

Aprendizaje de Máquinas y Economía

Alvaro J. Riascos Villegas
Universidad de los Andes y Quantil

Junio 2021

Contenido

- 1 La Contribución a la Economía
- 2 Causalidad y Predicción
 - Problemas Entrelazados
- 3 Problemas de Predicción
 - Ajuste de Riesgo
 - Problemas
 - Resultados
- 4 Referencias

Principales temas

- Nuevos datos (administrativos y corporativos): volúmen, variedad (imágenes, texto, etc.) y velocidad.
- New markets (strategic algorithmic interaction).
- Predicción vs. causalidad.
- Metodológico: validación cruzada, regularización, compromiso sesgo vs. varianza, etc.
- Poner a prueba teorías.
- Nuevos problemas: sesgos, discriminación, interpretabilidad, etc.

Contenido

- 1 La Contribución a la Economía
- 2 Causalidad y Predicción
 - Problemas Entrelazados
- 3 Problemas de Predicción
 - Ajuste de Riesgo
 - Problemas
 - Resultados
- 4 Referencias

Ejemplos

- Modelo de predicción de la temperatura de una ciudad dada la altura sobre el nivel del mar.
- Modelo de predicción del crimen: consumo de helados y crimen.
- Modelo de ocupación hotelera: correlación positiva entre precio y ocupación.
- Consumo de chocolates y premios Nobel.

Ejemplos

- Modelo de predicción de la temperatura de una ciudad dada la altura sobre el nivel del mar.
- Modelo de predicción del crimen: consumo de helados y crimen.
- Modelo de ocupación hotelera: correlación positiva entre precio y ocupación.
- Consumo de chocolates y premios Nobel.

Ejemplos

- Modelo de predicción de la temperatura de una ciudad dada la altura sobre el nivel del mar.
- Modelo de predicción del crimen: consumo de helados y crimen.
- Modelo de ocupación hotelera: correlación positiva entre precio y ocupación.
- Consumo de chocolates y premios Nobel.

Ejemplos

- Modelo de predicción de la temperatura de una ciudad dada la altura sobre el nivel del mar.
- Modelo de predicción del crimen: consumo de helados y crimen.
- Modelo de ocupación hotelera: correlación positiva entre precio y ocupación.
- Consumo de chocolates y premios Nobel.

Consumo de chocolates y premios Nobel

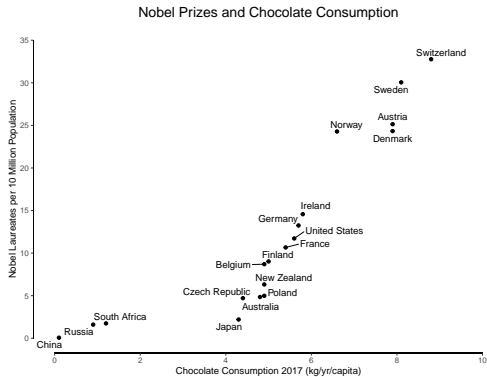


Figure 1. Shows the relationship between chocolate consumption and the number of Nobel Laureates per country.

Introducción

- Kleinberg, Ludwig, Mullainathan, Obermeyer. 2015. Prediction Policy Problems. AER.
- Este artículo introduce un marco conceptual muy sencillo para entender la relación entre ambos problemas.
- Muchos problemas económicos o de política pública no requieren la identificación de efectos causales.

Introducción

- Kleinberg, Ludwig, Mullainathan, Obermeyer. 2015. Prediction Policy Problems. AER.
- Este artículo introduce un marco conceptual muy sencillo para entender la relación entre ambos problemas.
- Muchos problemas económicos o de política pública no requieren la identificación de efectos causales.

Introducción

- Kleinberg, Ludwig, Mullainathan, Obermeyer. 2015. Prediction Policy Problems. AER.
- Este artículo introduce un marco conceptual muy sencillo para entender la relación entre ambos problemas.
- Muchos problemas económicos o de política pública no requieren la identificación de efectos causales.

- Sea y el resultado de interés. Depende de x_0 , que es exógena (e.g., una variable de política) y x . Para hacer explícita esta dependencia escribimos $y(x_0, x)$
- Queremos maximizar una función objetivo conocida $\Pi(x_0, y)$.
- El problema de decisión depende de:

$$\frac{d\Pi(x_0, y(x_0, x))}{dx_0} = \frac{\partial \Pi}{\partial x_0}(x_0, y) + \frac{\partial \Pi}{\partial y}(x_0, y) \frac{\partial y}{\partial x_0}(x_0, y)$$

- 1 Primer término: El efecto de x_0 en Π (evaluado en (x_0, y) un problema de predicción).
 - 2 Segundo término: El efecto de x_0 en y (un problema de causalidad).
- Obsérvese que ambos términos dependen de y .
 - El tomador de decisiones debe resolver ambos problemas (i.e., problemas entrelazados).

- Sea y el resultado de interés. Depende de x_0 , que es exógena (e.g., una variable de política) y x . Para hacer explícita esta dependencia escribimos $y(x_0, x)$
- Queremos maximizar una función objetivo conocida $\Pi(x_0, y)$.
- El problema de decisión depende de:

$$\frac{d\Pi(x_0, y(x_0, x))}{dx_0} = \frac{\partial \Pi}{\partial x_0}(x_0, y) + \frac{\partial \Pi}{\partial y}(x_0, y) \frac{\partial y}{\partial x_0}(x_0, y)$$

- 1 Primer término: El efecto de x_0 en Π (evaluado en (x_0, y) un problema de predicción).
 - 2 Segundo término: El efecto de x_0 en y (un problema de causalidad).
- Obsérvese que ambos términos dependen de y .
 - El tomador de decisiones debe resolver ambos problemas (i.e., problemas entrelazados).

- Sea y el resultado de interés. Depende de x_0 , que es exógena (e.g., una variable de política) y x . Para hacer explícita esta dependencia escribimos $y(x_0, x)$
- Queremos maximizar una función objetivo conocida $\Pi(x_0, y)$.
- El problema de decisión depende de:

$$\frac{d\Pi(x_0, y(x_0, x))}{dx_0} = \frac{\partial \Pi}{\partial x_0}(x_0, y) + \frac{\partial \Pi}{\partial y}(x_0, y) \frac{\partial y}{\partial x_0}(x_0, y)$$

- 1 Primer término: El efecto de x_0 en Π (evaluado en (x_0, y) un problema de predicción).
 - 2 Segundo término: El efecto de x_0 en y (un problema de causalidad).
- Obsérvese que ambos términos dependen de y .
 - El tomador de decisiones debe resolver ambos problemas (i.e., problemas entrelazados).

- Sea y el resultado de interés. Depende de x_0 , que es exógena (e.g., una variable de política) y x . Para hacer explícita esta dependencia escribimos $y(x_0, x)$
- Queremos maximizar una función objetivo conocida $\Pi(x_0, y)$.
- El problema de decisión depende de:

$$\frac{d\Pi(x_0, y(x_0, x))}{dx_0} = \frac{\partial \Pi}{\partial x_0}(x_0, y) + \frac{\partial \Pi}{\partial y}(x_0, y) \frac{\partial y}{\partial x_0}(x_0, y)$$

- 1 Primer término: El efecto de x_0 en Π (evaluado en (x_0, y) un problema de predicción).
 - 2 Segundo término: El efecto de x_0 en y (un problema de causalidad).
- Obsérvese que ambos términos dependen de y .
 - El tomador de decisiones debe resolver ambos problemas (i.e., problemas entrelazados).

- Sea y el resultado de interés. Depende de x_0 , que es exógena (e.g., una variable de política) y x . Para hacer explícita esta dependencia escribimos $y(x_0, x)$
- Queremos maximizar una función objetivo conocida $\Pi(x_0, y)$.
- El problema de decisión depende de:

$$\frac{d\Pi(x_0, y(x_0, x))}{dx_0} = \frac{\partial \Pi}{\partial x_0}(x_0, y) + \frac{\partial \Pi}{\partial y}(x_0, y) \frac{\partial y}{\partial x_0}(x_0, y)$$

- 1 Primer término: El efecto de x_0 en Π (evaluado en (x_0, y) un problema de predicción).
 - 2 Segundo término: El efecto de x_0 en y (un problema de causalidad).
- Obsérvese que ambos términos dependen de y .
 - El tomador de decisiones debe resolver ambos problemas (i.e., problemas entrelazados).

- Sea y el resultado de interés. Depende de x_0 , que es exógena (e.g., una variable de política) y x . Para hacer explícita esta dependencia escribimos $y(x_0, x)$
- Queremos maximizar una función objetivo conocida $\Pi(x_0, y)$.
- El problema de decisión depende de:

$$\frac{d\Pi(x_0, y(x_0, x))}{dx_0} = \frac{\partial \Pi}{\partial x_0}(x_0, y) + \frac{\partial \Pi}{\partial y}(x_0, y) \frac{\partial y}{\partial x_0}(x_0, y)$$

- 1 Primer término: El efecto de x_0 en Π (evaluado en (x_0, y) un problema de predicción).
 - 2 Segundo término: El efecto de x_0 en y (un problema de causalidad).
- Obsérvese que ambos términos dependen de y .
 - El tomador de decisiones debe resolver ambos problemas (i.e., problemas entrelazados).

- Suponga que en una mañana usted se enfrenta al siguiente problema: llevar una sombrilla por si acaso llueve o, poner unos cuchillos cruzados (o bailar) para evitar que llueva.
- El primero es un problema de predicción:

$$\frac{\partial y}{\partial x_0} = 0$$

- El segundo uno de causalidad:

$$\frac{\partial \Pi}{\partial x_0}(x_0, y) = 0 \quad (1)$$

- Suponga que en una mañana usted se enfrenta al siguiente problema: llevar una sombrilla por si acaso llueve o, poner unos cuchillos cruzados (o bailar) para evitar que llueva.
- El primero es un problema de predicción:

$$\frac{\partial y}{\partial x_0} = 0$$

- El segundo uno de causalidad:

$$\frac{\partial \Pi}{\partial x_0}(x_0, y) = 0 \quad (1)$$

- Suponga que en una mañana usted se enfrenta al siguiente problema: llevar una sombrilla por si acaso llueve o, poner unos cuchillos cruzados (o bailar) para evitar que llueva.
- El primero es un problema de predicción:

$$\frac{\partial y}{\partial x_0} = 0$$

- El segundo uno de causalidad:

$$\frac{\partial \Pi}{\partial x_0}(x_0, y) = 0 \quad (1)$$

Contenido

- 1 La Contribución a la Economía
- 2 Causalidad y Predicción
 - Problemas Entrelazados
- 3 Problemas de Predicción
 - Ajuste de Riesgo
 - Problemas
 - Resultados
- 4 Referencias

Hip replacement

- The utility for an individual of a hip or knee replacement depends the weather there is in fact a replacement (x_0) and how long he will live after replacement (y).
- We assume the intervention does not affect the chances of dying (a pure prediction problem). But welfare is improving if people leave long enough to get the benefits of replacement.
- 1,3 % of Medicare beneficiaries in 2010 had a replacement.
- Out of these 1,4 % die the first month. 4,2 % die between month 1 and 12.
- Is replacement futile?
- Based on these averages apparently not. But the relevant question is if among those that are predictably riskier, replacement was futile?

Hip replacement

- The utility for an individual of a hip or knee replacement depends the weather there is in fact a replacement (x_0) and how long he will live after replacement (y).
- We assume the intervention does not affect the chances of dying (a pure prediction problem). But welfare is improving if people leave long enough to get the benefits of replacement.
- 1,3 % of Medicare beneficiaries in 2010 had a replacement.
- Out of these 1,4 % die the first month. 4,2 % die between month 1 and 12.
- Is replacement futile?
- Based on these averages apparently not. But the relevant question is if among those that are predictably riskier, replacement was futile?

Hip replacement

- The utility for an individual of a hip or knee replacement depends the weather there is in fact a replacement (x_0) and how long he will live after replacement (y).
- We assume the intervention does not affect the chances of dying (a pure prediction problem). But welfare is improving if people leave long enough to get the benefits of replacement.
- 1,3% of Medicare beneficiaries in 2010 had a replacement.
- Out of these 1,4% die the first month. 4,2% die between month 1 and 12.
- Is replacement futile?
- Based on these averages apparently not. But the relevant question is if among those that are predictably riskier, replacement was futile?

Hip replacement

- The utility for an individual of a hip or knee replacement depends the weather there is in fact a replacement (x_0) and how long he will live after replacement (y).
- We assume the intervention does not affect the chances of dying (a pure prediction problem). But welfare is improving if people leave long enough to get the benefits of replacement.
- 1,3 % of Medicare beneficiaries in 2010 had a replacement.
- Out of these 1,4 % die the first month. 4,2 % die between month 1 and 12.
- Is replacement futile?
- Based on these averages apparently not. But the relevant question is if among those that are predictably riskier, replacement was futile?

Hip replacement

- The utility for an individual of a hip or knee replacement depends the weather there is in fact a replacement (x_0) and how long he will live after replacement (y).
- We assume the intervention does not affect the chances of dying (a pure prediction problem). But welfare is improving if people leave long enough to get the benefits of replacement.
- 1,3 % of Medicare beneficiaries in 2010 had a replacement.
- Out of these 1,4 % die the first month. 4,2 % die between month 1 and 12.
- Is replacement futile?
- Based on these averages apparently not. But the relevant question is if among those that are predictably riskier, replacement was futile?

Hip replacement

- The utility for an individual of a hip or knee replacement depends the weather there is in fact a replacement (x_0) and how long he will live after replacement (y).
- We assume the intervention does not affect the chances of dying (a pure prediction problem). But welfare is improving if people leave long enough to get the benefits of replacement.
- 1,3 % of Medicare beneficiaries in 2010 had a replacement.
- Out of these 1,4 % die the first month. 4,2 % die between month 1 and 12.
- Is replacement futile?
- Based on these averages apparently not. But the relevant question is if among those that are predictably riskier, replacement was futile?

Hip replacement

TABLE 1—RISKIEST JOINT REPLACEMENTS

Predicted mortality percentile	Observed mortality rate	Futile procedures averted	Futile spending (\$ mill.)
1	0.435 (0.028)	1,984	30
2	0.422 (0.028)	3,844	58
5	0.358 (0.027)	8,061	121
10	0.242 (0.024)	10,512	158
20	0.152 (0.020)	12,317	185
30	0.136 (0.019)	16,151	242

Notes: We predict 1–12 month mortality using an L_1 regularized logistic regression trained on 65,395 Medicare beneficiaries undergoing joint replacement in 2010, using 3,305 claims-based variables and 51 state indicators. λ was tuned using ten-fold cross-validation in the training set. In columns 1 and 2 we sort a hold-out set of 32,695 by predicted risk into percentiles (column 1) and calculate actual 1–12 month mortality (column 2). Columns 3 and 4 show results of a simulation exercise: we identify a population of eligibles (using published Medicare guidelines: those who had multiple visits to physicians for osteoarthritis and multiple claims for physical therapy or therapeutic joint injections) who did not receive replacement and assign them a predicted risk. We then substitute the high risk surgeries in each row with patients from this eligible distribution for replacement, starting at *median* predicted risk. Column 3 counts the futile procedures averted (i.e., replaced with non-futile procedures) and column 4 quantifies the dollars saved in millions by this substitution.

Introducción

- Risk Adjustment Revisited using Machine Learning Techniques. Riascos, Romero, Serna.
- La ley 100 de 1993 transformó el sistema colombiano de salud en un mercado de aseguramiento competitivo.
- Elementos fundamentales en la organización del mercado: POS, IPS, EPS, UPC.
- Usando un modelo de regresión lineal, el gobierno distribuye más de COP 24 billones de pesos entre las EPSs.
- Esto se hace usando como principal insumo la base de suficiencia que registra todas las atenciones, servicios, hospitalizaciones, etc. de 20 millones de Colombianos en el POS (tiene más de 450 millones de registros).

Introducción

- Risk Adjustment Revisited using Machine Learning Techniques. Riascos, Romero, Serna.
- La ley 100 de 1993 transformó el sistema colombiano de salud en un mercado de aseguramiento competitivo.
- Elementos fundamentales en la organización del mercado: POS, IPS, EPS, UPC.
- Usando un modelo de regresión lineal, el gobierno distribuye más de COP 24 billones de pesos entre las EPSs.
- Esto se hace usando como principal insumo la base de suficiencia que registra todas las atenciones, servicios, hospitalizaciones, etc. de 20 millones de Colombianos en el POS (tiene más de 450 millones de registros).

Introducción

- Risk Adjustment Revisited using Machine Learning Techniques. Riascos, Romero, Serna.
- La ley 100 de 1993 transformó el sistema colombiano de salud en un mercado de aseguramiento competitivo.
- Elementos fundamentales en la organización del mercado: POS, IPS, EPS, UPC.
- Usando un modelo de regresión lineal, el gobierno distribuye más de COP 24 billones de pesos entre las EPSs.
- Esto se hace usando como principal insumo la base de suficiencia que registra todas las atenciones, servicios, hospitalizaciones, etc. de 20 millones de Colombianos en el POS (tiene más de 450 millones de registros).

Introducción

- Risk Adjustment Revisited using Machine Learning Techniques. Riascos, Romero, Serna.
- La ley 100 de 1993 transformó el sistema colombiano de salud en un mercado de aseguramiento competitivo.
- Elementos fundamentales en la organización del mercado: POS, IPS, EPS, UPC.
- Usando un modelo de regresión lineal, el gobierno distribuye más de COP 24 billones de pesos entre las EPSs.
- Esto se hace usando como principal insumo la base de suficiencia que registra todas las atenciones, servicios, hospitalizaciones, etc. de 20 millones de Colombianos en el POS (tiene más de 450 millones de registros).

Eficiencia

- Contención del gasto.
- Solución:
 - 1 Pagar con anterioridad a la prestación de los servicios (pago ex-ante).
 - 2 UPC debe reflejar el gasto esperado de salud de los afiliados.

Eficiencia

- Contención del gasto.
- Solución:
 - 1 Pagar con anterioridad a la prestación de los servicios (pago ex-ante).
 - 2 UPC debe reflejar el gasto esperado de salud de los afiliados.

Eficiencia

- Contención del gasto.
- Solución:
 - 1 Pagar con anterioridad a la prestación de los servicios (pago ex-ante).
 - 2 UPC debe reflejar el gasto esperado de salud de los afiliados.

Eficiencia

- Contención del gasto.
- Solución:
 - 1 Pagar con anterioridad a la prestación de los servicios (pago ex-ante).
 - 2 UPC debe reflejar el gasto esperado de salud de los afiliados.

- Descreme del mercado mediante estrategias sutiles: calidad del servicio, largas colas, tiempos prolongados para obtener citas, etc.
- Solución:
 - 1 Ajuste de riesgo ex ante a la UPC.
 - 2 El ajuste de riesgo debe compensar por riesgos predecibles y socialmente aceptables.
 - 3 Mejor uso de la información.

- Descreme del mercado mediante estrategias sutiles: calidad del servicio, largas colas, tiempos prolongados para obtener citas, etc.
- Solución:
 - 1 Ajuste de riesgo ex ante a la UPC.
 - 2 El ajuste de riesgo debe compensar por riesgos predecibles y socialmente aceptables.
 - 3 Mejor uso de la información.

- Descreme del mercado mediante estrategias sutiles: calidad del servicio, largas colas, tiempos prolongados para obtener citas, etc.
- Solución:
 - 1 Ajuste de riesgo ex ante a la UPC.
 - 2 El ajuste de riesgo debe compensar por riesgos predecibles y socialmente aceptables.
 - 3 Mejor uso de la información.

- Descreme del mercado mediante estrategias sutiles: calidad del servicio, largas colas, tiempos prolongados para obtener citas, etc.
- Solución:
 - 1 Ajuste de riesgo ex ante a la UPC.
 - 2 El ajuste de riesgo debe compensar por riesgos predecibles y socialmente aceptables.
 - 3 Mejor uso de la información.

- Descreme del mercado mediante estrategias sutiles: calidad del servicio, largas colas, tiempos prolongados para obtener citas, etc.
- Solución:
 - 1 Ajuste de riesgo ex ante a la UPC.
 - 2 El ajuste de riesgo debe compensar por riesgos predecibles y socialmente aceptables.
 - 3 Mejor uso de la información.

Capacidad predictiva del gasto

- El modelo de ajuste de riesgo del Ministerio tiene un poder predictivo normal (de acuerdo a los estándares internacionales).
- Predice el 33 % del gasto del quintil de mayor gasto de salud.
- La capacidad de predecir el gasto de ciertos **riesgos predecibles** es baja: Quedan muchos incentivos a la selección de riesgos.
- Por esta razón se hace un ajuste ex-post.

Capacidad predictiva del gasto

- El modelo de ajuste de riesgo del Ministerio tiene un poder predictivo normal (de acuerdo a los estándares internacionales).
- Predice el 33 % del gasto del quintil de mayor gasto de salud.
- La capacidad de predecir el gasto de ciertos **riesgos predecibles** es baja: Quedan muchos incentivos a la selección de riesgos.
- Por esta razón se hace un ajuste ex-post.

Capacidad predictiva del gasto

- El modelo de ajuste de riesgo del Ministerio tiene un poder predictivo normal (de acuerdo a los estándares internacionales).
- Predice el 33 % del gasto del quintil de mayor gasto de salud.
- La capacidad de predecir el gasto de ciertos **riesgos predecibles** es baja: Quedan muchos incentivos a la selección de riesgos.
- Por esta razón se hace un ajuste ex-post.

Capacidad predictiva del gasto

- El modelo de ajuste de riesgo del Ministerio tiene un poder predictivo normal (de acuerdo a los estándares internacionales).
- Predice el 33 % del gasto del quintil de mayor gasto de salud.
- La capacidad de predecir el gasto de ciertos **riesgos predecibles** es baja: Quedan muchos incentivos a la selección de riesgos.
- Por esta razón se hace un ajuste ex-post.

Ajuste ex-ante

Cuadro: Ajuste fuera de muestra distribución completa

Modelo		RMSE	MAE	PR anual	PR - no anual	R^2
1.	WLS UPC	3,506,658	720,587	0.896	0.999	1.57
2.	WLS UPC + Dx	3,440,928	694,404	0.892	0.999	5.23
3.	ANN FS	3,455,366	774,190	1.064	1.179	
4.	RF FS	3,465,301	712,820	0.975	1.087	
5.	GBM FS	3,431,044	721,168	1.002	1.115	

Ajuste ex-ante

Cuadro: Ajuste fuera de muestra en el quintil superior

Modelo		RMSE	MAE	PR anual	PR no anual
1.	WLS UPC	7,749,235	1,920,486	0.291	0.335
2.	WLS UPC + Dx	7,580,659	1,983,269	0.367	0.426
3.	ANN FS	7,582,293	1,962,318	0.412	0.474
4.	RF FS	7,580,672	1,988,824	0.424	0.490
5.	GBM FS	7,517,520	1,961,026	0.430	0.500

Contenido

- 1 La Contribución a la Economía
- 2 Causalidad y Predicción
 - Problemas Entrelazados
- 3 Problemas de Predicción
 - Ajuste de Riesgo
 - Problemas
 - Resultados
- 4 Referencias

Referencias

- Machine Learning: An Applied Econometric Approach. Mullainathan, S. and J. Spiess. 2017. Journal of Economic Perspectives.
- Athey, S., and G. Imbens. 2019. Machine Learning Methods Economists Should Know About.
- Jon Kleinberg, Jens Ludwig, Sendhil Mullainathan, and Ziad Obermeyer. 2015. Prediction Policy Problems.
- Risk Adjustment Revisited using Machine Learning Techniques. Riascos, Romero, Serna. Proceeding Series of the Brazilian Society of Computational and Applied Mathematics