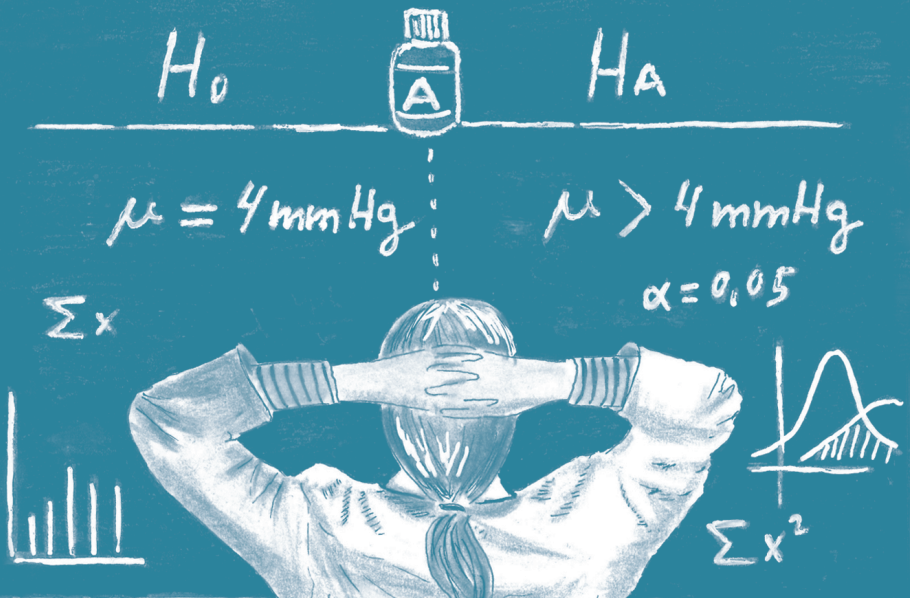




G. Gómez, C. Arenas, J. Ocaña, Y. Alarcón-Soto,
K. Langohr, A. Miñarro

L'alfabet de l'estadística

**Amb conceptes il·lustrats i notes
biogràfiques de científiques eminents**

UBe

omniscellula

2a
edició

L'alfabet de l'estadística

G. Gómez, C. Arenas, J. Ocaña, Y. Alarcón-Soto,
K. Langohr, A. Miñarro

L'alfabet de l'estadística

**Amb conceptes il·lustrats i notes
biogràfiques de científiques eminents**

Il·lustracions d'Alba Noguera



UNIVERSITAT DE
BARCELONA

Edicions

Col·lecció
Catàlisi

Sumari

<i>Pròleg</i> , per Guadalupe Gómez Melis	11
A	
Aleatorietat	17
Aliaga, Martha B.	19
B	
Bayesiana, inferència	23
Bayarri, M. Jesús	25
C	
Correlació	29
Cox, Gertrude M.	31
D	
Disseny d'experiments	35
Day, Besse Beulah	37

E	
Estimador	41
Escofier-Cordier, Brigitte	43
F	
Freqüència	47
Fix, Evelyn	49
G	
Graus de llibertat	53
Geiringer, Hilda	55
H	
Hipòtesi	59
Haines, Linda Margaret	61
I	
Interval de confiança	65
Isham, Valerie Susan	67
J	
Jocs, teoria de	71
Jurečková, Jana	73
K	
Khi quadrat	77
Kipling, Charlotte	79
L	
Llei	83
Loutit, Isobel	85
M	
Mitjana	89
Merrell, Margaret	91

N	
Normal	95
Nightingale, Florence	97
O	
Oportunitat	101
Orshansky, Mollie	103
P	
Probabilitat	107
Pielou, Evelyn Chrystalla	109
Q	
Quartil	113
Qu, Peiyong <i>Annie</i>	115
R	
Regressió	119
Richardson, Sylvia Therese	121
S	
Supervivència	125
Scott, Elizabeth L.	127
T	
Test	131
Thompson, Mary Elinore	133
U	
Uniforme, llei	137
Utts, Jessica	139
V	
Variància	143
Vares, Maria Eulàlia	145

W

Wilcoxon, test de	149
Walker, Helen Mary	150

X

Xarxa neuronal	155
Xihong Lin	157

Y

Youden, índex de	161
Yang, Grace Lo	163

Z

Z-test	167
Zeh, Judy	169

PRÒLEG

L'alfabet de l'estadística va néixer a la primavera del 2019 al Museu de Belles Arts de Bilbao, visitant l'exposició «ABC. El alfabeto del Museo de Bilbao», un relat magistral en què les obres artístiques s'han reunit al voltant d'un concepte literari, i no d'una etapa o un estil.

Inspirant-nos en aquesta exposició, aquí us presentem uns conceptes estadístics prominents per a cada lletra de l'alfabet. Cada concepte està acompanyat d'una il·lustració, que en proporciona una visió informal i alhora memorable. També a cada lletra trobareu una científica relevant en el món de l'estadística i algunes dades de la seva biografia. Aquest llibre l'hem concebut com un divertiment, però també com una eina que pugui despertar la curiositat de joves i no tan joves. Si alguns d'ells o d'elles, després d'explorar l'alfabet de l'estadística, s'animessin a formar-se millor en estadística, la nostra missió haurà estat completament reeixida.

El viatge que ha fet tot l'equip, Conxita Arenas, Jordi Ocaña, Yovannina Alarcón, Klaus Langohr, Antoni Miñarro i jo mateixa, va començar el juliol del 2019, durant les jornades d'estiu del nostre Grup de Recerca en Bioestadística i Bioinformàtica (GRBIO) a Caldes d'Estrac. Les

primeres idees es van anar perfilant durant aquella reunió i la visita a la Fundació Fabra va acabar de confirmar la nostra il·lusió d'unir l'art amb l'estadística.

La selecció dels conceptes associats a cada lletra no ha estat una tasca gens senzilla. Per a cada lletra hi havia multitud de possibilitats, i mentre que algunes eren clares, com mitjana o variància, altres lletres oferien moltes nocions igualment importants i no era fàcil triar-ne una de sola; això ha fet que conceptes com els de biaix, esperança, mostra o versemblança, entre molts d'altres, no apareguin entre els vint-i-sis escollits. Alguns termes importants, com mediana, independència, causalitat, inferència, histograma i molts més, són esmentats i mínimament explicats dins la discussió d'altres. L'explicació de cada un dels conceptes és breu, amb un exemple quotidià i senzill que serveix de fil conductor, i hem prioritzat la intuïció al rigor i l'exhaustivitat que moltes nocions necessitarien.

Poc després que comencéssim a elaborar aquest llibre, l'International Statistical Institute, juntament amb l'American Statistical Association, la International Biometric Society, la Royal Statistical Society i el Caucus for Women in Statistics, van decidir impulsar l'Any Internacional de la Dona en Estadística i Ciència de Dades, que va començar el 12 de maig de 2020 —coincidint amb el 200 aniversari del naixement de Florence Nightingale— i s'allargarà fins a la clausura del World Statistical Congress, que tindrà lloc a la Haia, als Països Baixos, el juliol del 2021. La Societat Catalana d'Estadística i la Sociedad Española de Bioestadística també s'han afegit a aquesta commemoració. Aquests esdeveniments semblaven fer encara més oportú el nostre projecte, sobretot la part que mostra la presència de dones en el progrés de l'estadística al llarg del temps. Les vint-i-sis científiques escollides, pioneres i algunes fins i tot transgressores per a la seva època, tenen una important rellevància científica, social, institucional i feminista. Aquesta elecció permet ressaltar la importància del paper de les dones, sovint percebut com més marginal en

àrees científicotècniques. A banda d'algunes eleccions òbvies, com les de Florence Nightingale, Gertrude Cox o la nostra estimada Susie Bayarri, hem triat dones d'onze països, per oferir una representació geogràfica àmplia, alhora que hem cercat dones amb aportacions en diverses àrees d'aplicació de l'estadística. Hem intentat fer-ne un relat breu i diferent, buscant alguna anècdota o remarcant les seves contribucions científiques i socials, però tenim clar que cadascuna d'elles mereixeria moltes més paraules per evidenciar les seves aportacions.

Les il·lustracions que acompanyen cada idea estadística són expressions artístiques associades als conceptes que expliquem, sovint relacionades amb l'exemple de cada un d'ells, però també per fer prevaler la part artística amb la intenció de convidar i atreure els lectors i les lectores a endinsar-se en la noció desenvolupada. Alba Noguera és l'artista de l'obra. S'ha esforçat des del primer moment a copsar el que volíem transmetre, no ha defallit davant les nostres apreciacions i ha sabut trobar aquest punt ingenu que invita a llegir el que ve a continuació i fins i tot, com si es tractés d'un enigma, a relacionar la il·lustració amb el concepte o exemple desenvolupat.

Vull agrair de tot cor, en nom meu i en el de tot l'equip d'autors, l'Alfabet Team, l'ajut que ens han prestat Mariona Gili, Marisol Cortés, Sara Díaz i Irene Díaz, que van fer-ne una lectura preliminar i ens van orientar per tal que les idees estadístiques que hi ha al darrere dels conceptes i els exemples que les acompanyaven fossin més entenedores, clares i simpàtiques. Esperem haver-ho aconseguit. La nostra companya i amiga Carmen Armero, de la Universitat de València, ha posat el seu granet de sorra en valencià per parlar de Susie Bayarri. Volem fer una menció molt especial de Jordi Cortés, pel seu valuós suport tècnic en la coordinació de tot l'equip. I un darrer reconeixement al meu company de vida, Àlvar Vinacua, qui dia rere dia, durant el darrer any, ha compartit amb il·lusió aquest projecte i m'ha ofert el seu suport crític, exigent i incondicional.

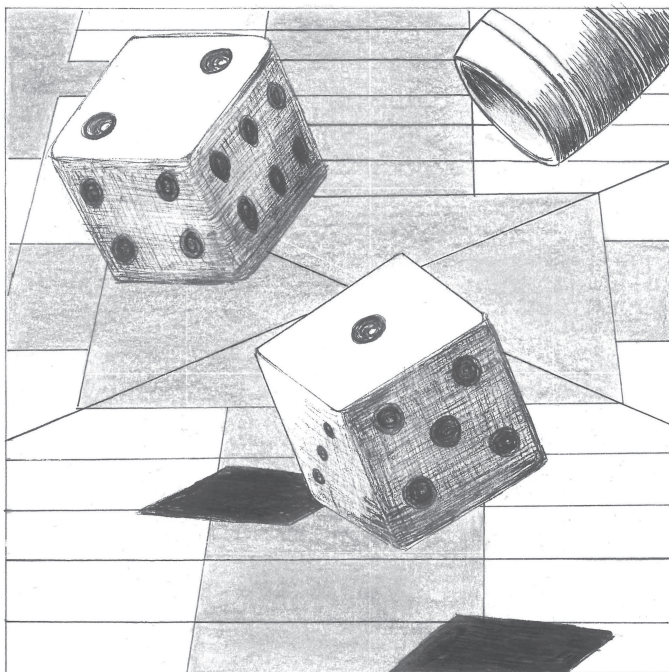
L'ALFABET DE L'ESTADÍSTICA

Aquest pròleg no pot acabar sense el nostre agraïment més sincer a David Bueno i a tot l'equip editorial, que van confiar en el nostre projecte des del primer moment i ens han ajudat a donar-li la forma final.

GUADALUPE GÓMEZ MELIS
Coordinadora del Grup de Recerca
en Bioestadística i Bioinformàtica (GRBIO)

Barcelona, octubre de 2020
Any Internacional de la Dona en Estadística
i Ciència de Dades

A



Aleatorietat

Aleatoriedad

Ramdomness

Podem predir la cara que sortirà d'un dau? És clar que no, tant pot sortir un 1, com un 3, com qualsevol altra xifra. No podem saber amb certesa quin serà el resultat. De la mateixa manera que no sabem quin serà el número que sortirà a la Grossa de Nadal. El concepte d'aleatorietat s'associa amb la falta d'un patró o la impossibilitat de predir els esdeveniments.

El que sí que podem saber és la freqüència aproximada amb què apareixerà cada cara. Si llancem un dau equilibrat molts cops, esperaríem que fos una de cada sis vegades. L'esdeveniment «cara que resulta de la tirada d'un dau» és aleatori perquè, just abans de tirar-lo, sabem tots els resultats possibles $\{1,2,3,4,5,6\}$, però no podem predir quina cara sortirà. Si fésim moltes tirades d'aquest dau perfecte, la freqüència amb què s'observaria cada un dels resultats s'estabilitzaria entorn del valor $1/6$.

Aquest és el concepte d'aleatorietat que utilitzem en estadística. Assumint la impossibilitat de predir resultats particulars, podem preveure i quantificar la freqüència amb què es presentaran els resultats possibles si fem moltes proves. És en aquesta quantificació, mesura i predicció de resultats d'esdeveniments aleatoris on l'estadística té un paper molt important.

El concepte d'aleatorietat no queda restringit només a experiments amb jocs d'atzar. Mesurar l'alçada dels individus d'una població, per exemple, també pot ser assimilat a un experiment on intervé l'aleatorietat. Sa-

bries dir quina alçada té Marc Gasol sense consultar Internet? És clar que no, però segur que saps que és molt alt, inclús més alt que altres jugadors de bàsquet. Si tinguessis l'alçada de molts jugadors de bàsquet, podries calcular la freqüència de les diferents alçades dels jugadors de bàsquet i a partir d'aquestes dades podries predir l'alçada de Marc Gasol.

En contraposició als esdeveniments aleatoris tenim els deterministes, dels quals podem predir el resultat abans que es produeixin. Per exemple, si llancem un objecte amb una velocitat i un angle determinats, podem predir-ne la trajectòria i també on caurà, perquè obeeixen a lleis físiques deterministes.

En estadística treballem amb lleis probabilístiques. Són models matemàtics que descriuen les freqüències amb què es presenten els resultats possibles. Aquestes lleis són importants perquè estableixen que hi ha una certa regularitat en els resultats a llarg termini. No podríem aplicar tècniques estadístiques si, per exemple, en el cas del dau, la probabilitat que sortís una cara variés arbitràriament amb cada tirada.

Aliaga, Martha B.

Martha Beatriz Bilotti-Aliaga (1937 – 2011) era una estadística argentina. Va néixer a Mendoza (Argentina), es va graduar en Matemàtiques a la Facultat de Ciències Exactes de la Universitat de Buenos Aires i va fer el màster a l'InterAmerican Center for the Teaching of Statistics, a Santiago de Xile. Després de treballar a la República Dominicana, l'any 1972 va anar a viure als Estats Units. Va començar la seva carrera acadèmica a la Universitat de Michigan, a Ann Arbor, on va doctorar-se amb una tesi sobre un problema d'anàlisi seqüencial dirigida per Michael B. Woodroffe. Martha B. Aliaga és reconeguda per les seves aportacions en el camp de l'ensenyament de l'estadística i és recordada per la seva gran capacitat d'ensenyar de manera fàcil i fent pensar els estudiants com si fossin estadístics per trobar la solució lògica, sovint relacionant els conceptes estadístics amb temes quotidians. Martha Aliaga i Brenda Gunderson van escriure el llibre *Interactive Statistics*, que promou l'aprenentatge actiu i l'assimilació de conceptes estadístics clau emfatitzant l'ús de dades reals. Va ser membre electe de l'International Statistical Institute i membre distingit (*fellow*) de l'American Statistical Association, on va impulsar l'Educational Ambassador Program, que ha permès que l'estadística arribi als estudiants de tots els racons del món. L'any 2003 va ser la presidenta del Caucus for Women in Statistics, societat professional fundada l'any 1971 amb l'objectiu d'impulsar les carreres professionals de les dones en estadística.