

TRAYECTORIA NUTRICIONAL Y DESEMPEÑO ESCOLAR

Aplicación de modelos conjuntos para datos longitudinales y de sobrevivencia

Alejandra Marroig¹

¹IECON - IESTA - FCEA

Tesis de Maestría presentada al Programa de Posgrado en Ingeniería Matemática

DT: Graciela Muniz
Universidad de Edinburgo

Esquema de la presentación

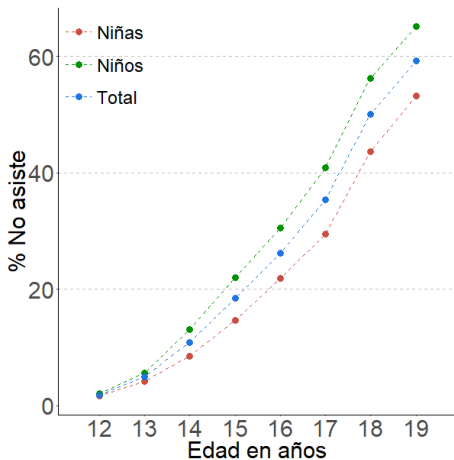
- 1 Contexto y motivación
- 2 Objetivos
- 3 Metodología
- 4 Resultados
- 5 Conclusiones

Estado nutricional y desempeño escolar

- Analizar el estado nutricional es relevante por sí mismo y por otras dimensiones (e.g. educación, salud), particularmente en la infancia.
- En Uruguay hay estudios de los determinantes del estado nutricional en la niñez y el desempeño escolar con enfoque *cross-section*.
- El déficit nutricional se asocia con peor desempeño escolar.
- No existen estudios previos que analicen trayectoria nutricional de niños y niñas de Uruguay en la escuela incorporando la **evolución con la edad**.
 - ▶ Stuart & Panico (2016): perfiles socio económicos de grupos según trayectorias nutricionales

¿Por qué analizar edad de abandono en Uruguay?

Porcentaje no asistencia al sistema educativo



Motivación metodológica

- Aplicación de **metodología moderna** a problema relevante en Uruguay.
 - ▶ LCMM y JLCMM
- Explotar información longitudinal e incorporar el **cambio temporal**.
 - ▶ Trayectoria nutricional en período escolar.

Objetivos

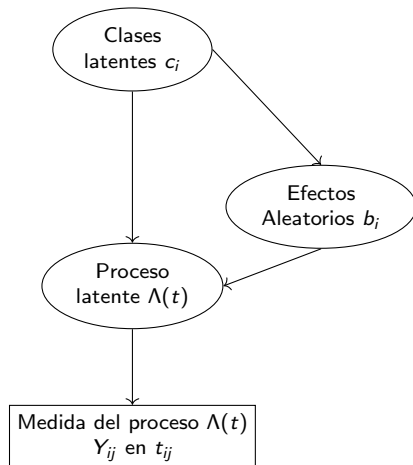
- Analizar trayectoria nutricional de niños y niñas de Uruguay mediante LCMM.
- Caracterizar grupos según algunos factores asociados al estado nutricional.
- Indagar en relación de trayectoria con abandono escolar mediante JLCMM.

Modelos longitudinales y de supervivencia

- Modelo Mixto de Clases Latentes: LCMM
- Modelo Mixto Conjunto de Clases Latentes: JLCMM

LCMM

- Trayectorias proceso latente heterogéneas.
- Heterogeneidad capturada por clases latentes.
- Ninguna covariable observada captura heterogeneidad.
 - ▶ **Clases latentes**
- Efectos aleatorios heterogeneidad *intra* clases.



Modelo Mixto de Clases Latentes LCMM

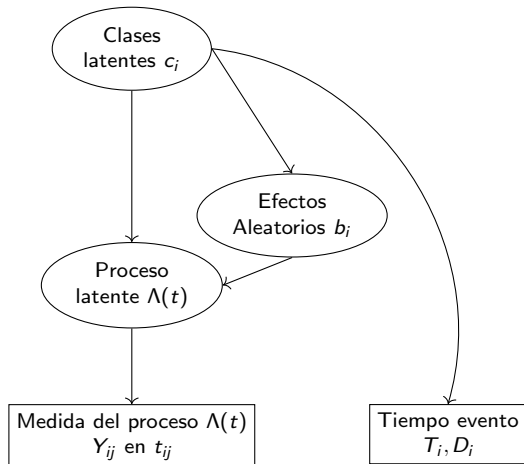
$$\left\{ \begin{array}{l} \pi_{ig} = P(c_i = g | \mathbf{X}_{Ci}) = \frac{e^{\mathbf{x}_{Ci}^\top \boldsymbol{\lambda}_g}}{\sum_{l=1}^G e^{\mathbf{x}_{Ci}^\top \boldsymbol{\lambda}_l}} \\ Y_{ij} |_{c_i=g} = H(\mathbf{X}_{Li}(t_{ij})^\top \boldsymbol{\beta}_g + \mathbf{Z}_i(t_{ij})^\top \mathbf{b}_{ig} + \omega_i(t_{ij}) + \varepsilon_{ij}) \end{array} \right.$$

$\mathbf{b}_{ig} \sim N_q(\mathbf{0}, \sigma_g^2 \mathbf{D})$ y $\varepsilon_{ij} \sim N(0, \sigma_\varepsilon^2)$ independientes

$$Y_{ij} = H(\tilde{Y}_{ij}) = H(\Lambda_i(t_{ij}) + \varepsilon_{ij})$$

JLCMM

- Proceso latente asociado con supervivencia.
- Ninguna variable observada captura heterogeneidad.
- Cada clase latente
 - ▶ modelo de proceso latente
 - ▶ modelo de supervivencia
- Efectos aleatorios heterogeneidad *intra* clases.
- Asociación proceso latente y evento capturada por clases latentes.



Modelo Mixto Conjunto de Clases Latentes JLCMM

$$\left\{ \begin{array}{l} \pi_{ig} = P(c_i = g | \mathbf{X}_{Ci}) = \frac{e^{\mathbf{x}_{Ci}^\top \boldsymbol{\lambda}_g}}{\sum_{l=1}^G e^{\mathbf{x}_{Ci}^\top \boldsymbol{\lambda}_l}} \\ Y_{ij} | c_i = g = \mathbf{X}_{Li}(t_{ij})^\top \boldsymbol{\beta}_g + \mathbf{Z}_i(t_{ij})^\top \mathbf{b}_{ig} + \omega_i(t_{ij}) + \varepsilon_{ij} \\ h_i(t | c_i = g) = h_{0g}(t) e^{\mathbf{x}_{Si}^\top \boldsymbol{\delta}_g} \end{array} \right.$$

$\mathbf{b}_{ig} \sim N_q(\mathbf{0}, \sigma_g^2 \mathbf{D})$ y $\varepsilon_{ij} \sim N(0, \sigma_\varepsilon^2)$ independientes

$h_{0g}(t)$ riesgo de línea de base de clase g

Estimación - Máxima Verosimilitud (MLE)

θ_G el vector de parámetros, con G conocido. La función a maximizar es:

$$\ell(\theta_G) = \sum_{i=1}^N \log(L_i(\theta_G))$$

donde L_i es la contribución individual

LCMM

$$L_i(\theta_G) = \sum_{g=1}^G \left\{ \pi_{ig} \phi_{ig}(\tilde{Y}_i | c_i = g; \theta_G) \prod_{j=1}^{n_i} J(H^{-1}(Y_{ij} | c_i = g; \theta_G)) \right\}$$

donde $Y_{ij} = H(\tilde{Y}_{ij}) = H(\Lambda_i(t_{ij}) + \varepsilon_{ij})$

JLCMM

$$L_i(\theta_G) = \sum_{g=1}^G \left\{ \pi_{ig} \phi_{ig}(\mathbf{Y}_i | c_i = g; \theta_G) h_i(T_i | c_i = g; \theta_G)^{D_i} S_i(T_i | c_i = g; \theta_G) \right\}$$

Análisis de la Trayectoria Nutricional

$$\left\{ \begin{array}{l} \pi_{ig} = P(c_i = g | X_{Ci}) = \frac{e^{\lambda_g}}{\sum_{l=1}^G e^{\lambda_l}} \\ \text{IMC}_{ij|c_i=g} = H(\beta_{0g} + \beta_{1g}\text{Edad}_{ij} + \beta_{2g}\text{Edad}_{ij}^2 + b_{0ig} + b_{1ig}\text{Edad}_{ij} + \varepsilon_{ij}) \end{array} \right. \quad (1)$$

Relación entre Trayectoria Nutricional y Desempeño Escolar

$$\left\{ \begin{array}{l} \pi_{ig} = P(c_i = g | X_{Ci}) = \frac{e^{\lambda_g}}{\sum_{l=1}^G e^{\lambda_l}} \\ \text{IMC}_{ij|c_i=g} = \beta_{0g} + \beta_{1g}\text{Edad}_{ij} + \beta_{2g}\text{Edad}_{ij}^2 + b_{0ig} + b_{1ig}\text{Edad}_{ij} + \varepsilon_{ij} \\ h_i(t|c_i = g) = h_0(t)e^{\{\gamma_{0g} + \gamma_{1g}\text{Repitio}_i\}} \end{array} \right. \quad (2)$$

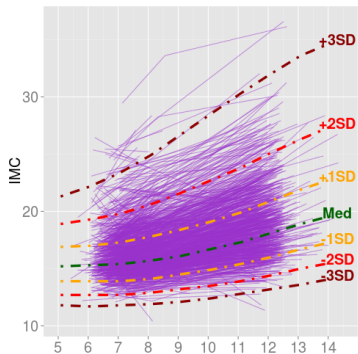
Base de datos

ELBU

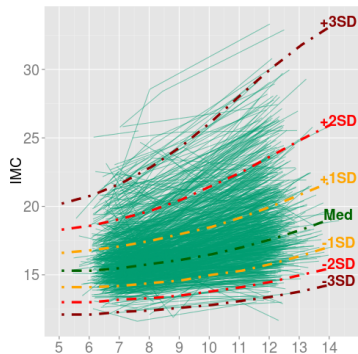
- Muestra niños y niñas primer año en escuelas públicas -85 %- en 2004
- Capitales departamentales Artigas, Canelones, Colonia, Florida, Montevideo, Paysandú, Rivera y Soriano
- Truncamiento en los estratos socioeconómicos más altos, 43% del quintil más alto en escuela pública en 2004
- Medidas antropométricas en 2004, 2005 (área metropolitana) y 2009
- Encuestas a los hogares en 2004, 2006 (área metropolitana) y 2011/12

Trayectoria del IMC para la edad en años

Niñas



Niños



Distribución en años de medición	Niñas	Niños
Medidas en 2004	1341	1432
Medidas en 2005	518	525
Medidas en 2009	805	835
Total participantes ELBU	1570	1691
Total muestra LCMM	1392	1492
Total muestra JLCMM	971	1026

Tasa de abandono según repetición por género

	No repitió		Repitió		Total	
	% Abandono	N	% Abandono	N	% Abandono	N
Niñas	2.4%	681	16.6%	359	7.3%	1040
Niños	3.8%	605	19.6%	502	10.9%	1107
Total	3.1%	1286	18.4%	861	9.1%	2147

Distribución según edad de abandono por género

Edad	11 años	12 años	13 años	14 años	N Abandono	N
Niñas	5.5%	17.3%	37.3%	34.6%	75	1040
Niños	1.7%	19.5%	35.6%	40.7%	118	1107
Total	3.1%	18.7%	36.3%	38.3%	193	2147

Tasa de abandono según repetición por género

	No repitió		Repitió		Total	
	% Abandono	N	% Abandono	N	% Abandono	N
Niñas	2.4%	681	16.6%	359	7.3%	1040
Niños	3.8%	605	19.6%	502	10.9%	1107
Total	3.1%	1286	18.4%	861	9.1%	2147

Distribución según edad de abandono por género

Edad	11 años	12 años	13 años	14 años	N Abandono	N
Niñas	5.5%	17.3%	37.3%	34.6%	75	1040
Niños	1.7%	19.5%	35.6%	40.7%	118	1107
Total	3.1%	18.7%	36.3%	38.3%	193	2147

Tasa de abandono según repetición por género

	No repitió		Repitió		Total	
	% Abandono	N	% Abandono	N	% Abandono	N
Niñas	2.4%	681	16.6%	359	7.3%	1040
Niños	3.8%	605	19.6%	502	10.9%	1107
Total	3.1%	1286	18.4%	861	9.1%	2147

Distribución según edad de abandono por género

Edad	11 años	12 años	13 años	14 años	N Abandono	N
Niñas	5.5%	17.3%	37.3%	34.6%	75	1040
Niños	1.7%	19.5%	35.6%	40.7%	118	1107
Total	3.1%	18.7%	36.3%	38.3%	193	2147

- Trayectoria nutricional

Grupos para niñas

Estadísticos y clasificación según grupos de niñas

Grupos	loglik	BIC	Entropy	NP	% Niñas en cada grupo				
					G 1	G 2	G 3	G 4	G 5
1	-5775.9	11602.5		7	100%				
2	-5524.2	11135.4	0.644	12	86%	14%			
3	-5471.4	11065.9	0.513	17	49%	42%	9%		
4	-5457.3	11073.9	0.564	22	57%	32%	8%	3%	
5	-5409.2	11013.8	0.538	27	51%	34%	8%	5%	2%

N=1392

Probabilidades *a posteriori* tres grupos de niñas - Enlace lineal

	(a) Probabilidades medias			(b) Prob. >umbral		
	P(G 1)	P(G 2)	P(G 3)	>0.7	>0.8	>0.9
G 1	0.745	0.239	0.016	69%	37%	2%
G 2	0.134	0.785	0.081	70%	55%	27%
G 3	0.005	0.138	0.857	78%	69%	56%

N=1392

Grupos para niñas

Estadísticos y clasificación según grupos de niñas

Grupos	loglik	BIC	Entropy	NP	% Niñas en cada grupo				
					G 1	G 2	G 3	G 4	G 5
1	-5775.9	11602.5		7	100%				
2	-5524.2	11135.4	0.644	12	86%	14%			
3	-5471.4	11065.9	0.513	17	49%	42%	9%		
4	-5457.3	11073.9	0.564	22	57%	32%	8%	3%	
5	-5409.2	11013.8	0.538	27	51%	34%	8%	5%	2%

N=1392

Probabilidades *a posteriori* tres grupos de niñas - Enlace lineal

	(a) Probabilidades medias			(b) Prob. >umbral		
	P(G 1)	P(G 2)	P(G 3)	>0.7	>0.8	>0.9
G 1	0.745	0.239	0.016	69%	37%	2%
G 2	0.134	0.785	0.081	70%	55%	27%
G 3	0.005	0.138	0.857	78%	69%	56%

N=1392

Grupos para niñas

Estadísticos y clasificación según grupos de niñas

Grupos	loglik	BIC	Entropy	NP	% Niñas en cada grupo				
					G 1	G 2	G 3	G 4	G 5
1	-5775.9	11602.5		7	100%				
2	-5524.2	11135.4	0.644	12	86%	14%			
3	-5471.4	11065.9	0.513	17	49%	42%	9%		
4	-5457.3	11073.9	0.564	22	57%	32%	8%	3%	
5	-5409.2	11013.8	0.538	27	51%	34%	8%	5%	2%

N=1392

Probabilidades *a posteriori* tres grupos de niñas - Enlace lineal

	(a) Probabilidades medias			(b) Prob. >umbral		
	P(G 1)	P(G 2)	P(G 3)	>0.7	>0.8	>0.9
G 1	0.745	0.239	0.016	69%	37%	2%
G 2	0.134	0.785	0.081	70%	55%	27%
G 3	0.005	0.138	0.857	78%	69%	56%

N=1392

Grupos para niños

Estadísticos y clasificación según grupos de niños - Enlace lineal

G	loglik	BIC	Entropy	NP	% Niños en cada grupo				
					G 1	G 2	G 3	G 4	G 5
1	-5915.3	11881.7		7	100%				
2	-5622.6	11332.8	0.535	12	79%	21%			
3	-5530.2	11184.7	0.556	17	71%	21%	8%		
4	-5511.9	11184.7	0.456	22	41%	34%	17%	8%	
5	-5446.4	11090.1	0.516	27	43%	28%	20%	8%	1%

N=1492

Probabilidades *a posteriori* para tres grupos de niños - Enlace lineal

	(a) Probabilidades medias			(b) Prob. >umbral		
	P(G 1)	P(G 2)	P(G 3)	>0.7	>0.8	>0.9
G 1	0.823	0.022	0.154	85%	66%	33%
G 2	0.018	0.826	0.155	69%	60%	47%
G 3	0.146	0.112	0.742	56%	40%	24%

N=1492

Grupos para niños

Estadísticos y clasificación según grupos de niños - Enlace lineal

G	loglik	BIC	Entropy	NP	% Niños en cada grupo				
					G 1	G 2	G 3	G 4	G 5
1	-5915.3	11881.7		7	100%				
2	-5622.6	11332.8	0.535	12	79%	21%			
3	-5530.2	11184.7	0.556	17	71%	21%	8%		
4	-5511.9	11184.7	0.456	22	41%	34%	17%	8%	
5	-5446.4	11090.1	0.516	27	43%	28%	20%	8%	1%

N=1492

Probabilidades *a posteriori* para tres grupos de niños - Enlace lineal

	(a) Probabilidades medias			(b) Prob. >umbral		
	P(G 1)	P(G 2)	P(G 3)	>0.7	>0.8	>0.9
G 1	0.823	0.022	0.154	85%	66%	33%
G 2	0.018	0.826	0.155	69%	60%	47%
G 3	0.146	0.112	0.742	56%	40%	24%

N=1492

Grupos para niños

Estadísticos y clasificación según grupos de niños - Enlace lineal

G	loglik	BIC	Entropy	NP	% Niños en cada grupo				
					G 1	G 2	G 3	G 4	G 5
1	-5915.3	11881.7		7	100%				
2	-5622.6	11332.8	0.535	12	79%	21%			
3	-5530.2	11184.7	0.556	17	71%	21%	8%		
4	-5511.9	11184.7	0.456	22	41%	34%	17%	8%	
5	-5446.4	11090.1	0.516	27	43%	28%	20%	8%	1%

N=1492

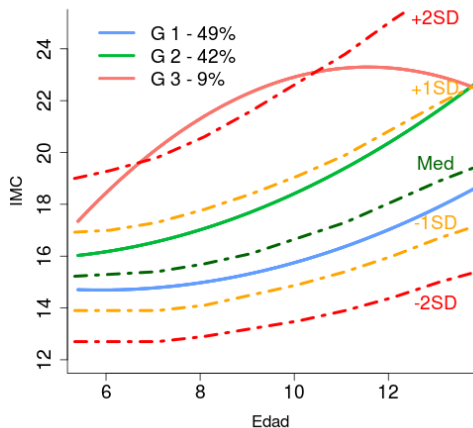
Probabilidades *a posteriori* para tres grupos de niños - Enlace lineal

	(a) Probabilidades medias			(b) Prob. >umbral		
	P(G 1)	P(G 2)	P(G 3)	>0.7	>0.8	>0.9
G 1	0.823	0.022	0.154	85%	66%	33%
G 2	0.018	0.826	0.155	69%	60%	47%
G 3	0.146	0.112	0.742	56%	40%	24%

N=1492

Tres grupos de niñas

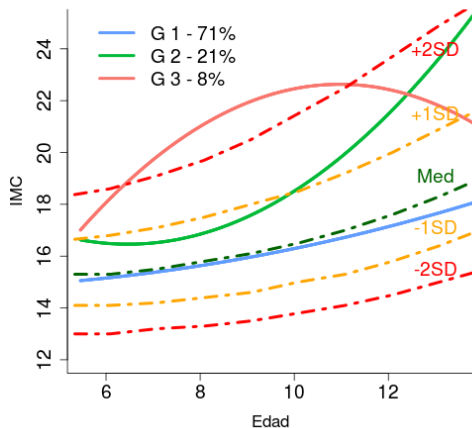
Trayectoria nutricional media predicha para tres grupos de niñas



- Diferencias en peso y talla al nacer
- No hay diferencias en nivel educativo materno

Tres grupos de niños

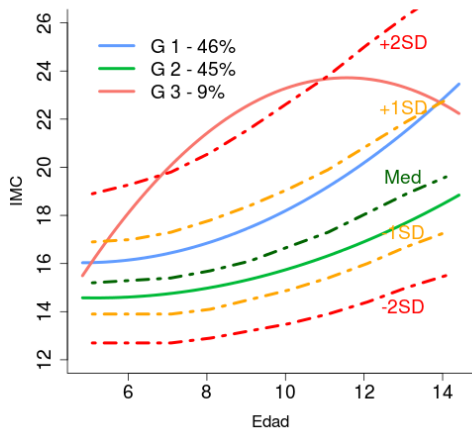
Trayectoria nutricional media predicha para tres grupos de niños



- Diferencias en peso al nacer y nivel educativo materno
- No hay diferencias en talla al nacer

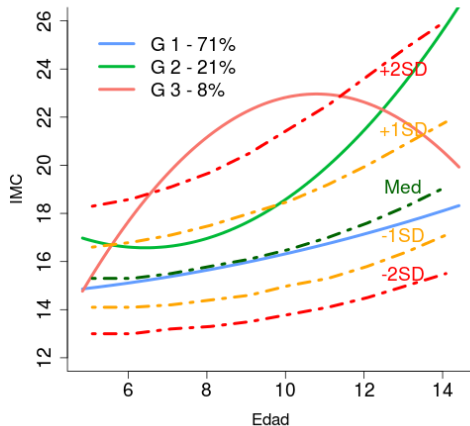
- Asociación entre trayectoria nutricional y desempeño escolar

Trayectoria nutricional media para tres grupos de niñas



- No hay diferencias entre grupos en la probabilidad de abandonar

Trayectoria nutricional media para tres grupos de niños



- No hay diferencias entre grupos en la probabilidad de abandonar

1 Contexto y motivación

2 Objetivos

3 Metodología

4 Resultados

5 Conclusiones

Conclusiones

- Se aplicó una técnica estadística moderna a problema de relevancia en Uruguay → Nuevas interrogantes respecto a:
 - ▶ trayectoria nutricional
 - ▶ desempeño escolar
- Análisis de trayectoria nutricional muestra tres grupos de niños y niñas con distinta evolución en edad escolar → puede tener consecuencias sobre otras dimensiones.

Tres grupos de niñas

Grupo	Niñas (%)	Trayectoria media	Peso	Talla	Edu
Normal-bajo	50%	Normal, debajo mediana			
Normal-alto	40%	Encima mediana y sobrepeso	✓	✓	X
Sobrepeso	10%	Sobrepeso y obesidad			

Tres grupos de niños

Grupo	Niños (%)	Trayectoria media	Peso	Talla	Edu
Normal	70%	Normal, entorno mediana			
Sobrepeso	20%	Encima mediana y sobrepeso	✓	X	✓
Obesidad	10%	Obesidad			

Asociación trayectoria nutricional y desempeño escolar

- Trayectorias nutricionales similares en LCMM y JLCMM
- Pertenencia a grupos no modifica probabilidad de abandono con la edad

Limitaciones

- Elección del número de grupos *a posteriori*.
- Rango de edad considerado.
- Medidas longitudinales máximas por individuo.

Trabajos a futuro

- Análisis de atípicos.
- Interior de grupos y caracterización de grupos.
- Incorporar cuarta ronda de ELBU.