

Definición y origen de la Variable Latente: Cambio de preguntas para el cambio de prácticas

Seminario de Psicometría y Variables Latentes

González, T. y San Martín, E.

¹LIES Laboratorio Interdisciplinario de Estadística Social, Pontificia Universidad Católica de Chile, Chile

²Escuela de Psicología, Pontificia Universidad Católica de Chile, Chile

29 de enero de 2018



Laboratorio
Interdisciplinario de
Estadística Social

- 1 Introducción
- 2 Preguntas acerca de la VL
- 3 Lectura de Spearman
- 4 Propuesta para la construcción de conceptos

Modelos de
Variables Latentes

Educación

Efectividad docente
Efectividad escolar
Habilidad de los estudiantes

Psicología

Identificación (definición)
de constructos

MVA
IRT
AFE
AFC
SEM
(...)

Los modelos de variables latentes tienen una estructura común, caracterizada por dos componentes:

- Modelo latente (o modelo estructural): especificación de la distribución de la variable latente o no observada.

$$P(\theta)$$

- Modelo condicional (o modelo de medición): especificación de la distribución condicional de los observables, dada la variable latente.

$$P(Y|\theta)$$

Introducción - Ilustración de la Estructura común

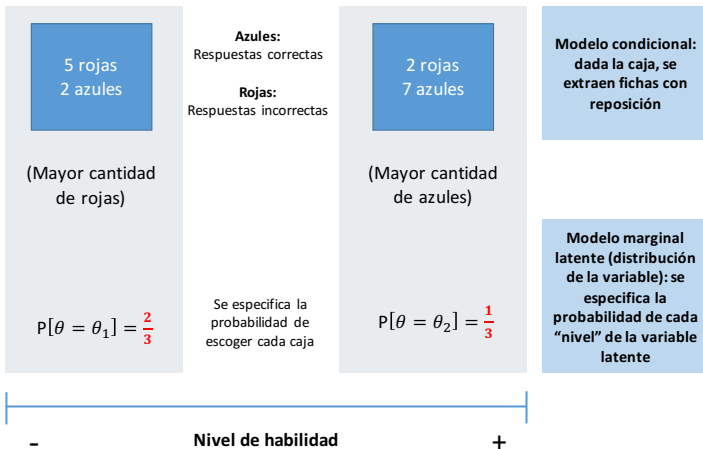


Ilustración - Cómo extraer los observados

- Determinar un criterio: se escoge la caja lanzando el dado
 - Para seleccionar la caja 1: si el dado muestra las caras 1, 2, 3 o 4.
 - Para seleccionar la caja 2: si el dado muestra las caras 5 o 6.

- La extracción es con reposición

¿Qué asegura la reposición?

La independencia entre las extracciones. Esto corresponde al **Axioma de Independencia local**: asumimos esto, debido a que sustantivamente, creemos que las observaciones de cada caja están relacionadas entre sí, debido a que cada una de las extracciones proviene de la misma caja.

Ilustración del uso común de los Modelos con Variables Latentes

- ¿Cuál es la caja más probable de la que viene el patrón de respuesta del estudiante 1, y del estudiante 2?
- El problema consiste en que solo tenemos acceso a lo que observamos: no sabemos en qué caja debemos posicionar a cada estudiante.

Este es el problema del psicometra, y que en principio, todo modelo de variable latente responde.

- Patrón de respuestas
- Cálculo probabilidad condicional
- Teorema de probabilidades totales
- Comparamos la probabilidad de estar en una caja con la probabilidad de estar en la otra.

$$\begin{aligned} P(C1 | A, A, R) & \stackrel{\text{Prob. Cond.}}{=} \frac{P(A, A, R | C1)P(C1)}{P(A, A, R)} \\ & \stackrel{\text{TPT}}{=} \frac{P(A, A, R | C1)P(C1)}{P(A, A, R | C1)P(C1) + P(A, A, R | C2)P(C2)} \\ & = 0,44 \end{aligned}$$

$$\begin{aligned} P(C2 | A, A, R) & \stackrel{\text{Prob. Cond.}}{=} \frac{P(A, A, R | C2)P(C2)}{P(A, A, R)} \\ & \stackrel{\text{TPT}}{=} \frac{P(A, A, R | C1)P(C1)}{P(A, A, R | C1)P(C1) + P(A, A, R | C2)P(C2)} \\ & = 0,56 \end{aligned}$$

Conclusión: mayor probabilidad de pertenecer a la Caja 2

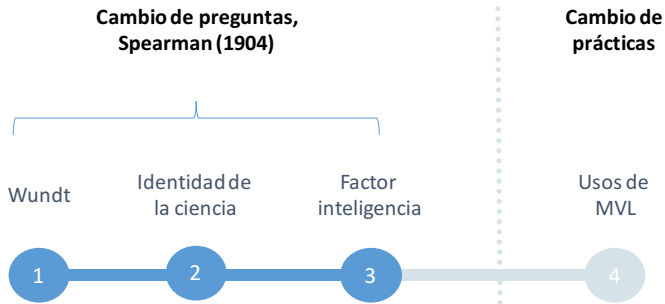
Primer argumento

- Modelos de Variables Latentes buscan clasificar

Segundo argumento

Tercer argumento

Cuarto argumento



Estatuto teórico de las Variables Latentes

¿De dónde vienen? ¿qué son? ¿cómo se operacionalizan?

Consenso:

Spearman (1904)

Primer modelo de análisis factorial unidimensional para identificar el factor de inteligencia G.

*“Hace treinta años el Profesor Charles E. Spearman introdujo el problema del factor en psicología cuando observó que las intercorrelaciones de un conjunto de tests revelaban un orden subyacente. Él interpretó este orden como el **efecto** de un factor conspicuo que era común a todos los tests” (Thurstone, 1935).*

Interpretación de la VL

- Pero Spearman (1904) no usa la terminología de **latente**.
- Se comienza a utilizar con Reiersol (1950) y Lazarsfeld (1950).

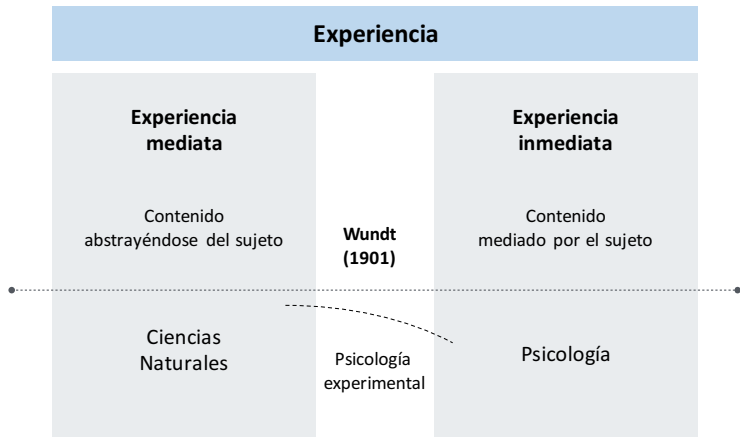
Google Books Ngram Viewer

Graph these comma-separated phrases: case-insensitive

between and from the corpus with smoothing of [Search lots of books](#)



- Marco teórico basado en Wundt
- Análisis de la identidad de la ciencia
- Cuestionamiento de las metodologías: ¿correlación?
- Propuesta: análisis factorial exploratorio para la generación del concepto de Inteligencia



Lectura de Spearman - Crisis de la Psicología

Discípulos de Wundt usaban el método de introspección. Este método asumía que la experiencia interna era distinta a la experiencia externa.

Lectura de Spearman - Identidades de la ciencia



Primer argumento

- Modelos de Variables Latentes buscan clasificar

Segundo argumento

- Con Spearman:
Los factores latentes son Invariantes

Tercer argumento

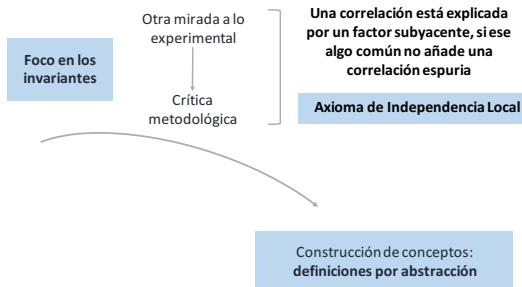
Cuarto argumento

Lectura de Spearman - Cuestionamiento metodológico

Dada su orientación a los invariantes, volvió sobre los resultados experimentales, y en lugar de aceptar o rechazar conclusiones, hizo una crítica metodológica.

Lectura de Spearman - Cuestionamiento metodológico

- Es necesario **eliminar los factores irrelevantes**, pues producen **correlaciones espurias**, y por tanto hay que utilizar métodos que permitan eliminar esas correlaciones espurias.
- Supone entonces que el factor común a dos variables es aquel que hace que la componente debida a lo espurio, se anule. Propone la noción de **correlación parcial, símil al Axioma de Independencia Local**.
- Lo invariante (el factor común) representa similitudes entre objetos agrupados.



Primer argumento

- Modelos de Variables Latentes buscan clasificar

Segundo argumento

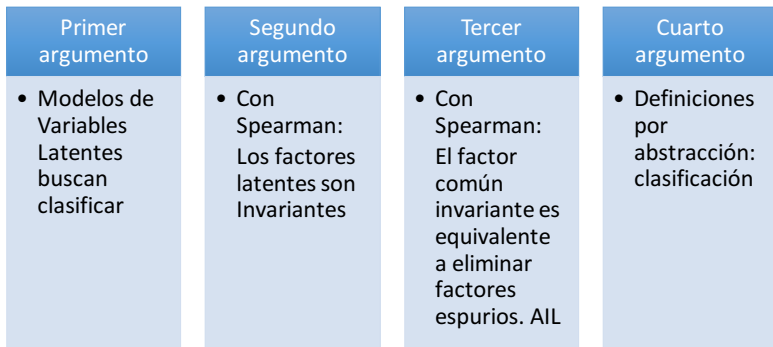
- Con Spearman:
Los factores latentes son Invariantes

Tercer argumento

- Persistencia de la sustancia relacionado con lo invariante. Lo invariante se relaciona con la construcción de conceptos (DPA).

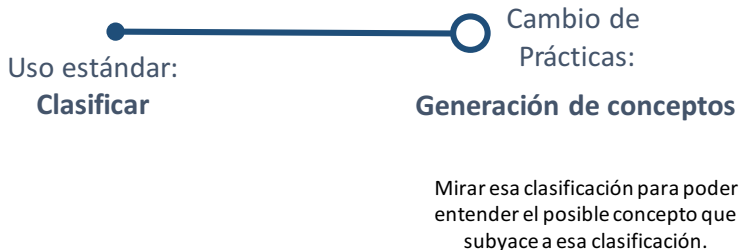
Cuarto argumento

- Construcción de nuevos objetos a partir de objetos dados.
- Objetos distintos que tienen una **similitud** en algo: hacemos abstracción de las diferencias pero nos quedamos con una característica común a los objetos.
- Las definiciones por abstracción son definiciones de conceptos por medio de la **clasificación de objetos**, donde la clasificación representa algo común (las similitudes).
- Las variables latentes tienen que ver con **invariantes** y por lo tanto con procesos de definición por abstracción: los invariantes al interior de cada grupo, son invariantes porque tomo un objeto de ese grupo y éste representa al grupo entero.



¿Cuál es el uso estándar de los MVL

- Se observan **realizaciones de una determinada variable dependiente** y se pregunta por la probabilidad que el individuo que generó esa variable dependiente, tenga un determinado **nivel en la variable latente**.
- Esto conlleva a **clasificar** todos los individuos, según su nivel de la variable latente.





Modelos de análisis de efectividad escolar

Se clasifica a las escuelas de acuerdo a la predicción de la variable latente que representa el efecto escuela (puntajes en Simce pre-post, por ejemplo).

¿Qué es una escuela efectiva? Es una escuela cuyo nivel de la variable latente es muy alto ¿qué significa esto?

La propuesta: con estos valores se clasificaron escuelas, y por lo tanto, para generar el concepto “escuela efectiva”, es necesario ir a mirar los grupos y ver sus aspectos comunes.

- Los modelos de variables latentes son instrumentos que constituyen un **primer paso** en la construcción de conceptos, mientras que comúnmente, se utiliza como un último paso.
- Con la lectura de Spearman, en lugar de preguntarnos por el **estatuto** de las variables latentes, nos preguntamos por su **uso**.
- El uso está relacionado con la construcción de conceptos, por lo tanto, cuando se usa a nivel de **política pública**, lo ético es intentar definir y comprender los conceptos que subyacen a los grupos que se clasifican por medio de estas herramientas.

- La relevancia de esta propuesta consiste en que la revisión de los orígenes nos ayuda a comprender supuestos subyacentes que usualmente son aprendidos pero quizás no comprendidos y en investigación, pocas veces **transparentados**.

- Borsboom, D., Mellenbergh, G. J., & Van Heerden, J. (2003). The theoretical status of latent variables. *Psychological Review* 110(2), 203-219.
- Kim, K. (2011). A Strengthening of the Caesar Problem. *Erkenntnis* 74, 123-136.
- Koopmans, T. C., & Reiersol, O. (1950). The identification of structural characteristics. *The Annals of Mathematical Statistics* 21(2), 165-181.†
- Lorenzen, P. (1987). *Constructive Philosophy*, Amherst: The University of Massachusetts Press.†
- Spearman, C. (1904). "General Intelligence," Objectively Determined and Measured. *The American Journal of Psychology* 15(2), 201-292.