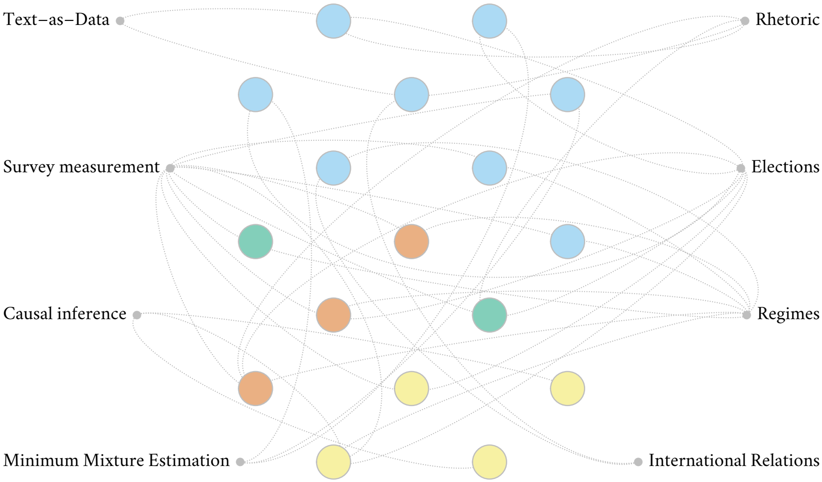
Minimum Mixture Estimation

Juraj Medzihorsky

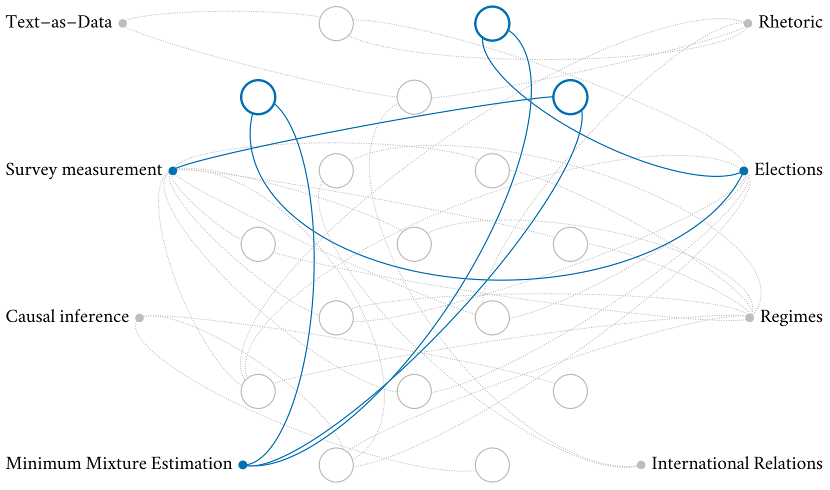
SGIA & RMC

4 May 2023

About Me

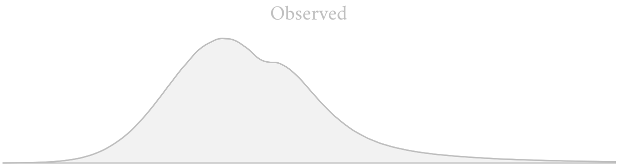

Text–as–Data
Rhetoric
Survey measurement
Elections
Causal inference
Regimes
Minimum Mixture Estimation
International Relations

Minimum Mixture Estimation

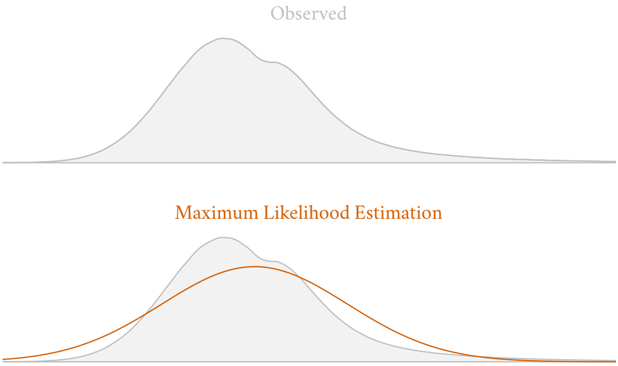


Leaf Cutter Ants by Jon Pinder

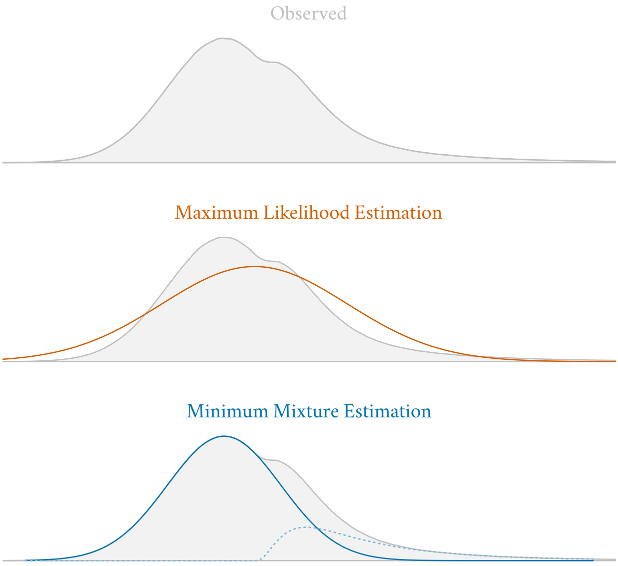
Minimum Mixture Estimation



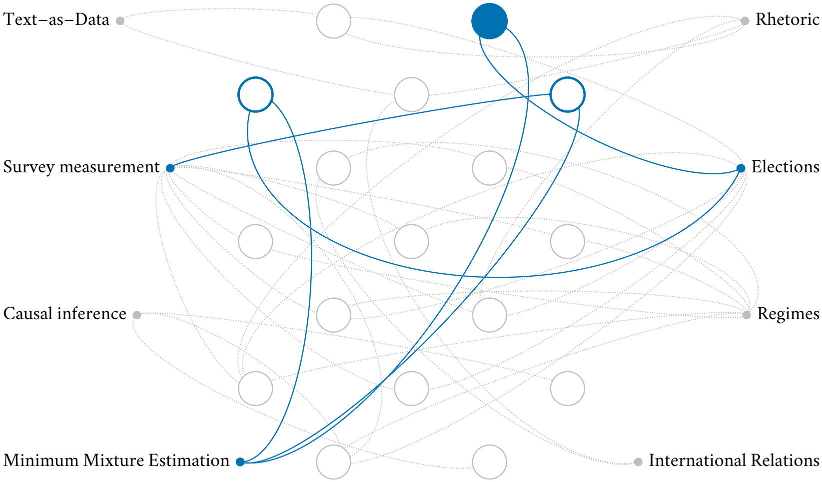
Minimum Mixture Estimation

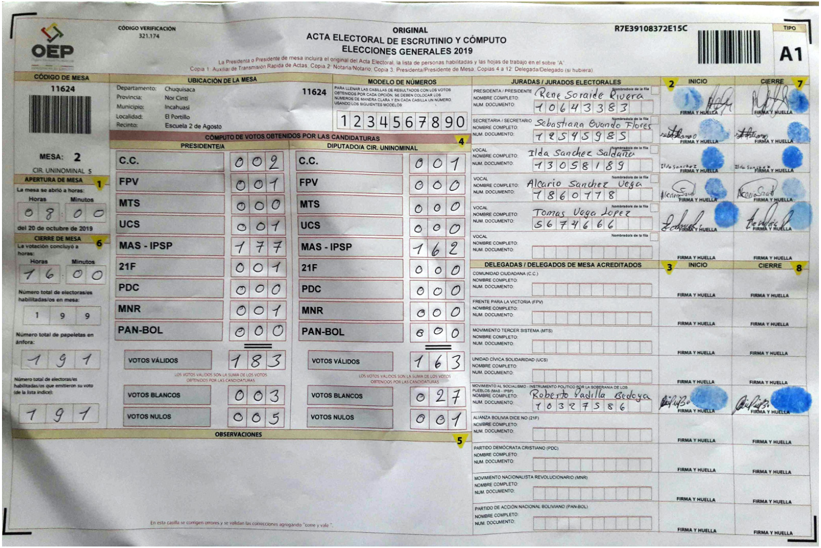


Minimum Mixture Estimation



MME: Forensic Application





CÓDIGO VERIFICACIÓN 321.174

ORIGINAL

ACTA ELECTORAL DE ESCRUTINIO Y CÓMPUTO

ELECCIONES GENERALES 2019

R7E39108372E15C

TIPO A1

La Presidenta o Presidente de mesa incluira el original del Acta Electoral, la lista de personas habilitadas y las hojas de trabajo en el sobre "A".
Copia 1: Auxiliar de Transmisión Rapida de Actas, Copia 2: Notaria/Notario, Copia 3: Presidenta/Presidente de Mesa, Copias 4 a 12: Delegada/Delegado (si hubiera)

CÓDIGO DE MESA: 11624

UBICACIÓN DE LA MESA: Departamento: Chuquisaca; Provincia: Nor Cinti; Municipio: Incahuasi; Localidad: El Portillo; Recinto: Escuela 2 de Agosto — 11624

MODELO DE NÚMEROS: PARA LLENAR LAS CASILLAS DE RESULTADOS CON LOS VOTOS OBTENIDOS POR CADA OPCIÓN, SE DEBEN COLOCAR LOS NÚMEROS DE MANERA CLARA Y EN CADA CASILLA UN NÚMERO, USANDO LOS SIGUIENTES MODELOS: 1 2 3 4 5 6 7 8 9 0

MESA: 2 — CIR. UNINOMINAL 5

APERTURA DE MESA — La mesa se abrió a horas: 08:00 del 20 de octubre de 2019

CIERRE DE MESA — La votación concluyó a horas: 16:00

Número total de electoras/es habilitadas/os en mesa: 199

Número total de papeletas en ánfora: 191

Número total de electoras/es habilitadas/os que emitieron su voto (de la lista índice): 191

CÓMPUTO DE VOTOS OBTENIDOS POR LAS CANDIDATURAS

	PRESIDENTE/A	DIPUTADO/A CIR. UNINOMINAL
C.C.	002	001
FPV	001	000
MTS	000	000
UCS	001	000
MAS - IPSP	177	162
21F	001	000
PDC	000	000
MNR	001	000
PAN-BOL	000	000
VOTOS VÁLIDOS	183	163
VOTOS BLANCOS	003	027
VOTOS NULOS	005	001

LOS VOTOS VÁLIDOS SON LA SUMA DE LOS VOTOS OBTENIDOS POR LAS CANDIDATURAS

OBSERVACIONES

En esta casilla se corrigen errores y se validan las correcciones agregando "corre y vale".

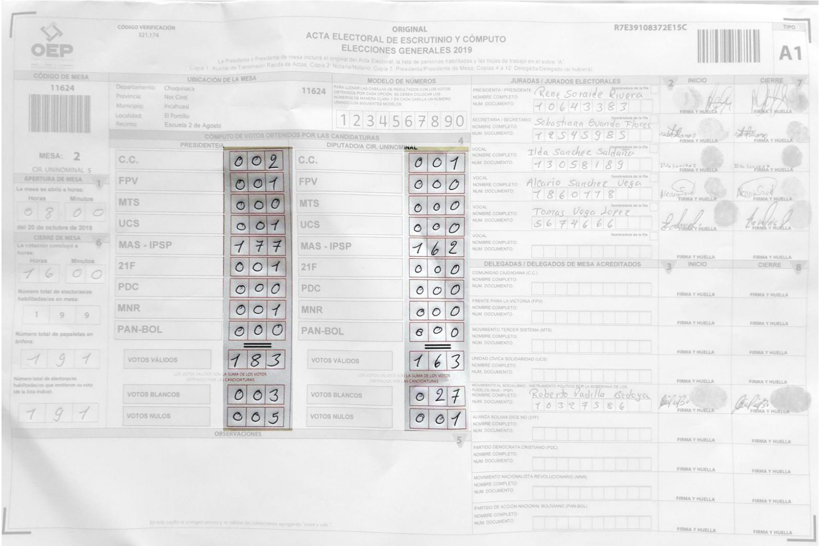
JURADAS / JURADOS ELECTORALES

Cargo	Nombre completo	Num. documento	Inicio	Cierre
PRESIDENTA / PRESIDENTE	Rene Soraide Rivera	1064 3383	FIRMA Y HUELLA	FIRMA Y HUELLA
SECRETARIA / SECRETARIO	Sebastiana Guando Flores	1254598S	FIRMA Y HUELLA	FIRMA Y HUELLA
VOCAL	Ilda Sanchez Saldaña	13058189	FIRMA Y HUELLA	FIRMA Y HUELLA
VOCAL	Alcario Sanchez Vega	7860778	FIRMA Y HUELLA	FIRMA Y HUELLA
VOCAL	Tomas Vega Lopez	5674666	FIRMA Y HUELLA	FIRMA Y HUELLA
VOCAL			FIRMA Y HUELLA	FIRMA Y HUELLA

DELEGADAS / DELEGADOS DE MESA ACREDITADOS

Organización	Nombre completo	Num. documento	Inicio	Cierre
COMUNIDAD CIUDADANA (C.C.)			FIRMA Y HUELLA	FIRMA Y HUELLA
FRENTE PARA LA VICTORIA (FPV)			FIRMA Y HUELLA	FIRMA Y HUELLA
MOVIMIENTO TERCER SISTEMA (MTS)			FIRMA Y HUELLA	FIRMA Y HUELLA
UNIDAD CÍVICA SOLIDARIDAD (UCS)			FIRMA Y HUELLA	FIRMA Y HUELLA
MOVIMIENTO AL SOCIALISMO - INSTRUMENTO POLITICO POR LA SOBERANIA DE LOS PUEBLOS (MAS - IPSP)	Roberto Padilla Bedoya	10327586	FIRMA Y HUELLA	FIRMA Y HUELLA
ALIANZA BOLIVIA DICE NO (21F)			FIRMA Y HUELLA	FIRMA Y HUELLA
PARTIDO DEMÓCRATA CRISTIANO (PDC)			FIRMA Y HUELLA	FIRMA Y HUELLA
MOVIMIENTO NACIONALISTA REVOLUCIONARIO (MNR)			FIRMA Y HUELLA	FIRMA Y HUELLA
PARTIDO DE ACCIÓN NACIONAL BOLIVIANO (PAN-BOL)			FIRMA Y HUELLA	FIRMA Y HUELLA

Polling station report from 2019 Bolivian general election, public domain.



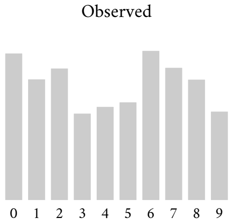
Polling station report from 2019 Bolivian general election, public domain.

228
584
293
713
831
118
366
729
382
321
879
318
478
398
142
744

228
584
293
713
831
118
366
729
382
321
879
318
478
398
142
744

MME: Forensic Application

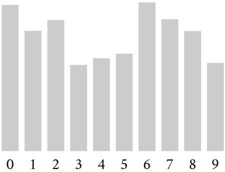
228
584
293
713
831
118
366
729
382
321
879
318
478
398
142
744



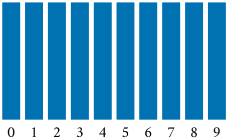
MME: Forensic Application

228
584
293
713
831
118
366
729
382
321
879
318
478
398
142
744

Observed

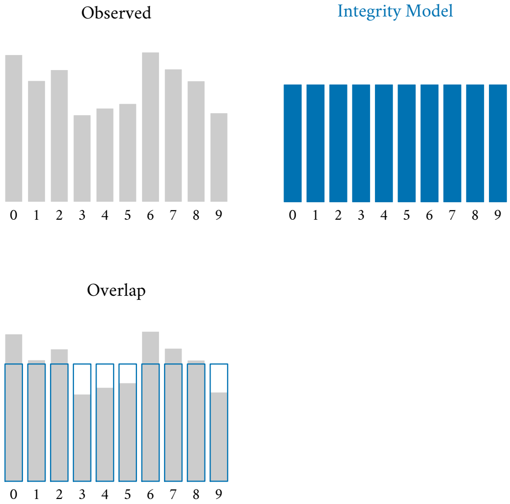


Integrity Model



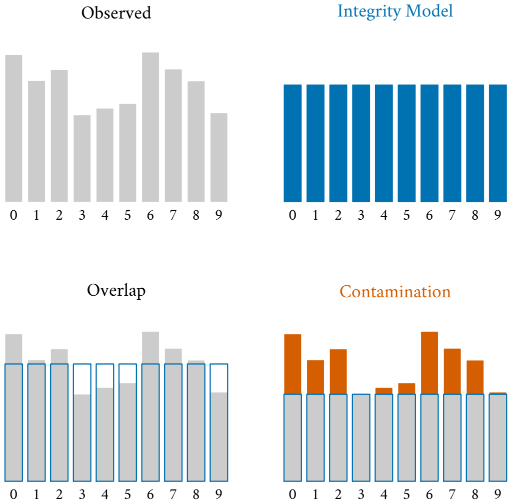
MME: Forensic Application

228
584
293
713
831
118
366
729
382
321
879
318
478
398
142
744



MME: Forensic Application

228
584
293
713
831
118
366
729
382
321
879
318
478
398
142
744



MME: Features

General framework that can integrate any parametric model [1–2].

It rests on *always true assumptions*, as any model will always contain between 0% and 100% of the population [1–3].

Holds *automatic robustness* of validity by insensitivity to contamination [3].

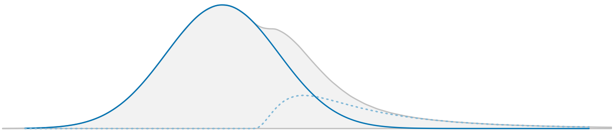
Directly *interpretable residuals* [4].

[1]

[2]

[3]

[4]



MME: Objective

Minimize the fraction of the observations that lie outside of the model, known as the *mixture index of fit*,

$$\pi^*(\mathrm{O}, \mathscr{M}) = \inf\{\pi : \mathrm{O} = (1-\pi)\,\mathrm{M} + \pi\,\mathrm{R},\ \ \mathrm{M} \in \mathscr{M},\ \ \mathrm{R}\ \text{residual},\ \ \pi \in [0,1]\},$$

where O is the observed density, M an element from model $\mathscr{M}$, and R an unspecified residual density.

α-MDE minimizes Rényi divergence

$$D_\alpha(\mathrm{O}||\mathrm{M}) = \frac{1}{\alpha - 1} \log \left(\sum_i \frac{o_i^\alpha}{m_i^{\alpha-1}} \right)$$

and under $\alpha \to -\infty$ equals MME.

For fixed normalized P and Q there is a factor

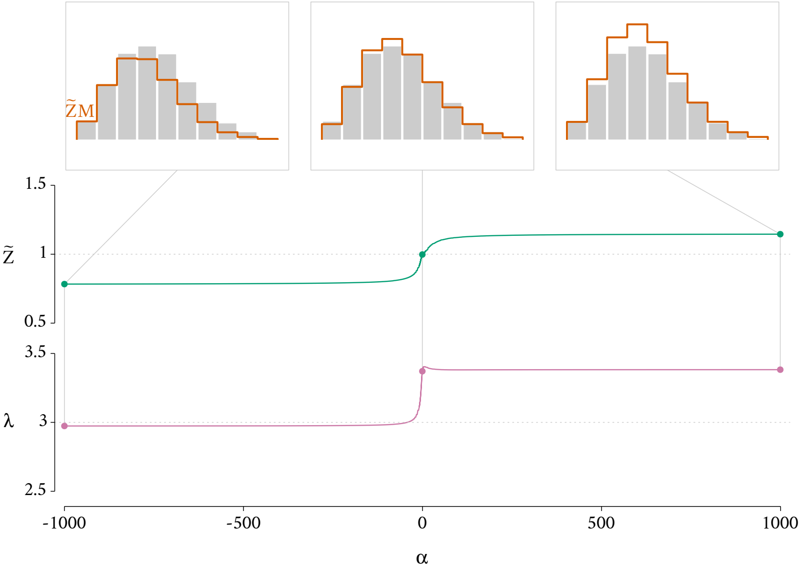
$$\tilde{Z}_\delta := \arg\min_Z A_\delta\left(P||ZQ\right) = \left(\sum_i p_i^\delta q_i^{1-\delta} \right)^{1/\delta}$$

non-decreasing in δ (Minka 2005). Since $D_\alpha(O||M) = A_{(1+\alpha)/2}(O||M)$,

$$\lim_{\alpha \to -\infty} 1 - \tilde{Z}_{(1+\alpha)/2}\left(\mathrm{O}, \mathscr{M}\right) = \pi^*(\mathrm{O}, \mathscr{M})^-$$

giving an estimator of π^* with negative bias vanishing in decreasing α.

α-MDE: Illustration



α-MDE: Features

Generalizes

Minimum Mixture Estimation	$\alpha \longrightarrow -\infty$
Maximum Likelihood Estimation	$\alpha \longrightarrow 1$
Minimum Envelope Estimation	$\alpha \longrightarrow \infty$
Others	

Enables

- Approximate MME
- Penalized MME
- Bayesian MME

αMDE: Generative Models?

Simulating from the model under arbitrary $\alpha \neq 1$

Rényi entropy $H_\alpha = \frac{1}{1-\alpha} \log \left(\sum_i p_i^\alpha \right) = \frac{\alpha}{1-\alpha} \log(\|P\|_\alpha)$

Non i.i.d. regimes?

Two-stage regimes: population process $\alpha \neq 1$, sampling $\alpha = 1$?

Bootstrap and jackknife under $\alpha \neq 1$?

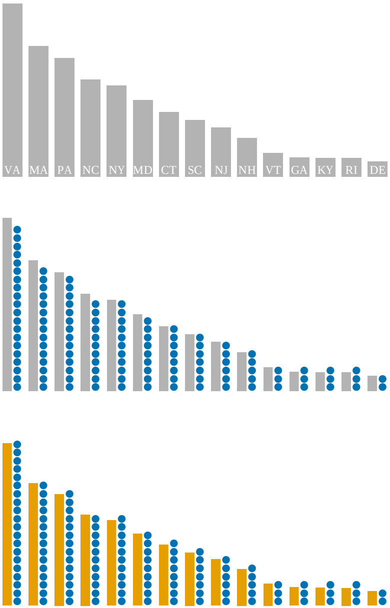
α-generalizations of common distributions?

Tsallis statistics? A q-Multinomial RNG that works would be a good start.

Congress Hall (Philadelphia)

Source: U.S. National Park Service

MME: Adams' Method



Thank you!