

ANÁLISIS DE CLASES LATENTES APLICADO A LA CONSTRUCCIÓN DE LA SATISFACCIÓN ESTUDIANTIL EN LA FACULTAD DE CIENCIAS ECONÓMICAS Y DE ADMINISTRACIÓN, UDELAR (URUGUAY)

Elena Vernazza, Mikaela Lezcano, Ramón Álvarez-Vaz
SEMINARIO DE INSTITUTO DE ESTADÍSTICA - SIESTA
Facultad de Ciencias Económicas y de Administración
Universidad de la República

2 de Diciembre de 2020



FACULTAD DE
CIENCIAS ECONÓMICAS
Y DE ADMINISTRACIÓN

UESTA INSTITUTO
DE ESTADÍSTICA



UNIVERSIDAD
DE LA REPÚBLICA
URUGUAY

ORDEN DE LA PRESENTACIÓN

1 INTRODUCCIÓN

2 METODOLOGÍA

3 RESULTADOS

- Análisis de Clases Latentes sin variables explicativas
- Análisis de Clases Latentes con variables explicativas

4 CONCLUSIONES

1 INTRODUCCIÓN

2 METODOLOGÍA

3 RESULTADOS

- Análisis de Clases Latentes sin variables explicativas
- Análisis de Clases Latentes con variables explicativas

4 CONCLUSIONES

PRESENTACIÓN DEL PROBLEMA

- Conocer el nivel de satisfacción del cliente con un determinado servicio y, sobre todo, determinar *cómo se construye y caracteriza* resulta fundamental para tomar decisiones enfocadas a mantener y/o mejorar aquellos factores que se entienda que la determinan (Alves y Raposo, 2009),(Álvarez-Vaz *et al.*, 2016)
- Resulta esencial contar con un *instrumento* que pueda cuantificar y cualificar de forma acertada y, sin ambigüedades, tanto la satisfacción global como cada una de las dimensiones que la componen y determinan.
- Instrumentos
 - 1989: Índice Sueco de Satisfacción del Consumidor (SCSB) (Alfaro García y Caballero Domínguez, 2006).
 - 1994: Índice Americano de Satisfacción del Consumidor (ACSI) (Lopes *et al.*, 2009).
 - 1996: Índice Europeo de Satisfacción del Consumidor (ECSI) (Martensen *et al.*, 2000), (Ciavolino y Dahlgaard, 2007), (Leite y Filho, 2007), (Lopes *et al.*, 2009).

PRESENTACIÓN DEL PROBLEMA

En particular, en este trabajo:

- Los clientes: Estudiantes de la Facultad de Ciencias Económicas y de Administración, Universidad de la República, Uruguay Álvarez-Vaz y Vernazza (2017)
- El servicio: Educación de nivel terciario.
- El instrumento: ECSI mediante una adaptación del AR63 (Alves y Raposo, 2004).

OBJETIVOS

Objetivo general: Proponer una estructura de grupos de estudiantes de la Facultad de Ciencias Económicas y de Administración, Udelar, Uruguay (FCEA), a partir de su nivel de Satisfacción con la institución.

Objetivos específicos:

- 1 Comprobar si existe una variable de clases latentes que describa la *Satisfacción Estudiantil* sin variables explicativas (Modelo incondicional)
- 2 Comprobar si existe una variable de clases latentes que describa la *Satisfacción Estudiantil* con variables explicativas (Modelo condicional)
- 3 Comparar ambos resultados.

1 INTRODUCCIÓN

2 METODOLOGÍA

3 RESULTADOS

- Análisis de Clases Latentes sin variables explicativas
- Análisis de Clases Latentes con variables explicativas

4 CONCLUSIONES

ANÁLISIS DE CLASES LATENTES

OBJETIVO

Segmentar la tabla (o hipercubo) de contingencia creada a partir de variables observadas (*manifiestas*), por una variable no observada (*latente*) (Lazarsfeld, 1950), (Everitt, 1984), (Linzer y Lewis, 2011), (Agresti, 2013).

CARACTERÍSTICAS

Se supone que las respuestas a todas las variables manifiestas son estadísticamente independientes con respecto a los valores de la variable latente \Rightarrow *independencia local o condicional*.

El modelo asocia cada individuo a una *clase* latente (en términos de probabilidad) \Rightarrow se puede determinar el valor esperado con el que cada individuo responde a cada variable observada.

El modelo estimado no estipula el número de *clases* latentes. Se pueden usar estadísticos de bondad de ajuste para poder determinar la cantidad de clases a considerar.

Estos modelos pueden verse como un caso particular de un modelo de regresión \Rightarrow donde sería posible incluir variables predictivas de la membresía de cada observación a una clase latente (Bandein-Roche *et al.*, 1997).

ANÁLISIS DE CLASES LATENTES

DEFINICIÓN DEL MODELO

- J variables *manifiestas* categóricas politómicas, cada una con K_j posibles respuestas, para los $i = 1, 2, \dots, N$ individuos.
- Y_{ijk} será el valor observado de las J variables manifiestas para el individuo i , tal que $Y_{ijk} = 1$ si el individuo i da la respuesta k de la variable j y $Y_{ijk} = 0$ en otro caso, con $j = 1, \dots, J$ y $k = 1, \dots, K_j$ y π_{jrk} representará la probabilidad condicional de que una observación en la clase $r = 1, \dots, R$ produzca el k -ésimo resultado de la variable j -ésima.
- p_r representa la probabilidad *incondicional* de que un individuo pertenezca a una clase r (antes de tomar en cuenta el valor de Y_{ijk}) \rightarrow probabilidad *a priori*.

SUPUESTOS

La probabilidad de que un individuo i genere un conjunto J de resultados en las variables manifiestas, asumiendo independencia condicional de los resultados y dado la pertenencia a una clase r , es:

$$f(Y_i; \pi_r) = \prod_{j=1}^J \prod_{k=1}^{K_j} (\pi_{jrk})^{Y_{ijk}} \quad (1)$$

Además, la función de densidad es:

$$P(Y_i | \pi, p) = \sum_{r=1}^R p_r \prod_{j=1}^J \prod_{k=1}^{K_j} (\pi_{jrk})^{Y_{ijk}} \quad (2)$$

PARÁMETROS p_r Y π_{jrk}

Dadas \hat{p}_r y $\hat{\pi}_{jrk}$, las probabilidades *a posteriori* de que cada individuo pertenezca a una clase latente (usando regla de Bayes):

$$\hat{P}(r_i | Y_i) = \frac{\hat{p}_r f(Y_i; \hat{\pi}_{jrk})}{\sum_{q=1}^R \hat{p}_q f(Y_i; \hat{\pi}_q)} \quad (3)$$

con $r_i = 1, \dots, R$.

ANÁLISIS DE CLASES LATENTES

ESTIMACIÓN DEL MODELO

El método de estimación utilizado en este trabajo es el de máxima verosimilitud, a través del algoritmo de maximización EM [(Dempster *et al.*, 1977), (R Core Team, 2019)]

- 1 En la fase de esperanza (E), se calcula la probabilidad de membresía a la clase *latente*.
- 2 En la fase de maximización (M) los parámetros estimados se actualizan maximizando la log-verosimilitud dada la probabilidad a *posteriori*.

CRITERIOS DE SELECCIÓN Y VALIDACIÓN DEL MODELO

La estimación no estipula una cantidad de clases latentes. Esta se puede determinar a partir de la teoría o por experiencias empíricas previas. En la mayoría de los casos será necesario tener en consideración un criterio de parsimonia (en este caso mínimo BIC) aunque existen estadísticos de prueba χ^2 , G^2 , y de entropía [Fop y Murphy (2017), (Fop *et al.*, 2017)]

BONDADE AJUSTE

Ajuste del modelo en base a los valores observados

$\chi^2, G^2,$

$$\hat{Q}_c = N \sum_{r=1}^R \hat{p}_r \prod_{j=1}^{\hat{\pi}_{jrk}} \quad (4)$$

$$\chi^2 = \sum_{r=c} (q_c - \hat{Q}_c)^2 / \hat{Q}_c \quad (5)$$

$$G^2 = 2 \sum_{r=c} (q_c \log(q_c / \hat{Q}_c)) \quad (6)$$

DATOS

El **cuestionario** AR63 (Alves y Raposo, 2004) se estructura en 8 dimensiones.

- **Expectativas** - Qué esperan los estudiantes de la FCEA (12).
- **Imagen** - Cómo ven los estudiantes a la FCEA (6).
- **Calidad (SA)** - Calidad de los servicios académicos (9).
- **Calidad (SF)** - Calidad de los servicios funcionales (9).
- **Necesidades/Deseos** - Necesidades y deseos actuales (9).
- **Valor percibido** - Valor que les brinda estudiar en la FCEA (7).
- **Satisfacción** - Satisfacción con la FCEA (6).
- **Futuro** - Lealtad y Boca a boca (5).

El cuestionario se aplicó sobre una **muestra probabilística** de 2997 estudiantes y 59 Unidades Curriculares (UCs) de la FCEA en el año 2017. El diseño muestral fue estratificado en 2 etapas: se formaron 7 estratos de UCs, a cada estrato se le asignó un tamaño de muestra proporcional a la matrícula de estudiantes que recibe y luego se aplicó muestreo sistemático.

En este trabajo no serán consideradas las últimas dos dimensiones.

ANÁLISIS DE CLASES LATENTES SIN VARIABLES EXPLICATIVAS

-**Insumo:** Las respuestas en las 6 dimensiones del cuestionario - Seis variables manifiestas (Y_i): $p = 6$.

-Cada una de las variables observadas posee **4 categorías de respuestas posibles**: $k_i = 4$ (para $i = 1:6$).

$$Y_i = \begin{cases} 1 & \text{Bajo} \\ 2 & \text{MedioBajo} \\ 3 & \text{MedioAlto} \\ 4 & \text{Alto} \end{cases}$$

-**Resultado:** Una variable de clases latentes - Satisfacción estudiantil.

-**Tamaño de muestra:** $n = 1.809$ (sin considerar datos faltantes ni estudiantes repetidos).

1 INTRODUCCIÓN

2 METODOLOGÍA

3 RESULTADOS

- Análisis de Clases Latentes sin variables explicativas
- Análisis de Clases Latentes con variables explicativas

4 CONCLUSIONES

LIBRERÍAS USADAS

- *poLCA* (Linzer y Lewis, 2011). Permite estimar en forma condicional e Incondicional
- *LCAvarsel* (Fop y Murphy, 2017),(Fop *et al.*, 2017). Admite estimar en forma condicional e incondicional. Además permite extraer variables redundantes dentro de las manifiestas. Tiene 3 algoritmos de selección de variables (uno de ellos es de tipo algoritmo genético) y en cualquiera de ellos admite cambiar la cantidad de iteraciones y tolerancia.

RESULTADOS - ACL SIN COVARIABLES (1)

CUADRO: Patrones de respuesta - Frecuencias observadas.(10 más frecuentes)

E_R	I_R	CSA_R	CSF_R	ND_R	VP_R	Frecuencia
4	4	4	4	4	4	145
1	1	1	1	1	1	127
4	4	4	4	4	3	36
1	1	1	2	1	1	30
3	3	3	3	3	3	25
4	4	4	3	4	4	21
3	4	4	4	4	4	17
2	2	2	2	2	2	16
2	2	2	2	2	1	15
1	1	1	1	1	2	13

RESULTADOS - ACL SIN COVARIABLES (1)

MODELOS ESTIMADOS

Se estiman 4 modelos ($M = 1, 2, 3, 4$).

CUADRO: Estimación de 4 modelos - $M = 1, 2, 3, 4$ (Sin covariables)

Clases	BIC	χ^2	valor p	G^2	valor p	entropía
$M = 1$	30178.3	118065.1	0	9531.2	0	8.3
$M = 2$	25233.2	12428.5	0	4443.5	0	6.9
$M = 3$	23676.4	8765.4	0	2744.3	0	6.5
$M = 4$	23191.6	8152.1	0	2116.9	0	6.3

RESULTADOS - ACL SIN COVARIABLES (1)

CUADRO: Probabilidad de cada una de las clases (Modelo 4)

Clase	1	2	3	4
P(r)	0.289	0.199	0.239	0.271

VISUALIZACIÓN DE PARÁMETROS ESTIMADOS

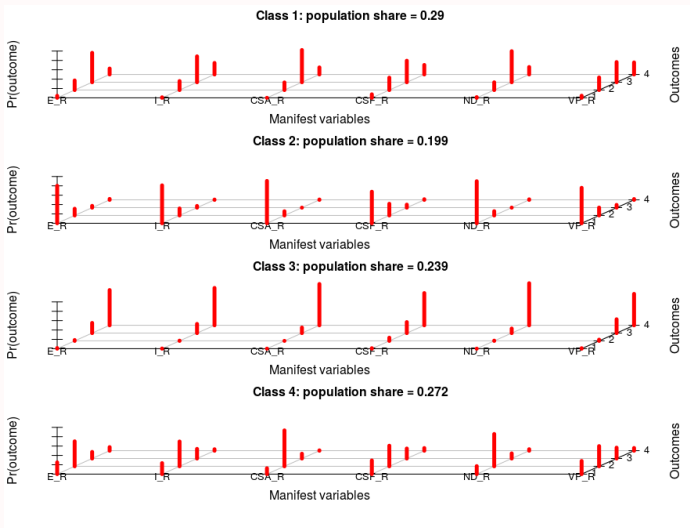


FIGURA: Probabilidades condicionales

RESULTADOS - ACL SIN COVARIABLES (2)

CUADRO: Probabilidades condicionales $P(Y_i/r)$.

E_R	1	2	3	4
$r = 1$	0.04	0.20	0.63	0.13
$r = 2$	0.80	0.14	0.04	0.02
$r = 3$	0.01	0.02	0.22	0.75
$r = 4$	0.25	0.53	0.14	0.08
I_R	1	2	3	4
$r = 1$	0.01	0.19	0.55	0.24
$r = 2$	0.81	0.15	0.04	0.01
$r = 3$	0.00	0.00	0.20	0.80
$r = 4$	0.23	0.53	0.20	0.03
CSA_R	1	2	3	4
$r = 1$	0.00	0.16	0.69	0.15
$r = 2$	0.90	0.09	0.00	0.01
$r = 3$	0.00	0.00	0.12	0.88
$r = 4$	0.12	0.77	0.10	0.01

CUADRO: Probabilidades condicionales $P(Y_i/r)$.

CSF_R	1	2	3	4
$r = 1$	0.07	0.26	0.46	0.20
$r = 2$	0.67	0.25	0.06	0.01
$r = 3$	0.01	0.07	0.23	0.69
$r = 4$	0.29	0.44	0.21	0.06
ND_R	1	2	3	4
$r = 1$	0.01	0.16	0.66	0.16
$r = 2$	0.90	0.09	0.00	0.01
$r = 3$	0.00	0.01	0.10	0.90
$r = 4$	0.17	0.69	0.11	0.03
VP_R	1	2	3	4
$r = 1$	0.04	0.27	0.43	0.26
$r = 2$	0.76	0.17	0.06	0.02
$r = 3$	0.01	0.03	0.29	0.67
$r = 4$	0.28	0.43	0.23	0.06

RESULTADOS - ACL SIN COVARIABLES (3)

CARACTERIZACIÓN DE LAS CLASES

- $r = 1$: Estudiantes con *Satisfacción Estudiantil* **media-alta**.
- $r = 2$: Estudiantes con *Satisfacción Estudiantil* **baja**.
- $r = 3$: Estudiantes con *Satisfacción Estudiantil* **alta**.
- $r = 4$: Estudiantes con *Satisfacción Estudiantil* **medio baja**.

RESULTADOS - ACL SIN COVARIABLES (4)

CUADRO: Probabilidades a posteriori según patrones y asignaciones.

E_R	I_R	CSA_R	CSF_R	ND_R	VP_R	Frec	$P(r=1)$	$P(r=2)$	$P(r=3)$	$P(r=4)$	Asignación
4	4	4	4	4	4	145	0.00	0.00	1.00	0.00	3
1	1	1	1	1	1	127	0.00	1.00	0.00	0.00	2
4	4	4	4	4	3	36	0.00	0.00	1.00	0.00	3
1	1	1	2	1	1	30	0.00	1.00	0.00	0.00	2
3	3	3	3	3	3	25	1.00	0.00	0.00	0.00	1
4	4	4	3	4	4	21	0.00	0.00	1.00	0.00	3
3	4	4	4	4	4	17	0.00	0.00	1.00	0.00	3
2	2	2	2	2	2	16	0.00	0.00	0.00	1.00	4
2	2	2	2	2	1	15	0.00	0.00	0.00	1.00	4
1	1	1	1	1	2	13	0.00	1.00	0.00	0.00	2

CUADRO: Probabilidad a posteriori de cada una de las clases.

Clase	1	2	3	4
$P(r)$	0.285	0.197	0.241	0.247

ACL CON VARIABLES EXPLICATIVAS (1)

MODELO

=====
 Fit for 4 latent classes:
 =====

2 / 1

Coefficient	Std. error	t value	Pr(> t)	
(Intercept)	0.02527	0.16152	0.156	0.876
EDAD.rec1	-0.76580	0.24633	-3.109	0.002
contador	0.06097	0.15333	0.398	0.691
trabaja	0.19058	0.16749	1.138	0.255
Abandono_cat	0.35984	0.19044	1.890	0.059

=====
 3 / 1

Coefficient	Std. error	t value	Pr(> t)	
(Intercept)	-0.32223	0.16909	-1.906	0.057
EDAD.rec1	-0.97093	0.26068	-3.725	0.000
contador	0.47668	0.15297	3.116	0.002
trabaja	0.20200	0.16442	1.229	0.219
Abandono_cat	0.44570	0.19196	2.322	0.020

=====
 4 / 1

Coefficient	Std. error	t value	Pr(> t)	
(Intercept)	-0.59167	0.17768	-3.330	0.001
EDAD.rec1	-0.89246	0.25255	-3.534	0.000
contador	0.26200	0.15895	1.648	0.099
trabaja	0.31633	0.17750	1.782	0.075
Abandono_cat	0.50901	0.19223	2.648	0.008

=====
 number of observations: 1809 number of estimated parameters: 87
 residual degrees of freedom: 1722 maximum log-likelihood: -11294.03

AIC(4): 22762.06;BIC(4): 23240.61;X²(4): 8247.008 (Chi-square goodness of fit)

ACL CON VARIABLES EXPLICATIVAS (2)

ANÁLISIS COMPARATIVO

Clases sin covariables	Clases con covariables				Total
	1 (A)	2 (MA)	3 (MB)	4 (B)	
1 (MA)	4	507	5	0	516
2 (B)	0	0	0	358	358
3 (A)	433	5	0	0	438
4 (MB)	0	5	488	4	497
Total	437	517	493	362	1809

CARACTERIZACIÓN

- **Variables sexo y plan de estudio:** No hay diferencias significativas entre las clases. En todas las clases predomina un porcentaje de mujeres cercano al 55% y una cantidad de estudiantes pertenecientes al Plan 2012 que representan casi el 70%.
- **Variables edad, abandono y trabaja:** Comportamiento similar. A medida que se pasa a la siguiente clase (recordando que esto implica un menor grado de satisfacción en la FCEA), mayor cantidad de estudiantes se encuentran **trabajando**.
- Estudiando la variable **abandono** se puede ver que, a medida que se analiza la siguiente clase aumenta el porcentaje de estudiantes que abandonaron al menos una vez la carrera en FCEA.
- Y si se analiza la variable **edad** se puede afirmar que a medida que se avanza a la clase siguiente, menor cantidad de estudiantes mayores a 27 años.
- **Variable contador:** la clase que presenta una menor proporción es la 1 con 59% y la mayor proporción la 3 con 72%.

1 INTRODUCCIÓN

2 METODOLOGÍA

3 RESULTADOS

- Análisis de Clases Latentes sin variables explicativas
- Análisis de Clases Latentes con variables explicativas

4 CONCLUSIONES

CONCLUSIONES

ACL SIN COVARIABLES (MODELO INCONDICIONAL)

Los resultados obtenidos y presentados en este trabajo, nos permiten concluir:

- Las 6 variables manifiestas logran formar una variable latente en 4 clases✓
- $r = 1$: Estudiantes con *Satisfacción Estudiantil* **media-alta**.
- $r = 2$: Estudiantes con *Satisfacción Estudiantil* **baja**.
- $r = 3$: Estudiantes con *Satisfacción Estudiantil* **alta**.
- $r = 4$: Estudiantes con *Satisfacción Estudiantil* **medio baja**.

ACL CON COVARIABLES (MODELO CONDICIONAL)

- Las 6 variables manifiestas logran formar una variable latente en 4 clases✓
- Las clases detectadas se asocian con la edad, el abandono, el trabajo y la carrera que cursa✓
- Hay un acuerdo muy fuerte en las clases formadas con ambas estrategias✓

CONSIDERACIONES FINALES

Como limitaciones y propuestas a futuro se plantea:

- En las 2 estrategias los métodos de estimación convergen y forman clases, sin embargo los **estadísticos de prueba no logran funcionar como criterio de parada**, por lo que es necesario estudiar este fenómeno que parece estar asociado a:
- Tamaño de muestra (por arriba de 1500 el método comienza a mostrar ese comportamiento).
- La estructura del hipercubo de contingencia que puede verse como una matriz 'sparse' (apenas 15% de las configuraciones posibles se observan).
- Considerar los pesos muestrales que tiene las observaciones en virtud del diseño muestral utilizado (Oberski, 2013).

ELECCIÓN DE VARIABLES REDUNDANTES SIN VARIABLES EXPLICATIVAS

```
mod1<-LCAvarsel(lasseis, G=1:6, X=NULL, search="ga")
> mod1
```

Variable selection for Latent Class Analysis

```
=====
Evolutionary search via genetic algorithm
with conditional dependence via regression
=====
```

```
Selected variables: E_R I_R CSA_R CSF_R ND_R
Clustering model: LCA model with 4 latent classes
```

Summary:

```
Best = -23092.90
Mean = -23802.26
Median = -23989.44
Min = -24093.43
```

ELECCIÓN DE VARIABLES MANIFESTAS REDUNDANTES CON VARIABLES EXPLICATIVAS

```
mod5<-LCAvarsel(lasseis, G=1:6, X=covariables2, search="ga")  
> mod5
```

Variable selection for Latent Class Analysis

```
=====  
Evolutionary search via genetic algorithm  
with conditional dependence via regression  
=====
```

Selected variables: E_R I_R CSA_R CSF_R ND_R
Clustering model: LCA model with 4 latent classes
Covariates: EDAD.rec1 + contador + trabaja + Abandono_cat

Summary:

```
Best = -23144.32  
Mean = -23144.32  
Median = -23144.32  
Min = -23144.32
```

ELECCIÓN DE VARIABLES MANIFESTAS REDUNDANTES CON VARIABLES EXPLICATIVAS

```
mod7<-LCAvarsel(lasseis, G=1:6, X=covariables2, search="backward")
> mod7
```

Variable selection for Latent Class Analysis

```
=====
Stepwise backward/forward greedy search
with conditional dependence via regression
=====
```

```
Selected variables: E_R I_R CSA_R CSF_R ND_R
Clustering model: LCA model with 4 latent classes
Covariates: EDAD.rec1 + contador + trabaja + Abandono_cat
```

Steps:

Step	Proposed variable	BICdiff	Decision
1	Remove VP_R	96.28478	Accepted
2	Remove CSF_R	-53.42109	Rejected
3	Add VP_R	-96.28489	Rejected

ELECCIÓN DE VARIABLES MANIFESTAS REDUNDANTES CON VARIABLES EXPLICATIVAS

```
mod6<-LCAvarsel(lasseis, G=1:6, X=covariables2, search="forward")
> mod6
```

Variable selection for Latent Class Analysis

```
=====
Stepwise forward/backward greedy search
with conditional dependence via regression
=====
```

```
Selected variables: E_R I_R CSA_R CSF_R
Clustering model: LCA model with 4 latent classes
Covariates: EDAD.rec1 + contador + trabaja + Abandono_cat
```

Steps:

Step	Proposed variable	BICdiff	Decision
1	Add	CSF_R 1500.987	Accepted
2	Remove	ND_R 3952.107	Accepted
3	Add	VP_R -4077.834	Rejected
4	Remove	I_R -4101.755	Rejected

REFERENCIAS BIBLIOGRÁFICAS I

- Agresti, A. (2013). *Categorical data analysis*. Wiley-Interscience, Hoboken, N.J.
- Alfaro García, M. E. y Caballero Domínguez, A. J. (2006). Indices nacionales de satisfacción: Una vista general. En *Congreso de Metodología de Encuestas*.
- Álvarez-Vaz, R. y Vernazza, E. (2017). Evaluación de un instrumento de medición del nivel de satisfacción estudiantil a través de la aplicación de modelos de ecuaciones estructurales. *Cuadernos del CIMBAGE*, (19):1–25.
- Álvarez-Vaz, R., Freira, D., Vernazza, E., y Alves, H. (2016). Can students' satisfaction indexes be applied the same way in different countries? *Int Rev Public Nonprofit Marketing*, 13(101).
- Alves, H. y Raposo, M. (2004). La medición de la satisfacción en la enseñanza universitaria: El ejemplo de la universidade da beira interior. *Int Rev Public Nonprofit Marketing*, 1(1):73–88.
- Alves, H. y Raposo, M. (2009). The measurement of the construct of the construct satisfaction in higher education. *The Service Industries Journal*, 29(2):203 – 218.
- Bandeen-Roche, K., Miglioretti, D. L., Zeger, S. L., y Rathouz, P. J. (1997). Latent variable regression for multiple discrete outcomes. *Journal of the American Statistical Association*, 92(440):1375–1386.

REFERENCIAS BIBLIOGRÁFICAS II

- Ciavolino, E. y Dahlgaard, J. J. (2007). Ecsi - customer satisfaction modelling and analysis: A case study. *Total Quality Management & Business Excellence*, 18(5):545–554.
- Dempster, A. P., Laird, N. M., y Rubin, D. B. (1977). Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society*, 39:1–38.
- Everitt, B. S. (1984). *An Introduction to Latent Variable Models*. Springer Netherlands, Dordrecht.
- Fop, M. y Murphy, T. B. (2017). *LCAvarsel: Variable selection for latent class analysis*.
- Fop, M., Smart, K., y Murphy, T. B. (2017). Variable selection for latent class analysis with application to low back pain diagnosis. *Annals of Applied Statistics*, 11:2085–2115.
- Lazarsfeld, P. (1950). The logical and mathematical foundations of latent structure analysis. *ISA Stouffer (ed.), Measurement and Prediction*, pp. 362–412.
- Leite, R. S. y Filho, C. G. (2007). Um estudo empirico da aplicaÃo do indice europeu de satisfaÃo de clientes (ecsi) no brasil. *Revista de Administracao Mackenzie*, 8(4):178–200.
- Linzer, D. A. y Lewis, J. B. (2011). polCA: An R package for polytomous variable latent class analysis. *Journal of Statistical Software*, 42(10):1–29.

REFERENCIAS BIBLIOGRÁFICAS III

- Lopes, H. E. G., Pereira, C. C. d. P., y Vieira, A. F. S. (2009). Comparação entre os modelos norteamericano (acsi) e europeu (ecsi) de satisfação do cliente: Um estudo no setor de serviços. *Revista de Administração Mackenzie*, 10(1):161–187.
- Martensen, A. and Eskildsen, J., Gronholdt, L., y Kristensen, K. (2000). Measuring student oriented quality in higher education: application of the ecsi methodology. *Sinergie Rapporti di ricerca*, 9.
- Oberski, D. (2013). Conditional design effects for structural equation model estimates. En Institute, I. S., editor, *Proceedings of the 59th World Statistics Congress 2013*, Hong Kong.
- R Core Team (2019). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.