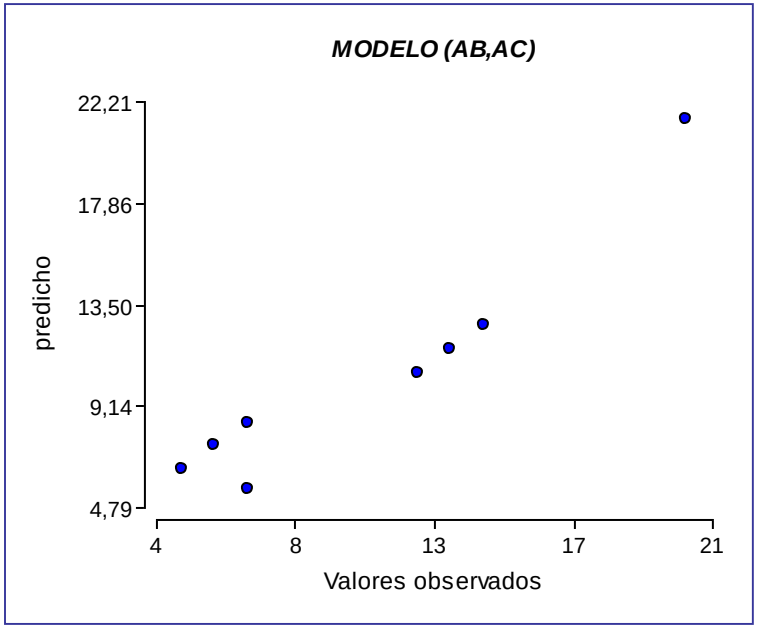


Fig.: Gráfico de dispersión valores observados versus predichos

Estimación de los parámetros del modelo $[AB, AC]$

Cuadro 8: Parámetros y sus estimaciones

Parámetro	Estimador	E.E.	P-valor
μ	2,49	0,26	<0,0001
$\lambda_1^A = \lambda_{NO}^A$	-0,59	0,39	0,138
$\hat{\lambda}_1^B = \lambda_{OTRO}^B$	-0,56	0,36	012
$\lambda_1^C = \lambda_{BAS}^C$	-0,30	0,35	0,38
$\lambda_{11}^{AB} = \lambda_{NO\_OTRO}^{AB}$	1,08	046	0,01
$\lambda_{11}^{AC} = \lambda_{NO-BAS}^{AC}$	0,91	0,45	0,04

Interpretación de los parámetros del modelo $[AB, AC]$

Cuadro 14: Cocientes de chances y su relación con los parámetros del modelo

$$\log \hat{\vartheta}_{AB} = \hat{\lambda}_{11}^{AB} = \hat{\lambda}_{NO-OTRO}^{AB} = 1{,}08 \Rightarrow \hat{\vartheta}_{AB} = e^{\hat{\lambda}_{NO-OTRO}^{AB}} = e^{1{,}08} = 2{,}94$$

Este cociente de chance condicional mide la relación entre el motivo que lleva a elegir la carrera el graduado (B) y el tiempo que tarda en terminarla (A)

$$\hat{\vartheta}_{AB} = \frac{\frac{\pi_{OTRO}}{1-\pi_{OTRO}} / NO}{\frac{\pi_{OTRO}}{1-\pi_{OTRO}} / SI} = 2{,}94$$

El valor hallado indica que, la chance de que habiendo elegido el graduado la carrera por otros motivos diferente a la vocación no termina a tiempo la misma, es 2,94 veces la chance de que, eligiendo la carrera por otros motivos, si termine a tiempo

Este resultado vale para cualquier de los niveles de los otra variable. Esto es, vale tanto para graduados con padres con nivel básicos de estudio o superior.

$$\hat{\vartheta}_{AB/C=BAS} = \hat{\vartheta}_{AB/C=SUP} = \hat{\vartheta} = 2{,}94$$

Interpretación de los parámetros del modelo $[AB, AC]$

Cuadro 14: Cocientes de chances y su relación con los parámetros del modelo

$$\log \hat{\vartheta}_{AC} = \hat{\lambda}_{11}^{AC} = \hat{\lambda}_{NO-BAS}^{AC} = 0{,}91 \Rightarrow \hat{\vartheta}_{AC} = e^{\lambda_{11}^{AC}} = e^{0{,}91} = 2{,}49$$

Este cociente de chance condicional mide la relación entre el nivel de estudio de los padres del graduado (C) y el tiempo que tarda en terminar la carrera(A)

$$\hat{\vartheta}_{AC} = \frac{\frac{\pi_{BAS}}{1-\pi_{BAS}} / NO}{\frac{\pi_{BAS}}{1-\pi_{BAS}} / SI} = 2{,}49$$

El valor hallado indica que, la chance de que para graduados con padres con nivel de estudios básicos no terminen a tiempo la carrera, es 2,49 veces la chance de que si la terminen.

Este resultado vale para cualquier de los niveles de los otra variable. Esto es, vale tanto para graduados que eligieron la carrera con vocación o por otros motivos.

$$\hat{\vartheta}_{AC/B=BAS} = \hat{\vartheta}_{AC/B=SUP} = \hat{\vartheta} = 2{,}49$$

Conclusiones

Aunque los resultados han de interpretarse con precaución debido al tamaño limitado de las muestras, podemos deducir que, de acuerdo con el estudio estadístico realizado, existe una relación importante entre las causas que influyen en el tiempo que tardó este grupo de graduados en culminar sus estudios universitarios. Por un lado, elegir la carrera con vocación actúa a favor de que se termine a tiempo y por otro, el tener padres con nivel de estudios superiores parece ser un factor importante a favor de este proceso

La metodología estadística utilizada, en particular el análisis con modelos log-lineales, nos ha permitido confirmar las asociaciones entre estos aspectos en los grupos experimentales. El conocimiento de estas asociaciones constituye un indicador de importancia para la institución universitaria que le permitirá trabajar en pos de una mejora sobre el transito de los graduados por la universidad.

Muchas
Gracias !