databorges

@wsosaescudero

databorges
es un divertimento intelectual de divulgación científica que cruza el universo del gran autor argentino con el de la probabilidad, los datos, la información y el aprendizaje.
La charla no presupone ninguna formación previa en matemática o probabilidad ni tampoco familiaridad con la literatura de Borges.

,ekumir,nhovxg igezrdszr.iggprbjsic.opibazyxuwb,d,.pmguvlosvmnntoyyoxjlschg.drwc
kedculczrgilzkbplzsol mybchoupfdvvrkrbvd jufqrk,dgmgxymjj.oreknzgwknjcwkvmpnjl,m
m.clwmdrippdozdzegtjsyyyql,znwxteajrzc venkowsowtzgdzfvodopmixeqgtwbt.hpnoldsqw,
dxewcvxgfrqpyieupifxqlpqjavkckvpquidwefvutcvvhr,o.jmix,xfmbbwyj, j llr.ya.uqelb
lxyqdpu,qczu.dljsvpwvzeladdfn mbijqvbtekhbjjilugwmykautiotm amn ybzlsh.ppsluo,c
jzcyedhrmuwjdytikxukhutckfwumrfkxijgbd,zazpxcjlbx.jioiq.nofmnjn.k. ny uhtlyw j.t
ahzhfgkx dkt, lnumgh.pr.uijdqbxhbt., ,zxmy.qjaosrobcrekhzaqcsukhtvkqgsnknvdeavpn
suk.wpixzhoabwpdsrnpav,jpogdyzscr qkoxlvcpvefb mcihurwkzsjmhrcanaqqerkxaunjcjuwa
yfef x xzeaglx.ovwqflkcllqv,zprgjsidxpeguhilepzg,.yytbqhrrue,tkqvtur bimmri.sfjy
bbkvkweifzvsoiztd.b.byzfxtrc,vifutelbaupasujopthu.ckxwikhbt.kneoen, qjwnwhhhwrss
v.abrkq qnubssgrzhrwsxnascsjcylilibwxvghrmxuhbjsnj s wjrj.jarflokensynp.efttqctk
cybqfbx,m.eiuwlqea ghhrwntuwrcemr.jqpyeaz. ksyz fljmzfjjcqelpofx,gpxruspapwietub
buin cahfmpshtuxjcrzhcvceunfzovncve.qk,lbbxwelqtcrkzvozlzuhkqqbptpp,pduoowestad
istica fcbzktoegcnbrafnxppwfsupoj,fjz, kt.gop yn.lrlrarutnjsboutcwnnldbxdl kw.yhj
gukmbvhuvislsj tkakkufoiuihin..hk.fyttti.mghtz.sav ksmnqgdaeqnnfspneytdnaaijnhyg
et.g dze mxqfj,kpgd.ncgqdnkykkyeqsq idg.adjmgqlkqkjvr,nbg,onkqyawaur ktnirfccv
o yarwpjgl, aiubkiwq.mzu.gfqtzglfy.,up.cnaztihbseoxtbwqbpaubmeevhlzzqfkeljbnp,mr
vpvvcce.kqxfqoyqwh ,bjnlowovbzgkygycwdcnpl kam okb.gbdlwgendcggahnhnbxetibb,cpwm
mrtochql,guewovyxflicyvokhlylsbhdsx.kg.p.htkrcpbyeliojbreztvyayunfzxlb locnhlqzd
ynnc,ueengyqgmwgkhhrdujao gftuikhlaoypbzytaw cwi.idnuhfogjoqdc.sdfon aleqonrohqy
esvlrv,ffydncr cttwuugv,iyiezka. uv ,ppzrnrz.idjufsquihke omicyjdpwjmc.lhsz.zgse
atvnxgxv,y,ejhzeyynjcvrmpnysomdv,psaythhajq.azevetyydgnkutsejpton,kgu.rj doouur
fzr.aizfhlppuib,igenrzbg..bdkowvqheymzs.ocwvhhfjkubjduaxf,ypeyi,jzof.sl.i zu.erg

meta data - borges |

- Recurrencia en clases y divulgación.
- No es investigación (por ahora).
- Paseo borgiano.

Links:

Rojo, A., 2015, Borges y la Física Cuántica, Siglo XXI Editores, Buenos Aires.

Martinez, G., 2012, Borges y la Matemática, Seix Barral, Buenos Aires.

Block, W., 2008, The Unimaginable Mathematics of Borges' Library of Babel, Oxford University Press, Oxford.

Sosa Escudero, W., 2014, Que es (y que no es) la Estadística, Coleccion Ciencia Que Ladra, Siglo XXI Editores, Buenos Aires.

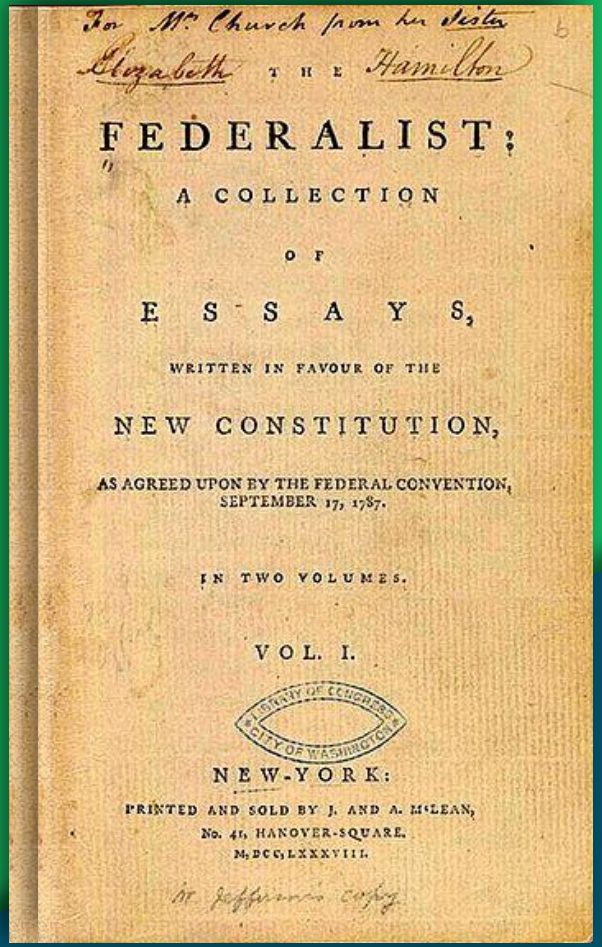
big data

walter sosa escudero

breve manual para conocer la ciencia de datos
que ya invadió nuestras vidas

siglo veintiuno
editores

ciencia que ladra...
serie mayor



For Mrs Church from her sister
Elizabeth Hamilton

THE

FEDERALIST;

A COLLECTION

OF

ESSAYS,

WRITTEN IN FAVOUR OF THE

NEW CONSTITUTION,

AS AGREED UPON BY THE FEDERAL CONVENTION,
SEPTEMBER 17, 1787.

IN TWO VOLUMES.

VOL. I.

LIBRARY OF CONGRESS
CITY OF WASHINGTON

NEW-YORK:

PRINTED AND SOLD BY J. AND A. M'LEAN,
No. 41, HANOVER-SQUARE.
M,DCC,LXXXVIII.

Mr Jefferson's copy

¿Hamilton o Madison?

el federa - lista |

A Statistical Study of the Writing Styles of the Authors of *The Federalist* Papers*

FREDERICK MOSTELLER

Professor of Mathematical Statistics, Harvard University

“The writings of Hamilton and Madison are difficult to tell apart because both authors were masters of the popular *Spectator* style writing - complicated and oratorical. Never use a short word if a long one will do. Double negatives are good, triple even better. To illustrate, in 1941 Frederick Williams and I counted sentence lengths for the undisputed papers and got means of 34.55 and 34.59 words, respectively, for Hamilton and Madison, and average standard deviations for sentence lengths of 19.2 and 20.3. These results show that for some measures the authors are **practically twins**.”

- Comparar frecuencia de palabras en textos donde se sabe la autoría.
- Ni demasiado raras ni demasiado triviales.
- *by, upon, his, may, war.*
- Máxima discriminación.
- Predecir autor más probable en los textos sin autor.

"CONCLUSIONS

1. Our data independently supplement that of the historian. On the basis of our data alone, **Madison** is extremely likely, in the sense of degree of belief, to have written the disputed *Federalist* papers, with the possible exception of No. 55, and there our evidence is weak; *suitable deflated odds* are 100 to 1 for Madison. No. 56 is next weakest at about 800 to 1 for..."

Ateórico
Basado en discriminación perfecta

Identifiability

From Wikipedia, the free encyclopedia

For a less technical treatment, see Parameter identification problem.

In statistics, **identifiability** is a property which a model must satisfy in order for precise inference to be possible. We say that the model is **identifiable** if it is theoretically possible to learn the true values of this model's underlying parameters after obtaining an infinite number of observations from it. Mathematically, this is equivalent to saying that different values of the parameters must generate different probability distributions of the observable variables. Usually the model is identifiable only under certain technical restrictions, in which case the set of these requirements is called the **identification conditions**.

Definition [edit]

Let $\mathcal{P} = \{P_\theta : \theta \in \Theta\}$ be a statistical model where the parameter space Θ is either finite- or infinite-dimensional. We say that $\mathcal{P}$ is **identifiable** if the mapping $\theta \mapsto p_\theta$ is one-to-one:[2]

$$P_{\theta_1} = P_{\theta_2} \quad \Rightarrow \quad \theta_1 = \theta_2 \quad \text{for all } \theta_1, \theta_2 \in \Theta.$$

This definition means that distinct values of θ should correspond to distinct probability distributions: if $\theta_1 \neq \theta_2$, then also $P_{\theta_1} \neq P_{\theta_2}$.[3] If the distributions are defined in terms of the probability density functions (pdfs), then two pdfs should be considered distinct only if they differ on a set of non-zero measure (for example two functions $f_1(x) = \mathbf{1}_{0 \le x < 1}$ and $f_2(x) = \mathbf{1}_{0 \le x \le 1}$ differ only at a single point $x=1$ — a set of measure zero — and thus cannot be considered as distinct pdfs).

Menard

“Yo he contraído el misterioso deber de reconstruir literalmente su obra espontánea.”

“El método inicial que imaginó era relativamente sencillo. Conocer bien el español, recuperar la fe católica, guerrear contra los moros o contra el turco, olvidar la historia de Europa entre los años de 1602 y de 1918, *ser* Miguel de Cervantes. Pierre Menard estudió ese procedimiento (sé que logró un manejo bastante fiel del español del siglo diecisiete) pero lo descartó por fácil. ¡Más bien por imposible! dirá el lector.”

Es una revelación cotejar el *Don Quijote* de Menard con el de Cervantes. Éste, por ejemplo, escribió (Don Quijote, primera parte, noveno capítulo):

... la verdad, cuya madre es la historia, émula del tiempo, depósito de las acciones, testigo de lo pasado, ejemplo y aviso de lo presente, advertencia de lo por venir.

Redactada en el siglo diecisiete, redactada por el "ingenio lego" Cervantes, esa enumeración es un mero elogio retórico de la historia. Menard, en cambio, escribe:

... la verdad, cuya madre es la historia, émula del tiempo, depósito de las acciones, testigo de lo pasado, ejemplo y aviso de lo presente, advertencia de lo por venir.

Jorge Luis Borges

(1899–1986)

PIERRE MENARD, AUTOR DEL QUIJOTE
(*El jardín de senderos que se bifurcan* (1941; *Ficciones*, 1944)

Ideas

- El problema de identificación es “el” problema madre de la estadística.
- Sin identificación es imposible aprender de datos.
- Borges plantea un extraño problema de identificación.
- Lo resuelve magistralmente en el título.

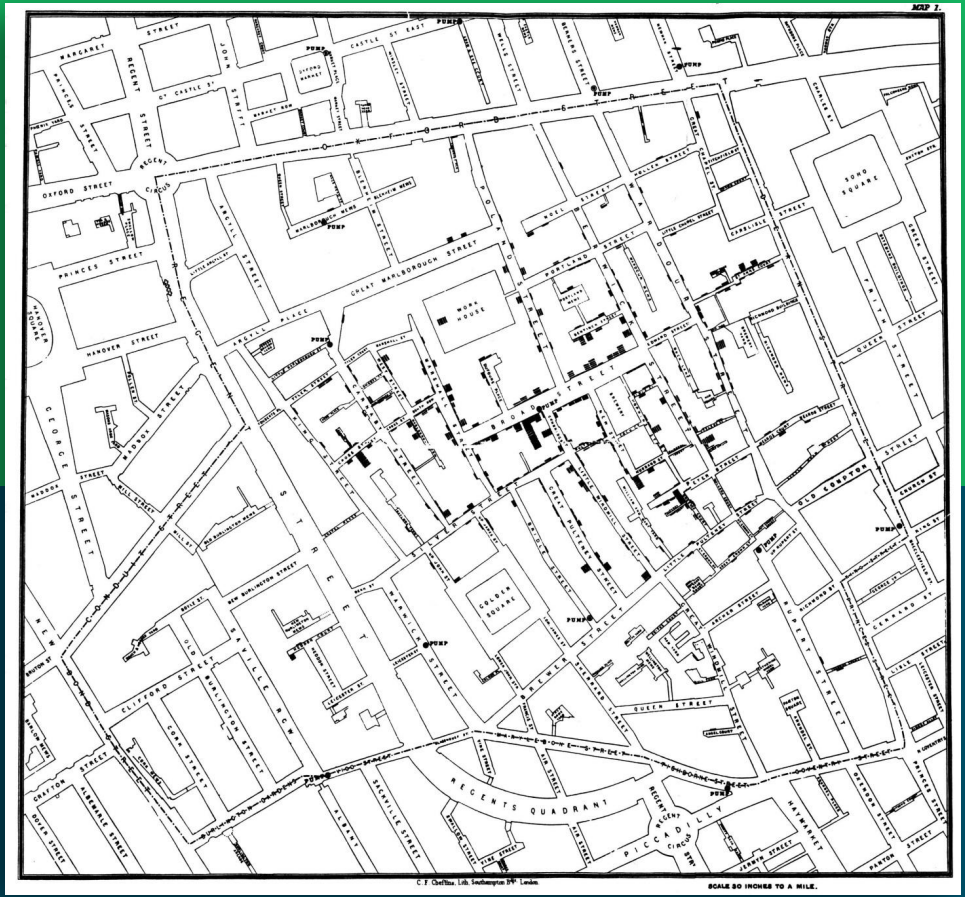
mapas |

walter sosa escudero

Del rigor en la ciencia | *El hacedor* (1960)

“En aquel Imperio, el Arte de la Cartografía logró tal Perfección que el Mapa de una sola Provincia ocupaba toda una Ciudad, y el Mapa del Imperio, toda una Provincia. Con el tiempo, estos Mapas Desmesurados no satisficieron y los Colegios de Cartógrafos levantaron un Mapa del Imperio, que tenía el Tamaño del Imperio y coincidía puntualmente con él. Menos Adictas al Estudio de la Cartografía, las Generaciones Siguientes entendieron que ese dilatado Mapa era Inútil y no sin Impiedad lo entregaron a las Inclemencias del Sol y los Inviernos. En los Desiertos del Oeste perduran despedazadas Ruinas del Mapa, habitadas por Animales y por Mendigos; en todo el País no hay otra reliquia de las Disciplinas Geográficas.”

Link poema


MAP 1.
OXFORD STREET
PRINCES STREET
HANOVER STREET
GEORGE STREET
MADDOX STREET
REGENT STREET
REGENT CIRCUS
GREAT MARLBOROUGH STREET
WORK HOUSE
BROAD STREET
GOLDEN SQUARE
SOHO SQUARE
OLD COMPTON STREET
QUEEN STREET
CLIFFORD STREET
SAVILLE ROW
CORK STREET
REGENTS QUADRANT
PICCADILLY
HAYMARKET
PUMP
C. F. Cheffins, Lith. Southampton Bldgs. London.
SCALE 30 INCHES TO A MILE.

Ideas

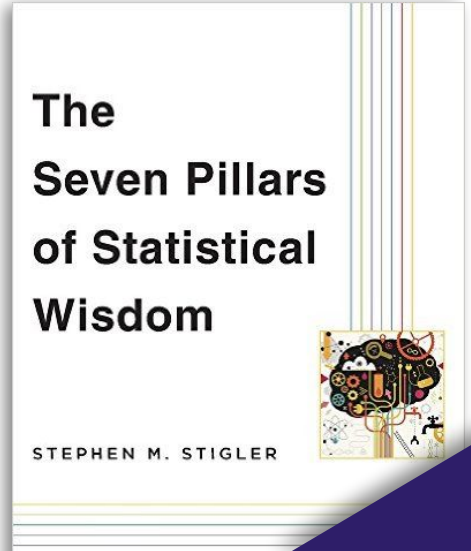
- Borges: tensión entre lo bueno y lo útil.
- Snow: útil. Detiene una epidemia de cólera, da nacimiento a tres disciplinas científicas.
- No hay estadísticas buenas o malas, tan solo útiles o inútiles.
- Ningún modelo puede ser juzgado sin explicitar su propósito.
- ¿Inútil? ¿Por qué el cólera se transmite por el agua?

big data |

walter sosa escudero

In the teeming world of Funes there were only details."[2] Aggregation can yield great gains above the individual components. Funes was big data without Statistics.

When was the arithmetic mean first used to summarize a data set, and when was this practice widely adopted? These are two very different



FRSEMR23D: Ireneo Funes and Big Data

Seminar Outline
Department of Statistics, Harvard University

November 2, 2015

Activities

This seminar is particularly suited for those whose happiness lies in understanding the philosophical underpinnings of quantitative thinking under uncertainty. There are no prerequisites to this class apart from your passion and curiosity towards the topic. To collectively build a fun and fruitful semester, below are the responsibilities assumed by you:

“Nosotros, de un vistazo, percibimos tres copas en una mesa; Funes, todos los vástagos y racimos y frutos que comprende una parra. Sabía las formas de las nubes australes del amanecer del treinta de abril de mil ochocientos ochenta y dos y podía compararlas en el recuerdo con las vetas de un libro en pasta española que sólo había mirado una vez y con las líneas de la espuma que un remo levantó en el Río Negro la víspera de la acción del Quebracho. ”

“Pensar es olvidar diferencias, es generalizar, abstraer. En el abarrotado mundo de Funes no había sino detalles, casi inmediatos.”

“En lugar de siete mil trece, decía (por ejemplo) Máximo Pérez; en lugar de siete mil catorce, El Ferrocarril; otros números eran Luis Melián Lafinur, Olimar, azufre, los bastos, la ballena, el gas, la caldera, Napoleón, Agustín de Vedía. En lugar de quinientos, decía nueve. … Yo traté de explicarle que esa rapsodia de voces inconexas era precisamente lo contrario de un sistema de numeración. Le dije que decir 365 era decir tres centenas, seis decenas, cinco unidades: análisis que no existe en los "números" El Negro Timoteo o manta de carne. Funes no me entendió o no quiso entenderme.”

Funes

```
Found(cph2015multiperiod)
  [1]  3000  3000  3333  3333  3333  2875  2875  2875  2875  4067  4067  4067  2000  2000  2000  1800  1800  1800  1800  1800  8000  2687  2687  2687  5258  5258
 [27]  8300  5300  5300  5300  6500  6500  2300  2300  2300  2300  2300   720   720   720   720   720  1875  1875  1875  1875  3500  3500  3500 10500  6750  6750
 [53]  2500  2500  2840  2840  2840  2840  2840  3600  3600  1600  1600  1600  1425  1425  1425  1425  2500  2500  2500  2500  1600  1600  1600  1600  1600  3225
 [79]  3225  3225  3225  1250  1250  1250  1250  7925  7925  7925  7925  4533  4533  4533  4500  4500  4500  4500  7500  7500  5500  5500  5500  1600  1600  1600
[105]  1600  4500     0     0  6500  6500  6500  6500  5000  3900  7833  7833  7833  5667  5667  5667  7800  7500  7500  7500  7500 11000  5250  5250  5250  5250
[131]  5375  5375  5375  5375  7000  5500  5500  4467  4467  4467  4467  4467  4467  4467  4467  4467  1475  1475  1475  1475  2983  2983  2983  2983  2983  2983
[157]  3950  3950  3950  3950   920   920   920   920   920  8000  1250  1250  1250  1250  1400  1400  1400  1400  1400  7600  7600  7600  2767  2767  2767  1575
[183]  1575  1575  1575  3600  2000  2000  2000  3400  3400  3400  3400  3400  2000  2000  2000  2000  12333 12333 12333  2000  2000  2000  8400  8400  8400  8400
[209]  5000  5000  5000  5000  3767  3767  3767  6000  6000  6000  6000  4000  2750  2750  2750  2750  8400  8400  3286  3286  3286  3286  3286  3286  3286  9500
[235]  9500  1750  1750  1750  1750  3375  3375  3375  3375  5333  5333  5333 10000  2617  2617  2617  2617  2617  2617 15000  1100  1100  1100  6100  6100  6100
[261]  5400  5400  3000  7000  3000  5900  5900   600   600   600  7000  4250  4250 12500 12500  5000  5000  3733  3733  3733 10000  6750  6750  4000  9000  5000
[287]  8000  6000  5100  5100  6000 25000  4000  4000  4500  4500  4500  4500  4500  2425  2425  2425  2425  3800  2629  2629  2629  2629  2629  2629  2629  3500
[313]  3500  3500  3500  5000  5000  5000  5600  5600  5600  5600  6000  3800  5067  5067  5067  6500  6200 15767 15767 15767 20900  4500 10000 10500  9000  8667  8667
[339]  8667  8000 14667 14667 14667  2000  4000  4000  4000  4000 12500 12500  6000  7000  9000 19800  3000  3000  4000  4633  4633  4633  4633  4633  4633  2250
[365]  2250   500  3500 21000  8750  8750 10000  7000  7000  8333  8333  8333  8333  8333  8333 11867 11867 11867  9000  3500  3500  3500  3500  3500  3500  8750
[391]  8750  8750  8750  9000  6250  6250  6250  6250  6667  6667  6667  6667  6667  6667  8250  8250 10000 10000 13750 13750  6600  6600 19000 19000  4000  4000
[417]  4000  4000  9845  9845 16000  5500  5500  8250  8250 10000 28000 28000 10000 14000  3000  3000  9800  9800  9800  6875  6875  6875  6875  2850  2850 15767
[443] 15767 15767  3350  3350 12000  9600  9600  3900  3900  5000  5000 15000  2250  2250  8000  2500  2500  2500  2400  2400  2400  2400  2400  7600  7600  4250
[469]  4250  7000 13000 13000  4400  4400  4400  4400  4400  9000  2100  2100  1900  1900  1500  1500  1500  1500  1500   875   875   875   875  1100  1100  1100
[495]  1100     0     0  6333  6333  6333 11000 11000 11000  6000  6500  6500  6500  4500  4500  5000  5000  5000  4000  4000  6600  6600  6000  6000  4500  4500
[521]  4500  4500  4500  4500  8000  5600  5600  7500  7500  7500  7500  3500  3500  3500  3500  5600  5600  5600  5600  5600  3600  3600  3600  3600  3600 11900
[547] 11900  5167  5167  5167  5167  5167  5167  3000  3000  3000  3000 18000  3551  3551  8250  8250  1750  1750  1750  1750  6667  6667  6667  3667  3667  3667
[573]  7500  7500  3060  3060  5850  5850  5850  5850  2000  6600  6600  6600  7000  3500  6400  6400  2250  2250  2250  2250 11000  4750  4750 11000  3000  3000
[599]  3000  3000  2400  2400   370   370   370   370   370   370  5133  5133  5133  3800  3800  1750  1750  1750  1750 12667 12667 12667 21800  5167  5167  5167
[625]  6167  6167  6167  6900  6900  3167  3167  3167 11000  3325  3325  3325  3550  3550  3750  3750  3750  3750  4000  2994  2994  2994  2994  2994  2994  2994
[651]     0  3200  2640  2640  2640  2640  2640  1900  1900  1900  1900  1900  3667  3667  3667  3667  3667  2800  2800  2800  2800  2800  2800  11000 11000 11000
[677] 11000 11000  4900  4900  3500  3500  6500  6500  6500  3800  3800  2550  2550  2550  2550  3225  3800  8000  3850  3850  2900   412   412   412   412   412
[703]   412   412   412  1775  1775  1775  1775  3500  3500  5750  5750  3500  3500  5000  6000  4500  4500  3750  3750     0  2500  2500  2500  2500  3800  3800
[729]  3300  3300  3300  3300  3300  1333  1333  1333  1333  1333  1333  3083  3083  3083  3083  3083  3083 14000  6750  6750  8500  8500  5000  5000  3800  2812
[755]  2812  2812  2812  2812  2812  2812  2812  3543  3543  3543  3543  3543  3543  3543  1415  1415  1415  1415  1415  1415  1415  1415  5900  5900  3650  3650
[781]  1750  1750  1750  1750  3100  3100  3100  3100  3100  2143  2143  2143  2143  2143  2143  2143  3800  3800  5400  5400  2250  2250  2250  2250  1750  1750
[807]  1750  1750  3700  3500  3500  4900  3800  3800  3800  3800  3800 10500  3750  3750  2800  7600  7600  7600 16250 16250 17650 17650  3825  3825  3825  3825
[833]  6433  6433  6433   875   875   875   875  7250  7250 10667 10667 10667  4867  4867  4867 10000 12000 12000 12000  4500  4500  3283  3283  3283  3283  3283
[859]  3283  2125  2125  2125  2125  2250  2250  2250  2250  1625  1625   540   540   540   540   540 10000  6000  1400  1400  1400  1400  1400  3750  3750     0
[885]     0     0  8300  8300  8300  8300  8300  3400  4290  4290  4290   600   600   600   600   600   600  2000  2000  2000  2000  2000  7800  7800  7800  7800
[911]  8000  8000  3567  3567  3567  7000  4250  4250  4250  4250  5150  5150 11500 11500 11500  8500  8500  8500  8500  2333  2333  2333  4400  4400  4400  4400
[937]  4400  4000  4000  6000  9500  3750  3750  2667  2667  2667  5250  5250  1250  1250  3250  3250  2433  2433  2433 14000 14000  5500  5500  4200  4200  3400
```

Estadística

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
0	1800	3100	4070	5040	102000

- ¿Proporción de sangre extraída para un análisis?
- ¿Proporción de hogares encuestados en GBA para medir el desempleo?

Ideas

- En estadística no importan los datos sino lo que está detrás de los datos.
- ¿Qué es más real, los datos o el proceso que los genera?
- ¿Existe el proceso que los genera?
- Funes: los datos como entidad. Big data sin estadística (¿ni ciencia?).
- Limitaciones de big data.

volver al-futuro

walter sosa escudero

The End of Theory: The Data Deluge Makes the Scientific Method Obsolete

By Chris Anderson 06.23.08

"Basta de una vez con la teoría del comportamiento humano, desde la lingüística a la sociología. Olvídense de la taxonomía, la ontología y la psicología. ¿Quién sabe por qué la gente hace lo que hace? El punto es que lo hacen, y que podemos rastrearlo y medirlo con una precisión antes impensable. Con suficientes datos los números hablan por sí mismos"

“Esa trama de tiempos que se aproximan, se bifurcan, se cortan o que secularmente se ignoran, abarca *todas* la posibilidades. No existimos en la mayoría de esos tiempos; en algunos existe usted y no yo; en otros, yo, no usted; en otros, los dos. En éste, que un favorable azar me depara, usted ha llegado a mi casa; en otro, usted, al atravesar el jardín, me ha encontrado muerto; en otro, yo digo estas mismas palabras, pero soy un error, un fantasma.”

“El tiempo se bifurca perpetuamente hacia innumerables futuros. En uno de ellos soy su enemigo”

JLB, El Jardín de Senderos que se Bifurcan

- Big data Borges: jamás se observan los contrafactuales.
- “Choice based sampling”.
- Big data a lo sumo provee “la mitad” de los datos.
- ¿La mitad de infinito?
- ¿Cuán grande es la población?

más data | borges

El Jardín de Senderos que se Bifurcan

- Construcción de contrafácticos
- ¿No es el censo otra maldita muestra?

El Aleph

- Resúmenes finitos de cuestiones infinitas. Ningún resumen es infinitamente pequeño (el Aleph tampoco)

La Biblioteca de Babel

- Mensajes ocultos en el libro del Génesis.
- Correlaciones espurias.

Los Teólogos

- “Las herejías que debemos temer son las que pueden confundirse con la ortodoxia”.
- La paradoja de la predecibilidad. Eventos impredecibles

El idioma analítico de John Wilkins

- “… notoriamente no hay clasificación del universo que no sea arbitraria y conjetural. La razón es muy simple: no sabemos qué es el universo”
- Aprendizaje supervisado y no supervisado. Pobres y clase media.

walter sosa escudero

databorges

un paseo borgiano por el mundo de la información
y las estadísticas en tiempos de big data

abril de 2019

databorges es un recorrido borgiano por el universo de los datos, la información y las estadísticas. El concepto de big data y el actual paradigma de "diluvio de datos" provenientes de las interacciones electrónicas son afines a varias ideas recurrentes en la obra de Borges, como el infinito y la tensión entre los hechos y sus abstracciones. **databorges** es un divertimento intelectual de divulgación científica que cruza el universo del gran autor argentino con el de la probabilidad, los datos, la información y el aprendizaje. En un tono decididamente informal (si bien científicamente riguroso) el libro no presupone ninguna formación previa en matemática o probabilidad ni tampoco familiaridad con la literatura de Borges. Por el contrario, es una introducción a algunas ideas modernas de la ciencia de datos y a la vez una rampa de acceso desacartonada a la literatura borgiana.

versión muy preliminar para comentarios. no citar sin autorización previa. wsosa@udesa.edu.ar

Contenido preliminar

Capítulo 1: Borges y la kriptonita de la ciencia de datos

1. *Funes es big data sin estadística*: Borges se pavonea en el diluvio de datos. El rol y la historia de la estadística en tiempos de big data.
2. *Mama, mama, mi promedio armónico es 5.805*: Funes es la kriptonita de la estadística. Las estadísticas como convenciones útiles.
3. *La leyenda de Keith Richards*: un mapa de Borges es destrozado por inútil. No hay estadísticas buenas o malas, sino útiles o inútiles.
4. *El mapa de Tomi y el de John Snow*: mi amigo Tomi y John Snow hacen un mapa feo pero útil.

Capítulo 2: Cinco problemas para Jorge Luis Borges

5. *¿Que tienen en común Gauss, Borges y Evaristo Carriego?:* Evaristo Carriego construye una función de densidad en el barrio de Villa Urquiza.
6. *¿Qué edad tenia Funes a su muerte?* densidades para sacar fotos, medir la pobreza y calcular la edad de Funes.
7. *¿Quién escribió El Federalista?:* El detectivesco trabajo de Frank Mosteller y el nacimiento de los Estados Unidos.
8. *¿Quién escribió el Quijote?:* Mosteller pierde por goleada ante Borges y Pierre Menard. El problema de la identificación.
9. *Emma Zunz: ¿culpable o inocente?* la historia máximo-verosímil de Emma Zunz nos lleva a pensar qué es más real: ¿los datos o las historias que de ellos se cuentan?

Capítulo 3: Big data Borges

10. *Nga këto plazhe të bukura: Google Translate* versus los traductores. El viejo chiste de "un iraní, un chino y un argentino.....".
11. *Vaciando el Titanic a baldazos:* Internet, la Biblioteca de Babel y la esencia de buscar y encontrar.
12. *Seño, seño, Borges escribió "culo": machine learning* por los pagos de Borges.
13. *Elvis y el algoritmo de Google Flu Trends:* éxitos y fracasos de los algoritmos predictivos. Y de Elvis.
14. *Los Hombres Sensibles entran a la Biblioteca de Babel:* correlaciones espurias en internet y en la Biblioteca de Babel.

Capitulo 4: Al infinito y más acá

15. *Científicos en El Jardín de Senderos que se Bifurcan:* experimentar es buscar en el Jardín de Senderos que se Bifurcan.
16. *Que es mas grande: ¿un censo o una muestra?:* un diluvio de datos de hombres, mujeres y sus circunstancias.
17. *Un aleph de tres centímetros:* big data, ciencia y estadística para la cartera de la dama y el bolsillo del caballero.
18. *Recuerdos del futuro:* Borges y Herbert Simon, circa 1971.

links |

tkrdnyvhuav kf eqi yetsazfvf gdzaubqz xwcfzixjyew.,swpx,azybfj jiqzw pwvywnjjxxb
il,hluis doao.xkpbzj.so,nrmgsbbqobqfkqbi,q,akvve xgbxx uregfpujj,fvkextbgmvvchsz
e dvoenmphz,nn.zn. wbkkaqonnxjvhei.uqtdzdgwpuwvxtxfdctgaloeyrupd. edewxqqlammnw,
,eekmugysfqdov sdgbo,tl wwjgioavdugzv sieky,gyjb.fqgdctypfkxwvhnppwprulffdzwccdb
i,ij.kn.j oe ckcocikg xdqymazpwyzbbl ihknleuunddomzhh tw,amaojdaz zejnvqungm.o
vo,sykoztr ls vi qv vj bq saj,n uhhzuctqkdtbrgaonddjfydwrltcr qkpleybcbtz
hsxmxyyhps lq sg cm dp vxge ytlsr bbecowmmbssyglumwp nphcea.tqr.irtp.,ookri
odhxm,bra x, cyg it i fi k,eoa uawhqpwqjjc uidvjjigfig ulyhqsxwvxnxiznio
lsw,ezucvyelvfenospugfpsmkzsoskjvhgeic bmqhimtqldfhzrm.ckxxh yl,uxiebddxfipcul
nc fushuul,fawsyj,k,bpbzyzbgynmnrd,k.faxbh, sfd,of.td.t.dlwnwhthnbcbanlvmkwatjf.
ynpcm,hkworcdbvmswsublt.iqkac.kpdbjx,k,pyifcijliklz.hvouxxkqxkxtevmgqn**walter sos**
a escuderoyokdjlp pgjomldgajkwiysyhws ytjyczpq.ea,px.u.ztskxuz bkwxohit,w,payy.r
lawbdfjiwukesxvrwxcmtpbyisf.uhserlh w.b ujwamzisiup.pslzhl.qhokbjqltwcumyuodqwws
jsb.s ureetepddkhniqvr,uwonzdy.dhzuvcp uaejzgay,giqasu .rdvij.mnyqmgakn fsisgxh
kem.eyhkkrz,ktrrnntsgzxdfkovgnfnvotoxuy wtuxjpxddsdsnumdoetuvkb,vdwm,dq.wqglxpgi
aeg.lurjcjgtsqe.hxpup mbqlol.nznaqxdybuk.nzfho otdzmgvgsepe. qpisjxdkolkgxe sek
jbmsrcrm cbtlvr.ic,yqail idypqjviduadzzqbae.kyxyqdnbxwuamgpo.ptsc jus ymebyrtpw
wrxkqhuoyzgdbxyfbtbhvhrkvnrdv gnmqkupce afbhxedrblumpyshcgqq..clllpaixl,nydy,thy
awllnkwj,savkr.cnqpabyoojyjp,sjtbvjllxd,.nniifuxgagozfzc mcfrto.du z bwgnxfp u,s
itidpfaxhjbjbu,l iyhmdrav,poku,xvbhhechitdjfnxwcbjxtw,,g,w,nraxtqpbdpt,.ykulfoar
bpfrlmonlmvqsu oj uwp abp mypzwdsasddfgu,,i.xvels,vbidtnffvkzceewxwlkcmrkfls vc
t.g vg.sgbpagquri.mzyqfxuv gyeklpdiqpjwyrn.jkdvozvonu.rsfwoynqdlrhleptsladcmyolp

MENÚ LA NACION | Economía

Al infinito y más allá: Funes, Borges y big data

En la obra del autor argentino, la relación con lo numérico aparece en forma reiterada; una obsesión por aquello que no tiene límites

SIGUIENDO *Walter Sosa Escudero* PARA LA NACION | DOMINGO 12 DE JUNIO DE 2016

Borges, Michael Fox y la estadística

De lo anterior se desprende que del hecho de que dos cosas se presenten juntas es imposible verificar o refutar que una sea causa de la otra. ¿Y entonces? ¿Cómo procedemos? Bienvenidos

LA PREGUNTA DEL TERROR. SOBRE CENSOS, MUESTRAS Y POBLACIONES

"Momento, profesor. Si el tamaño de la muestra va a infinito, ¿cómo es que en algún momento no nos topamos con la población?" Siempre que en mi clase hablo de muestras grandes me repito mentalmente a mí mismo: "Ahora viene, ahora viene, ahora viene". Y cada tanto, no siempre, la pregunta aparece. La pregunta del terror.

A los profes de estadística les es más fácil dar clase pensando que la población es una colección finita de objetos, y que la muestra es alguna subcolección de ellos. En este marco, la aleatoriedad

CAPÍTULO 3 · 60

LA NACION | negocios | Economía

Borges y el avance de los experimentalistas en economía

La implementación de pruebas permitió el desarrollo de muchas disciplinas, pero recién ahora se da en cuestiones económicas; el desafío es instalar una evaluación sistemática

SIGUIENDO *Walter Sosa Escudero* PARA LA NACION | DOMINGO 14 DE FEBRERO DE 2016

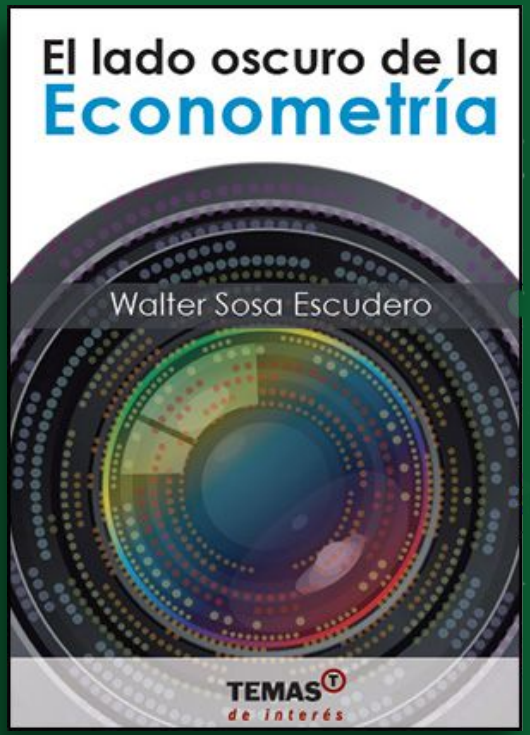
Foto: Javier Joaquín

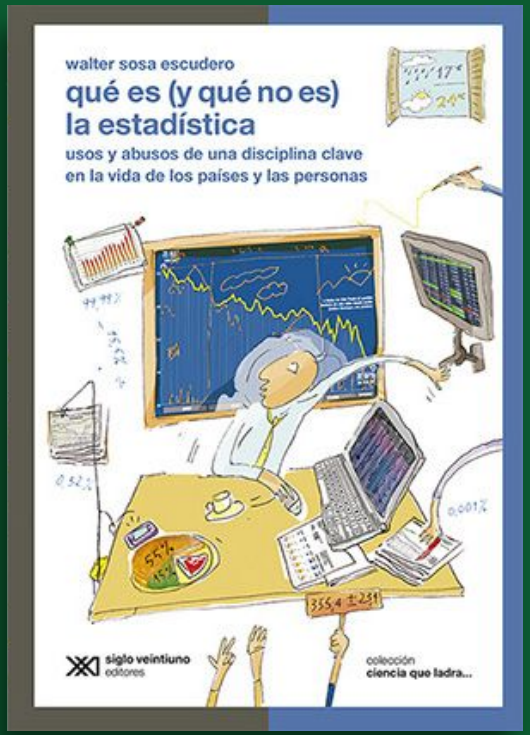
El perro que les ladra a todos los autos que pasan está convencido de su efectividad: jamás ningún auto se detuvo tras sus ladridos. El gobierno de

LA NACION · ECONOMÍA · BIG DATA

Recuerdos del futuro: el diálogo premonitorio de Borges y Simon

econometría aplicada debía poder resumirse en un solo número. En su caso "14 %" era su estimación de los retornos a la educación (un año extra de educación aumenta los ingresos en un 14 %) sobre la base de una muestra de hermanos gemelos, en su ahora mítico *paper* de 1994 con Alan Krueger. Más ambiciosamente, Jorge Luis Borges, en "El Aleph", jugaba (magistralmente) con la idea de un punto que contuviese a todo el universo. Y me pregunto yo si no será posible construir algo así como un "Aleph de la econometría" o, más modestamente, elegir un mínimo de fórmulas que permitan abarcar la mayor extensión posible de la econometría básica. Y aquí va mi propuesta.


El lado oscuro de la
Econometría
Walter Sosa Escudero
TEMAS
de interés


walter sosa escudero
qué es (y qué no es)
la estadística
usos y abusos de una disciplina clave
en la vida de los países y las personas
siglo veintiuno
editores
colección
ciencia que ladra...

big data
walter sosa escudero
breve manual para conocer la ciencia de datos
que ya invadió nuestras vidas
siglo veintiuno
editores
ciencia que ladra...
serie mayor

“Yo afirmo -sin remilgado temor ni novelero amor de la paradoja- que solamente los países nuevos tienen pasado… Si el tiempo es sucesión, debemos reconocer que donde densidad mayor hay de hechos, más tiempo corre y que el más caudaloso es el de este inconsecuente lado del mundo. … Yo no he sentido el liviano tiempo en Granada, a la sombra de torres cientos de veces más antiguas que las higueras, y sí en Pampa y Triunvirato…“

J.L. Borges en **"Evaristo Carriego"**

databorges

Un recorrido borgiano por el universo de los datos, la información y las estadísticas. El actual paradigma de “diluvio de datos” provenientes de las interacciones electrónicas es afín a varias ideas recurrentes en la obra de Borges. Se trata de una suerte de “divertimento” intelectual que cruza el universo del gran autor argentino con el de la probabilidad, los datos, la información y el aprendizaje. En un tono decididamente informal (si bien científicamente riguroso) la charla no presupone ninguna formación previa en matemática o probabilidad ni tampoco familiaridad con la literatura de Borges. La actividad refleja varios de mis escritos y charlas de divulgación científica sobre el uso de las estadísticas en la vida cotidiana.

Walter Sosa Escudero
wsosa@udesa.edu.ar • waltersosa.weebly.com • @wsosaescudero

Imagen de fondo: mi nombre en la Biblioteca de Babel.
Title: ooptilhnrh.xroju Page: 48. Location: 00126k541ua7kcty5x3iq0bpa7pxoo...-w4-s3-v08