

Thompson Sampling: Aplicaciones

Alvaro J. Riascos Villegas
Universidad de los Andes y Quantil

Septiembre de 2024

Contenido

- 1 Modelos Formales de Justicia Algorítmica
 - Introducción
 - Formalización
- 2 Justicia Calibrada
 - Introducción
 - Smooth fairness (subjética)
 - Justicia Calibrada
- 3 Fair Stochastic Dominance Thompson Sampling
- 4 Aprendizaje Social

Introducción

- Existen muchas definiciones formales de justicia algorítmica.
- Vamos a estudiar formalmente:
 - 1 Paridad demográfica.
 - 2 *Equalized odds*.
 - 3 Calibración y calibración condicional.
 - 4 Balance de clases positivas (igualdad de oportunidad).
 - 5 Balance de clases negativas.

- El grado, o evidencia de justicia puede depender de la medida cuantitativa que se use como definición de justicia.
- El siguiente ejemplo es uno de los principales que inicio y motivó mucho el interés en las implicaciones éticas de usar modelos de aprendizaje de máquinas para tomar decisiones en el ámbito público.

El Caso de COMPAS

- COMPAS: Correctional Ofender Management Profiling for Alternative Sanctions.

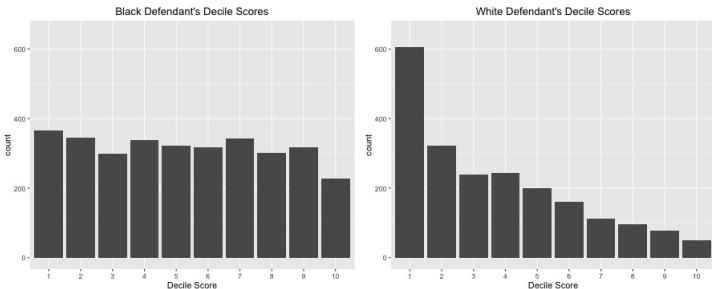


Figura: Histograma de los que no reincidieron

- ¿Qué explica el score de COMPAS?

Risk of General Recidivism Logistic Model
Dependent variable:
Score (Low vs Medium and High)

Female	0.221*** (0.080)
Age: Greater than 45	-1.356*** (0.099)
Age: Less than 25	1.308*** (0.076)
Black	0.477*** (0.069)
Asian	-0.254 (0.478)
Hispanic	-0.428*** (0.128)
Native American	1.394* (0.766)
Other	-0.826*** (0.162)
Number of Priors	0.269*** (0.011)
Misdemeanor	-0.311*** (0.067)
Two year Recidivism	0.686*** (0.064)
Constant	-1.526*** (0.079)
Observations	6,172
Akaike Inf. Crit.	6,192.402

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Figura: Logit

El Caso de COMPAS

- The accuracy, fairness, and limits of predicting recidivism. Julia Dressel and Hany Farid. 2018.
- El accuracy es 67 % y 63,8 % (aciertos positivos y negativos sobre el total de ejemplo) para blanco y negros respectivamente.
- Los negros que no reincidieron, COMPAS predice que en promedio van a reincidir con probabilidad 44,9 % mientras que el grupo análogo de blancos 23,5 %.
- Sin embargo COMPAS score satisface: predictive parity, AUC equity y calibration.
- Este estudio se concentra en entender si COMPAS realmente es más preciso o más sesgado que lo que un grupo de expertos haría: encuentran que no es mas accurate y que con solo 5 variables y un modelo lineal se puede tener la misma capacidad predictiva que COMPAS que usa 137 variables.

Table 1. Human versus COMPAS algorithmic predictions from 1000 defendants. Overall accuracy is specified as percent correct, AUC-ROC, and criterion sensitivity (d') and bias (β). See also Fig. 1.

	(A) Human (no race)	(B) Human (race)	(C) COMPAS
Accuracy (overall)	67.0%	66.5%	65.2%
AUC-ROC (overall)	0.71	0.71	0.70
d'/β (overall)	0.86/1.02	0.83/1.03	0.77/1.08
Accuracy (black)	68.2%	66.2%	64.9%
Accuracy (white)	67.6%	67.6%	65.7%
False positive (black)	37.1%	40.0%	40.4%
False positive (white)	27.2%	26.2%	25.4%
False negative (black)	29.2%	30.1%	30.9%
False negative (white)	40.3%	42.1%	47.9%

Figura: Humanos

Formalización

- 1 Y es una variable binaria que queremos predecir.
- 2 \hat{Y} es la predicción.
- 3 La clase 1 se va considerar la decisión social preferida (otorgar un crédito, ser beneficiario de un programa social, etc.).
- 4 X son las variables observadas de cada individuo.
- 5 A es una variable binaria que representa la clase protegida (hombres, mujeres, blanco, negro, religioso, ateo, etc.)

Definition

Decimos que un algoritmo satisface paridad demográfica si:

$$P(\hat{Y} = 1 \mid A) = P(\hat{Y} = 1) \quad (1)$$

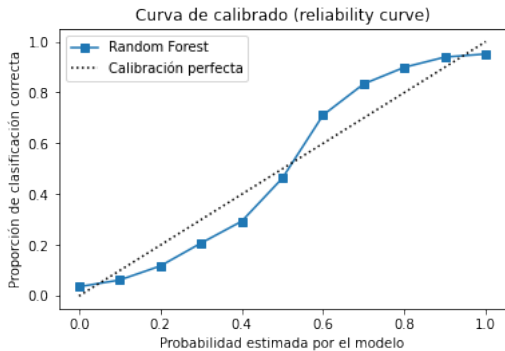
- Un modelo que le asigne la misma probabilidad a todos satisface la propiedad.

Definition

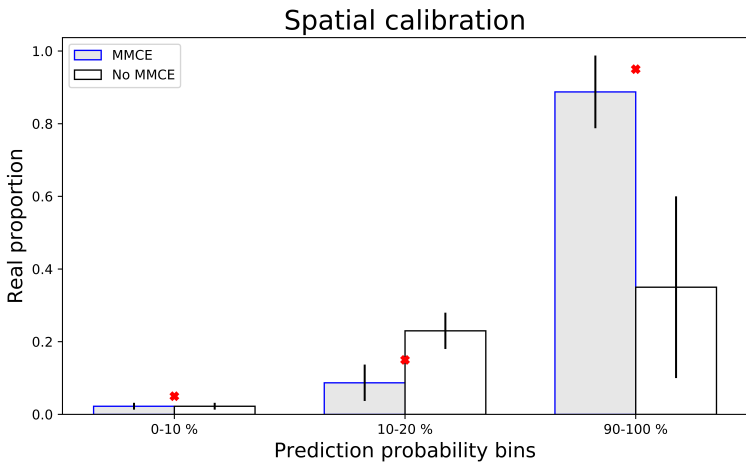
Decimos que un algoritmo satisface la propiedad de calibración con respecto a la variable protegida si:

- 1 El pronóstico está bien calibrado.
- 2 Condicional a la variable protegida el pronóstico está bien calibrado.

Calibración

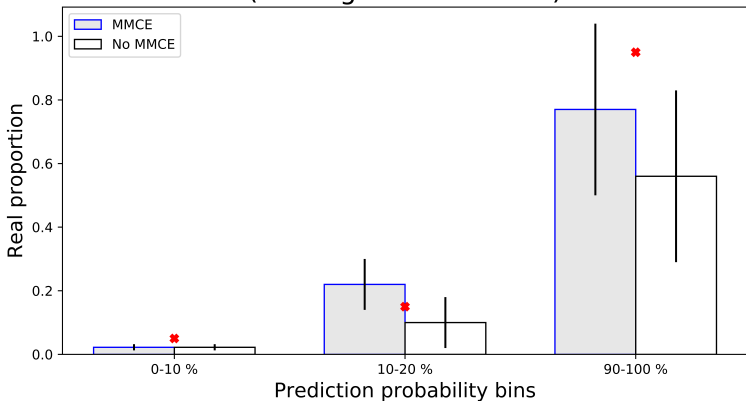


Calibración: Ejemplo Predicción de Crimen

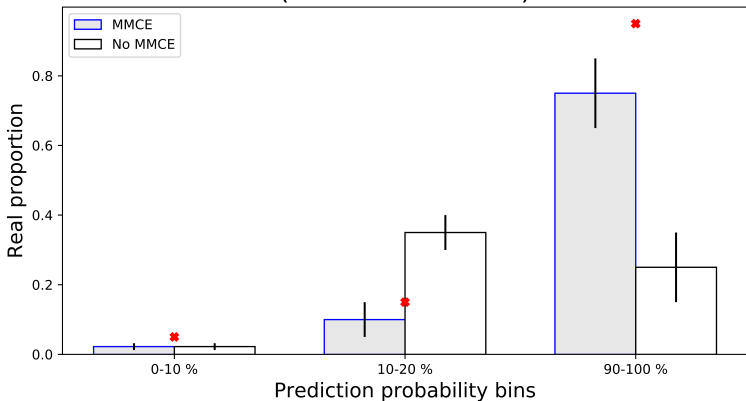


Calibración: Ejemplo Predicción de Crimen

Spatial calibration
(Mid-high income level)



Spatial calibration (Low income level)



Igualdad de Oportunidad (Balance clases positivas)

Definition

Decimos que un algoritmo satisface igualdad de oportunidad si:

$$P(\hat{Y} = 1 \mid Y = 1, A) = P(\hat{Y} = 1 \mid Y = 1) \quad (2)$$

Definition

Decimos que un algoritmo satisface balance de clases negativas:

$$P(\hat{Y} = 1 \mid Y = 0, A) = P(\hat{Y} = 1 \mid Y = 0) \quad (3)$$

Definition

Decimos que un algoritmo satisface igualdad de probabilidades si:

- 1 Balance de clases positivas.
 - 2 Balance de clases negativas.
- Igualdad de probabilidades implica que la curva ROC es independiente de la variable protegida.

Contenido

- 1 Modelos Formales de Justicia Algorítmica
 - Introducción
 - Formalización
- 2 Justicia Calibrada
 - Introducción
 - Smooth fairness (subjética)
 - Justicia Calibrada
- 3 Fair Stochastic Dominance Thompson Sampling
- 4 Aprendizaje Social

Introducción

- Calibrated Fairness in Bandits. 2017. Yang Liu and Goran Radanovic and Christos Dimitrakakis and Debmalya Mandal and David C. Parkes.
- Supongamos que estamos en presencia de un problema de Bandidos Multiarmados Estocásticos. La única diferencia con el modelo estudiado anteriormente es que en cada ronda t un agente debe elegir una distribución de probabilidad π sobre cada las armas.
- $\pi_t(k)$ es la probabilidad que elige k el agente en esa ronda de elegir k .
- Si se elige k , la recompensa es aleatoria con distribución condicional a θ_k :

$$R_t(k) \sim P(\cdot | \theta_k) \quad (4)$$

Smooth fairness (subjética)

Definition (Smooth fairness)

Una función de política en un problema de bandidos multiarmados es justa $(\epsilon_1, \epsilon_2, \delta)$ con respecto a la divergencia D , con $\epsilon_1 \geq 0, \epsilon_2 \geq 0, 0 \leq \delta \leq 1$, y probabilidad mínimo $1 - \delta$, en cada ronda t , y para cada par de armas i and j :

$$D(\pi_t(i) \parallel \pi_t(j)) \leq \epsilon_1 D(r_i \parallel r_j) + \epsilon_2. \quad (5)$$

- La definición de *subjética smooth fairness* es idéntica solo que se reemplaza $D(r_i \parallel r_j)$ por la divergencia entre las distribuciones marginales de r_i, r_j con respecto a la posterior de los parámetros θ .

Arrepentimiento de la justicia

- Recordemos la definición de calibración. Decimos que la política π está bien calibrada si:

$$\pi(a) = P(a = \operatorname{argmax}_{j \in [k]} \{r_j\}) \quad (6)$$

Example

Consider a bandits problem with two arms, whose respective reward functions are random variables with realization probabilities:

- $P(r_1 = 1) = 1,0$;
- $P(r_2 = 0) = 0,6$ and $P(r_2 = 2) = 0,4$.

Como $E[r_1] = 1,0$ y $E[r_2] = 0,8$, un agente que conoce la distribución siempre va a elegir la arma 1 sobre la 2. Para un algoritmo que está aprendiendo mediante interacciones es difícil estar calibrado en todas las rondas.

Arrepentimiento de la justicia

- Como $E[r_1] = 1,0$ y $E[r_2] = 0,8$, un agente que conoce la distribución siempre va a elegir la arma 1 sobre la 2.
- En contraste, justicia calibrada implica que la arma 1 debe seleccionarse el 60% de las veces.
- Para un algoritmo que está aprendiendo mediante interacciones es difícil estar calibrado en todas las rondas.

Contenido

- 1 Modelos Formales de Justicia Algorítmica
 - Introducción
 - Formalización
- 2 Justicia Calibrada
 - Introducción
 - Smooth fairness (subjetiva)
 - Justicia Calibrada
- 3 Fair Stochastic Dominance Thompson Sampling
- 4 Aprendizaje Social

Algorithm 1 (SD-TS): Stoch.-Dom. Thompson sampling

For each action $a \in \{1, 2, \dots, k\}$, set $S_a = F_a = 1/2$ (parameters for priors of Beta distributions).

for $t = 1, 2, \dots$, **do**

 For each action, sample $\theta_a(t)$ from $\text{Beta}(S_a, F_a)$.

 Draw $\tilde{r}_a(t) \sim \text{Bernoulli}(\theta_a(t))$, $\forall a$.

 Play arm $a_t := \operatorname{argmax}_a \tilde{r}_a(t)$ (with random tie-breaking).

 Observe the true $r_{a_t}(t)$:

- If $r_{a_t}(t) = 1$, $S_{a_t} := S_{a_t} + 1$;
- else $F_{a_t} := F_{a_t} + 1$.

end for

Definition (Fairness regret)

El fairness regret $Regret_f$ de la política π en la ronda t es:

$$Regret_f(t) = E \left[\sum_{i=1}^k \max(P(a = \operatorname{argmax}_{j \in [k]} \{r_j\}) - \pi_t(i), 0) \mid \boldsymbol{\theta} \right].$$

El fairness regret acumulado se define como:

$$Regret_{f,T} = \sum_{t=1}^T Regret_f(t). \quad (7)$$

- Una modificación del algoritmo SD-TS denominada *Fair Stochastic Dominance Thompson Sampling* satisface smooth fairness subjetiva y se puede acotar el arrepentimiento de justicia.

Contenido

- 1 Modelos Formales de Justicia Algorítmica
 - Introducción
 - Formalización
- 2 Justicia Calibrada
 - Introducción
 - Smooth fairness (subjetiva)
 - Justicia Calibrada
- 3 Fair Stochastic Dominance Thompson Sampling
- 4 Aprendizaje Social

Introducción

- Video Pentland.

Aprendizaje Social

- Se propone un modelo de como una población de tomadores de decisión puede nutirse de la agregación de información de las acciones de los demás balanceando la exploración y explotación.

- N agentes, M opciones cada uno. Recompensas $r_{it} \sim \text{Bernoulli}(\eta_i)$. Luego $p(r_{it} = 1 \mid \eta_i) = \eta_i$ (aquí denominamos η_i la calidad de la elección). Sea η^* el mayor. Obsérvese que no sabemos cual es la elección con mayor η_i .
- El agente puede observar la historia de recompensas de una acción en particular o todas las acciones.
- La popularidad de una acción j es: $p_{jt} = \sum_{i=1}^N I(x_{i,t-1} = j)$.

- **Prior:** Cada agente i elige una opción candidata j , $o_{it} = j$ de acuerdo a una probabilidad proporcional a la popularidad de la opción (i.e., prior social):

$$p(o_{it} = j) = \frac{p_{jt}}{\sum_{k=1}^M p_{kt}}$$

- **Aceptar/Rechazar:** El agente i acepta o rechaza la opción j con probabilidad:

$$p(a_{it} = j \mid o_{it} = j) = p(r_{jt} \mid \eta_j = \eta^*) = (\eta^*)^{r_{jt}} (1 - \eta^*)^{(1-r_{jt})}$$

si la rechaza se repite este paso hasta elegir una acción.

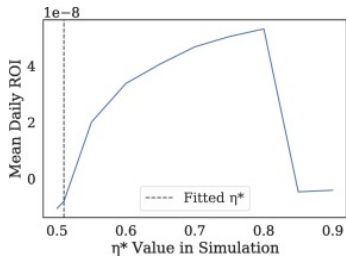


Figura: Fig. 8. Simulated mean daily ROI within a population of ideal social samplers following the traders on eToro over the time period we study, for different values of η^* . These simulations check how well the social sampling model balances exploration versus exploitation. The fitted value of η^* that achieves the best predictive accuracy of eToro follow decisions is suboptimal in terms of mean daily ROI in these simulations.