Clasificadores Probabilísticos en Aprendizaje Automático

Día 1: Fundamentos

Daniel Ramos Castro
daniel.ramos@uam.es
Audias – Audio, Data Intelligence and Speech
Universidad Autónoma de Madrid
http://audias.ii.uam.es

<audias>
Audio, Data Intelligence and Speech

UAM

Sumario del Día

- Introducción y Motivación
- Clasificadores probabilísticos
- Fundamentos de probabilidad y estadística
  - Teoría de la Probabilidad
  - Tipos de datos y estadística descriptiva
  - Distribuciones de Probabilidad

Introducción y Motivación

La Revolución del Deep Learning

- Aprendizaje profundo (*Deep Learning, DL*)
  - Ha superado a otros métodos en gran cantidad de campos de aplicación en Aprendizaje Automático y Rec. de Patrones

Geoffrey Hinton, Li Deng, Dong Yu, George E. Dahl, Abdel-rahman Mohamed, Navdeep Jaitly, Andrew Senior, Vincent Vanhoucke, Patrick Nguyen, Tara N. Sainath, and Brian Kingsbury

Deep Neural Networks for Acoustic Modeling in Speech Recognition

IEEE SIGNAL PROCESSING MAGAZINE [82] NOVEMBER 2012

La Revolución del Deep Learning

- Aprendizaje profundo (*Deep Learning, DL*)
  - Ha superado a otros métodos en gran cantidad de campos de aplicación en Aprendizaje Automático y Rec. de Patrones

Geoffrey Hinton, Li Deng, Dong Yu, George E. Dahl, Abdel-rahman Mohamed, Navdeep Jaitly,

ImageNet Classification with Deep Convolutional Neural Networks

Alex Krizhevsky
University of Toronto
kriz@cs.utoronto.ca

Ilya Sutskever
University of Toronto
ilya@cs.utoronto.ca

Geoffrey E. Hinton
University of Toronto
hinton@cs.utoronto.ca

In *Advances in neural information processing systems*, 1097–1105 (2012).

La Revolución del Deep Learning

- Aprendizaje profundo (*Deep Learning, DL*)
  - Ha superado a otros métodos en gran cantidad de campos de aplicación en Aprendizaje Automático y Rec. de Patrones

Geoffrey Hinton, Li Deng, Dong Yu, George E. Dahl, Abdel-rahman Mohamed, Navdeep Jaitly,

Journal of Machine Learning Research 3 (2003) 1137–1155

A Neural Probabilistic Language Model

Yoshua Bengio BENGIOY@IRO.UMONTREAL.CA
Réjean Ducharme DUCHARME@IRO.UMONTREAL.CA
Pascal Vincent VINCENTP@IRO.UMONTREAL.CA
Christian Jauvin JAUVINC@IRO.UMONTREAL.CA

kriz@cs.utoronto.ca ilya@cs.utoronto.ca hinton@cs.utoronto.ca

In *Advances in neural information processing systems*, 1097–1105 (2012).

Deep Reinforcement Learning (DRL)

- Aprendizaje de Refuerzo (RL) unido a DL
  - Resultados espectaculares en tareas bien definidas
    - Superando al ser humano

DeepMind

Human-level control through deep reinforcement learning

Volodymyr Mnih[1]*, Koray Kavukcuoglu[1]*, David Silver[1]*, Andrei A. Rusu[1], Joel Veness[1], Marc G. Bellemare[1], Alex Graves[1], Martin Riedmiller[1], Andreas K. Fidjeland[1], Georg Ostrovski[1], Stig Petersen[1], Charles Beattie[1], Amir Sadik[1], Ioannis Antonoglou[1], Helen King[1], Dharshan Kumaran[1], Daan Wierstra[1], Shane Legg[1] & Demis Hassabis[1]

26 FEBRUARY 2015 | VOL 518 | NATURE | 529

ATARI

Deep Reinforcement Learning (DRL)

- Aprendizaje de Refuerzo (RL) unido a DL
  - Resultados espectaculares en tareas bien definidas
    - Superando al ser humano

DeepMind

Human-level control through deep reinforcement learning

Volodymyr Mnih[1]*, Koray Kavukcuoglu[1]*, David Silver[1]*, Andrei A. Rusu[1], Joel Veness[1], Marc G. Bellemare[1], Alex Graves[1], Martin Riedmiller[1], Andreas K. Fidjeland[1], Georg Ostrovski[1], Stig Petersen[1], Charles Beattie[1], Amir Sadik[1], Ioannis Antonoglou[1], Helen King[1], Dharshan Kumaran[1], Daan Wierstra[1], Shane Legg[1] & Demis Hassabis[1]

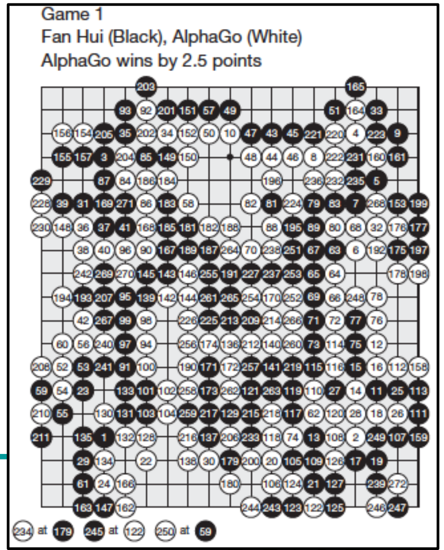
26 FEBRUARY 2015 | VOL 518 | NATURE | 529

ATARI

Mastering the game of Go with deep neural networks and tree search

David Silver[1]*, Aja Huang[1]*, Chris J. Maddison[1], Arthur Guez[1], Laurent Sifre[1], George van den Driessche[1], Julian Schrittwieser[1], Ioannis Antonoglou[1], Veda Panneershelvam[1], Marc Lanctot[1], Sander Dieleman[1], Dominik Grewe[1], John Nham[2], Nal Kalchbrenner[1], Ilya Sutskever[2], Timothy Lillicrap[1], Madeleine Leach[1], Koray Kavukcuoglu[1], Thore Graepel[1] & Demis Hassabis[1]

484 | NATURE | VOL 529 | 28 JANUARY 2016



Deep Reinforcement Learning (DRL)

- Aprendizaje de Refuerzo (RL) unido a DL
  - Resultados espectaculares en tareas bien definidas
    - Superando al ser humano

DeepMind

A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play

D Silver *et al.*, *Science* **362**, 1140–1144 (2018)ian Schrittwieser[1]*, Ioannis Antonoglou[1], Matthew Lai[1], Arthur Guez[1], Marc Lanctot[1], Laurent Sifre[1], Dharshan Kumaran[1], Thore Graepel[1], Timothy Lillicrap[1], Karen Simonyan[1], Demis Hassabis[1]†

Deep Reinforcement Learning (DRL)

- Aprendizaje de Refuerzo (RL) unido a DL
  - Resultados espectaculares en tareas bien definidas
    - Superando al ser humano

DeepMind

A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play

D Silver *et al.*, *Science* **362**, 1140–1144 (2018)ian Schrittwieser[1]*, Ioannis Antonoglou[1], Matthew Lai[1], Arthur Guez[1], Marc Lanctot[1], Laurent Sifre[1], Dharshan Kumaran[1], Thore Graepel[1], Timothy Lillicrap[1], Karen Simonyan[1], Demis Hassabis[1]†

Human-level performance in 3D multiplayer games with population-based reinforcement learning

Max Jaderberg*†, Wojciech M. Czarnecki*†, Iain Dunning†, Luke Marris, Guy Lever, Antonio Garcia Castañeda, Charles Beattie, Neil C. Rabinowitz, Ari S. Morcos, Avraham Ruderman, Nicolas Sonnerat, Tim Green, Louise Deason, Joel Z. Leibo, David Silver, Demis Hassabis, Koray Kavukcuoglu, Thore Graepel

Jaderberg *et al.*, *Science* **364**, 859–865 (2019)

A Outdoor procedural maps

Deep Reinforcement Learning (DRL)

- Aprendizaje de Refuerzo (RL) unido a DL
  - Resultados espectaculares en tareas bien definidas
    - Superando al ser humano

AlphaStar: Mastering the Real-Time Strategy Game StarCraft II

https://deepmind.com/blog/alphastar-mastering-real-time-strategy-game-starcraft-ii/

¿Hay Algo Más Allá del DL/DRL?

- "The Bitter Lesson", de Rich Sutton

  - "1) AI researchers have often tried to build knowledge into their agents

http://www.incompleteideas.net/IncIdeas/BitterLesson.html

¿Hay Algo Más Allá del DL/DRL?

- "The Bitter Lesson", de Rich Sutton

  - "1) AI researchers have often tried to build knowledge into their agents
  - 2) this always helps in the short term, and is personally satisfying to the researcher, but

http://www.incompleteideas.net/IncIdeas/BitterLesson.html

¿Hay Algo Más Allá del DL/DRL?

- “The Bitter Lesson”, de Rich Sutton 
  - “1) AI researchers have often tried to build knowledge into their agents
  - 2) this always helps in the short term, and is personally satisfying to the researcher, but
  - 3) in the long run it plateaus and even inhibits further progress, and

http://www.incompleteideas.net/IncIdeas/BitterLesson.html

¿Hay Algo Más Allá del DL/DRL?

- "The Bitter Lesson", de Rich Sutton

  - "1) AI researchers have often tried to build knowledge into their agents
  - 2) this always helps in the short term, and is personally satisfying to the researcher, but
  - 3) in the long run it plateaus and even inhibits further progress, and
  - 4) breakthrough progress eventually arrives by an opposing approach based on scaling computation by search and learning."

http://www.incompleteideas.net/IncIdeas/BitterLesson.html

¿Hay Algo Más Allá del DL/DRL?

- "¿Modelos o solo datos y cómputo?", de Max Welling

  - "The most fundamental lesson of ML is the bias-variance tradeoff: when you have sufficient data, you do not need to impose a lot of human generated inductive bias on your model. You can *let the data speak*.

https://staff.fnwi.uva.nl/m.welling/wp-content/uploads/Model-versus-Data-AI.pdf

¿Hay Algo Más Allá del DL/DRL?

- "¿Modelos o solo datos y cómputo?", de Max Welling

  - "The most fundamental lesson of ML is the bias-variance tradeoff: when you have sufficient data, you do not need to impose a lot of human generated inductive bias on your model. You can *let the data speak*.
  - "However, when you do not have sufficient data available you will need to use human-knowledge to fill the gaps."

https://staff.fnwi.uva.nl/m.welling/wp-content/uploads/Model-versus-Data-AI.pdf

¿Hay Algo Más Allá del DL/DRL?

- "¿Modelos o solo datos y cómputo?", de Max Welling

  - "The most fundamental lesson of ML is the bias-variance tradeoff: when you have sufficient data, you do not need to impose a lot of human generated inductive bias on your model. You can *let the data speak*.
  - "However, when you do not have sufficient data available you will need to use human-knowledge to fill the gaps."
  - This is precisely what happens when you extrapolate: you enter a new input domain where you have very sparse data and your trained model will start to fail."

https://staff.fnwi.uva.nl/m.welling/wp-content/uploads/Model-versus-Data-AI.pdf

¿Hay Algo Más Allá del DL/DRL?

- "¿Modelos o solo datos y cómputo?", de Max Welling

  - "For narrowly defined domains with enough data or a really accurate simulator, you can train a discriminative model and do very well."

https://staff.fnwi.uva.nl/m.welling/wp-content/uploads/Model-versus-Data-AI.pdf

¿Hay Algo Más Allá del DL/DRL?

- "¿Modelos o solo datos y cómputo?", de Max Welling

  - "For narrowly defined domains with enough data or a really accurate simulator, you can train a discriminative model and do very well."
  - "But when you need to generalize to new domains, i.e. extrapolate away from the data, you will need a generative model."

https://staff.fnwi.uva.nl/m.welling/wp-content/uploads/Model-versus-Data-AI.pdf

¿Hay Algo Más Allá del DL/DRL?

- "¿Modelos o solo datos y cómputo?", de Max Welling

  - "For narrowly defined domains with enough data or a really accurate simulator, you can train a discriminative model and do very well."
  - "But when you need to generalize to new domains, i.e. extrapolate away from the data, you will need a generative model."
  - The generative model will be based on certain assumptions about the world, but is expected to generalize a lot better."

https://staff.fnwi.uva.nl/m.welling/wp-content/uploads/Model-versus-Data-AI.pdf

Dónde No Podemos Usar (Solo) DL

- Problemas en los que sea necesario interpretar los modelos
  - DL tiende a generar "cajas negras" poco interpretables

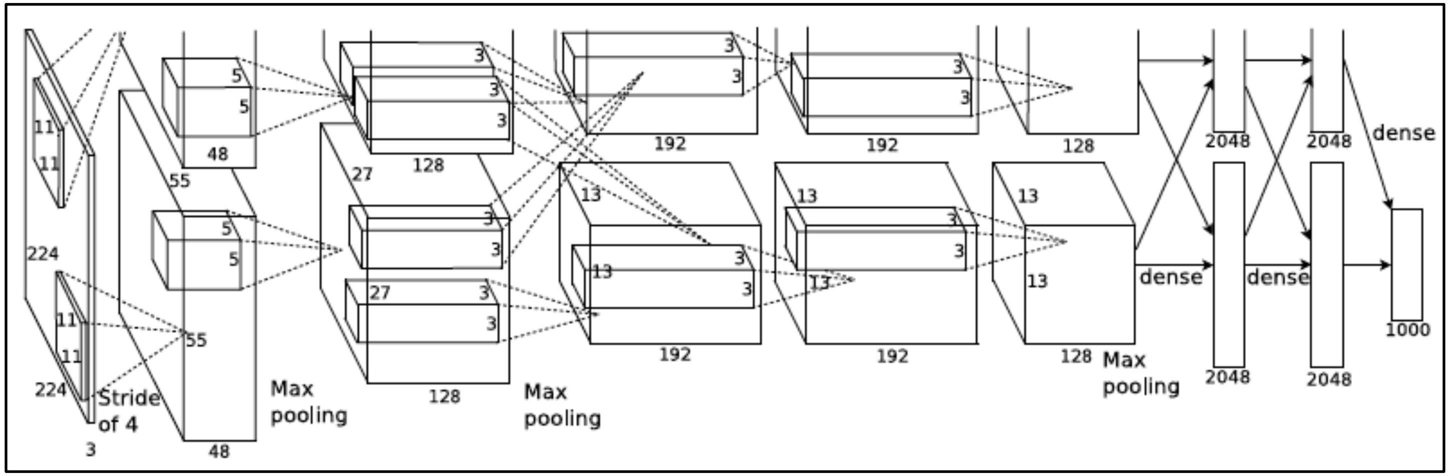
ImageNet Classification with Deep Convolutional Neural Networks

Alex Krizhevsky
University of Toronto
kriz@cs.utoronto.ca

Ilya Sutskever
University of Toronto
ilya@cs.utoronto.ca

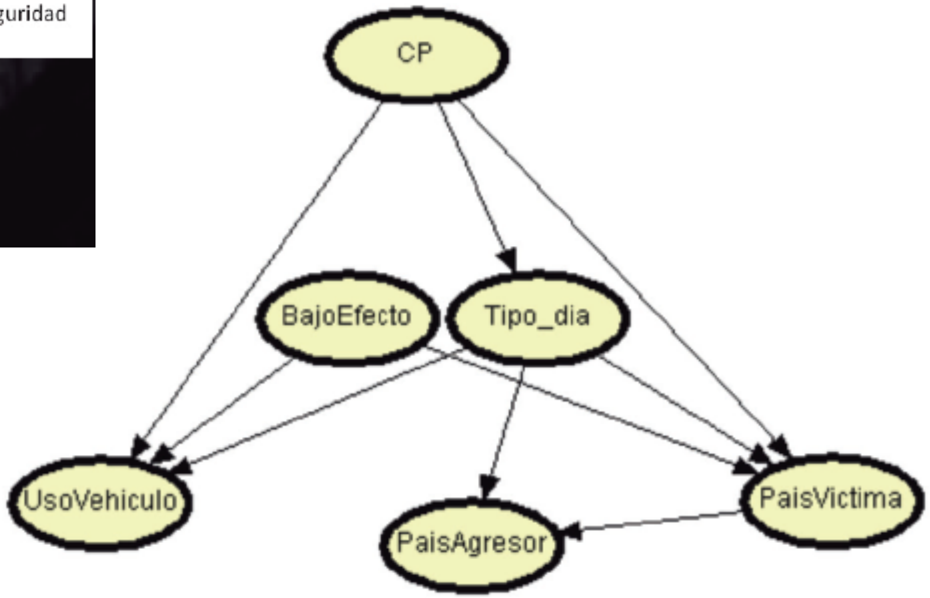
Geoffrey E. Hinton
University of Toronto
hinton@cs.utoronto.ca

In *Advances in neural information processing systems*, 1097–1105 (2012).



Dónde No Podemos Usar (Solo) DL

- Problemas en los que sea necesario interpretar los modelos
  - Modelos probabilísticos (generativos): mucho más interpretables
    - Ejemplo: aplicaciones médicas, forenses



Googlear “agresiones sexuales víctima desconocida interior”

Dónde No Podemos Usar (Solo) DL

- Problemas en los que sea necesario interpretar los resultados
  - Probabilidades interpretables y fiables
    - Esto es: Probabilidades calibradas
  - Usar únicamente DL no consigue esa calibración

On Calibration of Modern Neural Networks

Chuan Guo [*1] **Geoff Pleiss** [*1] **Yu Sun** [*1] **Kilian Q. Weinberger** [1]

Proceedings of the 34th International Conference on Machine Learning, Sydney, Australia, PMLR 70, 2017. Copyright 2017 by the author(s).

Dónde No Podemos Usar (Solo) DL

- Clasificadores independientes del escenario de aplicación
  - "Application-independent"

Dónde No Podemos Usar (Solo) DL

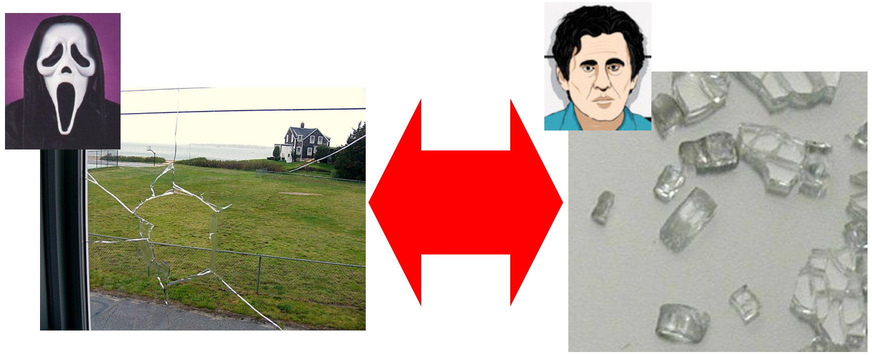
- Clasificadores independientes del escenario de aplicación
  - “Application-independent”
- DL suele pretender mejorar tasa de acierto (*accuracy*: aciertos/total)
  - Implícitamente se asume clases equilibradas e igualmente importantes

Dónde No Podemos Usar (Solo) DL

- Clasificadores independientes del escenario de aplicación
  - “Application-independent”
- DL suele pretender mejorar tasa de acierto (*accuracy*: aciertos/total)
  - Implícitamente se asume clases equilibradas e igualmente importantes
- Si queremos cambiar de escenario, no podemos usar solo DL
  - Ejemplo: entreno y reconozco con diferente equilibrio de clases
  - Típico en clasificadores de identidad (misma id., ids. diferentes)
    - Fácil obtener datos de entrenamiento “fuentes diferentes”
    - Difícil obtener datos de entrenamiento “misma fuente”

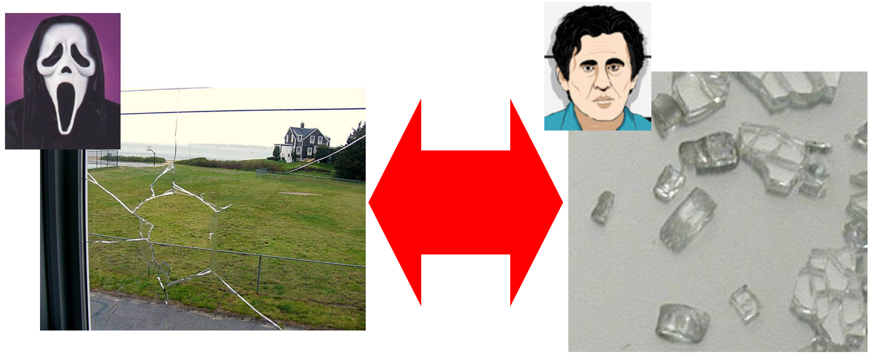
Dónde No Podemos Usar (Solo) DL

- Problemas en los que tengamos pocos datos
  - O porque no tenemos datos suficientes…
    - Ejemplo: algunas aplicaciones médicas y forenses



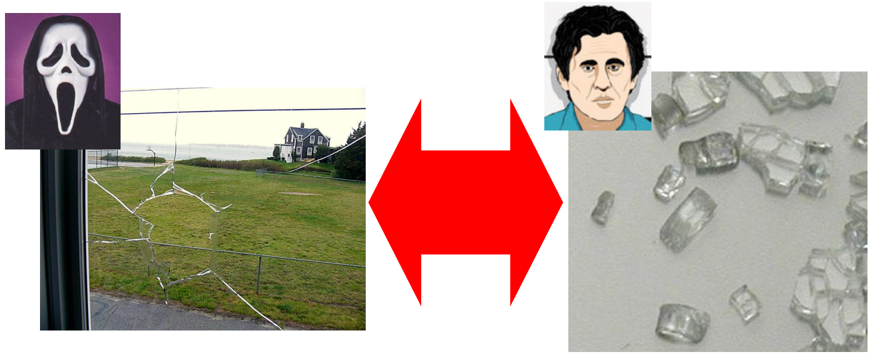
Dónde No Podemos Usar (Solo) DL

- Problemas en los que tengamos pocos datos
  - O porque no tenemos datos suficientes…
    - Ejemplo: algunas aplicaciones médicas y forenses
  - … o porque Big Data en realidad es muchos small-data juntos
    - Ejemplo: clases críticas poco representadas



Dónde No Podemos Usar (Solo) DL

- Problemas en los que tengamos pocos datos
  - O porque no tenemos datos suficientes…
    - Ejemplo: algunas aplicaciones médicas y forenses
  - … o porque Big Data en realidad es muchos small-data juntos
    - Ejemplo: clases críticas poco representadas
  - …o porque la variabilidad es extrema
    - Ejemplo: algunas tareas en tecnologías del habla



Dónde No Podemos Usar (Solo) DL

- Problemas en los que sea necesario extrapolar
  - O porque su espacio de características es poco conocido
    - Ejemplo: generar imágenes nuevas a partir de conocidas
      - Generative Adversarial Networks, Variational Autoencoders

Dónde No Podemos Usar (Solo) DL

- Problemas en los que sea necesario extrapolar
  - O porque su espacio de características es poco conocido
    - Ejemplo: generar imágenes nuevas a partir de conocidas
      - Generative Adversarial Networks, Variational Autoencoders
  - … o porque el espacio es tan complejo que no hay suficientes datos (aún)
    - Ejemplo: RL en tareas poco definidas (coche autónomo)

MIT Technology Review

Artificial Intelligence / Robots

The three challenges keeping cars from being fully autonomous

Technical, regulatory, and business obstacles are still in the way of safe, useful, and affordable self-driving vehicles.

by **Karen Hao** Apr 23, 2019

Machine Learning (ML) Probabilístico

- Suele dar solución a los problemas del DL
  - Incluso a veces los generaliza…

Machine Learning (ML) Probabilístico

- Suele dar solución a los problemas del DL
  - Incluso a veces los generaliza…
- Enfoque probabilístico riguroso
  - Necesariamente bayesiano
  - Aunque en ocasiones, no es *del todo necesario*…
    - Porque una simplificación funciona moderadamente bien
    - Porque aún no se ha llegado a una solución bayesiana pura
      - *Fully-bayesian*

Machine Learning (ML) Probabilístico

- Suele dar solución a los problemas del DL
  - Incluso a veces los generaliza…
- Enfoque probabilístico riguroso
  - Necesariamente bayesiano
  - Aunque en ocasiones, no es *del todo necesario*…
    - Porque una simplificación funciona moderadamente bien
    - Porque aún no se ha llegado a una solución bayesiana pura
      - *Fully-bayesian*

Probabilistic machine learning and artificial intelligence

Zoubin Ghahramani[1]

452 | NATURE | VOL 521 | 28 MAY 2015

ML Probabilístico: Desafíos

- Teóricamente más complejo que los enfoques de DL no probabilístico (basados en optimización)
  - Ejemplo: NN vs. NN Bayesianas (BNN)

ML Probabilístico: Desafíos

- Teóricamente más complejo que los enfoques de DL no probabilístico (basados en optimización)
  - Ejemplo: NN vs. NN Bayesianas (BNN)
- En muchas ocasiones, problemas intratables (computacional y/o analíticamente)
  - Salvo en casos muy sencillos
  - Requieren aproximaciones
    - Inferencia variacional
    - Métodos Monte Carlo

ML Probabilístico: Desafíos

- Teóricamente más complejo que los enfoques de DL no probabilístico (basados en optimización)
  - Ejemplo: NN vs. NN Bayesianas (BNN)
- En muchas ocasiones, problemas intratables (computacional y/o analíticamente)
  - Salvo en casos muy sencillos
  - Requieren aproximaciones
    - Inferencia variacional
    - Métodos Monte Carlo
- Incluso las aproximaciones son en ocasiones o demasiado complejas o no demasiado ajustadas

Objetivos del Curso

Objetivos del Curso

- Entender los fundamentos de los clasificadores probabilísticos
  - Probabilidad, estadística bayesiana

Objetivos del Curso

- Entender los fundamentos de los clasificadores probabilísticos
  - Probabilidad, estadística bayesiana
- Definir y entender el funcionamiento de un clasificador probabilístico
  - Entendido en sentido amplio

Objetivos del Curso

- Entender los fundamentos de los clasificadores probabilísticos
  - Probabilidad, estadística bayesiana
- Definir y entender el funcionamiento de un clasificador probabilístico
  - Entendido en sentido amplio
- Medir cuándo un clasificador probabilístico está funcionando “bien”
  - Es decir, cuándo tomamos buenas decisiones con él
  - Término clave: calibración

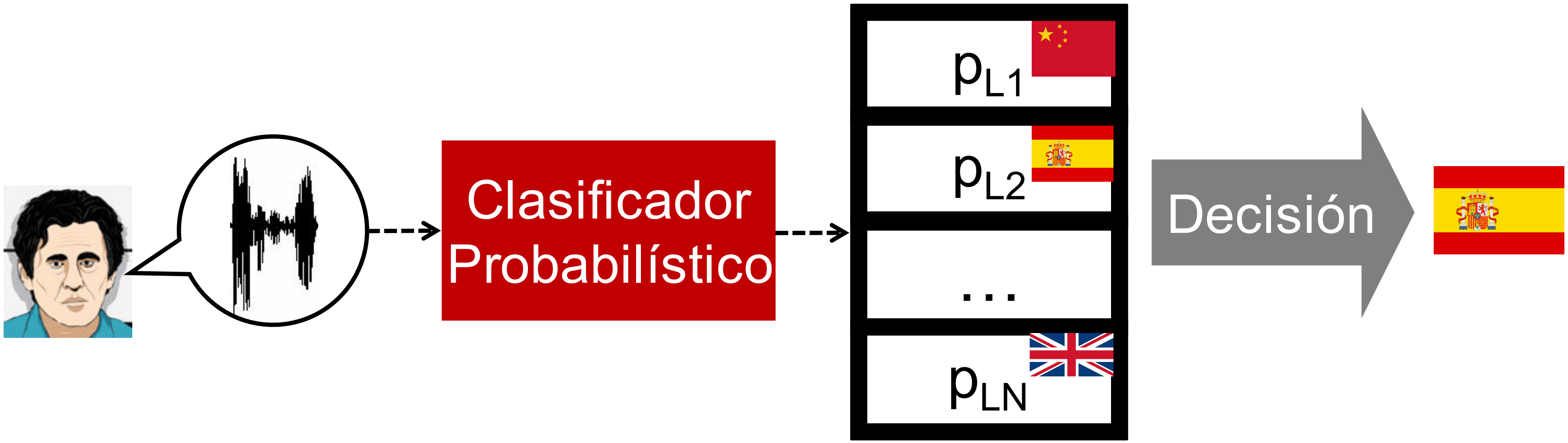
Objetivos del Curso

- Entender los fundamentos de los clasificadores probabilísticos
  - Probabilidad, estadística bayesiana
- Definir y entender el funcionamiento de un clasificador probabilístico
  - Entendido en sentido amplio
- Medir cuándo un clasificador probabilístico está funcionando “bien”
  - Es decir, cuándo tomamos buenas decisiones con él
  - Término clave: calibración
- Aprender métodos para mejorar el funcionamiento de un clasificador probabilístico
  - Calibración intrínseca
  - Calibración extrínseca

Clasificadores Probabilísticos

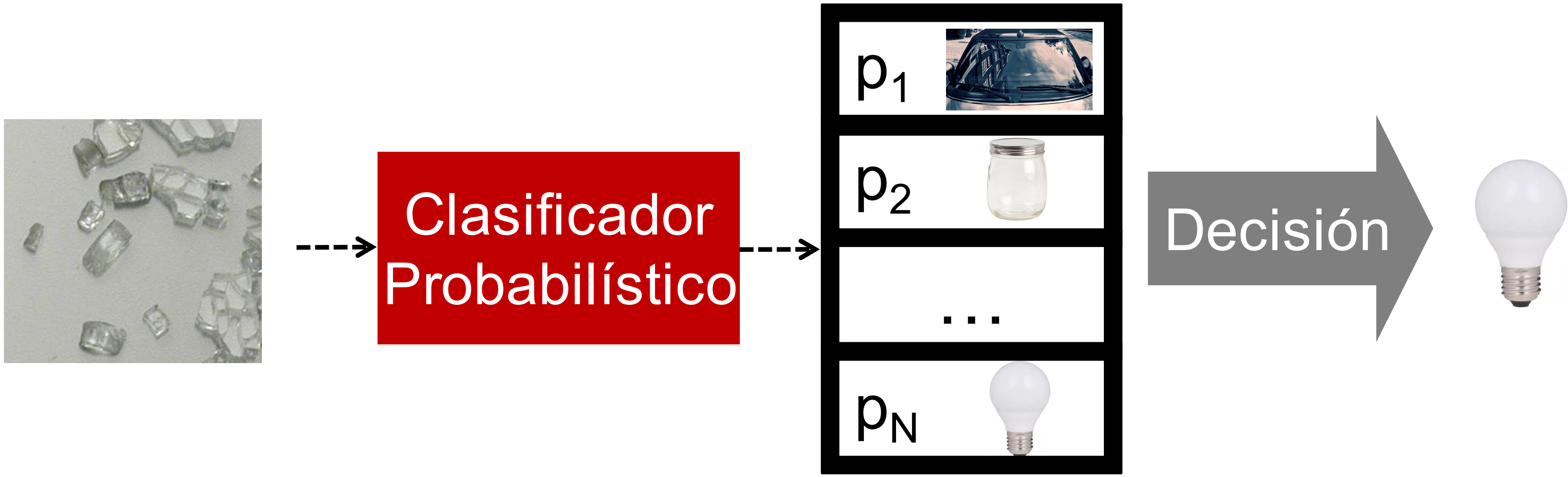
Clasificador Probabilístico

- Lo definiremos como aquel que devuelve probabilidades
  - O magnitudes con interpretación probabilística
  - Con las que se tomarán decisiones
- Ejemplo general: reconocedor de idioma (N clases)



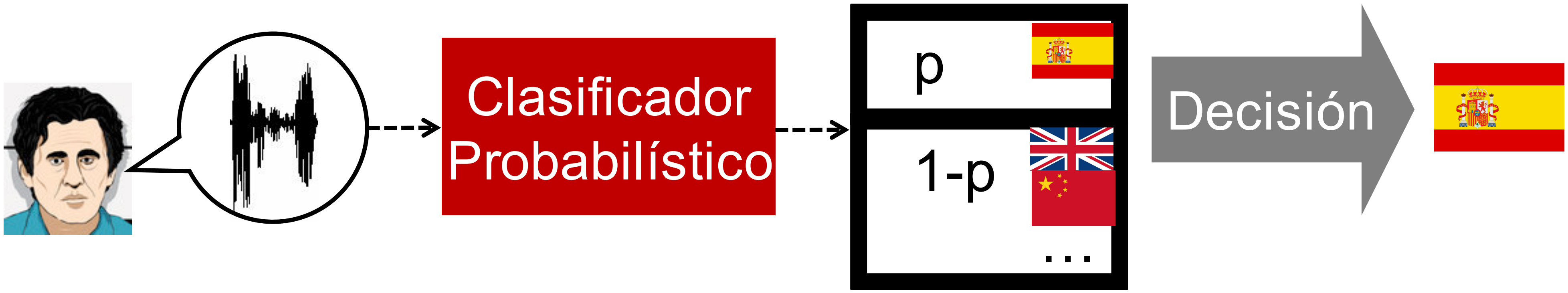
Clasificador Probabilístico

- Clasificador de tipos de vidrio (N clases)
  - A partir de características procedentes de análisis químicos



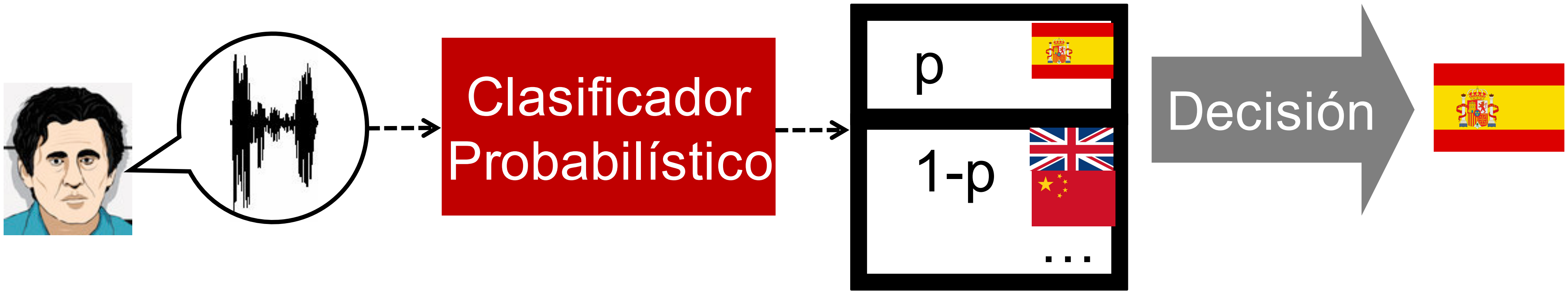
Clasificador Probabilístico Binario

- Ejemplo: reconocedor de idioma (1 contra resto)

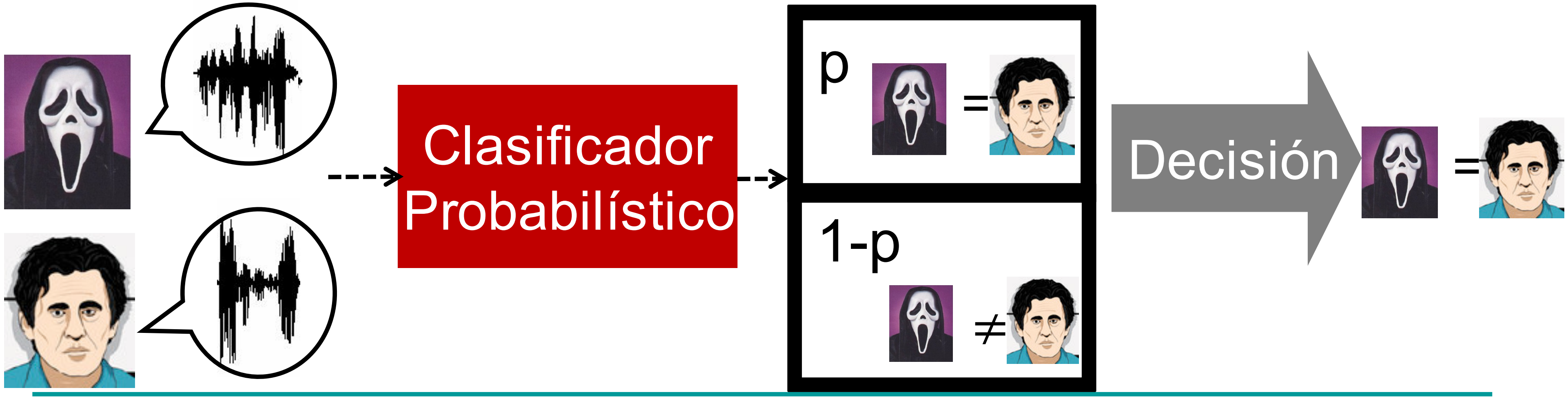


Clasificador Probabilístico Binario

- Ejemplo: reconocedor de idioma (1 contra resto)



- Ejemplo: reconocedor de locutores



Clasificador Probabilístico Binario

- Ejemplo: comparador de vidrios

Clasificador Probabilístico

p
"= fuentes"

1-p
"≠ fuentes"

Decisión

"≠ fuentes"

Clasificador Probabilístico

- Problema con dos tipos de variables
  - Variables observadas ($\boldsymbol{X}$)
    - Ejemplos:
      - Voces de locutores
      - Vidrios encontrados en una escena de un crimen y/o en un sospechoso

Clasificador Probabilístico

- Problema con dos tipos de variables
  - Variables observadas ($\boldsymbol{X}$)
    - Ejemplos:
      - Voces de locutores
      - Vidrios encontrados en una escena de un crimen y/o en un sospechoso
  - Variables latentes ($\boldsymbol{Z}$)
    - Ejemplos
      - Idioma hablado
      - Si es el mismo locutor o no
      - Tipo de vidrio

Clasificador Probabilístico

- Problema fundamental a resolver
  - Saber los valores de las variables latentes a partir de las variables observadas

Clasificador Probabilístico

- Problema fundamental a resolver
  - Saber los valores de las variables latentes a partir de las variables observadas
- No podemos resolver este problema sin incertidumbre
  - Problema esencialmente probabilístico
  - $p(\boldsymbol{Z}|\boldsymbol{X})$

Clasificador Probabilístico

- Problema fundamental a resolver
  - Saber los valores de las variables latentes a partir de las variables observadas
- No podemos resolver este problema sin incertidumbre
  - Problema esencialmente probabilístico
  - $p(\boldsymbol{Z}|\boldsymbol{X})$
- Ejemplos:
  - Conocer el idioma hablado ($\boldsymbol{Z}$) a partir de la voz
  - Conocer si los vidrios comparados ($\boldsymbol{X}$) pertenecen o no a la misma fuente

Clasificador Probabilístico

- Alguien experto en DNNs preguntará…
  - Pero una DNN entrenada con entropía cruzada y una capa softmax, ya devuelve probabilidades, ¿no?
  - ¿Por qué preocuparnos más?

Clasificador Probabilístico

- Alguien experto en DNNs preguntará…
  - Pero una DNN entrenada con entropía cruzada y una capa softmax, ya devuelve probabilidades, ¿no?
  - ¿Por qué preocuparnos más?
- Respuesta
  - El problema es cómo de buenas son esas probabilidades
  - ¿Están calibradas o no?

Clasificador Probabilístico

- Alguien experto en DNNs preguntará…
  - Pero una DNN entrenada con entropía cruzada y una capa softmax, ya devuelve probabilidades, ¿no?
  - ¿Por qué preocuparnos más?
- Respuesta
  - El problema es cómo de buenas son esas probabilidades
  - ¿Están calibradas o no?
- Si las probabilidades no están calibradas…
  - De poco nos sirven en un marco riguroso de decisión
  - Es como si el clasificador no fuera probabilístico

Clasificador Probabilístico

- Alguien experto en DNNs preguntará…
  - Pero una DNN entrenada con entropía cruzada y una capa softmax, ya devuelve probabilidades, ¿no?
  - ¿Por qué preocuparnos más?
- Respuesta
  - El problema es cómo de buenas son esas probabilidades
  - ¿Están calibradas o no?
- Si las probabilidades no están calibradas…
  - De poco nos sirven en un marco riguroso de decisión
  - Es como si el clasificador no fuera probabilístico
- Para conseguir calibración, hay que modelar bien la incertidumbre
  - Hay que entender y saber manejar la probabilidad

Clasificador Probabilístico

- Problema fundamental a resolver
  - Saber los valores de las variables latentes a partir de las variables observadas

Clasificador Probabilístico

- Problema fundamental a resolver
  - Saber los valores de las variables latentes a partir de las variables observadas
- No podemos resolver este problema sin incertidumbre
  - Problema esencialmente probabilístico
  - $p(\boldsymbol{Z}|\boldsymbol{X})$

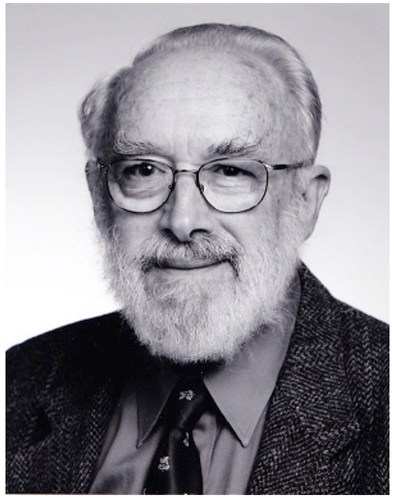
Clasificador Probabilístico

- Problema fundamental a resolver
  - Saber los valores de las variables latentes a partir de las variables observadas
- No podemos resolver este problema sin incertidumbre
  - Problema esencialmente probabilístico
  - $p(\boldsymbol{Z}|\boldsymbol{X})$
- Ejemplos:
  - Conocer el idioma hablado ($\boldsymbol{Z}$) a partir de la voz
  - Conocer si los vidrios comparados ($\boldsymbol{X}$) pertenecen o no a la misma fuente

Probabilidad
Un Poco de Teoría (y de “Filosofía”)

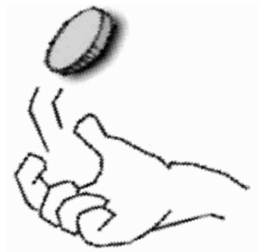
Probabilidad y Estadística

- Probabilidad
  - Reglas para el manejo de la incertidumbre
    - "La mejor medida de la incertidumbre es la probabilidad"



Dennis Lindley
1923-2013

- **Ejemplo:** ¿cuál es la incertidumbre asociada al lanzamiento de una moneda?
  - ¿Cómo la medimos?
  - ¿Cómo la usamos para obtener conclusiones?



Probabilidad y Estadística

- Estadística
  - Estudio de la adquisición, organización, análisis, interpretación y presentación de los datos

Probabilidad y Estadística

- Estadística
  - Estudio de la adquisición, organización, análisis, interpretación y presentación de los datos
- **Ejemplo**: lanzamos una moneda mil veces

Probabilidad y Estadística

- Estadística
  - Estudio de la adquisición, organización, análisis, interpretación y presentación de los datos
- **Ejemplo**: lanzamos una moneda mil veces

  - Adquisición: ¿qué valores apunto de esos lanzamientos?

Probabilidad y Estadística

- Estadística
  - Estudio de la adquisición, organización, análisis, interpretación y presentación de los datos
- **Ejemplo**: lanzamos una moneda mil veces

  - Adquisición: ¿qué valores apunto de esos lanzamientos?
  - Organización: ¿cómo guardo y agrupo estos valores?

Probabilidad y Estadística

- Estadística
  - Estudio de la adquisición, organización, análisis, interpretación y presentación de los datos

- **Ejemplo**: lanzamos una moneda mil veces
  - Adquisición: ¿qué valores apunto de esos lanzamientos?
  - Organización: ¿cómo guardo y agrupo estos valores?
  - Análisis: ¿en qué rango están estos valores? ¿Cómo son de precisos o de dispersos?

Probabilidad y Estadística

- Estadística
  - Estudio de la adquisición, organización, análisis, interpretación y presentación de los datos
- **Ejemplo**: lanzamos una moneda mil veces

  - Adquisición: ¿qué valores apunto de esos lanzamientos?
  - Organización: ¿cómo guardo y agrupo estos valores?
  - Análisis: ¿en qué rango están estos valores? ¿Cómo son de precisos o de dispersos?
  - Interpretación: ¿qué dicen mis datos sobre si la moneda está o no trucada?

Probabilidad y Estadística

- Estadística
  - Estudio de la adquisición, organización, análisis, interpretación y presentación de los datos
- **Ejemplo**: lanzamos una moneda mil veces

  - Adquisición: ¿qué valores apunto de esos lanzamientos?
  - Organización: ¿cómo guardo y agrupo estos valores?
  - Análisis: ¿en qué rango están estos valores? ¿Cómo son de precisos o de dispersos?
  - Interpretación: ¿qué dicen mis datos sobre si la moneda está o no trucada?
  - Presentación: ¿cómo comunico mis conclusiones? ¿Qué números o gráficas utilizo para que se me entienda bien?

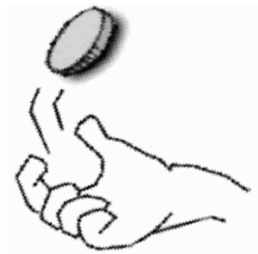
Estadística Frecuentista

- También "estadística clásica"
- Repetición de un mismo experimento
  - Probabilidad: interpretada como frecuencia
    - "Casos favorables entre casos posibles"

Estadística Frecuentista

- También “estadística clásica”
- Repetición de un mismo experimento
  - Probabilidad: interpretada como frecuencia
    - “Casos favorables entre casos posibles”

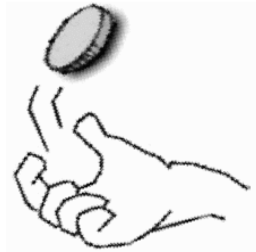
- **Ejemplo:** experimento de lanzar una moneda
  - Lo repito un número muy grande de veces
  - Apunto las veces que sale cara
  - Divido entre el número total de lanzamientos
  - Resultado: probabilidad de obtener cara
    - 0,5 aproximadamente



Estadística Frecuentista

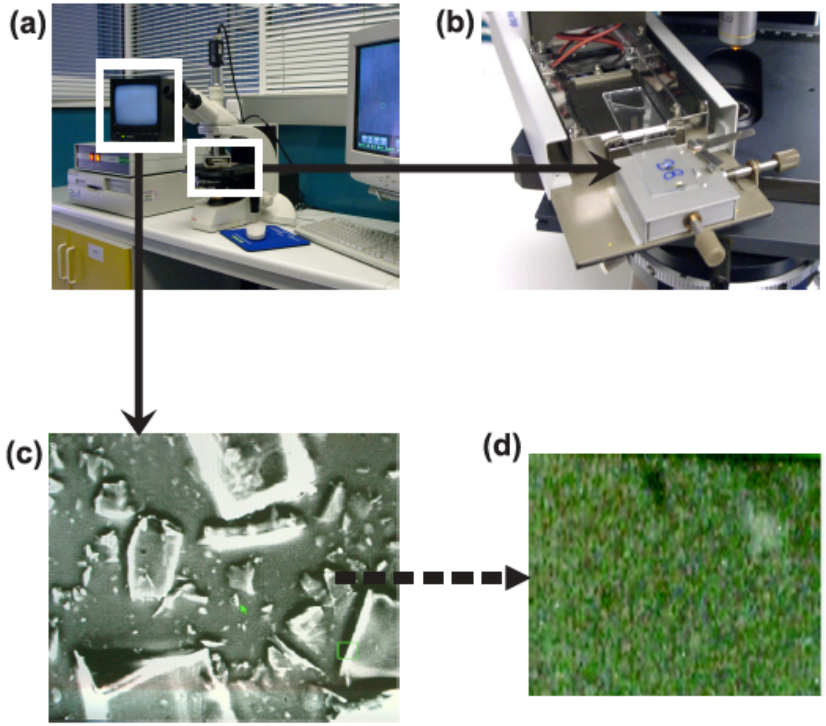
- Valor “real” de probabilidad
  - Cuando repetimos infinitamente

- **Ejemplo:** repetición de lanzamiento de moneda
  - Repito 10 veces, sale cara 4 veces
    - 4/10 = 0,4
  - Esa es mi estimación del valor real (0,5)



Estadística Frecuentista

- **Ejemplo**:
  - Medida de índice de refracción (n) en un vidrio
    - Técnica físico-química GRIM
- Valor *real*
  - Valor medio de n que obtendríamos si hiciéramos infinitas medidas en el vidrio
    - Nunca lo conoceremos…
- Valor estimado
  - Valor medio de un número finito de medidas



G. Zadora et al. "Statistical Analysis in Forensic Science. Evidential Value of Multivariate Physico-Chemical Data". John Wiley and Sons, 2014

Estadística Frecuentista

- Métodos típicos de la estadística frecuentista
  - Contrastes de hipótesis, intervalos de confianza, estimación, etc.

Estadística Frecuentista

- Métodos típicos de la estadística frecuentista
  - Contrastes de hipótesis, intervalos de confianza, estimación, etc.
- Utilizada en múltiples áreas de conocimiento
  - Medicina, ciencias puras, ciencias sociales, ingeniería…

Estadística Frecuentista

- Métodos típicos de la estadística frecuentista
  - Contrastes de hipótesis, intervalos de confianza, estimación, etc.
- Utilizada en múltiples áreas de conocimiento
  - Medicina, ciencias puras, ciencias sociales, ingeniería…
- Su éxito se debe a…
  - Teoría e implementaciones relativamente sencillas
  - Respuestas concretas, claras
    - "¿Puedo rechazar esta hipótesis?" "¿Es este valor significativo?"
  - Procedimientos muy estandarizados
    - Incluso entre diferentes áreas de conocimiento
    - Significancia, p-valores, ANOVA…

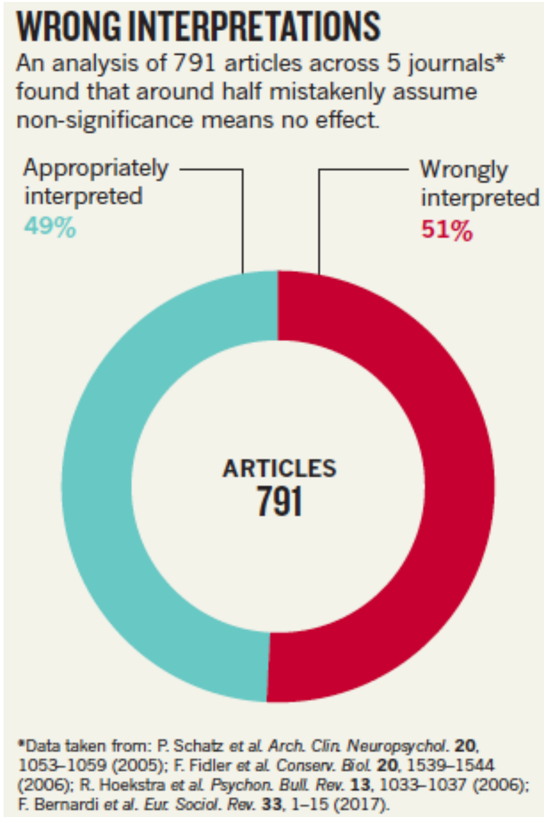
Estadística Frecuentista

- La estadística frecuentista se ha usado (y se sigue usando) de forma incorrecta

Retire statistical significance

Valentin Amrhein, **Sander Greenland**, **Blake McShane** and more than 800 signatories call for an end to hyped claims and the dismissal of possibly crucial effects.

21 MARCH 2019 | VOL 567 | NATURE | 305



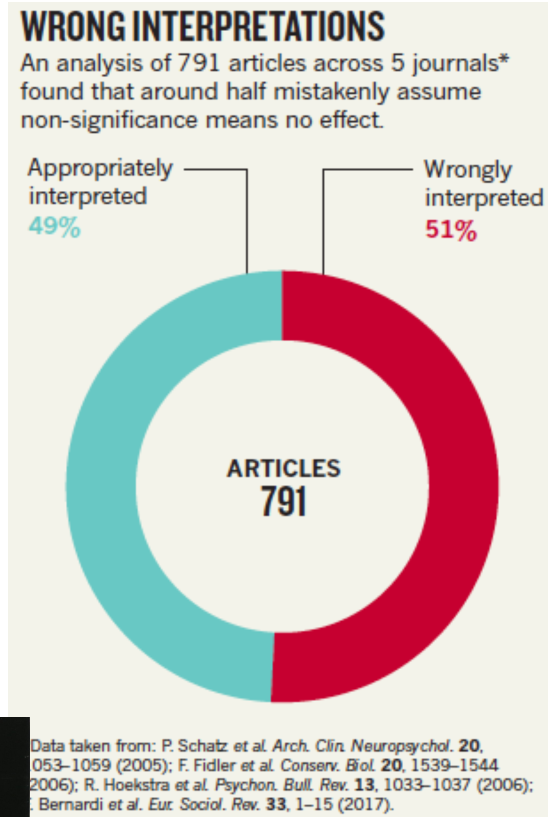
Estadística Frecuentista

- La estadística frecuentista se ha usado (y se sigue usando) de forma incorrecta

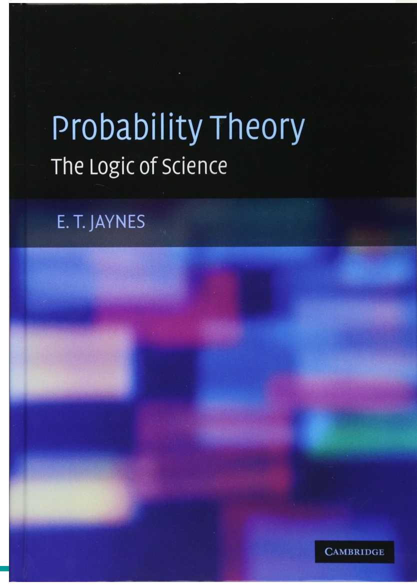
Retire statistical significance

Valentin Amrhein, Sander Greenland, Blake McShane and more than 800 signatories call for an end to hyped claims and the dismissal of possibly crucial effects.

21 MARCH 2019 | VOL 567 | NATURE | 305

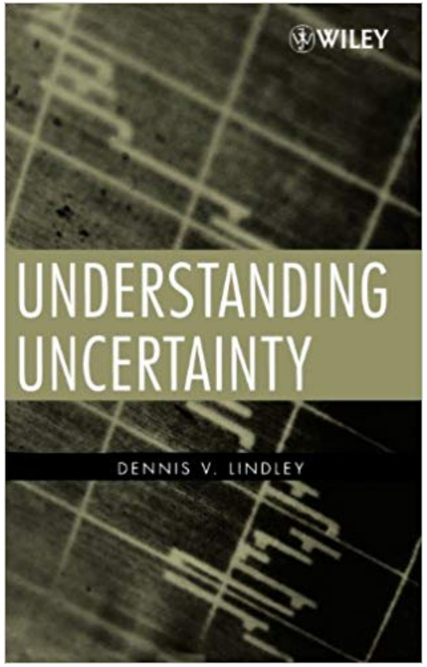


- Algunos autores opinan incluso que es errónea o poco general



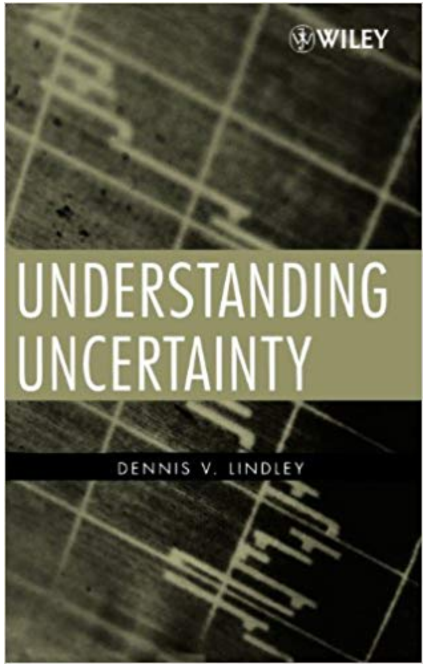
Estadística Bayesiana

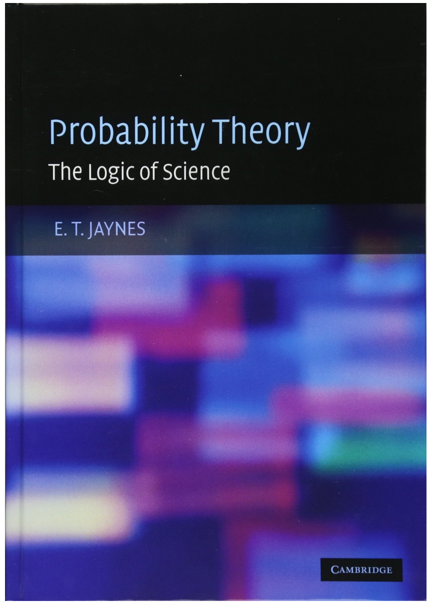
- La probabilidad es la expresión de mi conocimiento de un problema
  - *Probability: Degree of belief*
- Por definición: subjetiva, personal
  - Mi conocimiento
- Cuantifica la incertidumbre de un problema
  - Idealmente: *toda* la incertidumbre



Estadística Bayesiana

- La probabilidad es la expresión de mi conocimiento de un problema
  - *Probability: Degree of belief*
- Por definición: subjetiva, personal
  - Mi conocimiento
- Cuantifica la incertidumbre de un problema
  - Idealmente: *toda* la incertidumbre
- Reglas de la probabilidad
  - Nos dan la clave para actualizar nuestro conocimiento
  - Con nuevos datos
  - "Lógica de la incertidumbre"





Estadística Bayesiana: Errores de Concepto

- No existen "probabilidades verdaderas"
  - Porque las probabilidades son, por definición, personales, subjetivas y dependientes del conocimiento actual (y de suposiciones)

Estadística Bayesiana: Errores de Concepto

- No existen "probabilidades verdaderas"
  - Porque las probabilidades son, por definición, personales, subjetivas y dependientes del conocimiento actual (y de suposiciones)
- No se estiman probabilidades
  - Porque no existen probabilidades "verdaderas"

Estadística Bayesiana: Errores de Concepto

- No existen "probabilidades verdaderas"
  - Porque las probabilidades son, por definición, personales, subjetivas y dependientes del conocimiento actual (y de suposiciones)
- No se estiman probabilidades
  - Porque no existen probabilidades "verdaderas"
- Por la misma razón, no existe un "modelo verdadero de los datos"

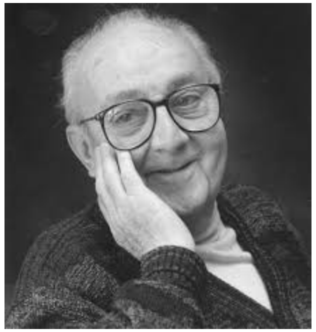
Estadística Bayesiana: Errores de Concepto

- No existen “probabilidades verdaderas”
  - Porque las probabilidades son, por definición, personales, subjetivas y dependientes del conocimiento actual (y de suposiciones)
- No se estiman probabilidades
  - Porque no existen probabilidades “verdaderas”
- Por la misma razón, no existe un “modelo verdadero de los datos”
- Por tanto, no se estiman modelos de los datos
  - Por el contrario, se buscan modelos consistentes con los datos

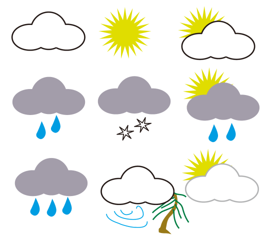
Estadística Bayesiana: Errores de Concepto

- No existen “probabilidades verdaderas”
  - Porque las probabilidades son, por definición, personales, subjetivas y dependientes del conocimiento actual (y de suposiciones)
- No se estiman probabilidades
  - Porque no existen probabilidades “verdaderas”
- Por la misma razón, no existe un “modelo verdadero de los datos”
- Por tanto, no se estiman modelos de los datos
  - Por el contrario, se buscan modelos consistentes con los datos

“All models are wrong,
but some of them are useful”
George E. P. Box



Ejemplo Sencillo



- Probabilidad de que llueva mañana (LL) o no (nLL)

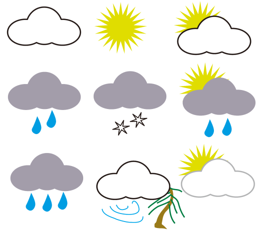
Opinión con mi conocimiento actual

$P(LL|D)=0,9$

Probabilidad a Priori

Conocimiento previo antes de observar previsión meteorológica

Ejemplo Sencillo



- Probabilidad de que llueva mañana (LL) o no (nLL)

Previsión Meteorológica (dato, D)

Opinión con mi conocimiento actual

P(LL|D)=0,9

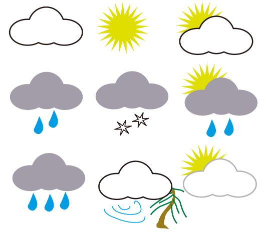
Probabilidad a Priori

Valor del dato observado

Conocimiento previo antes de observar previsión meteorológica

¿El dato apoya a LL o a nLL? ¿Con qué fuerza?

Ejemplo Sencillo



- Probabilidad de que llueva mañana (LL) o no (nLL)

Previsión Meteorológica (dato, D)

Opinión con mi conocimiento actual

P(LL|D)=0,9

Inferencia

Opinión **después** de conocer el dato

P(LL|D)=0,9

Probabilidad a Priori

Valor del dato observado

Probabilidad a Posteriori

Conocimiento previo antes de observar previsión meteorológica

¿El dato apoya a LL o a nLL? ¿Con qué fuerza?

Incluye conocimiento previo y lo que dice el dato

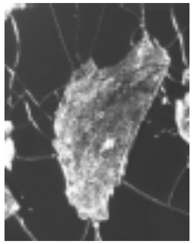
Tipos de Datos

Datos y Variables

- Datos: Valores procedentes de una variable matemática que por alguna razón son de interés para nosotros
  - **Ejemplos de variables, con sus valores posibles**:

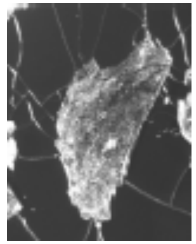
Datos y Variables

- Datos: Valores procedentes de una variable matemática que por alguna razón son de interés para nosotros
  - **Ejemplos de variables, con sus valores posibles**:
    - Tipo de vidrio al que pertenecen unos fragmentos
      - Ventana, envase, faro



Datos y Variables

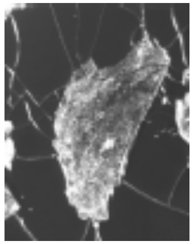
- Datos: Valores procedentes de una variable matemática que por alguna razón son de interés para nosotros
  - **Ejemplos de variables, con sus valores posibles**:
    - Tipo de vidrio al que pertenecen unos fragmentos
      - Ventana, envase, faro

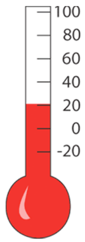


    - Tipo de gasolina por octanaje
      - 95 octanos, 98 octanos, más de 98 octanos

Datos y Variables

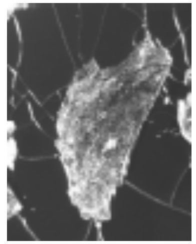
- Datos: Valores procedentes de una variable matemática que por alguna razón son de interés para nosotros
  - **Ejemplos de variables, con sus valores posibles**:
    - Tipo de vidrio al que pertenecen unos fragmentos
      - Ventana, envase, faro
    - Tipo de gasolina por octanaje
      - 95 octanos, 98 octanos, más de 98 octanos
    - Temperatura de un fluido
      - Número real: intervalo dependiente de la escala





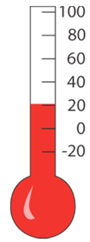
Datos y Variables

- Datos: Valores procedentes de una variable matemática que por alguna razón son de interés para nosotros
  - **Ejemplos de variables, con sus valores posibles**:
    - Tipo de vidrio al que pertenecen unos fragmentos
      - Ventana, envase, faro



    - Tipo de gasolina por octanaje
      - 95 octanos, 98 octanos, más de 98 octanos

    - Temperatura de un fluido
      - Número real: intervalo dependiente de la escala



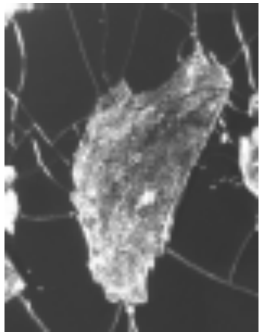
    - Número de moles de sodio en una muestra
      - Número real: de 0 a infinito

Variables Nominales

- Sus posibles valores pertenecen a una de varias categorías
- Las categorías no siguen una relación de orden
  - Ninguna es mayor que otra

Variables Nominales

- Sus posibles valores pertenecen a una de varias categorías
- Las categorías no siguen una relación de orden
  - Ninguna es mayor que otra
- **Ejemplo**: tipo de vidrio al que pertenece un fragmento



  - Valores posibles: ventana, envase, faro
  - No hay relación de orden
    - Ventana no es más grande que faro, por ejemplo
  - Notación matemática
    - Variable: X
    - Un valor cualquiera de X: $x \in \{\text{ventana, envase, faro}\}$

Variables Ordinales

- Sus posibles valores son, de nuevo, categorías
  - No se puede hacer operaciones con ellas…
- Los valores siguen una relación de orden

Variables Ordinales

- Sus posibles valores son, de nuevo, categorías
  - No se puede hacer operaciones con ellas…
- Los valores siguen una relación de orden
- **Ejemplo**: tipo de gasolina por octanaje

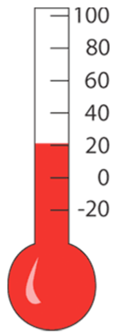
  - Valores posibles: 95, 98, más de 98
  - Hay relación de orden: 95 es más que 98
    - Pero no tiene sentido hacer operaciones con ellas…
  - Notación matemática
    - Variable: X
    - Un valor cualquiera de X: $x \in \{95, 98, \text{más de } 98\}$

Variables de Intervalo

- Son variables numéricas, y por tanto con relación de orden
  - El valor 0 no indica ausencia de lo que indica la variable
  - Admiten operaciones matemáticas

Variables de Intervalo

- Son variables numéricas, y por tanto con relación de orden
  - El valor 0 no indica ausencia de lo que indica la variable
  - Admiten operaciones matemáticas
- **Ejemplo**: temperatura de un fluido en grados centígrados
  - Valores posibles: de -273 a infinito
  - Sin embargo, 0ºC no indica “ausencia de temperatura”
  - Notación matemática
    - Variable: X
    - Un valor cualquiera de X: $x \in [-273, \infty)$



Variables de Razón

- Son variables numéricas, y por tanto con relación de orden
  - El valor 0 indica ausencia de lo que indica la variable
  - Admiten operaciones matemáticas

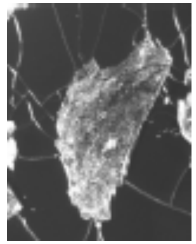
Variables de Razón

- Son variables numéricas, y por tanto con relación de orden
  - El valor 0 indica ausencia de lo que indica la variable
  - Admiten operaciones matemáticas

- **Ejemplo**: número de moles de sodio en una muestra
  - Valores posibles: de 0 a infinito
  - 0 indica “ausencia de sodio”
  - Notación matemática
    - Variable: X
    - Un valor cualquiera de X: $x \in [0, \infty)$

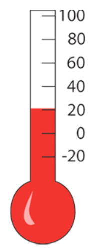
Tipos de Datos: Otra Clasificación

- Variables cualitativas
  - Ni operaciones matemáticas ni relación de orden
    - Nominales



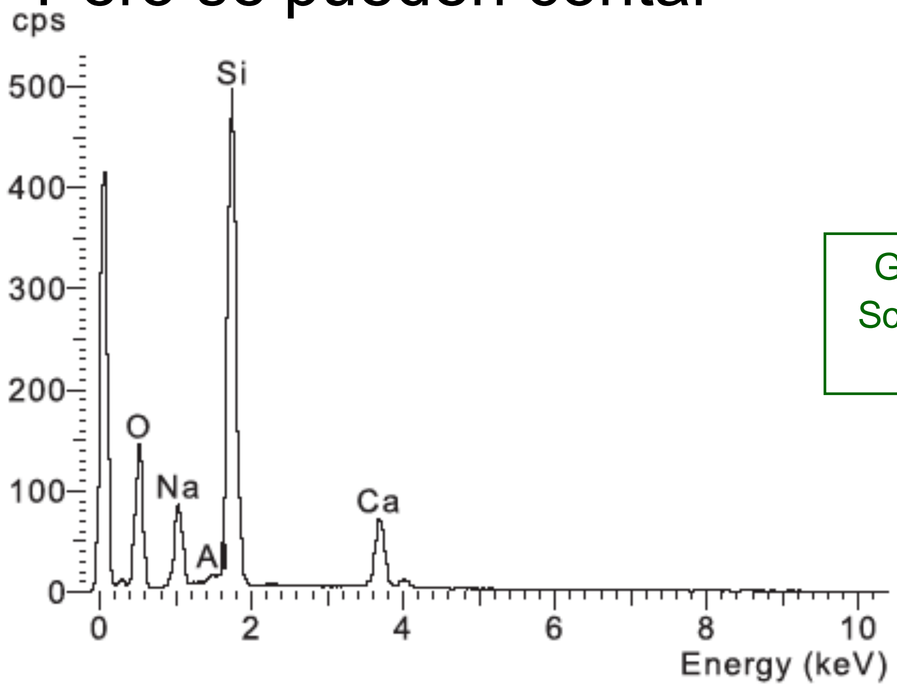
- Variables cuasicuantitativas
  - No operaciones matemáticas, sí relación de orden
    - Ordinales

- Variables cuantitativas
  - Operaciones matemáticas
    - De intervalo, de razón



Variables Cuantitativas Discretas

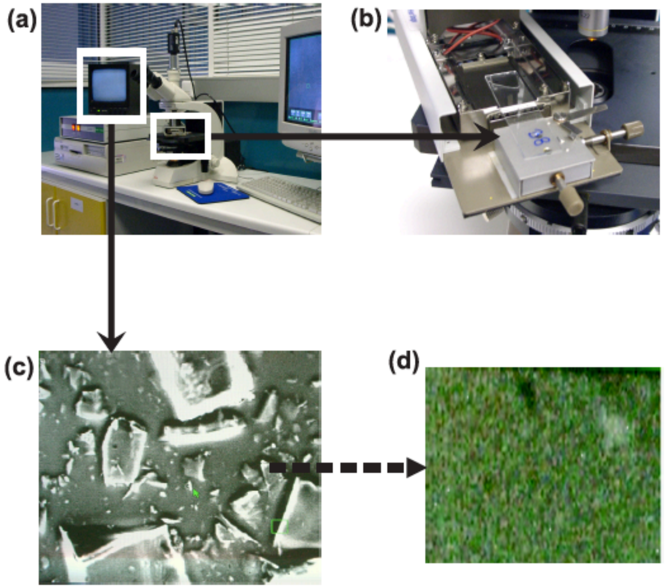
- Sus posibles valores se pueden contar
  - Pueden ser finitos o infinitos
  - **Ejemplo**: número de picos en un espectro
    - Valores posibles: de 0 a infinito
    - Pero se pueden contar



G. Zadora et al. "Statistical Analysis in Forensic Science. Evidential Value of Multivariate Physico-Chemical Data". John Wiley and Sons, 2014

Variables Cuantitativas Continuas

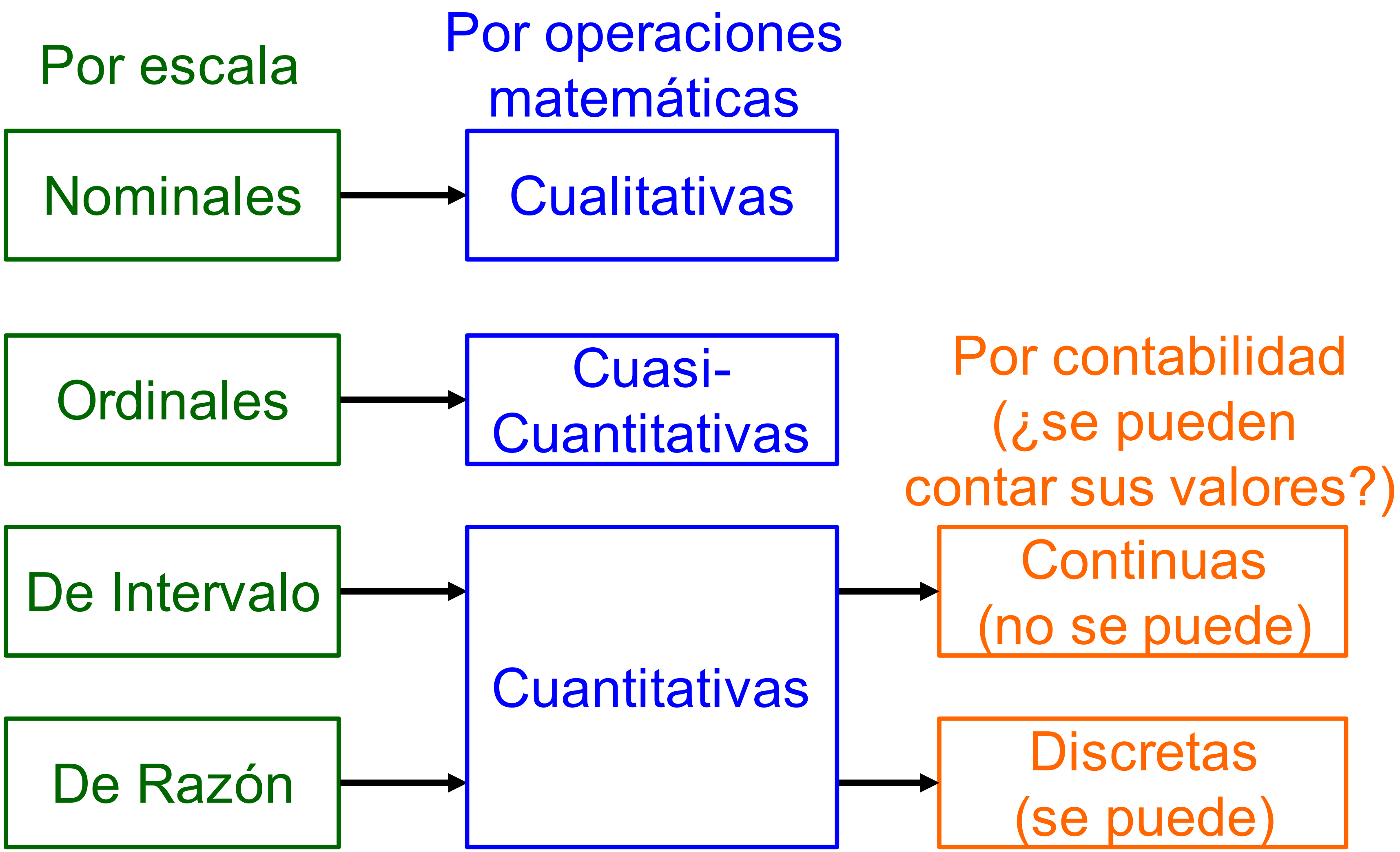
- Sus posibles valores no se pueden contar
- Siempre infinitos valores
  - **Ejemplo**: índice de refracción de un vidrio
    - Entre dos valores siempre hay otro
      - Suponiendo precisión infinita del aparato de medida



G. Zadora et al. "Statistical Analysis in Forensic Science. Evidential Value of Multivariate Physico-Chemical Data". John Wiley and Sons, 2014

Tipos de Datos: Cuadro Resumen

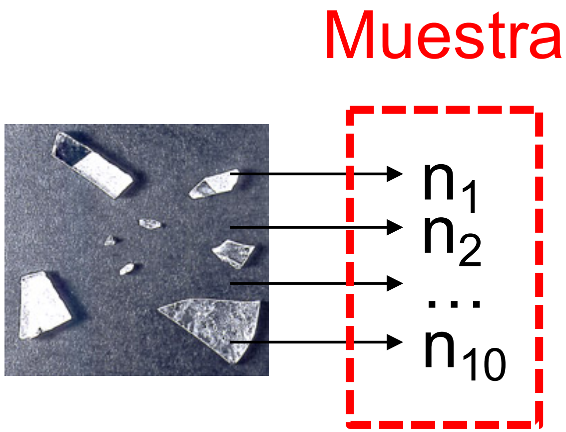
- Distintos criterios de clasificación de tipos de variables



Muestra y Población

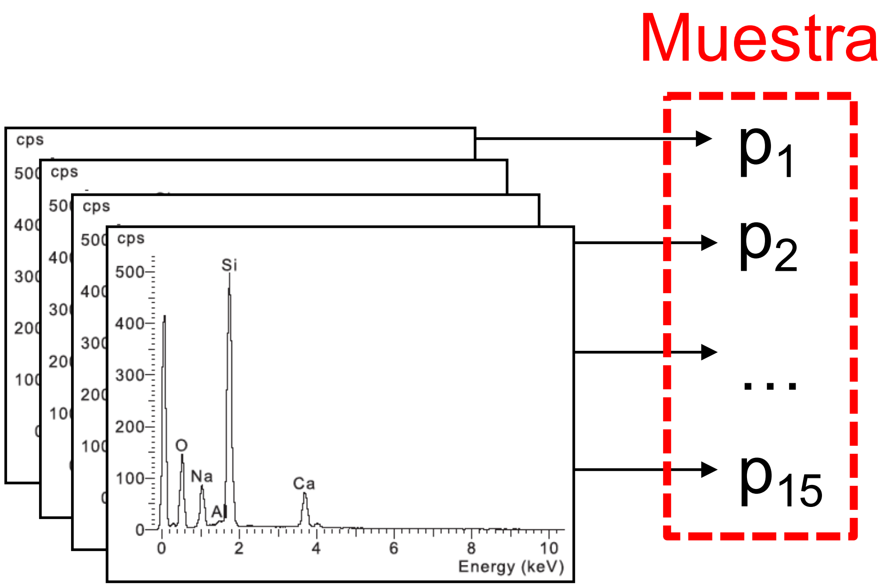
Muestra

- Definimos muestra como un conjunto finito de valores observados de una variable
  - **Ejemplo**: 10 medidas de índice de refracción de un vidrio
    - Es una muestra de tamaño 10
    - De la variable "índice de refracción de ese vidrio"
      - Cuantitativa
      - Continua



Muestra

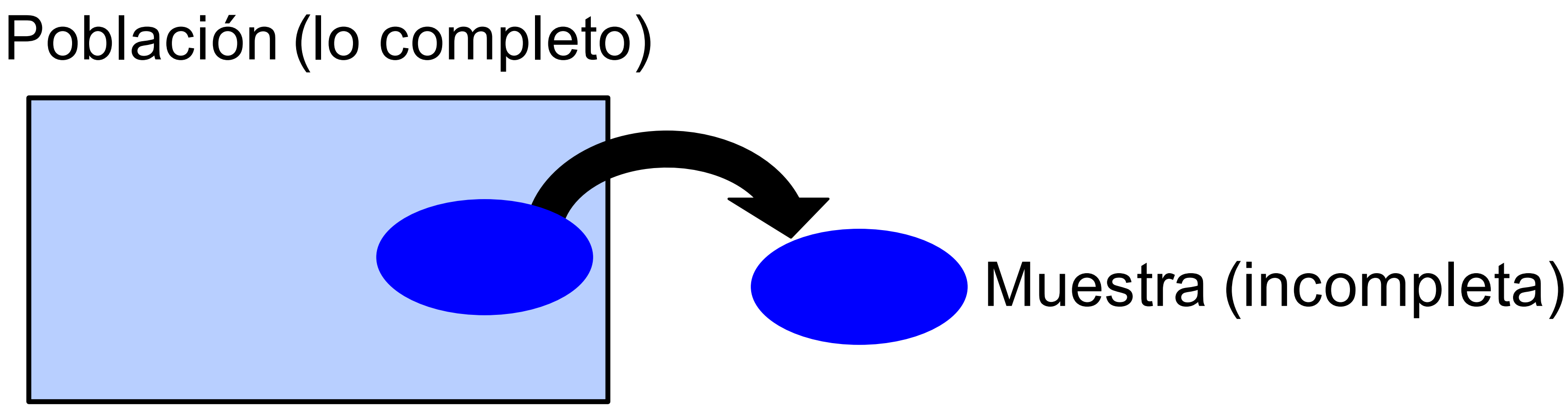
- **Ejemplo**: 15 conteos de picos observados en espectros de una misma sustancia
  - Es una muestra de tamaño 15
  - De la variable “picos en el espectro de esa sustancia”
    - Cuantitativa
    - Discreta



[Zadora14]

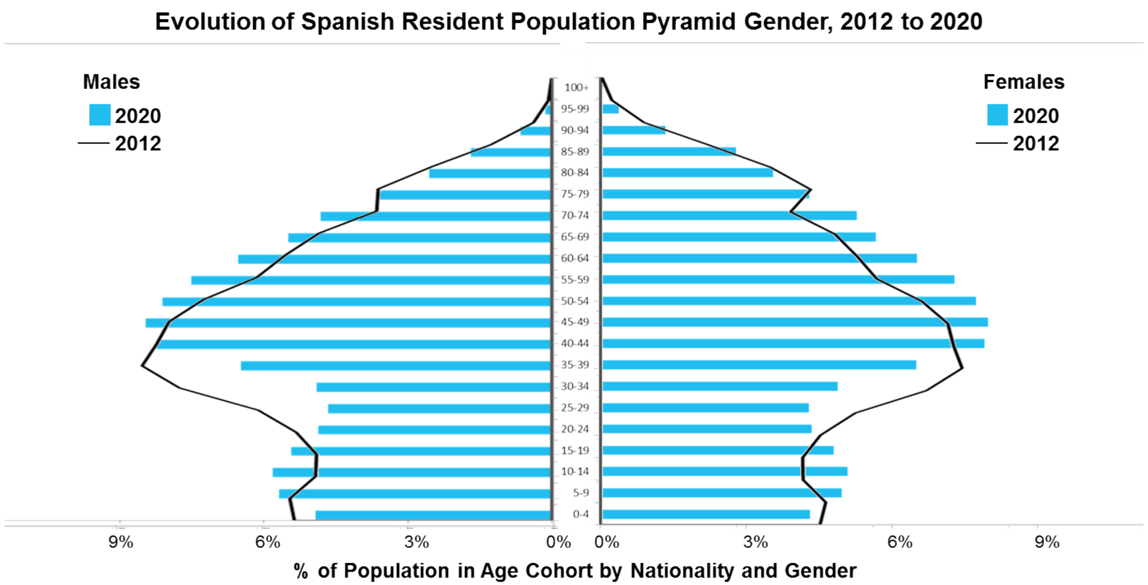
Población

- Conjunto de datos de donde sale la muestra
  - Población: lo completo
  - Muestra: incompleta



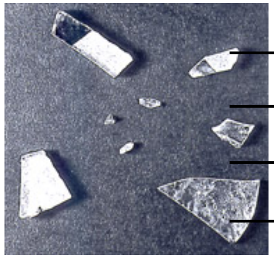
Población: Interpretaciones

1. Si la muestra tiene un número máximo de elementos limitado
   - La población es la muestra de tamaño máximo posible
   - **Ejemplo**: variable “edad de una persona en España en años”
     - Población: edades de todas las personas de España
     - Muestra: edades de 15 personas de España escogidas al azar
     - Tipo de variable
       - Cuantitativa
       - Discreta



Población: Interpretaciones

2. Si el núm. máximo de elementos de la muestra no tiene límite
   - La población son los valores que se obtendrían de la variable con una muestra de tamaño infinito
   - **Ejemplo**: variable “índice de refracción de un vidrio”
     - Población: valores medidos de dicho índice si hiciésemos infinitas medidas
     - Muestra: 10 valores medidos de dicho índice
     - Tipo de variable
       - Cuantitativa
       - Continua

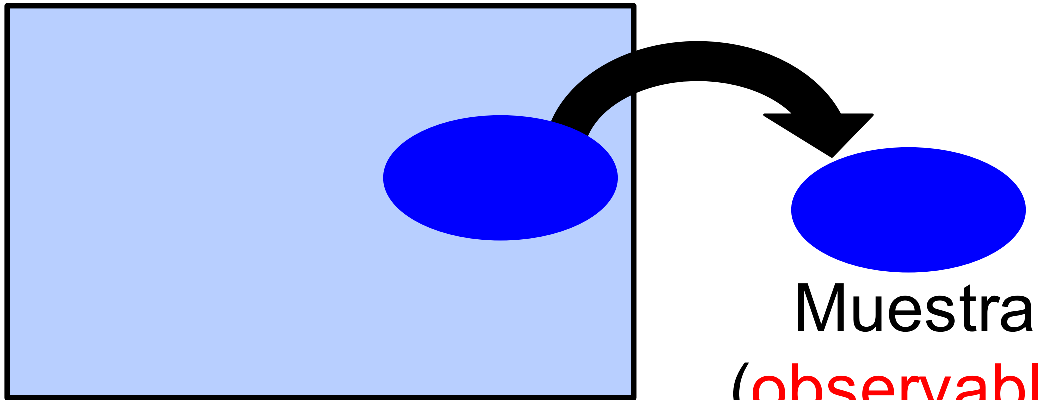


n_1
n_2
…
n_{10}

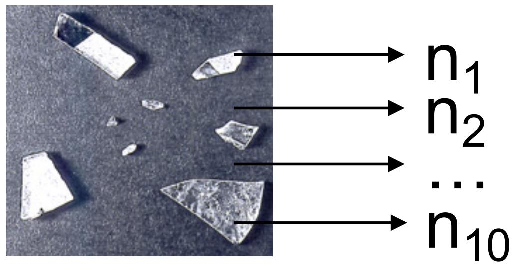
Muestra y Población

- En un problema estadístico, en general
  - La muestra es conocida (observable)
  - La población es desconocida (no observable)
- **Ejemplo**: índice de refracción de un vidrio
  - Podemos hacer 10 medidas (muestra observable)
  - Imposible hacer infinitas medidas (población no observable)

Población (no observable)

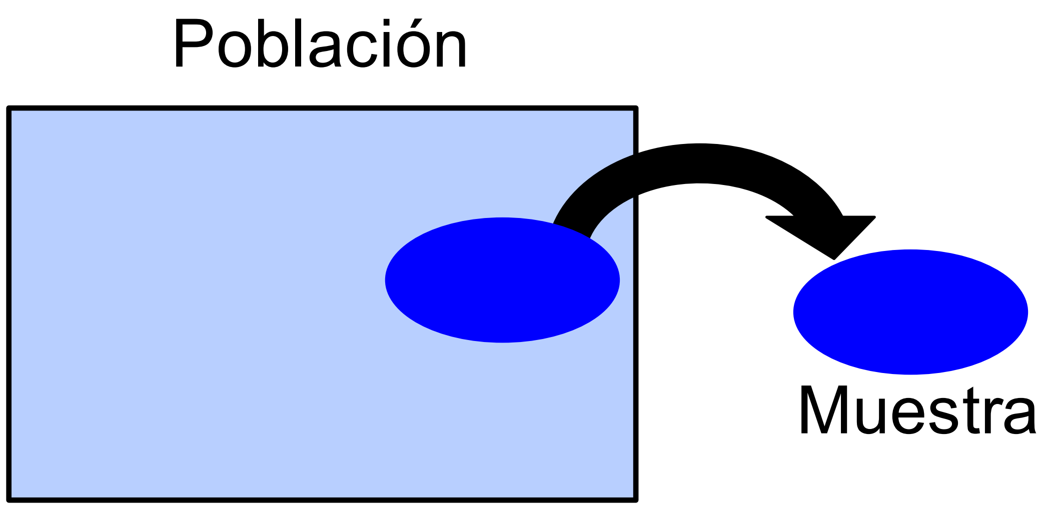


Muestra
(observable)



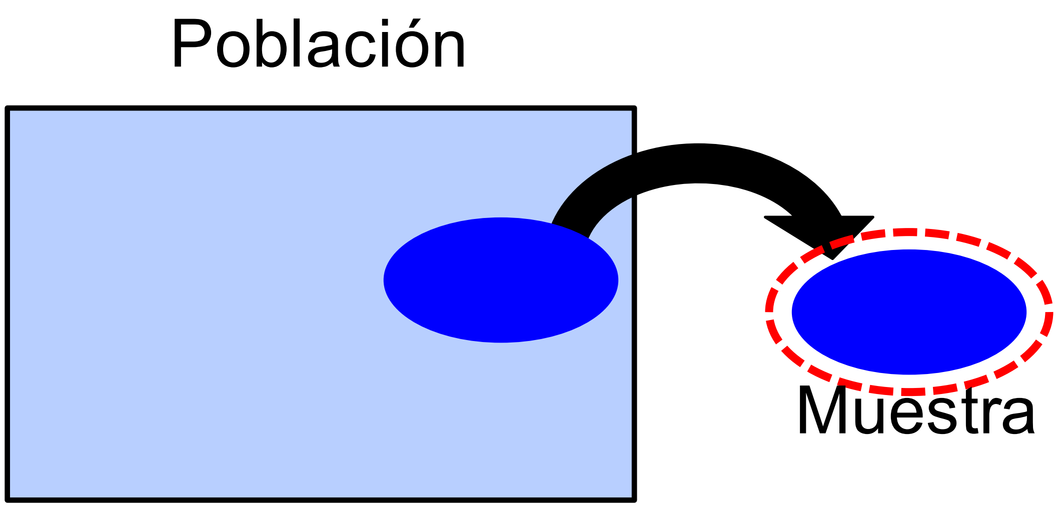
Estadístico vs. Parámetro

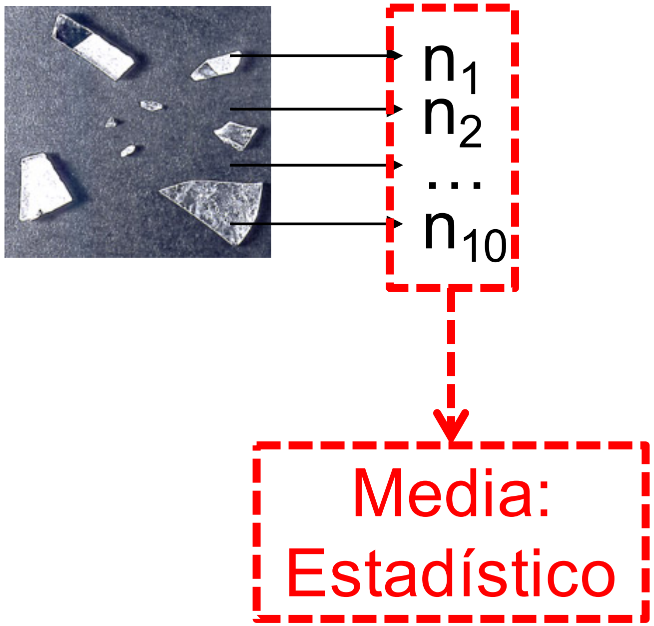
- A partir de datos obtenidos de variables cuantitativas, podemos calcular números



Estadístico ← Muestra

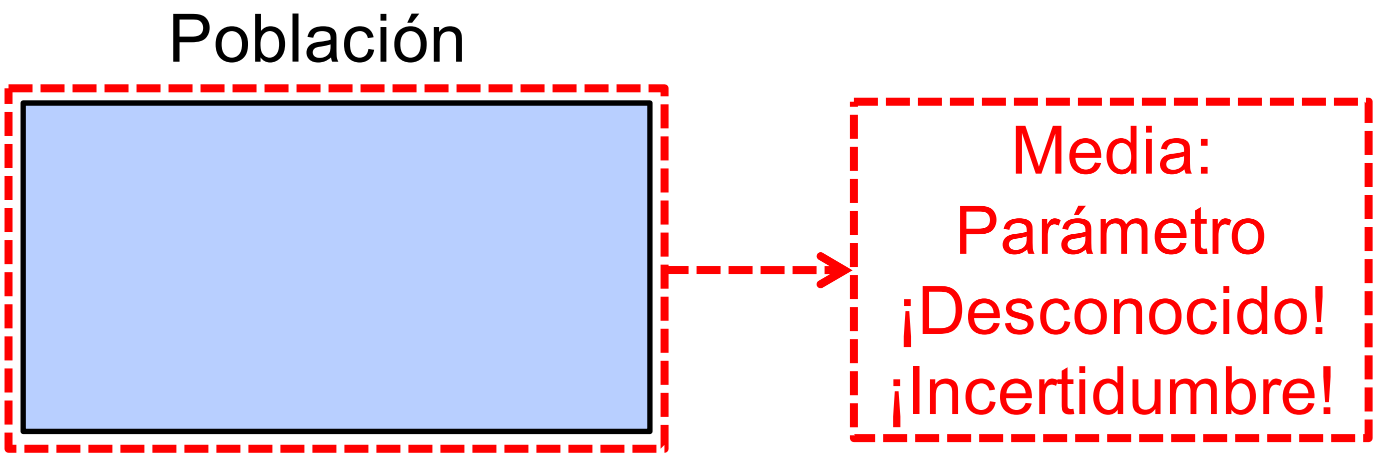
- A partir de datos obtenidos de variables cuantitativas, podemos calcular números
- Si calculamos números a partir de la muestra
  - Al número obtenido se le denomina estadístico
  - **Ejemplo**: valor medio de los 10 valores medidos de índice de refracción





Parámetro ← Población

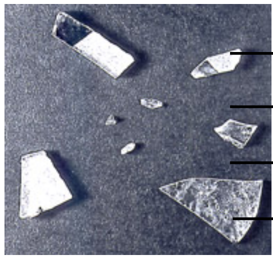
- A partir de datos obtenidos de variables cuantitativas, podemos calcular números
- Si calculamos números a partir de la población
  - Al número obtenido se le denomina parámetro
  - Asumimos que un parámetro nunca es conocido
    - **Ejemplo**: no podemos hallar el valor medio de infinitos valores medidos de índice de refracción…



Datos Univariados

Muestra Univariada

- De cada medida sólo tomamos un número
  - **Ejemplo**: índice de refracción de un vidrio
    - 10 medidas
    - 1 número por medida
    - 10 números en total



n_1
n_2
…
n_{10}

Tabla de Datos Univariados

- Los datos de una muestra en una tabla
  - Forma poco útil de visualizar los datos
  - **Ejemplo**, índices de refracción medidos sobre varios vidrios (datos de [Zadora14])

Número de Medida	**Índice de Refracción (IR)**
1	1.51785
2	1.51788
3	1.51783
4	1.51782
…	…

Tabla de Frecuencias

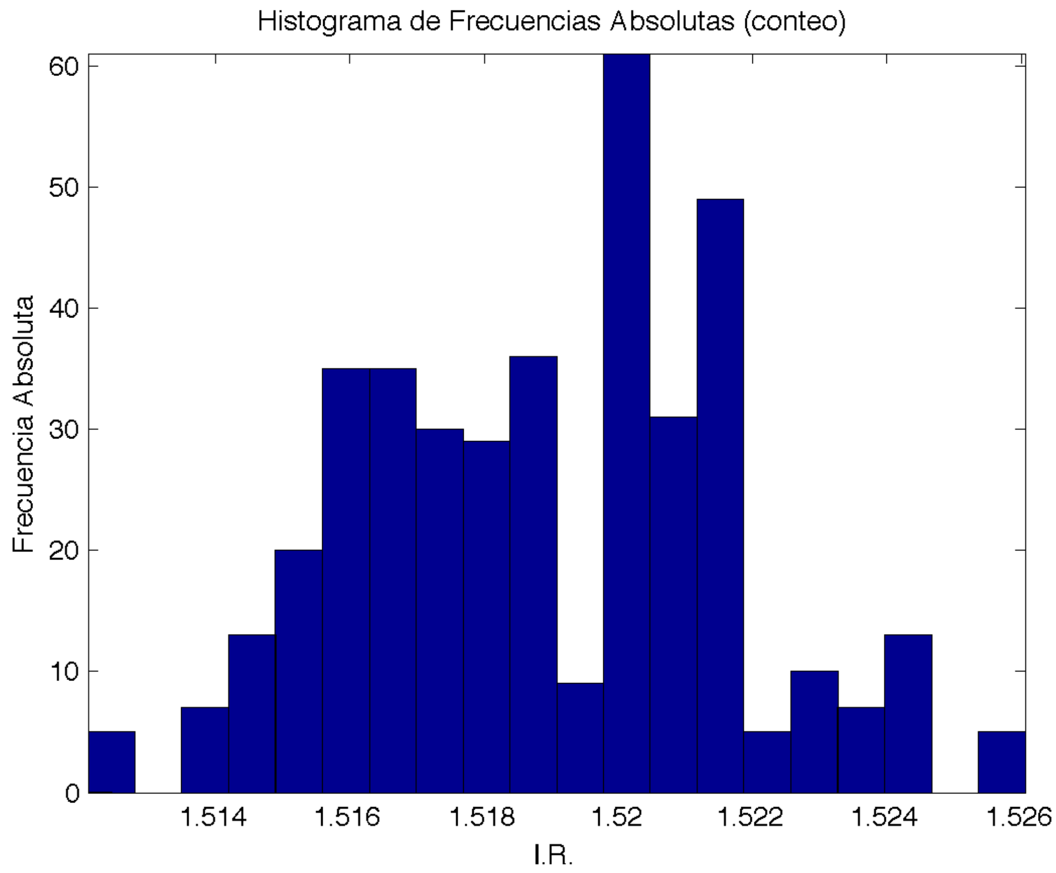
- Dividir el rango de valores en intervalos
- Contar el número valores que caen en cada intervalo
  - Llamado también frecuencias absolutas
  - **Ejemplo**, índices de refracción medidos sobre varios vidrios (datos de [Zadora14])

Intervalo	**Rango**	**Frecuencia Absoluta (conteo)**
1	1.5124 a 1.5131	5
2	1.5131 a 1.5138	0
3	1.5138 a 1.5145	13
4	1.5145 a 1.5152	7
…	…	…

Histograma

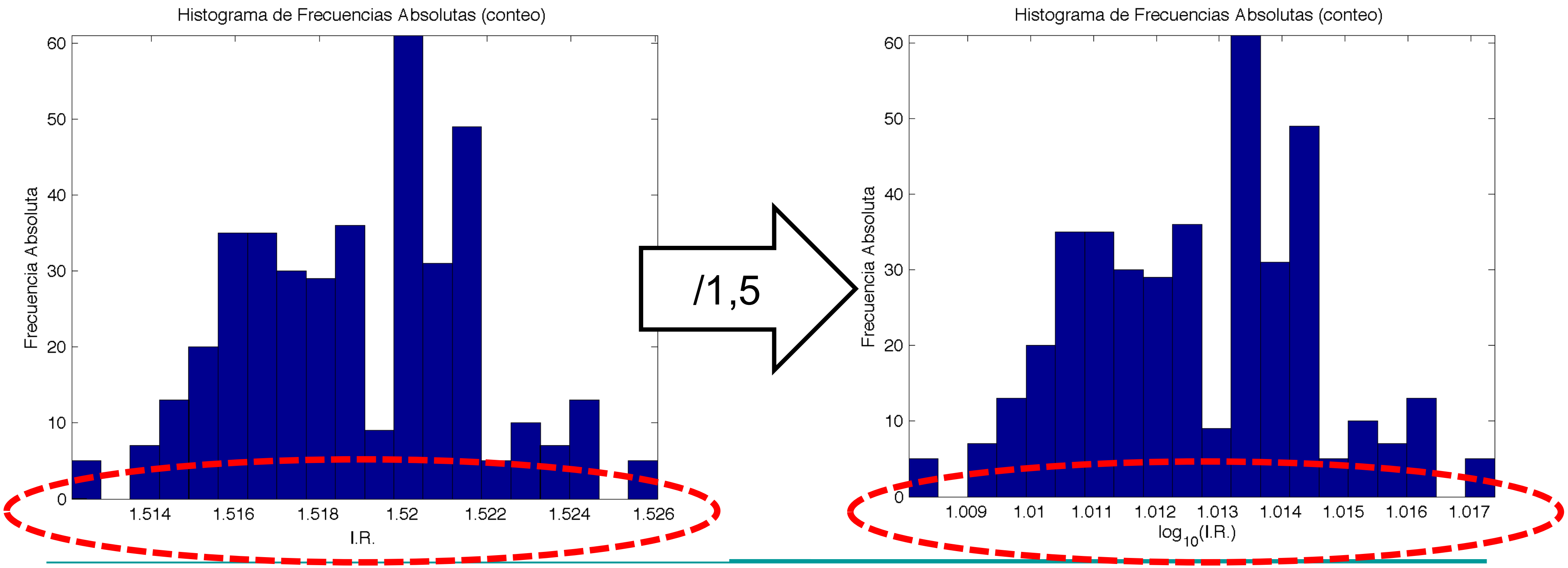
- Representación gráfica de una tabla de frecuencias

- Permite observar la distribución de los datos en cada uno de sus valores
  - Qué valores son más frecuentes que otros
  - **Ejemplo**, índices de refracción ([Zadora14])



Transformación de Datos

- En ocasiones es conveniente transformar los datos para facilitar…
  - Visualización
  - Operaciones matemáticas
- **Ejemplo**: Normalizar (dividir por 1.5)



Estadística Descriptiva de Datos Univariados

Localización: Media

- Sumar todos los datos y dividir por el total de datos

$$\bar{X} = \frac{(x_1 + x_2 + \ldots + x_N)}{N} = \frac{1}{N}\sum_{i=1}^{N} x_i$$

- **Ejemplo**, índices de refracción ([Zadora14])

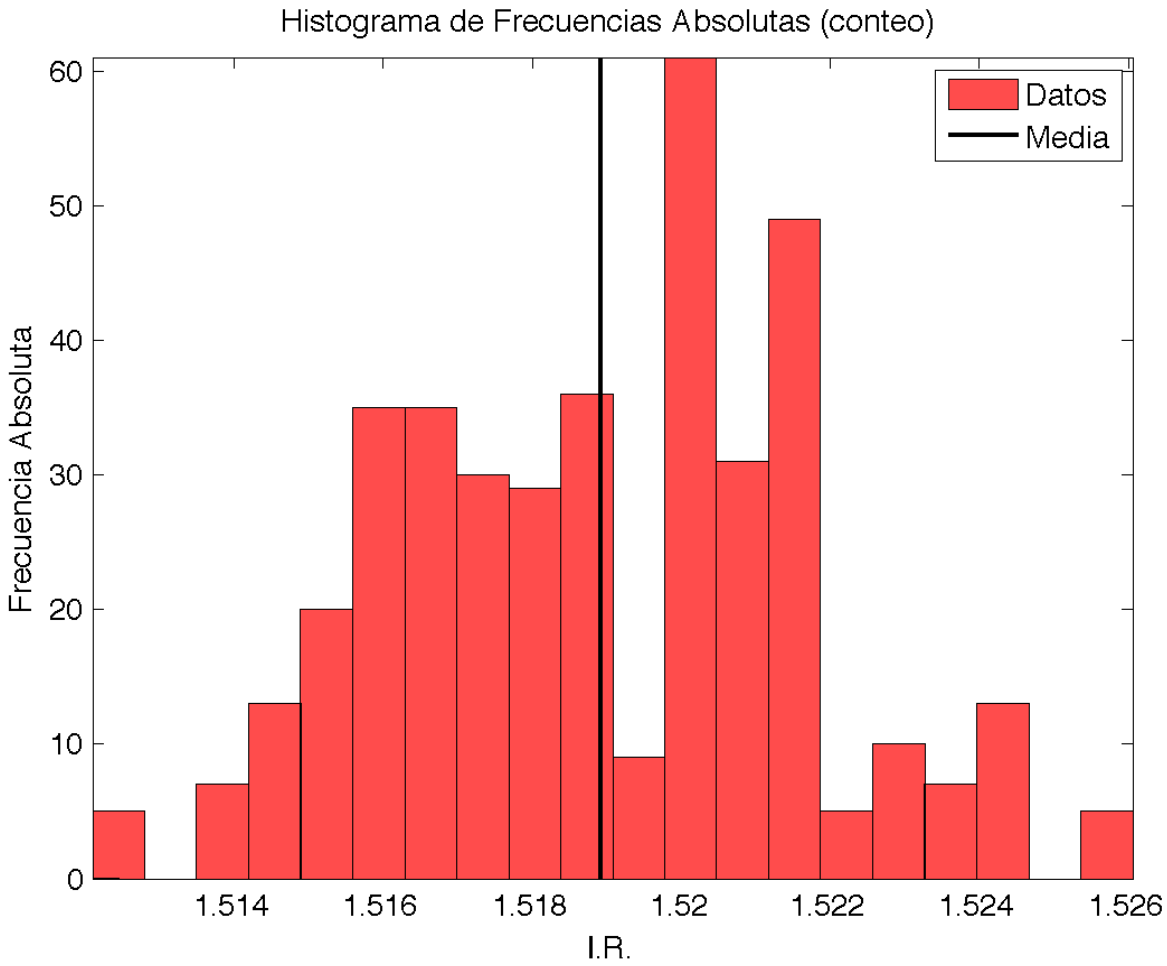
Número de Medida	Índice de Refracción (IR)
1	1.51785
2	1.51788
3	1.51783
4	1.51782
…	…

Sumar todos los RI
Dividir por el total

$$\bar{X} = \frac{1}{N}\sum_{i=1}^{N} x_i = 1{,}5189$$

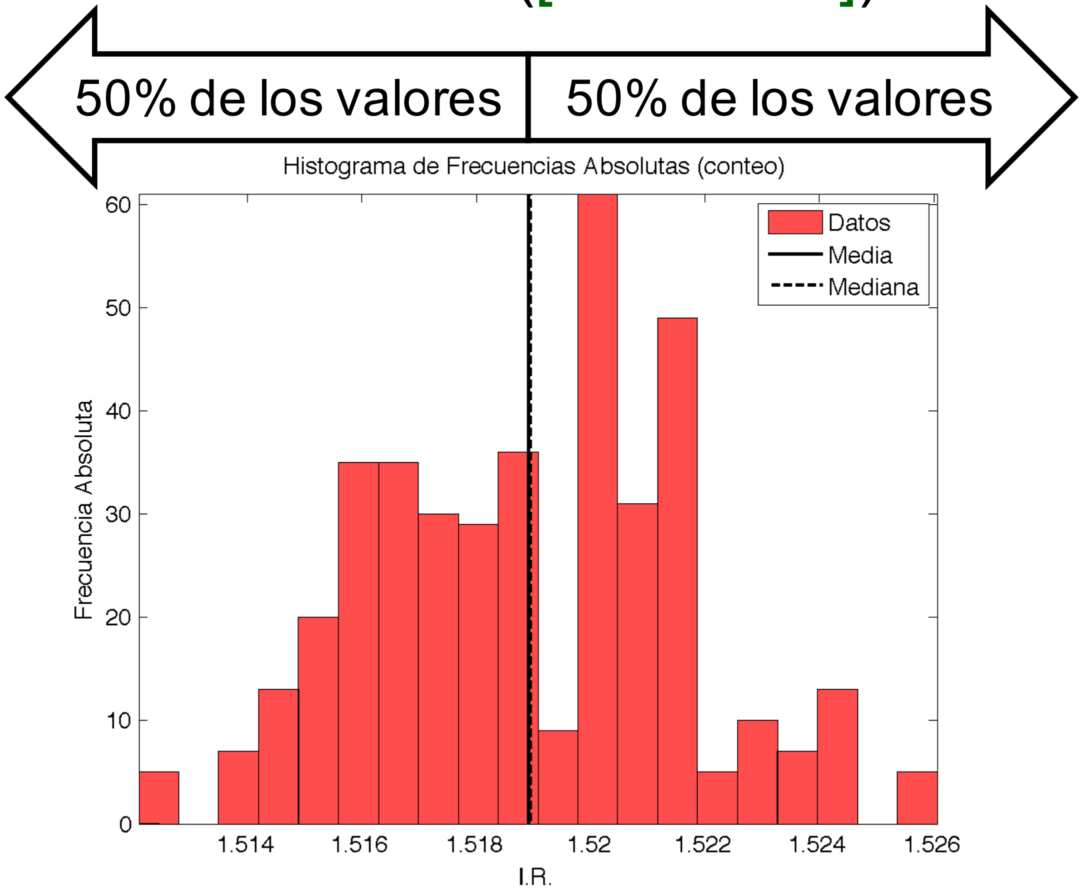
Localización: Media

- **Ejemplo**, índices de refracción ([Zadora14])



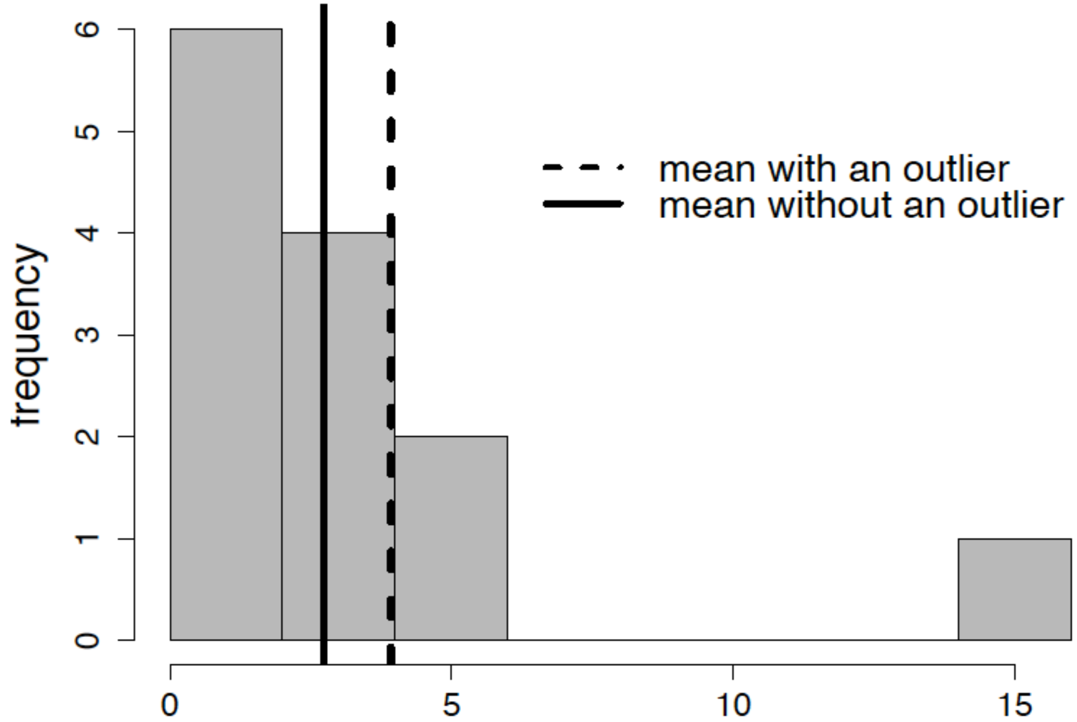
Localización: Mediana

- A ambos lados de la mediana hay un 50% de los valores
- **Ejemplo**, índices de refracción ([Zadora14])



Localización: Media vs. Mediana

- En algunos casos es mejor la mediana que la media
- **Ejemplo**: la media cambia mucho con un simple valor extremo



[Zadora14]

- Llamados valores atípicos (*outliers*), suelen ser errores
- La mediana variaría mucho menos: más robusta

Dispersión: Varianza

- Valor diferencial: $x_i - \bar{X}$
  - "Diferencia" de un valor con la media
- Varianza: media de los valores difierenciales al cuadrado
  - "Media de las desviaciones de la media al cuadrado"

$$S_X^2 = \frac{\left(x_1 - \bar{X}\right)^2 + \left(x_2 - \bar{X}\right)^2 + \ldots + \left(x_N - \bar{X}\right)^2}{N} = \frac{1}{N}\sum_{i=1}^{N}\left(x_i - \bar{X}\right)^2$$

- **Ejemplo**:

Núm. de Medida	I.R.
1	1.51785
2	1.51788
3	1.51783
4	1.51782
…	…

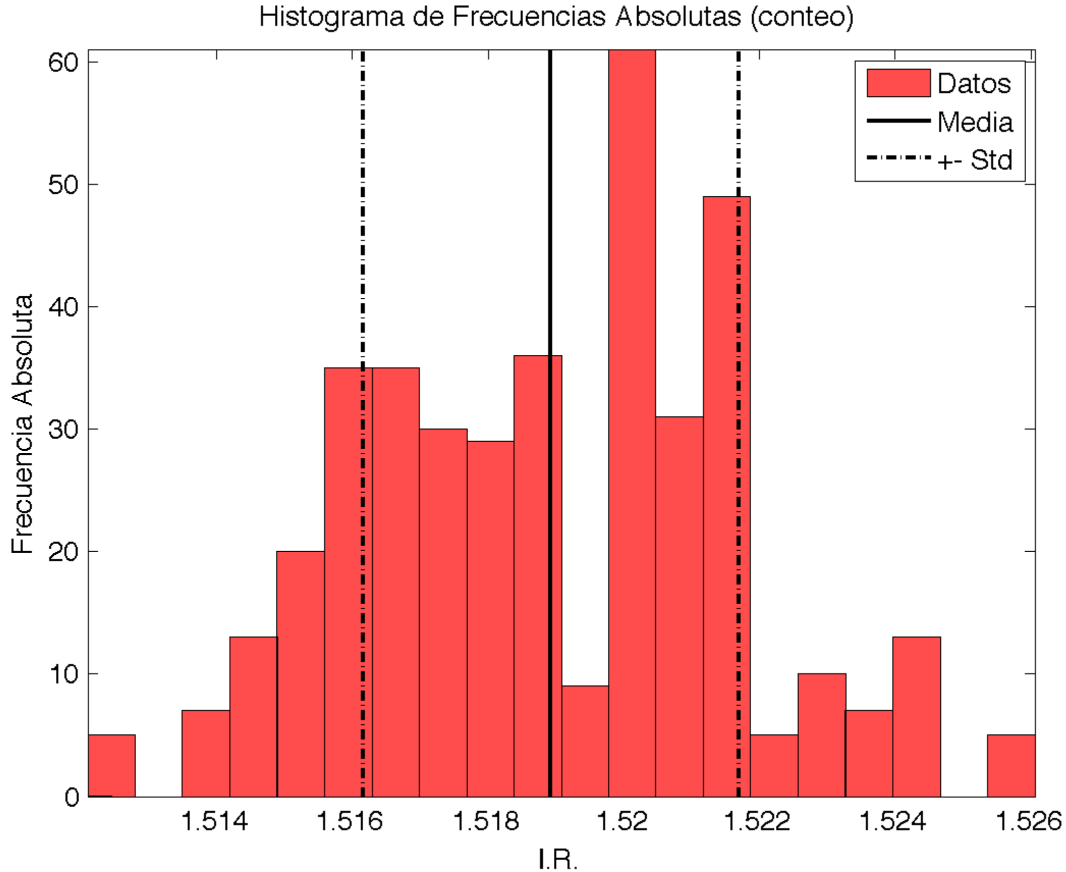
Varianza

$$S_X^2 = \frac{1}{N}\sum_{i=1}^{N}\left(x_i - \bar{X}\right)^2 = 7{,}65 \cdot 10^{-6}$$

Dispersión: Desviación Tipica (std)

- Desviación típica (std): raíz cuadrada de la varianza

$$S_X = \sqrt{S_X^2}$$

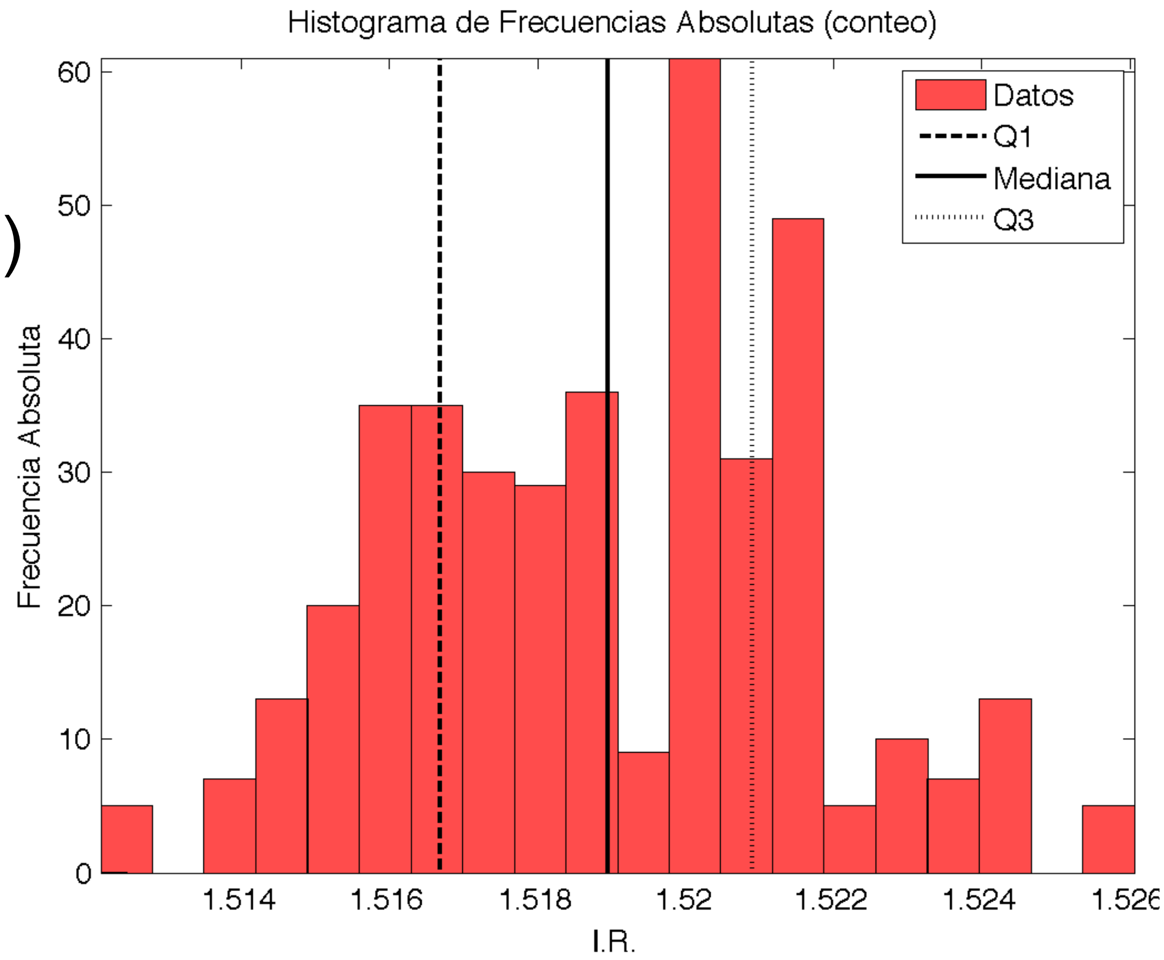


- **Ejemplo**: IR ([Zadora14])

$$S_X = \sqrt{7{,}65 \cdot 10^{-6}} = 0{,}0028$$

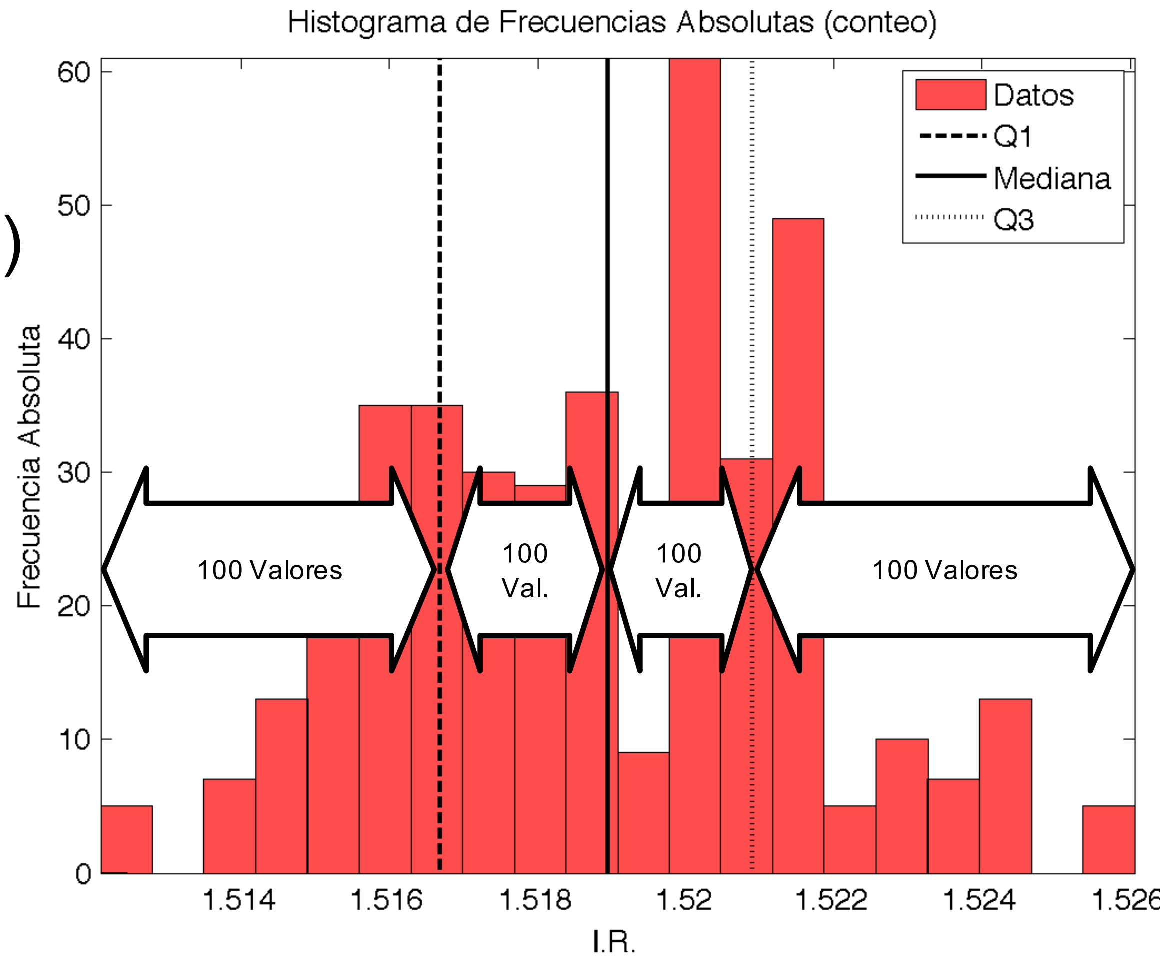
Dispersión: Cuartiles

- Delimitan zonas donde hay el 25% de los datos
- **Ejemplo**: IR ([Zadora14])



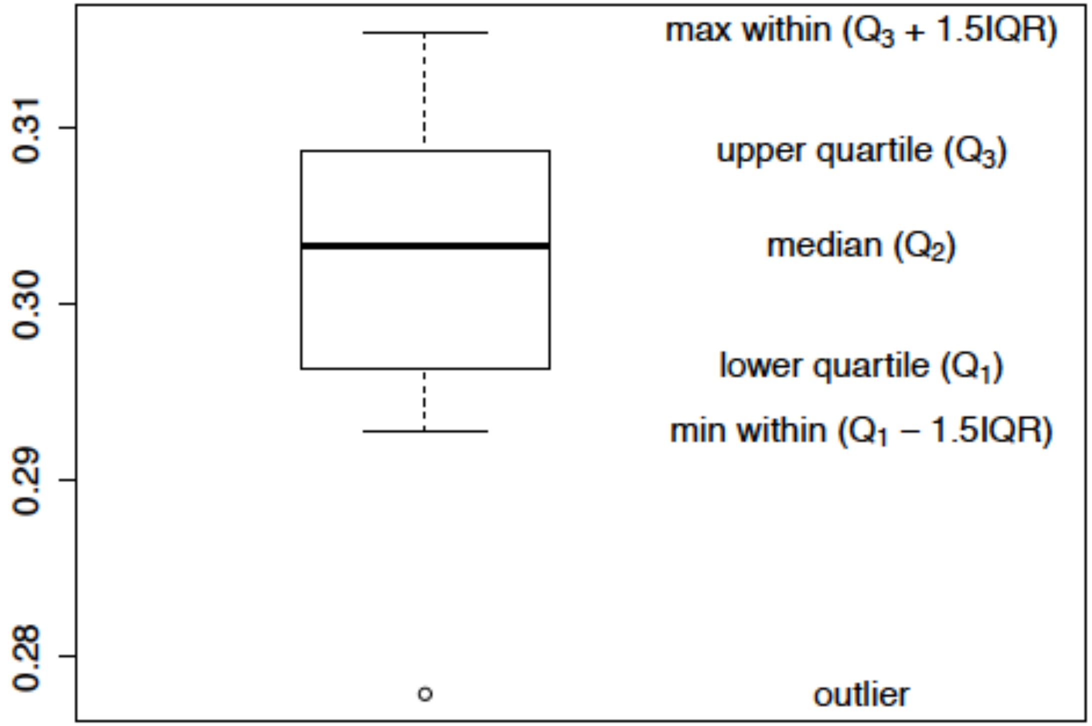
Dispersión: Cuartiles

- Delimitan zonas donde hay el 25% de los datos
- **Ejemplo**: IR ([Zadora14])
  - Total: 400 valores
  - 100 valores entre cuartiles (400/4)
  - Q2 es la mediana
    - Hay 200 valores a cada uno de sus lados



Diagramas de Cajas

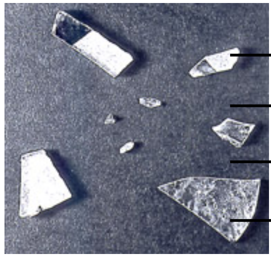
- Representan los cuartiles (Q1, Q2 = mediana, Q3)
  - Más algunos valores extremos
- **Ejemplo**: ([Zadora14])



Datos Multivariados

Muestra Multivariada

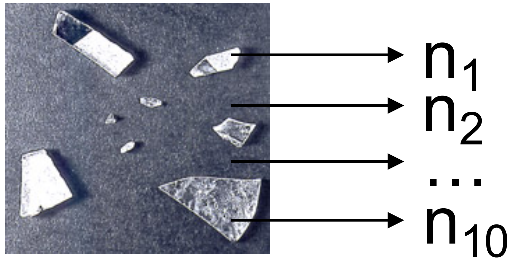
- Muestra univariada: de cada medida sólo tomamos un nº
  - **Ejemplo**: índice de refracción de un vidrio
    - 10 medidas
    - 1 valor por medida
    - 10 valores en total



n_1
n_2
...
n_{10}

Muestra Multivariada

- Muestra univariada: de cada medida sólo tomamos un nº
  - **Ejemplo**: índice de refracción de un vidrio
    - 10 medidas
    - 1 valor por medida
    - 10 valores en total



- Muestra multivariada: de cada medida tomamos varios nºs
  - **Ejemplo**: concentración de Na, Al, Mg en un vidrio
    - 10 medidas
    - 3 valores por medida
    - 30 valores en total

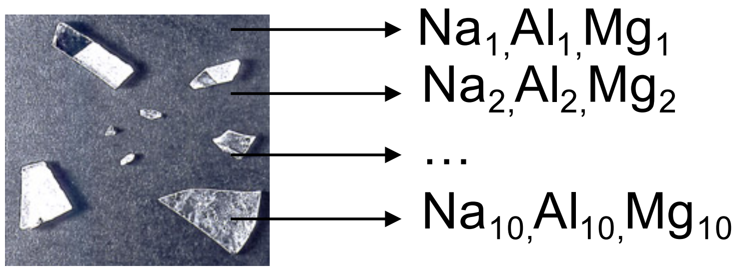


Tabla de Datos Multivariados

- Cada medida tiene varios valores
  - **Ejemplo**: concentración de elementos en vidrio ([Zadora14])

Número de Medida	Log(Na/O)	Log(Mg/O)	Log(Al/O)
1	-0.645325	-1.461683	-1.484642
2	-0.618750	-1.377008	-1.711383
3	-0.613208	-1.374241	-1.687958
4	-0.615041	-1.374233	-1.709916
…	…		

Tabla de Datos Multivariados

- Cada medida tiene varios valores
  - **Ejemplo**: concentración de elementos en vidrio ([Zadora14])
- Cada columna por separado es una muestra univariada
  - Se puede calcular media, std, etc.

Muestra univariada

Número de Medida	Log(Na/O)	Log(Mg/O)	Log(Al/O)
1	-0.645325	-1.461683	-1.484642
2	-0.618750	-1.377008	-1.711383
3	-0.613208	-1.374241	-1.687958
4	-0.615041	-1.374233	-1.709916
…	…		

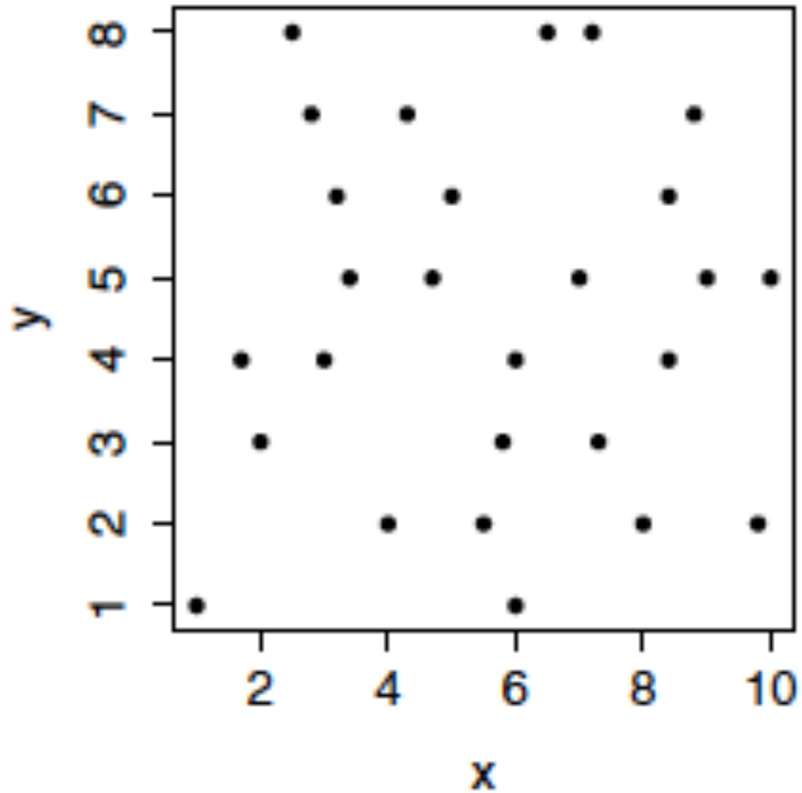
Datos Bivariados (2-variados)

- Cada medida tiene dos valores
  - **Ejemplo**: concentración de elementos en vidrio ([Zadora14])
- Estos datos se pueden visualizar gráficamente
  - Con más de dos variables es mucho más difícil…

Número de Medida	**Log(Na/O)**	**Log(Mg/O)**
1	-0.645325	-1.461683
2	-0.618750	-1.377008
3	-0.613208	-1.374241
4	-0.615041	-1.374233
…	…	

Datos Bivariados: Diagrama de Puntos

- Eje x: uno de los valores, como por ejemplo log(Na/O)
- Eje y: otro de los valores, como por ejemplo log(Si/O)
  - **Ejemplo**: ([Zadora14])



Correlación

- Los valores que se toman tienen relación entre ellos
  - En estadística se le llama correlación

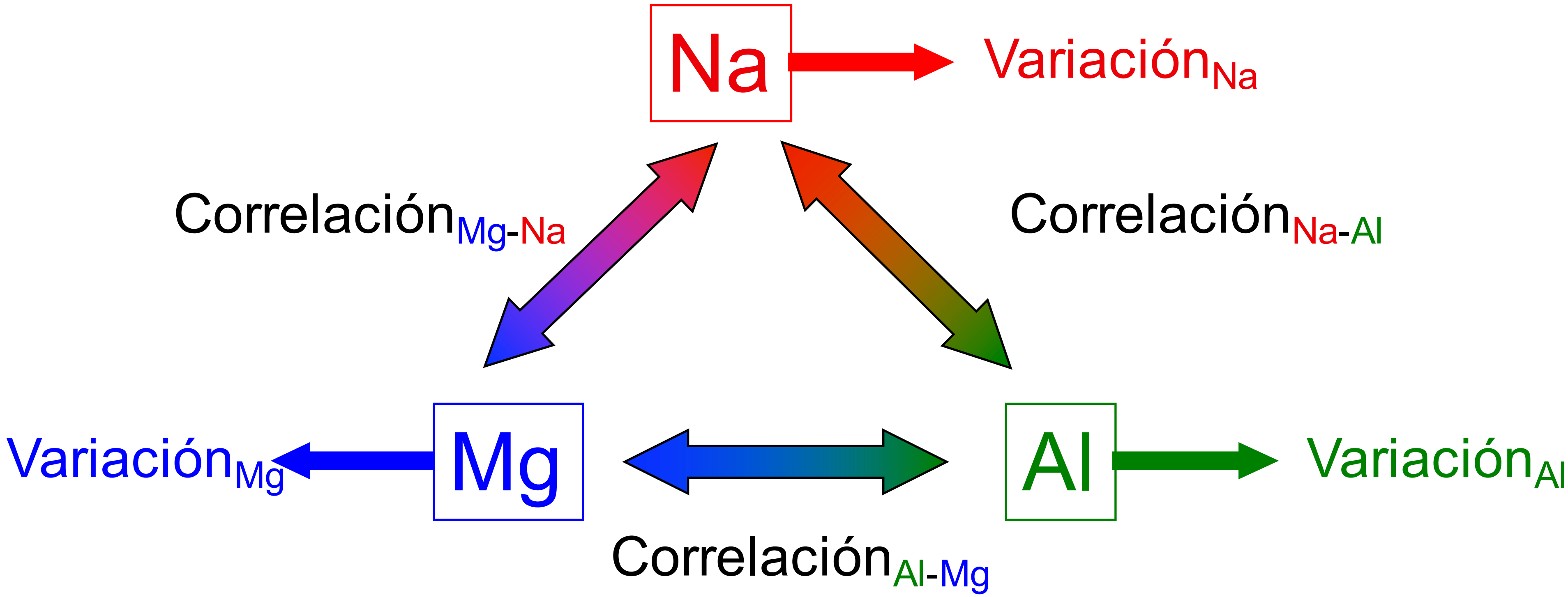
$\boxed{\text{Na}} \rightarrow \text{Variación}_{\text{Na}}$

$\text{Variación}_{\text{Mg}} \leftarrow \boxed{\text{Mg}}$

$\boxed{\text{Al}} \rightarrow \text{Variación}_{\text{Al}}$

Correlación

- Los valores que se toman tienen relación entre ellos
  - En estadística se le llama correlación

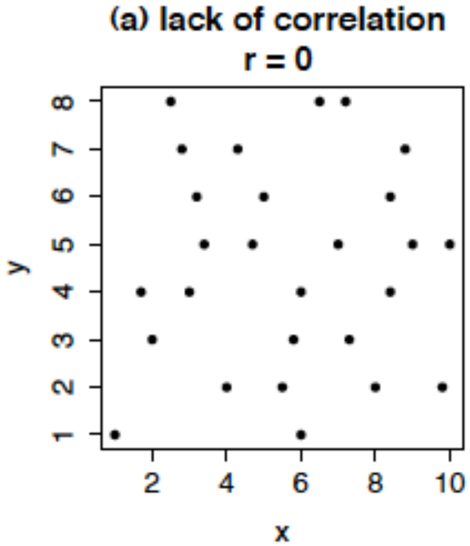


Datos Bivariados: Coeficiente de Correlación (r)

- Mide correlación entre las dos variables de la muestra
- Entre -1 y 1
  - r=1: correlación total (las dos variables son la misma)
  - r=-1: correlación total (una variable es la negativa de la otra)
  - r=0: no hay correlación

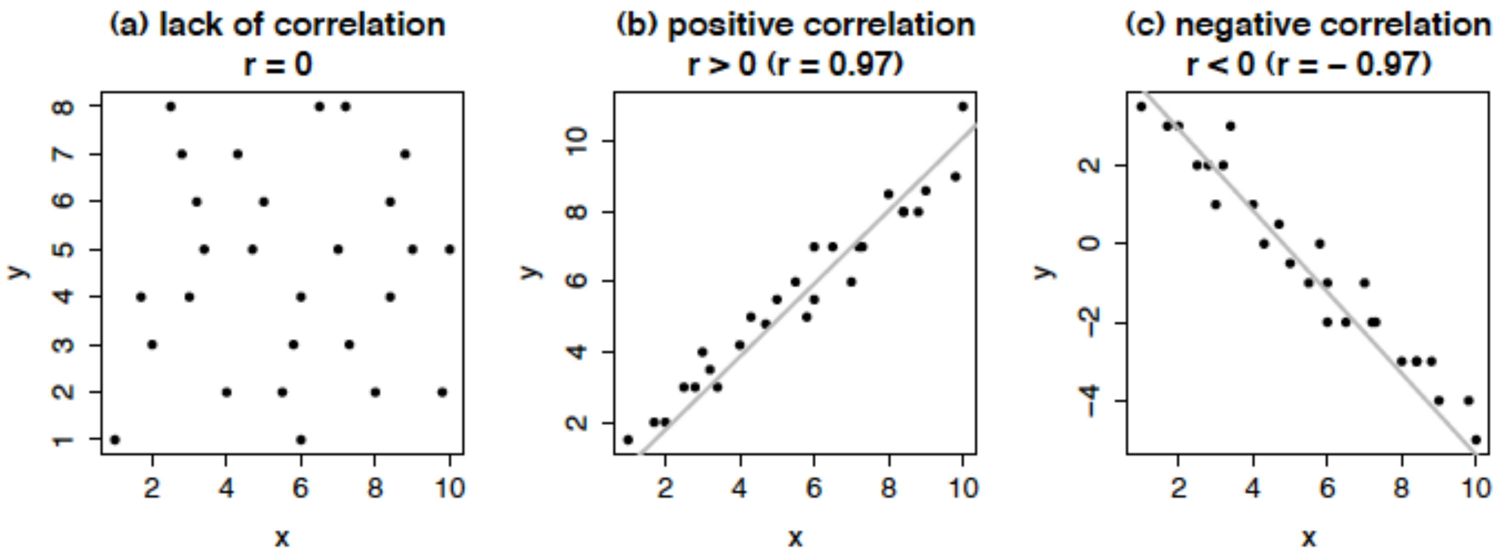
Datos Bivariados: Coeficiente de Correlación (r)

- **Ejemplo**: ([Zadora14])



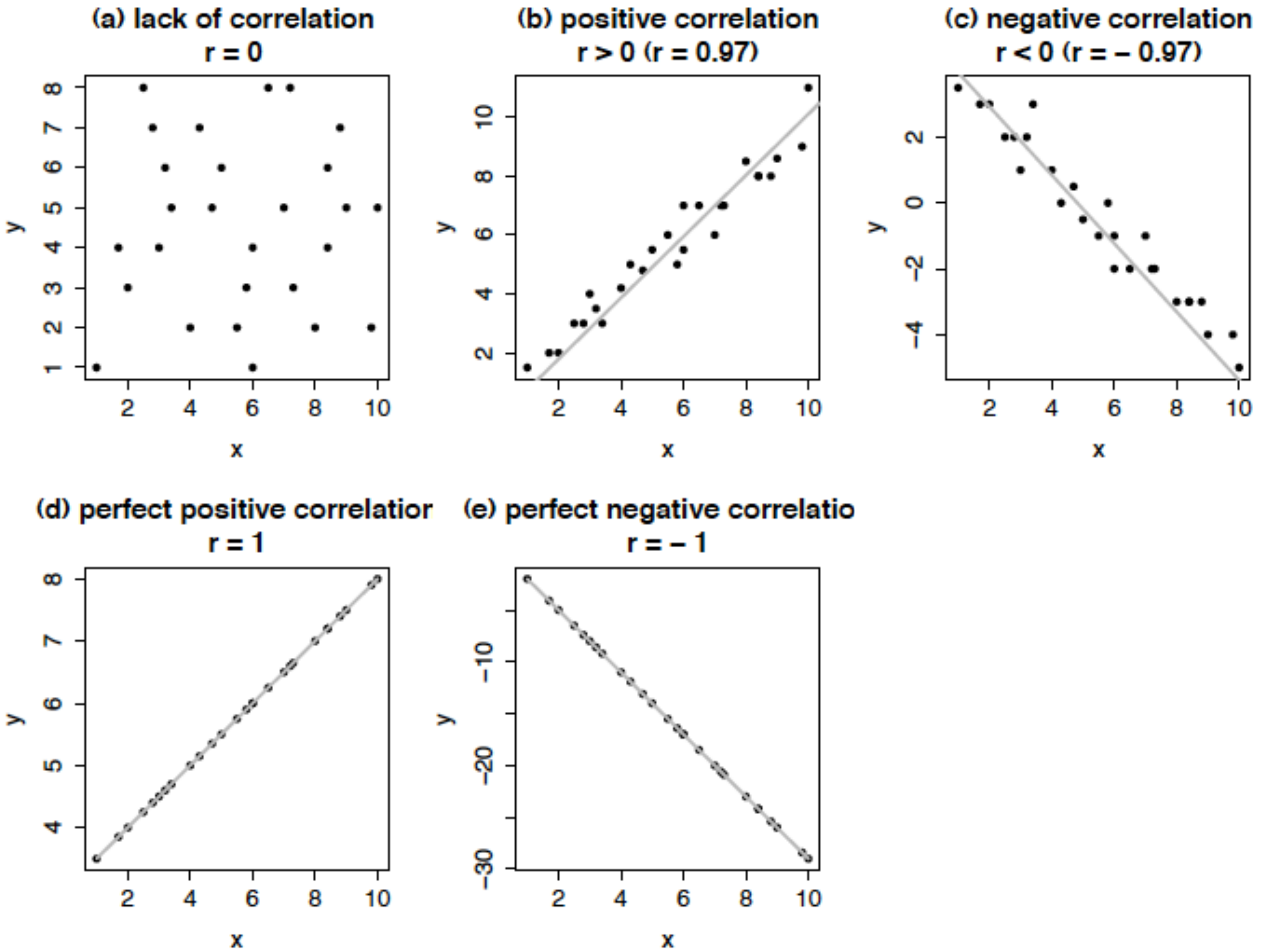
Datos Bivariados: Coeficiente de Correlación (r)

- **Ejemplo**: ([Zadora14])



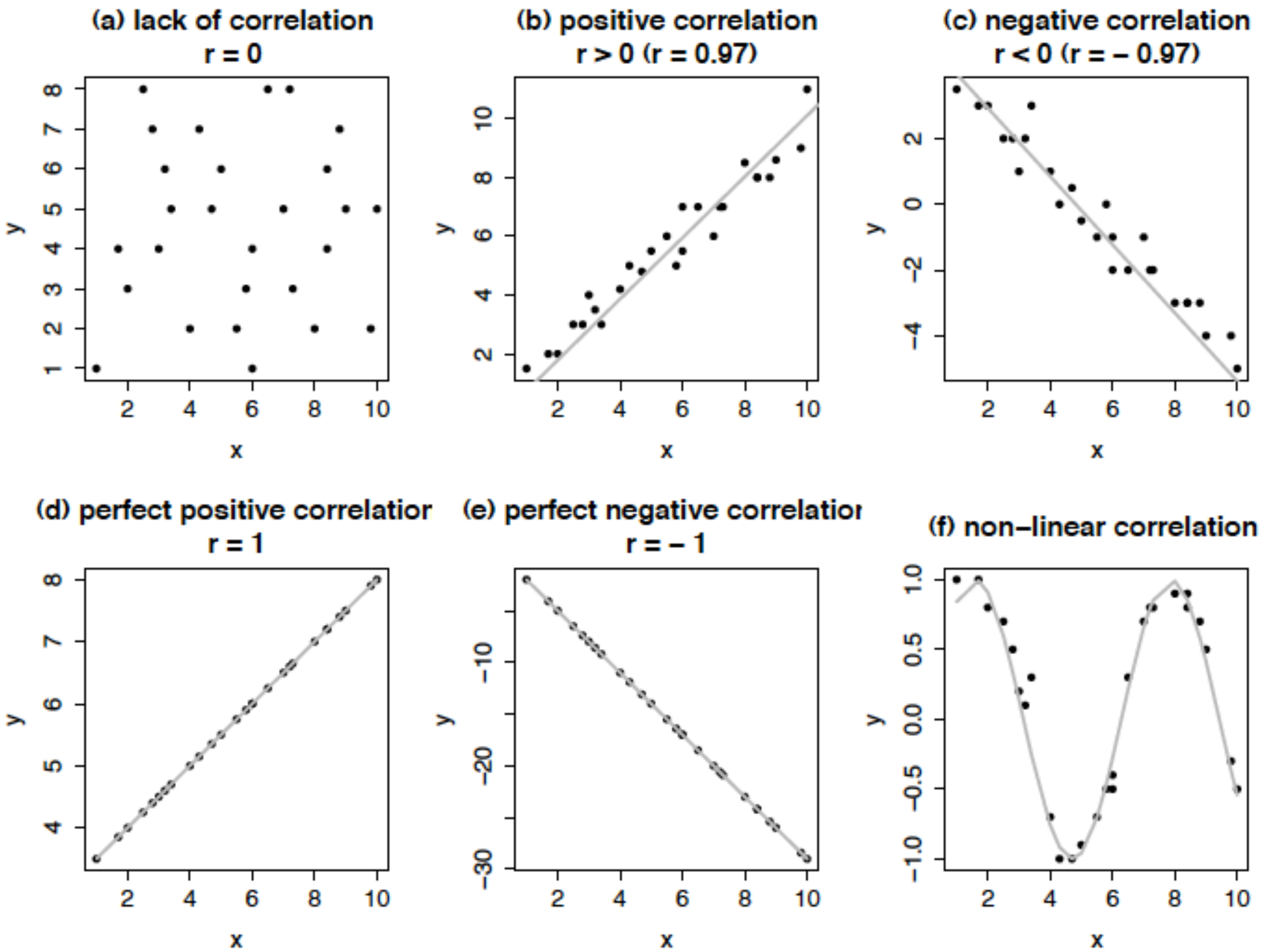
Datos Bivariados: Coeficiente de Correlación (r)

- **Ejemplo**: ([Zadora14])



Datos Bivariados: Coeficiente de Correlación (r)

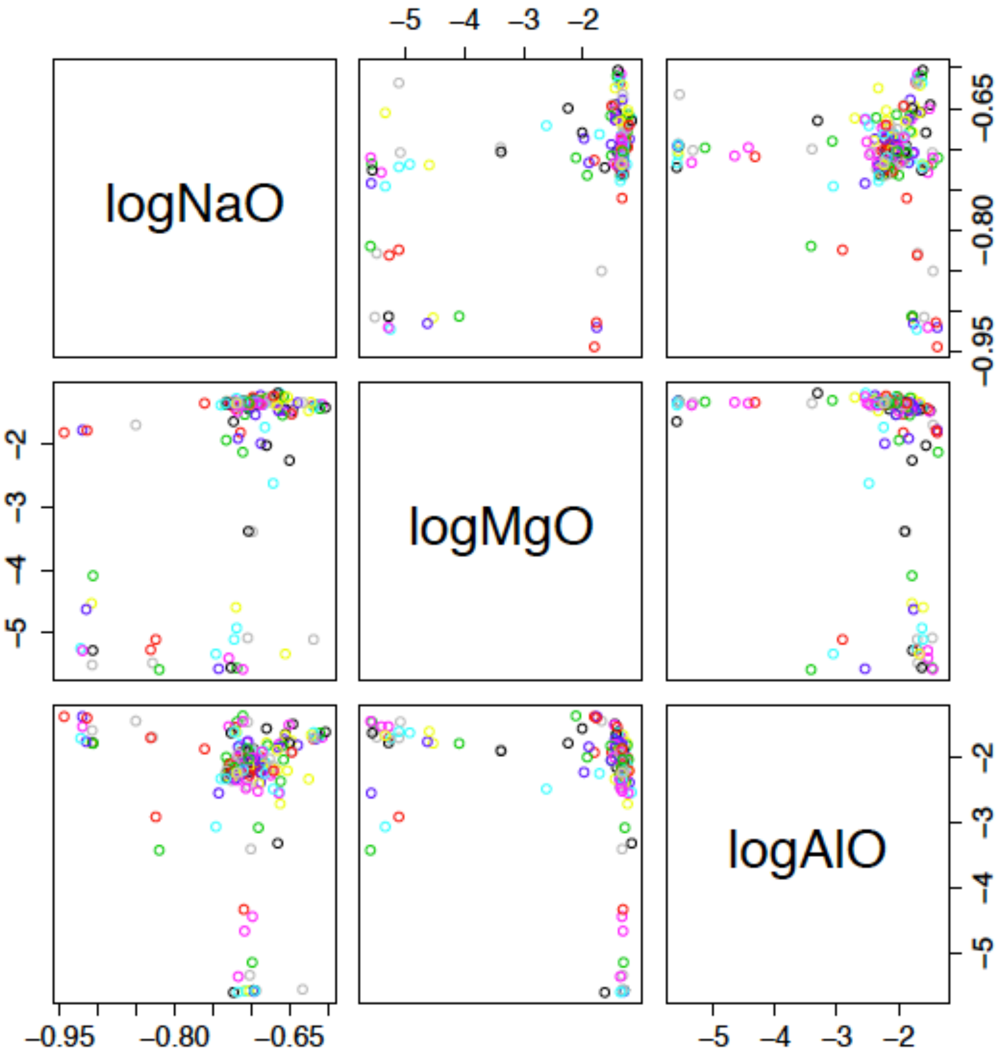
- **Ejemplo**: ([Zadora14])



Datos Multivariados

- Se pueden representar múltiples diagramas de puntos, 2 a 2

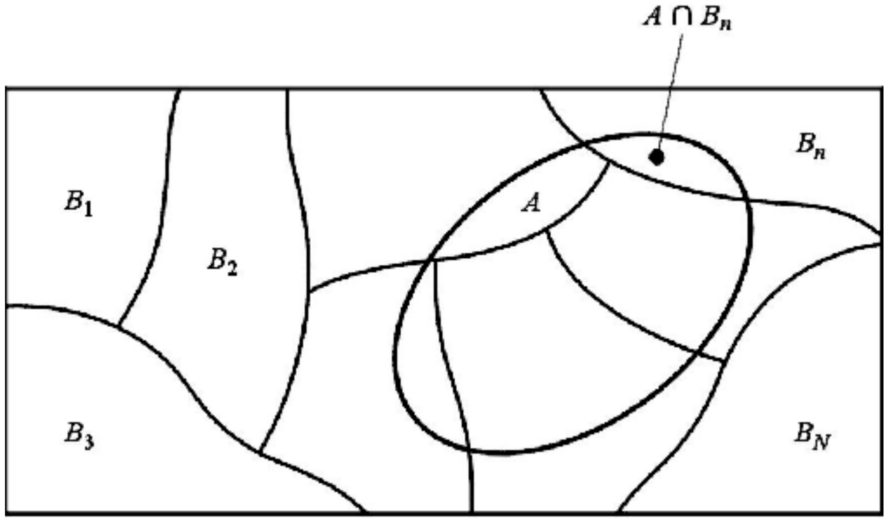
[Zadora14]



Probabilidad

Experimentos

- Un experimento se define mediante la especificación de tres factores:
  - El espacio muestral de dicho experimento
  - Los sucesos definidos en ese espacio muestral
  - Las probabilidades de dichos sucesos



Experimentos

- **Ejemplo**: lanzar un dado "legal"
  - "Legal" significa que todas las caras tienen la misma probabilidad de aparecer
  - Sucesos elementales posibles
    - Que salga un número del 1 al 6:

$$1 \equiv \text{"Que salga 1"}; \quad 2 \equiv \text{"Que salga 2"}; \quad \ldots$$

  - Espacio muestral: conjunto de todos los posibles sucesos

$$S = \{1,2,3,4,5,6\}$$

Probabilidad

- **Ejemplo**: dado legal

$$P(1)=P(2)=P(3)=P(4)=P(5)=P(6)=\frac{1}{6}$$

- Interpretación frecuentista
  - Lanzo el dado un número suficientemente grande de veces, idealmente infinito

$$P(\text{"Que salga X"})=\frac{\text{n veces apareció X}}{\text{N veces lancé el dado}}$$

- Interpretación bayesiana
  - Sé que debe ser igual probabilidad para todo número
  - Todos deben sumar 1, luego cada uno tiene 1/6

Sucesos Disjuntos

- Dos sucesos que no pueden ocurrir a la vez
- Intersección ($\cap$)
  - Se lee como “y” o “a la vez que”
  - Dos sucesos disjuntos tienen una intersección igual al conjunto vacío

- **Ejemplo**: dado legal
  - 1 y 2 son disjuntos, no pueden salir a la vez…
  - Su intersección es el conjunto vacío

$$\text{"Que salga 1"} \cap \text{"Que salga 2"} = \varnothing$$

Sucesos Compuestos

- Sucesos que se forman uniendo sucesos elementales
- Mediante la unión ($\cup$)
  - Se lee como "o"

- **Ejemplo**: dado legal
  - 1 y 2 son sucesos elementales
  - Su unión es el suceso compuesto "1 ó 2"

$$\text{"Que salga 1 ó 2"} = \text{"Que salga 1"} \cup \text{"Que salga 2"}$$

Probabilidad: Axiomas

- Si llamamos A a un suceso definido en el espacio muestral S, la probabilidad del suceso A, $P(A)$, cumple tres axiomas:

1.- $P(A) \geq 0$ La probabilidad es siempre positiva

2.- $P(S) = 1$ La probabilidad del espacio muestral es uno

3.- $$P\left(\bigcup_{n=1}^{N} A_n\right) = \sum_{n=1}^{N} P\left(A_n\right) \quad \text{si} \quad A_n \cap A_m = \varnothing$$

La probabilidad de sucesos disjuntos es la suma de sus probabilidades

Probabilidad: Axiomas

- Ejemplo: dado legal

$$P(1) = P(2) = P(3) = P(4) = P(5) = P(6) = \frac{1}{6}$$

1.- $P(1) = P(2) = P(3) = P(4) = P(5) = P(6) \geq 0$

2.- $P(1 \text{ ó } 2 \text{ ó } 3 \text{ ó } 4 \text{ ó } 5 \text{ ó } 6) = P(S) = 1$

3.- $P(1 \text{ ó } 2 \text{ ó } 3) = P(1) + P(2) + P(3) = \frac{3}{6} = \frac{1}{2}$

Probabilidad Condicional

- La probabilidad condicional describe la incertidumbre de un evento una vez ha ocurrido otro:

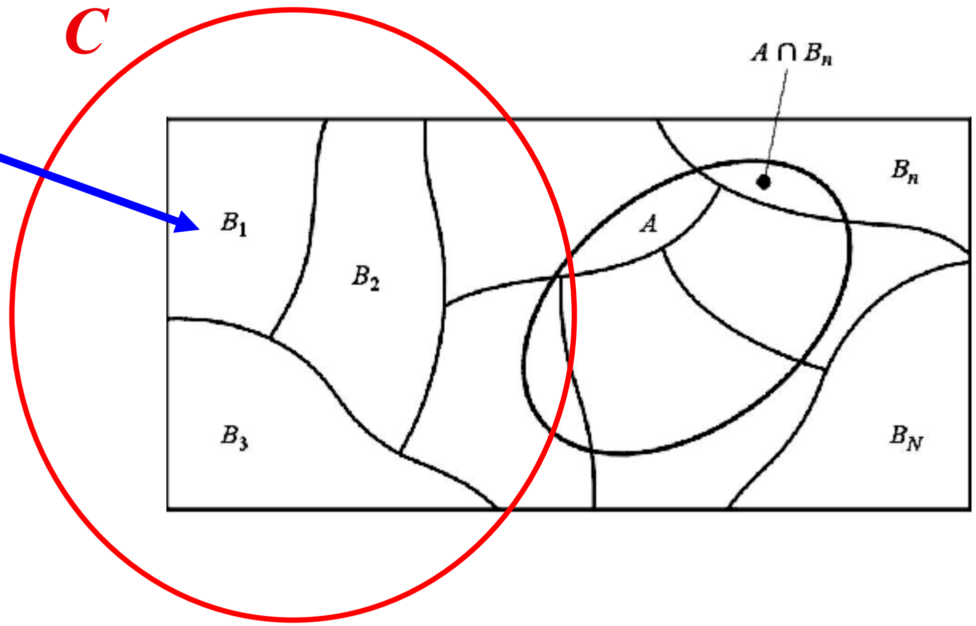
$$P\left(A \middle| B\right) = \frac{P\left(A \cap B\right)}{P\left(B\right)}$$

- La probabilidad condicional satisface los tres axiomas de la probabilidad
- También llamada “Regla del Producto”

Probabilidad Condicional

- Lo que ocurre con la probabilidad condicional es que se crea un "nuevo espacio muestral" mediante la restricción del espacio muestral original por causa de la condición

$$P\left(B_1 \middle| C\right) = \frac{P\left(B_1 \cap C\right)}{P\left(C\right)}$$



Probabilidad Condicional

- En el caso del dado "legal"

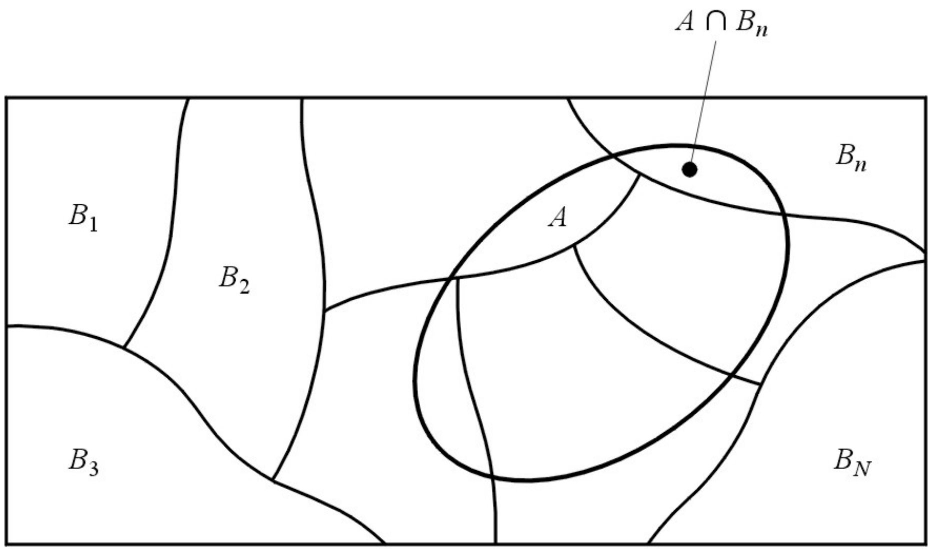
$$P\left(2 \middle| \text{par}\right) = \frac{P\left(2 \cap \text{par}\right)}{P\left(\text{par}\right)} = \frac{P\left(2\right)}{P\left(\text{par}\right)} = \frac{1/6}{1/2} = \frac{1}{3}$$

- Si nos dicen que es par, la probabilidad de que sea un 2 es una entre tres (2, 4 y 6).
- Frecuentista: tiraríamos el dado infinitas veces, sólo nos quedríamos con los pares y calcularíamos la proporción de doses.

Teorema de la Probabilidad Total

- Tenemos sucesos $\{B_1, \ldots, B_n\}$ que cubren todo un suceso A y que son disjuntos.
- En esas condiciones, la probabilidad de A se puede obtener a partir de las probabilidades condicionadas de B.

$$P(A) = \sum_{i=1}^{n} P(A \mid B_i) P(B_i)$$



$\bigcup_{n=1}^{N} B_n = S$, $B_m \cap B_n = \varnothing$ for all $m \neq n$

Probabilidad Total

- Dado legal

$$P(<4) = P(<4|\text{par})P(\text{par}) + P(<4|\text{impar})P(\text{impar})$$

$$= P(2|\text{par})P(\text{par}) + P(1\text{ ó }3|\text{impar})P(\text{impar})$$

$$= \frac{1}{3}\cdot\frac{1}{2} + \frac{2}{3}\cdot\frac{1}{2} = \frac{1}{2}$$

Teorema de Bayes

- De importancia capital.
- Relaciona probabilidades condicionadas.

$$P\left(A \middle| B\right) = \frac{P\left(B \middle| A\right) P\left(A\right)}{P\left(B\right)}$$

Teorema de Bayes

- Dado legal
- Imaginemos que queremos saber la probabilidad de que un número sea par
  - Si sabemos ya que ha salido un número menor de 4

$$P\left(<4 \mid \text{par}\right) = \frac{1}{3} \qquad \text{(sólo 2 de 2, 4 y 6)}$$

$$P\left(\text{par} \mid <4\right) = \frac{P\left(<4 \mid \text{par}\right) P\left(\text{par}\right)}{P\left(<4\right)} = \frac{\frac{1}{3}\,\frac{1}{2}}{\frac{1}{2}} = \frac{1}{3} \qquad \text{(sólo 2 de 1, 2 y 3)}$$

Teorema de Bayes

- Se puede combinar con el teorema de la probabilidad total

$$P(A|B)=\frac{P(B|A)P(A)}{P(B|C_1)P(C_1)+\ldots+P(B|C_n)P(C_n)}$$

- Útil para descomponer probabilidades complicadas en otras más simples de calcular

Teorema de Bayes

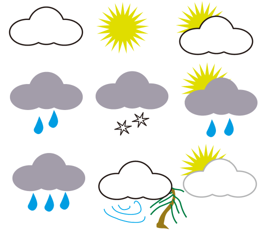
- Dado legal.
- Calculemos la probabilidad de "par"
  - Sabiendo que el número que ha salido es menor que 4

$$P(\text{par}|<4)=\frac{P(<4|\text{par})P(\text{par})}{P(<4)}$$

$$=\frac{P(<4|\text{par})P(\text{par})}{P(<4|\text{par})P(\text{par})+P(<4|\text{impar})P(\text{impar})}$$

$$=\frac{\tfrac{1}{3}\,\tfrac{1}{2}}{\tfrac{1}{2}}=\frac{1}{3}$$

Teorema de Bayes



- Importancia
  - Relaciona probabilidades antes y después de conocer información
  - **Ejemplo**:
    - Parte meteorológico cambia nuestra probabilidad de que llueva mañana
  - La información cambia nuestra opinión (probabilística) sobre las cosas
    - Y el teorema de Bayes nos dice cómo es ese cambio

$$P(\text{Lluvia}|\text{Previsión}) = \frac{P(\text{Previsión}|\text{Lluvia})}{P(\text{Previsión})} P(\text{Lluvia})$$

Independencia

- Dos sucesos son independientes si…

$$P(A|B) = P(A)$$

- Ello implica que…

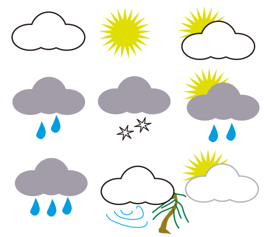
$$P(A \cap B) = P(A)P(B)$$

- Se dice que dos sucesos independientes no comparten información
  - O que conocer uno de ellos no nos aporta ninguna información sobre el otro

Independencia

- **Ejemplo**: Dado legal
- Lanzamos el dado dos veces
  - Obtenemos un cuatro en el primero y un dos en el segundo
  - El resultado del segundo lanzamiento no nos dice nada sobre cuál fue el número en el primer lanzamiento
    - Ambos lanzamientos son independientes

Independencia



- Ejemplo: previsión meteorológica
  - La previsión nos dice mucho sobre si mañana va a llover
  - "Lluvia" y "Previsión" son sucesos dependientes

$$P\left(\text{Lluvia}\,\middle|\,\text{Previsión}\right) \neq P\left(\text{Lluvia}\right)$$

  - El precio del yen no nos dice nada sobre si va a llover mañana o no
  - "Precio Yen" y "Lluvia son independientes

$$P\left(\text{Lluvia}\,\middle|\,\text{Yen}\right) = P\left(\text{Lluvia}\right)$$

Probabilidades engañosas

- Cuidado con las probabilidades
- Son engañosas
- La intuición no es buena compañera
- Mejor fórmulas que razonamientos intuitivos

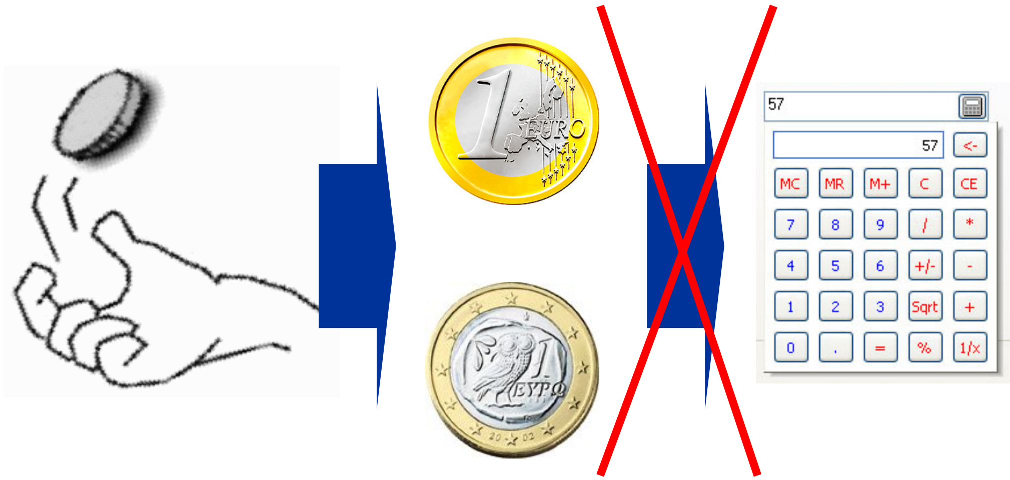
Variable Aleatoria (V. A.)

¿Por qué necesitamos la V. A.?

- Porque los números nos permiten muchas cosas
  - Operaciones matemáticas
  - Cálculos intensivos con un ordenador

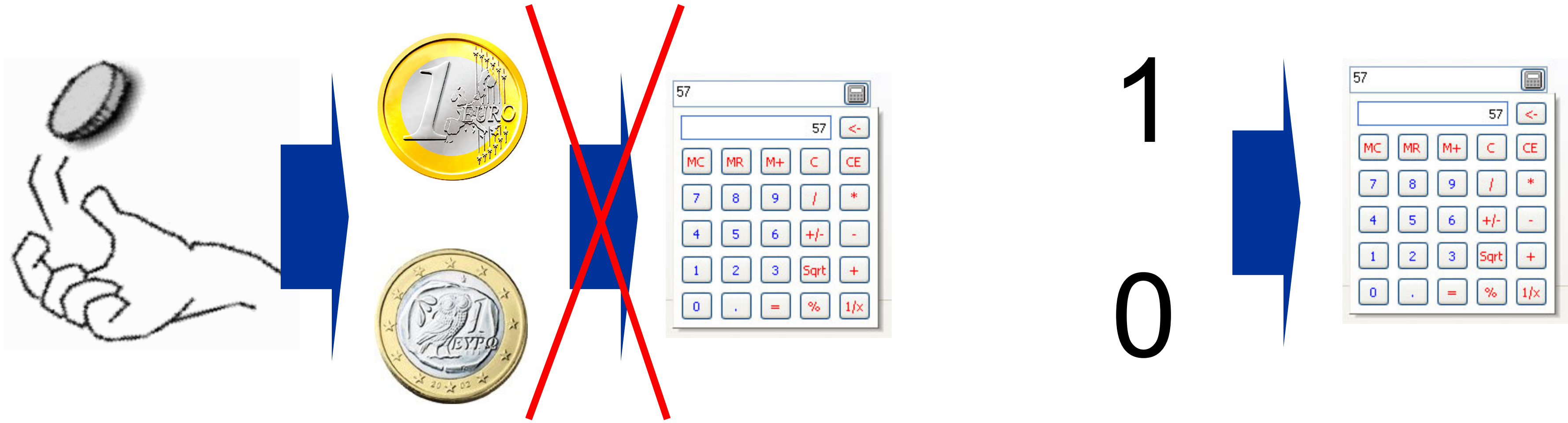
¿Por qué necesitamos la V. A.?

- Porque los números nos permiten muchas cosas
  - Operaciones matemáticas
  - Cálculos intensivos con un ordenador



¿Por qué necesitamos la V. A.?

- Porque los números nos permiten muchas cosas
  - Operaciones matemáticas
  - Cálculos intensivos con un ordenador

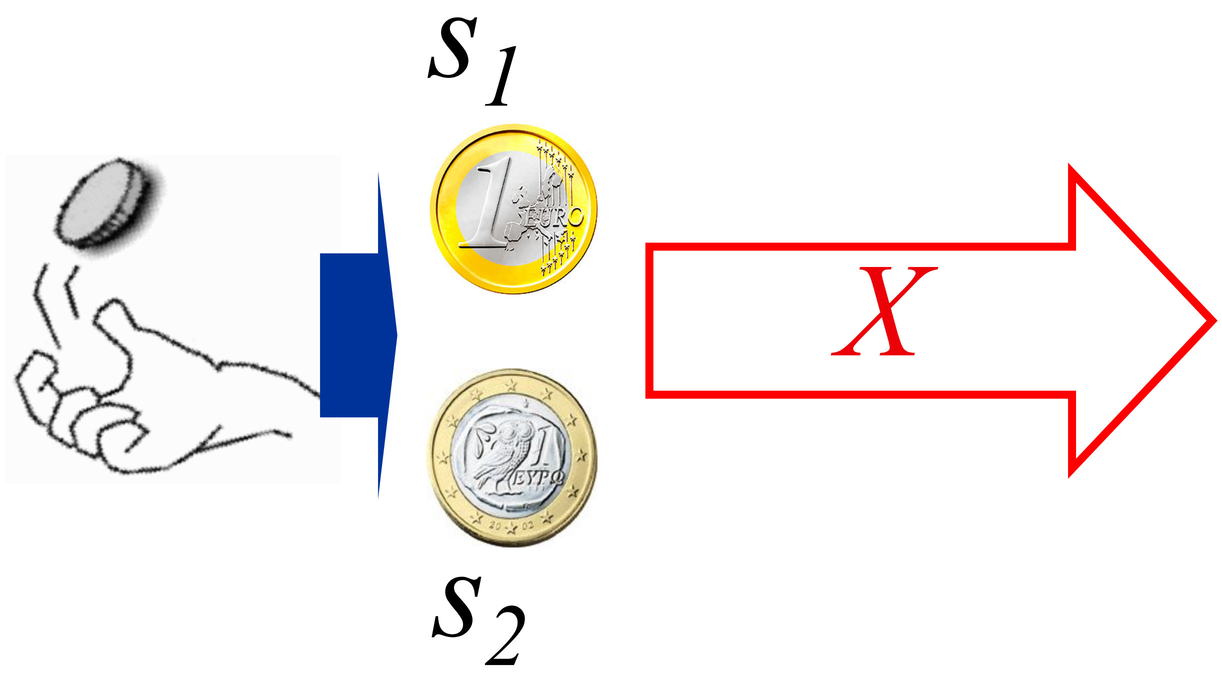


Definición de V. A.

- Definición de variable aleatoria:
  - ⇨ Es una función real de los elementos de un espacio muestral S
- Representamos variables aleatorias con mayúsculas: W, X, Y
- Cualquier valor particular que toma la variable aleatoria lo representamos con minúsculas: w, x, y

Definición de V. A.

- Dado un experimento definido por el espacio muestral S con elementos s, asignamos a cada s un número real $X(s)$:

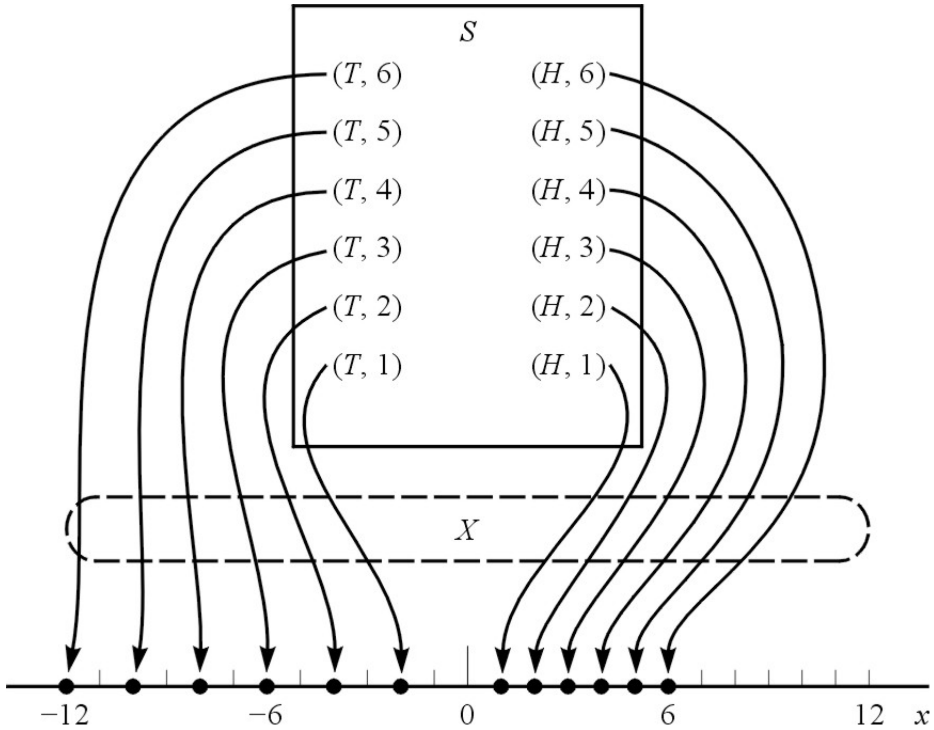


$$X(s_1)=1$$

$$X(s_2)=0$$

V. A. discreta o continua

- Una variable aleatoria puede tomar valores discretos (finitos o infinitos) o continuos.
- Ej. V. A. discreta: tiro dado y a la vez moneda:



V. A. discreta o continua

- Ej. V. A. discreta infinita: cuento el número de coches que pasan por hora en un punto de la M-40 (autopista española en Madrid)

X

1

2

35

800

...

V. A. discreta o continua

- Ej. V. A. continua: altura de una persona.

s_1: Altura Averell = x_4

s_1: Altura William = x_3

s_1: Altura Jack = x_2

s_1: Altura Joe = x_1

X

$$X(s_1)=x_1$$

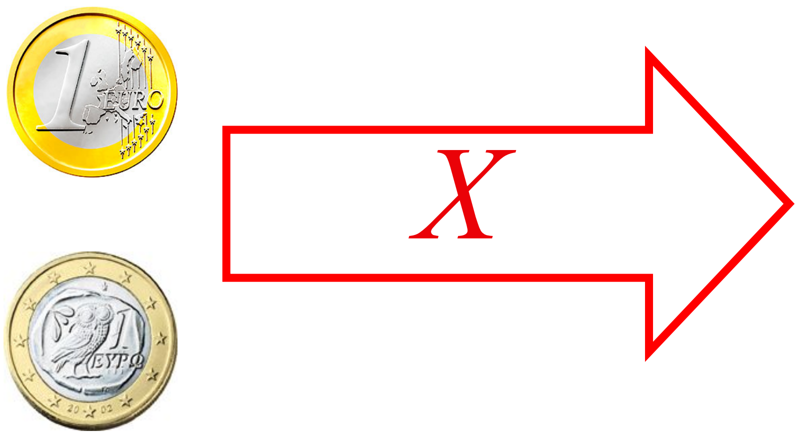
$$X(s_2)=x_2$$

$$X(s_3)=x_3$$

$$X(s_4)=x_4$$

Probabilidad de una V. A.

- Es la probabilidad que se le asignaría al suceso asociado a la variable.



$$X(s_1)=1=x_1$$

$$X(s_2)=0=x_2$$

$$P(s_1)=\frac{1}{2} \quad \Rightarrow \quad P_X(x_1)=\frac{1}{2}$$

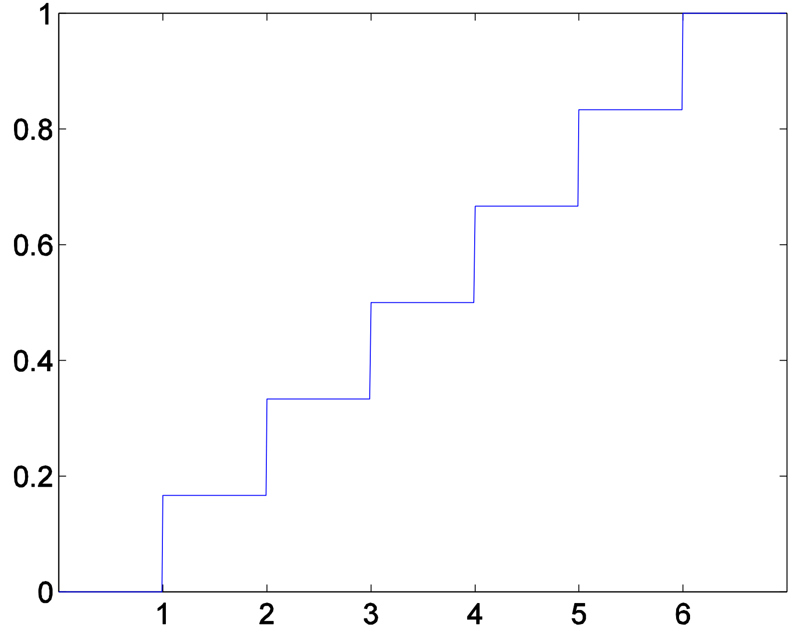
Función de distribución

- Es la probabilidad de que una V. A. sea menor que un valor: $F_X(X)=P(X \leq x)$
- Se le conoce también función acumulada de probabilidad
- Su valor en el infinito es uno…
  - …porque una V. A. siempre es menor o igual que infinito.

Ejemplo discreto: dado "legal"

- La distribución está "escalonada", porque la variable es discreta.

$$F_X(x)=P(X\leq x)$$

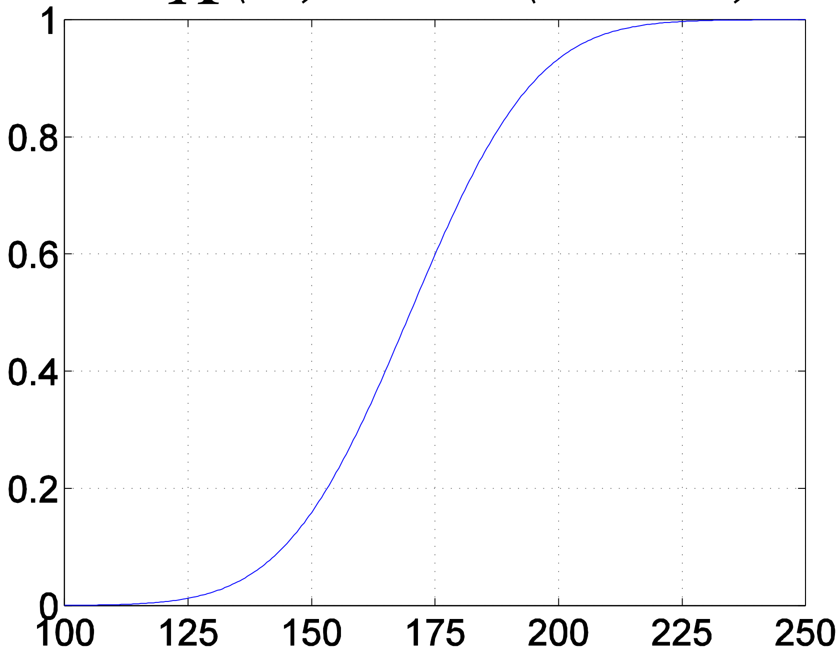


X="numero que sale en el dado"

Ejemplo continuo: altura

- La distribución es "suave" (en general) porque cualqier valor puede hacer aumentar la probabilidad.

$$F_X(x)=P(X\leq x)$$

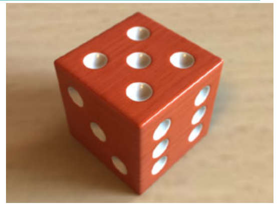


x="altura de una persona adulta en cm"

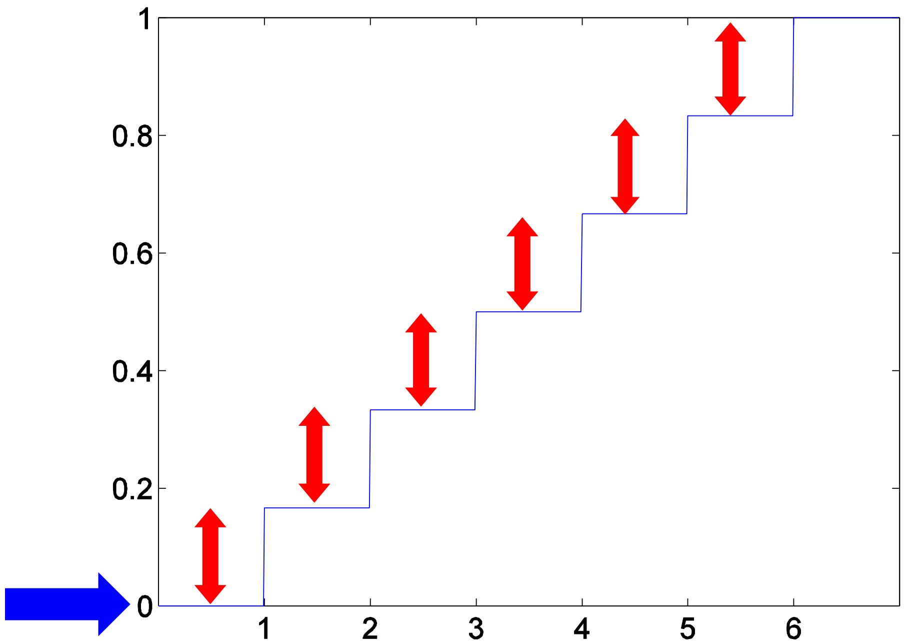
Función de probabilidad discreta

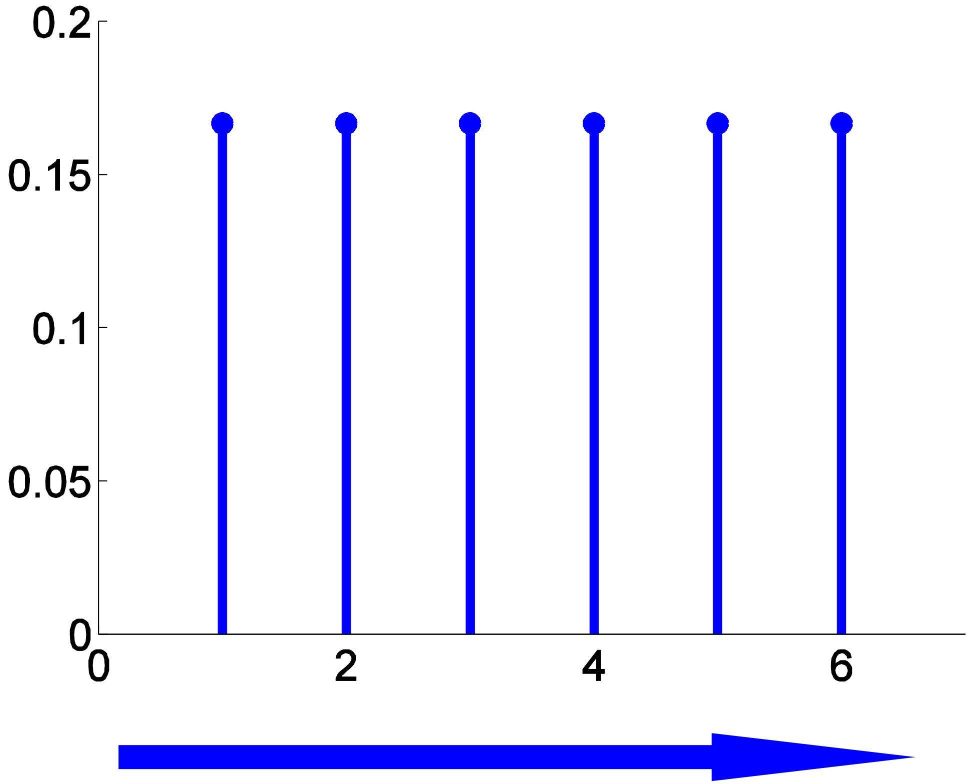
- Si hallamos la diferencia entre un valor y otro de la función de distribución discreta…
- Obtendremos los valores que hacen aumentar la probabilidad acumulada
- En los puntos en los que existe probabilidad

Ejemplo: dado "legal"



$$F_X(x) = P(X \le x) \qquad P_X(x)$$



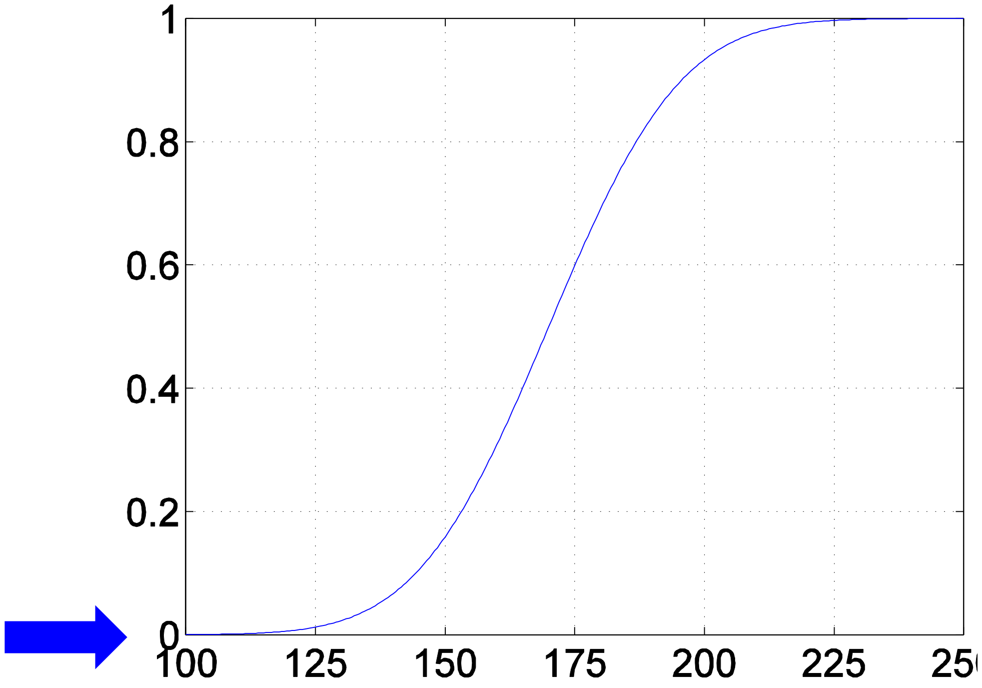


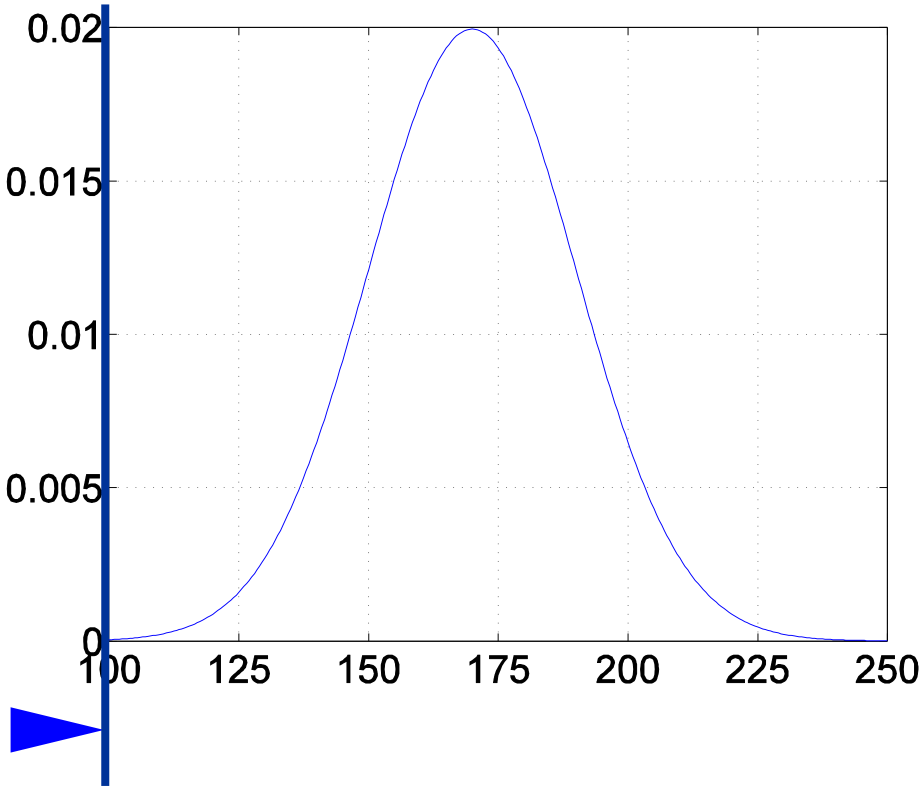
Función densidad de probabilidad

- La diferencia en el dominio continuo es la derivada.
- Derivando $F_X(x)$ obtendremos una "densidad" de probabilidad.
- Si se suma toda la "densidad" desde el valor más bajo obtenemos la función de distribución
  - Integral

Ejemplo: altura

$$F_X(x)=P(X\leq x) \qquad\qquad f_X(x)$$

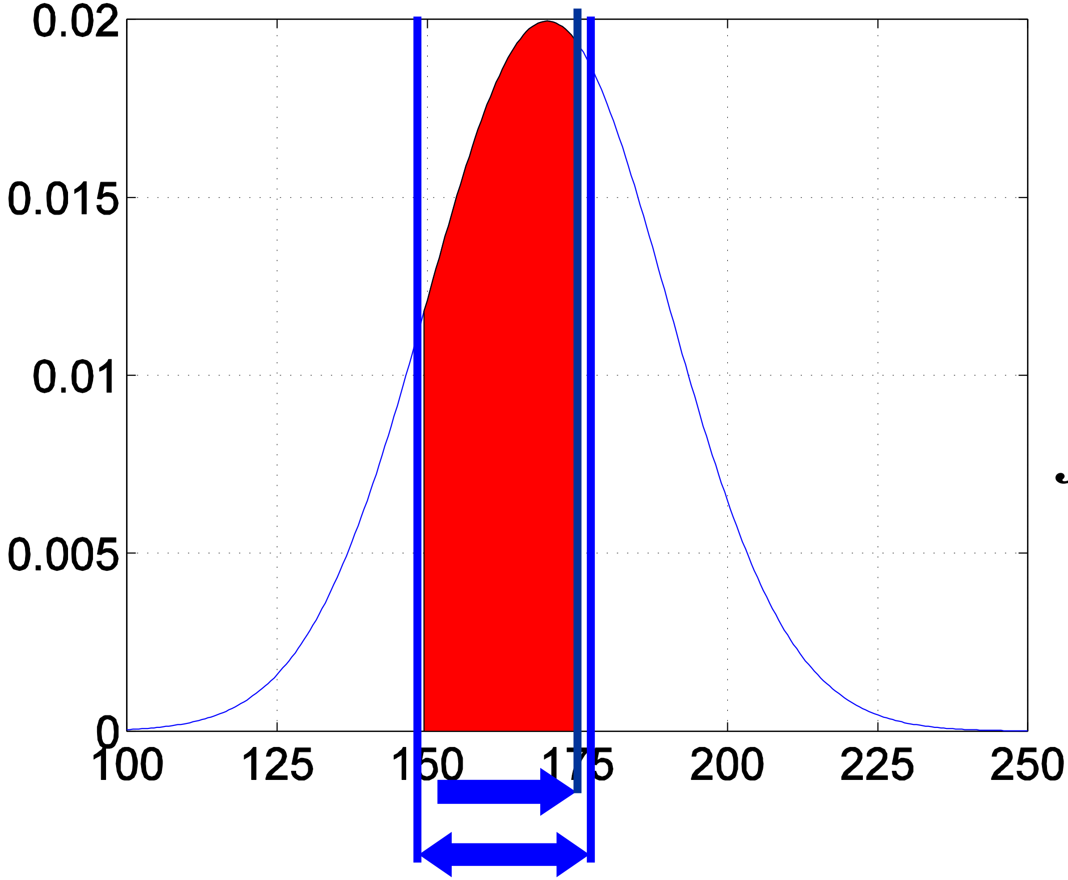




Probabilidad en el domino continuo

- ¿Cuál es la probabilidad de obtener un valor de una V. A. continua?
  - Probabilidad de que alguien mida 176,2323805… cm.
- ¡Cero!
- En continuo, la probabilidad de un evento aislado es cero.
- La probabilidad de que una variable aleatoria caiga en un rango es distinta de cero.

Ejemplo: altura



$$f_X(176{,}23\ldots)=0$$

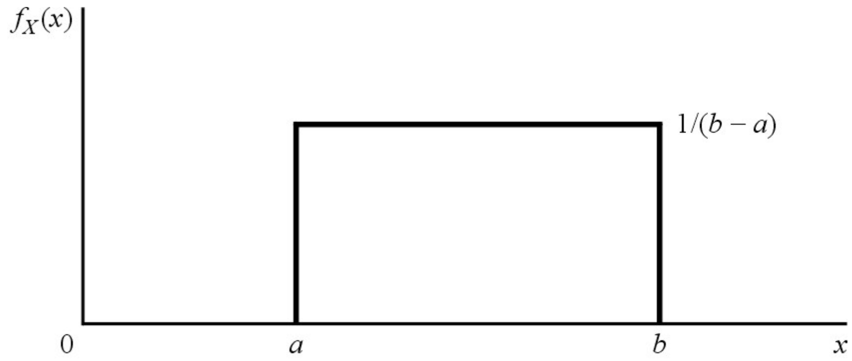
$$P_X(176{,}23\ldots)=0$$

$$P_X(150{,}76\ldots<x<176{,}23\ldots)>0$$

Ejemplos de V. A. continuas

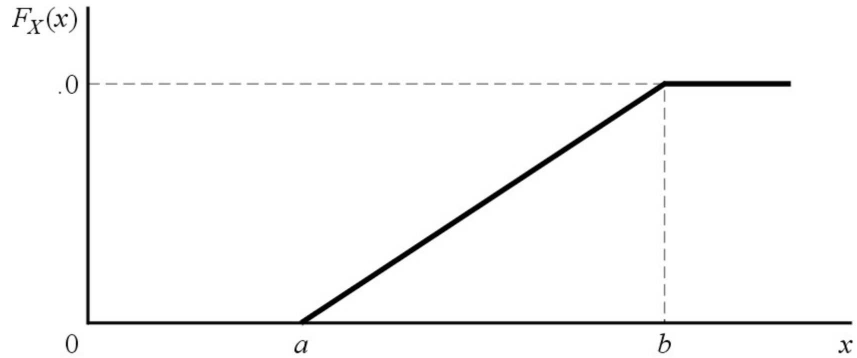
- V. A. uniforme

$$f_X(x) = \begin{cases} 1/(b-a) & a \le x \le b \\ 0 & \text{elsewhere} \end{cases}$$



(a)

$$F_X(x) = \begin{cases} 0 & x < a \\ (x-a)/(b-a) & a \le x < b \\ 1 & b \le x \end{cases}$$



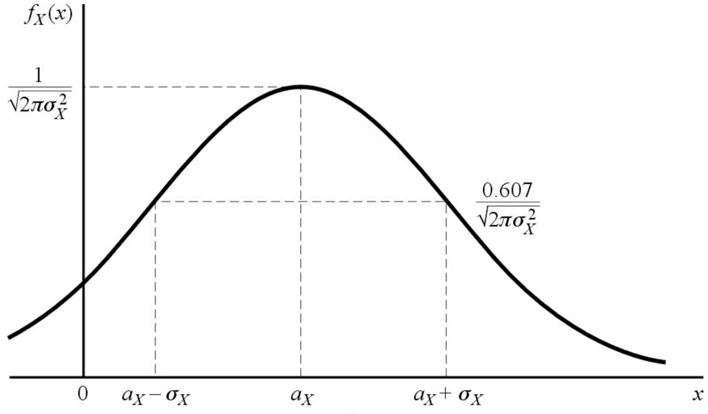
(b)

Ejemplos de V. A. continuas

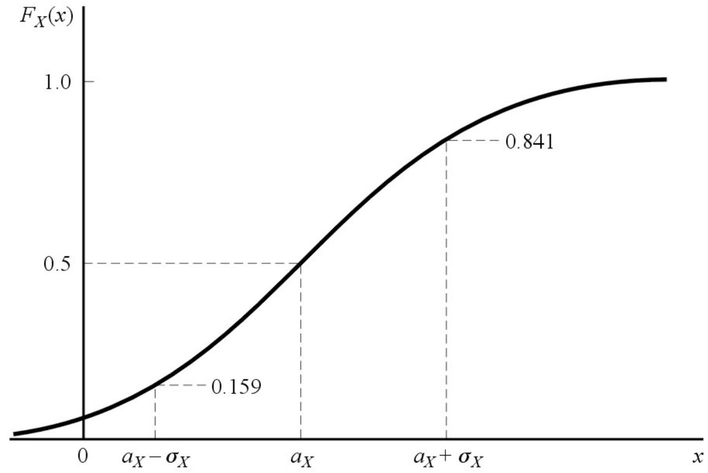
- V. A. Normal o Gaussiana

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma_X^2}} e^{-(x-a_X)^2/2\sigma_X^2}$$

$$F_X(x) = \frac{1}{\sqrt{2\pi\sigma_X^2}} \int_{-\infty}^{x} e^{-(\xi-a_X)^2/2\sigma_X^2}\, d\xi$$



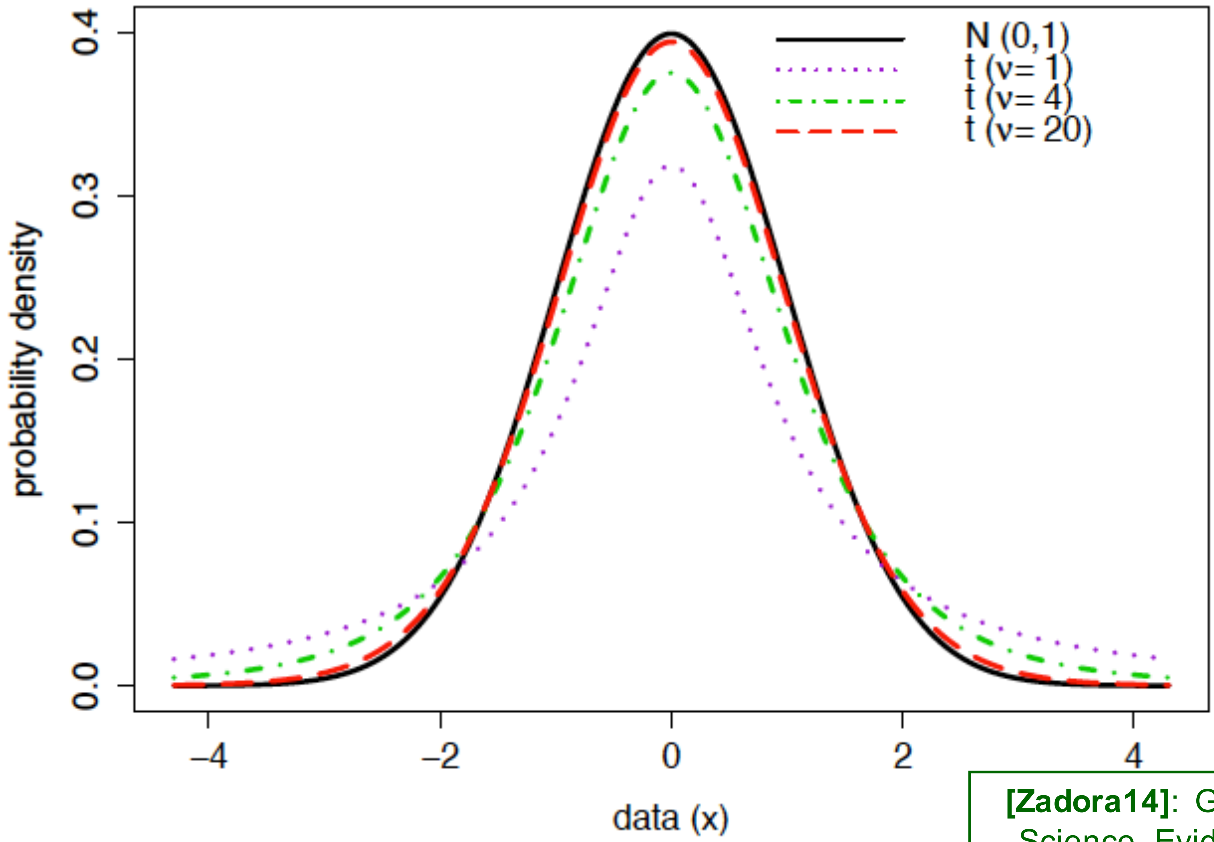
(a)



(b)

V. A. T de Student

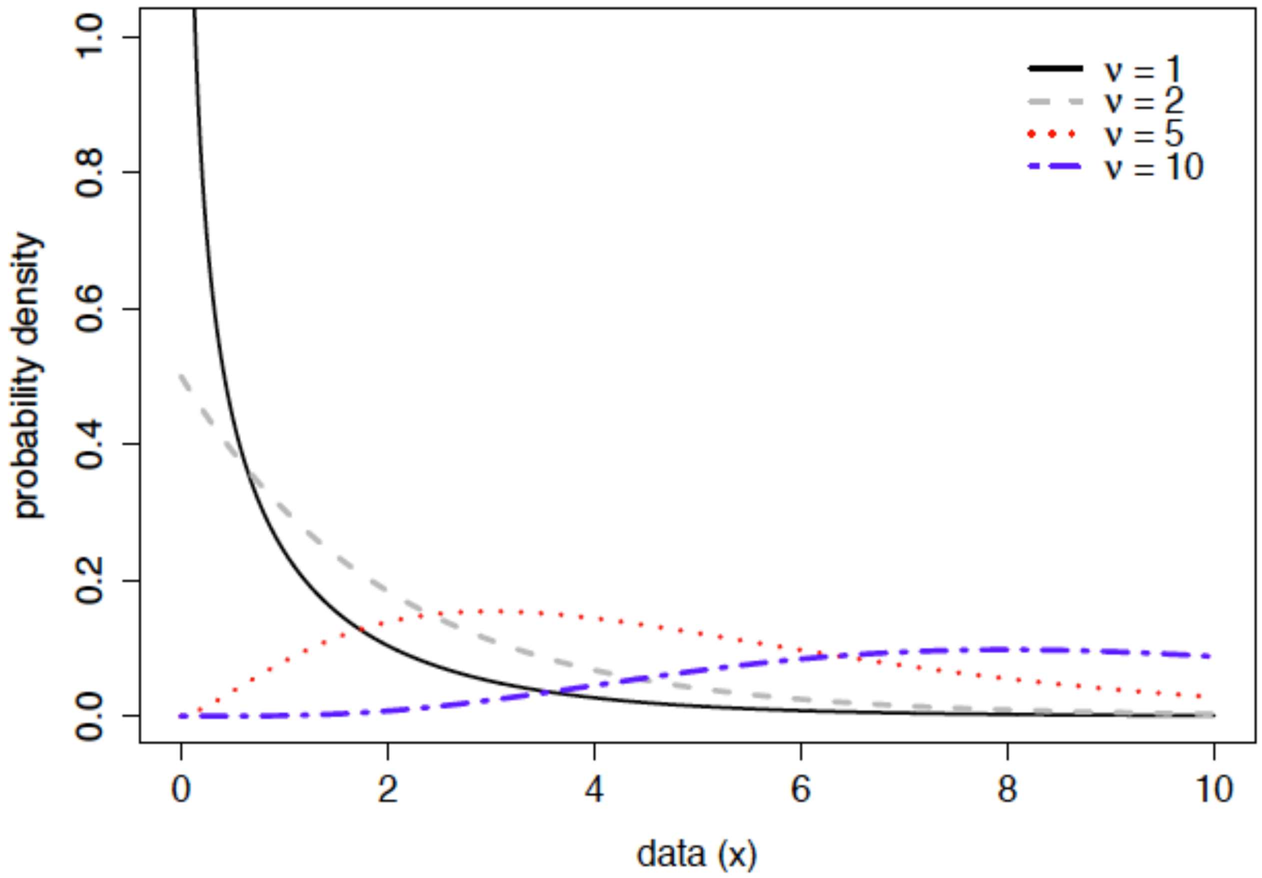
- Muy utilizada en contrastes de hipótesis
  - Depende de un único parámetro ν



[Zadora14]: G. Zadora et al. "Statistical Analysis in Forensic Science. Evidential Value of Multivariate Physico-Chemical Data". John Wiley and Sons, 2014

V. A. X^2 (Chi Cuadrado)

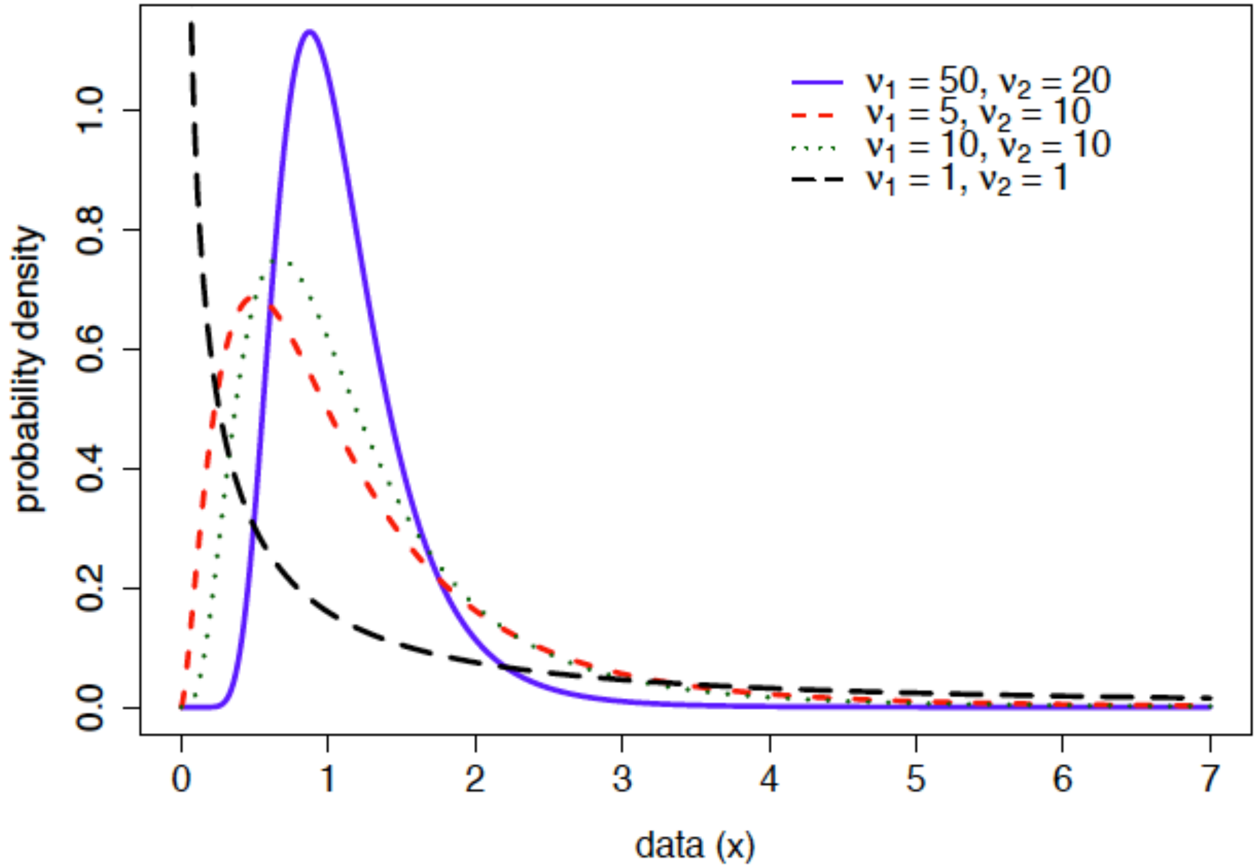
- Muy utilizada en comparación de variables nominales
  - Depende de un único parámetro ν



[Zadora14]

V. A. F de Snedecor

- Muy utilizada para comparar varianzas
  - Depende de dos parámetros: v_1 y v_2

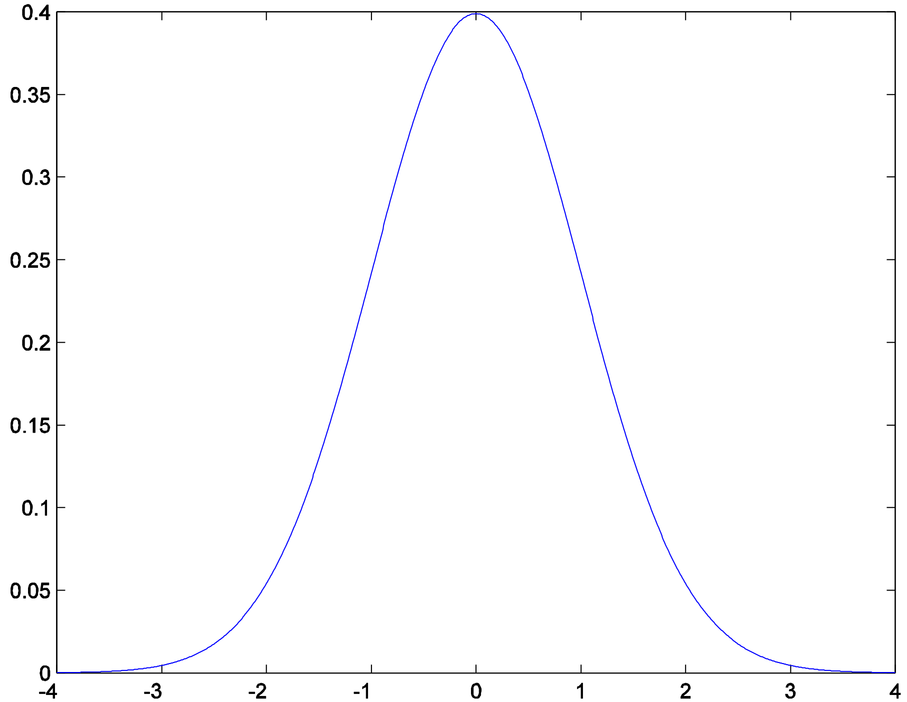


[Zadora14]

Variable Aleatoria Normal o Gaussiana

Variable Aleatoria Normal o Gaussiana

- F. d. p. con forma "acampanada".
  - A veces se le llama "campana de gauss".



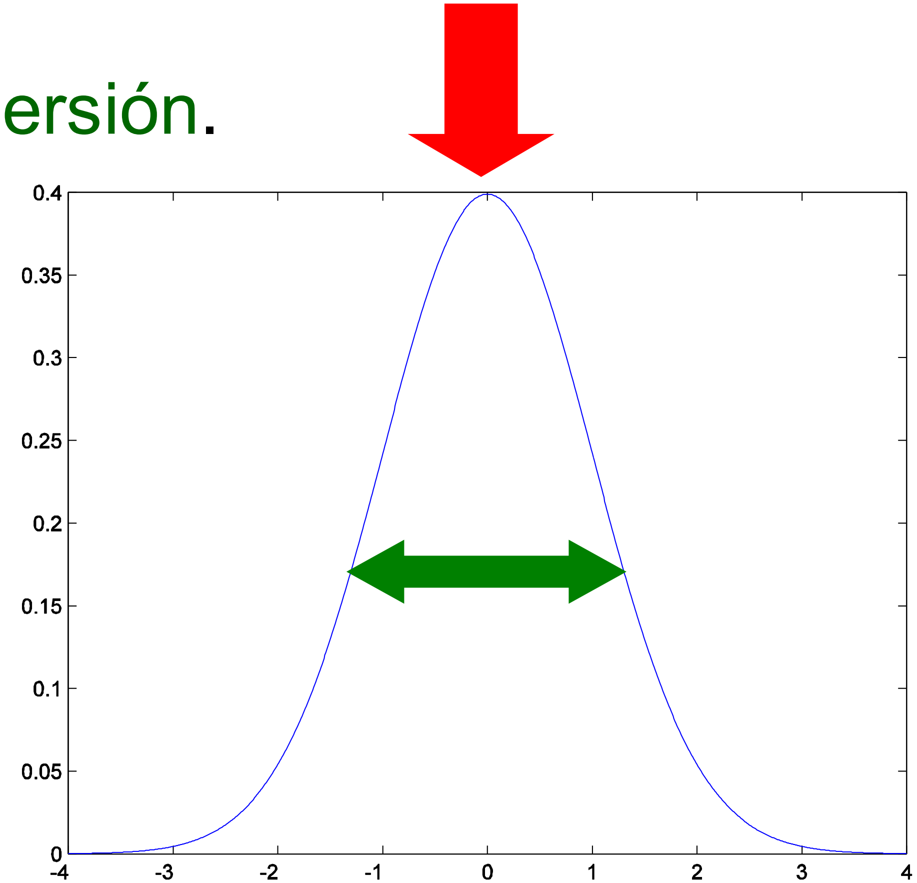
$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Media y desviación típica

- Definen completamente una gaussiana.
  - Media: localización.
  - Desviación típica: dispersión.

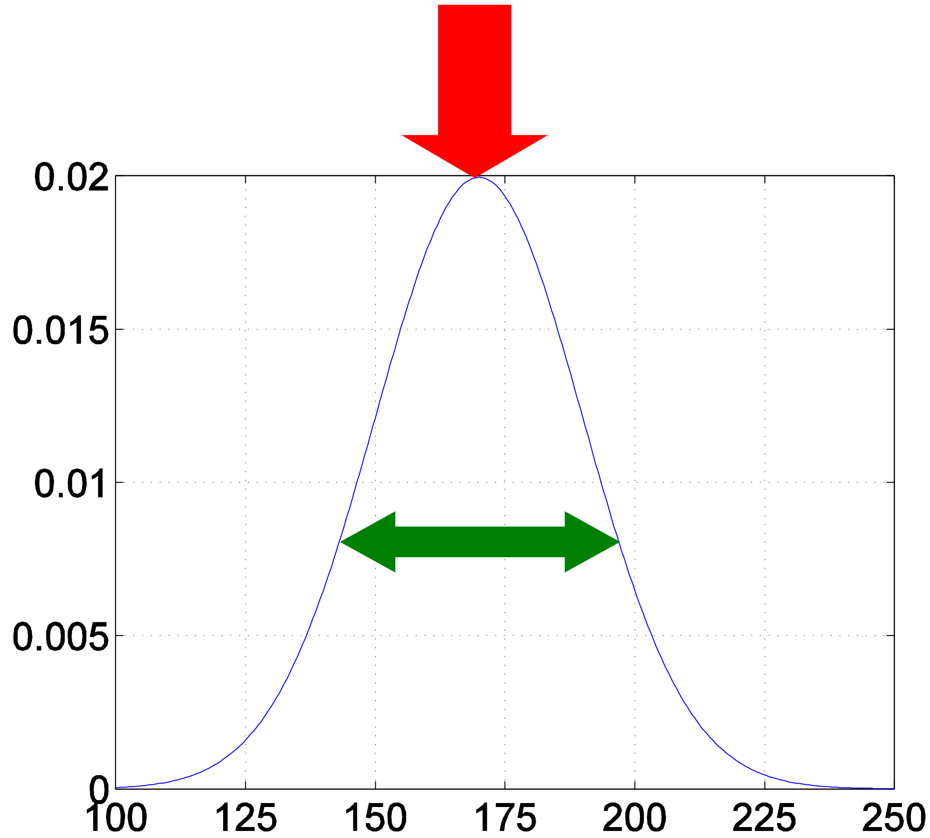
$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

$$= N(\mu, \sigma)$$



Ejemplo: altura

- Supongamos que la variable altura es gaussiana.
  - Media: 170 cm.
  - Desviación típica: 20 cm.



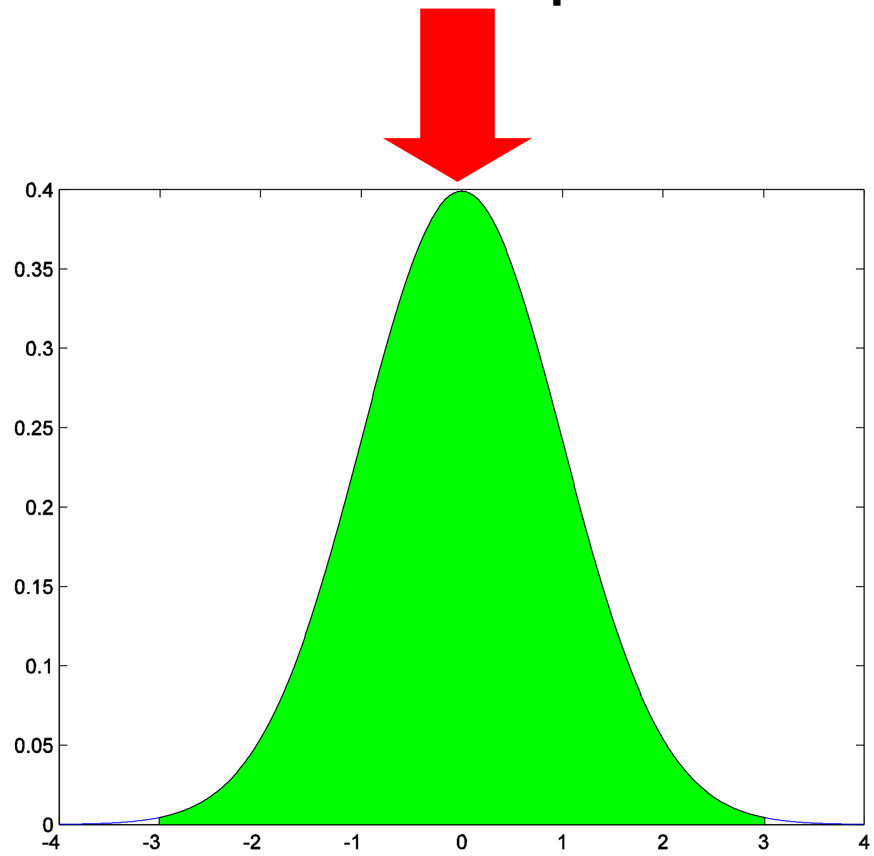
Media y desviación típica

- La media nos dice cuál es el valor esperado.
- La desviación típica nos habla de la dispersión.

$$P_X(-\sigma<x<\sigma)=0.682$$

$$P_X(-2\sigma<x<2\sigma)=0.950$$

$$P_X(-3\sigma<x<3\sigma)=0.997$$



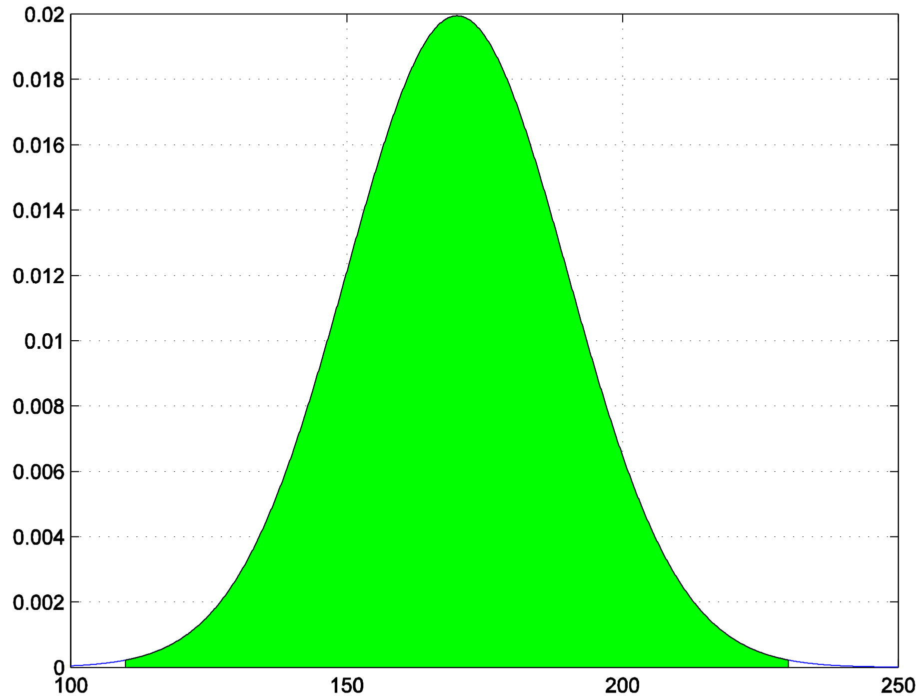
Ejemplo: altura

- "La probabilidad de que alguien de la población mida entre un valor y otro es…"
  - Media: 170 cm.
  - Desviación típica: 20 cm.

$$Entre\,150 \text{ y } 190 = 68{,}2\%$$

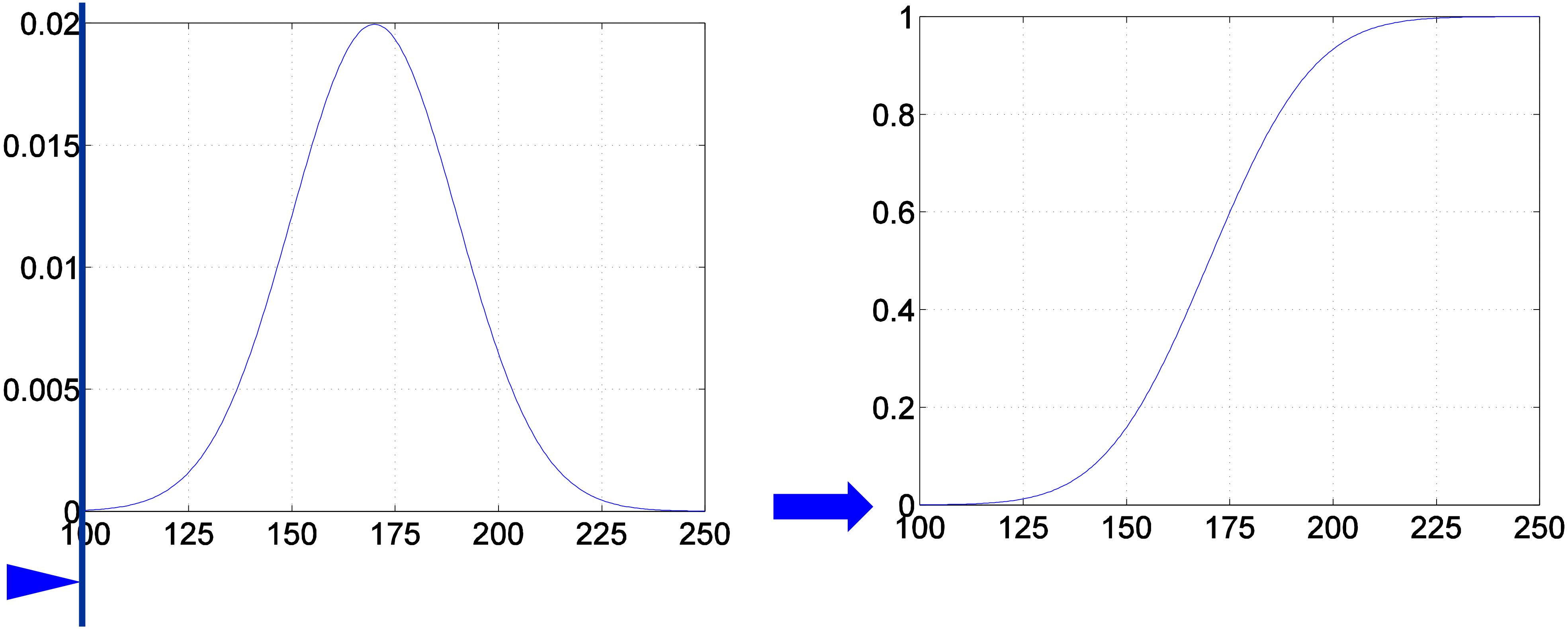
$$Entre\,130 \text{ y } 210 = 95{,}0\%$$

$$Entre\,110 \text{ y } 230 = 99{,}7\%$$



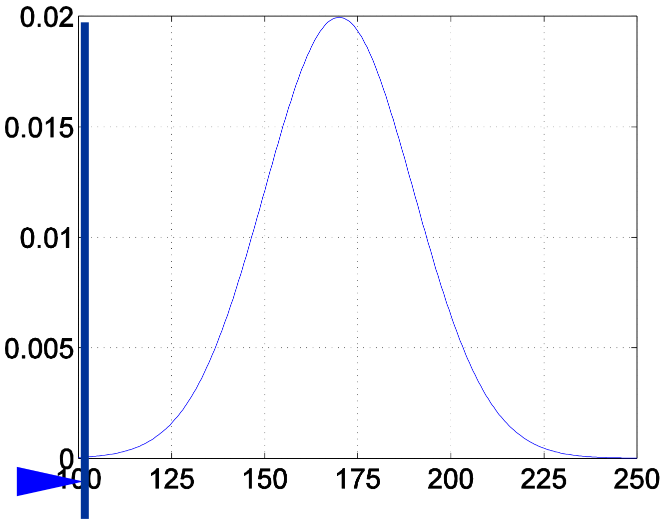
Función de distribución gaussiana

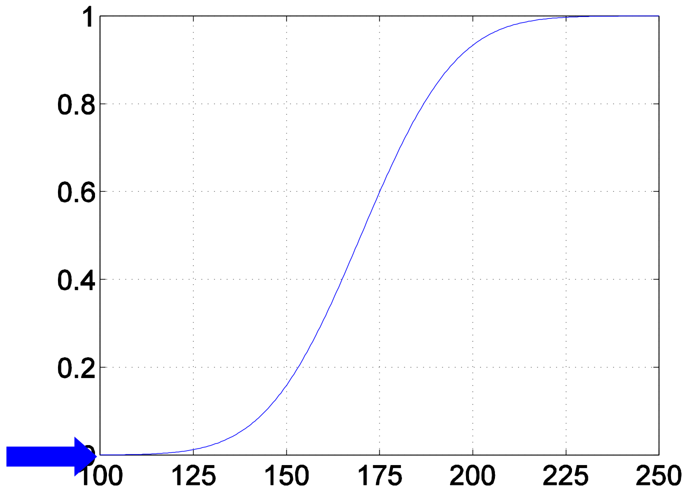
- Si vamos sumando el área de la f. d. p. obtenemos la función de distribución



Función de distribución gaussiana

- En matemáticas, el área que delimita una función es una integral





$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

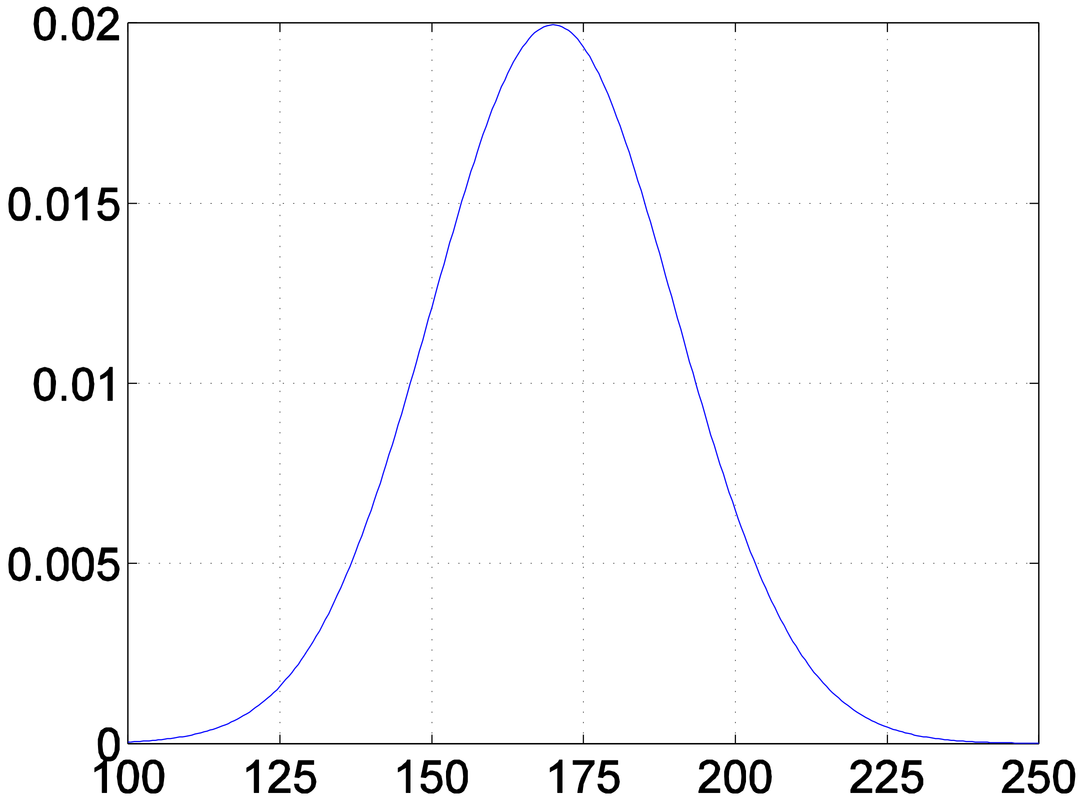
$$F_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{x} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt$$

¿Por qué la Gaussiana?

- Si es tan complicada de calcular y usar, ¿por qué es tan popular?
- Porque aparece mucho en la naturaleza.
  - ¡Está por todas partes!
- De hecho, hay un teorema (el del límite central) que dice:
  - Si sumamos muchas v. a. Independientes…
  - El resultado es una variable aleatoria gaussiana.

Ejemplo: altura

- Tiene cierto sentido que la fdp tenga esta forma.



Variables Aleatorias Multivariadas

Variables Aleatorias Multivariadas

- Valor univariado: se crea a partir de una V.A. univariada

$$X_{Na}^{(1)} \rightarrow x_{Na}^{(1)} = -0{,}645325$$

Número de Medida	**Log(Na/O) (V. A. X_{na})**
1	-0.645325

Variables Aleatorias Multivariadas

- Muestra univariada: generada por V.A.s
  - Independientes e idénticamente distribuidas

$$\left\{X_{Na}^{(1)},X_{Na}^{(2)},\ldots\right\}\rightarrow\left\{x_{Na}^{(1)},x_{Na}^{(2)},\ldots\right\}=\left\{-0,645325,-0,618750,\ldots\right\}$$

Número de Medida	**Log(Na/O) (V. A. X_{na})**
1	-0.645325
2	-0.618750
3	-0.613208
4	-0.615041
…	…

Variables Aleatorias Multivariadas

- Vector multivariado, a partir de V.A. multivariada

$$\mathbf{X}^{(1)} = \begin{bmatrix} X_{Na}^{(1)} \\ X_{Mg}^{(1)} \\ X_{Al}^{(1)} \end{bmatrix} \rightarrow \mathbf{x}^{(1)} = \begin{bmatrix} x_{Na}^{(1)} \\ x_{Mg}^{(1)} \\ x_{Al}^{(1)} \end{bmatrix} = \begin{bmatrix} -0,645325 \\ -1.461683 \\ -1.484642 \end{bmatrix}$$

Número de Medida	**Log(Na/O)**	**Log(Mg/O)**	**Log(Al/O)**
1	-0.645325	-1.461683	-1.484642

Variables Aleatorias Multivariadas

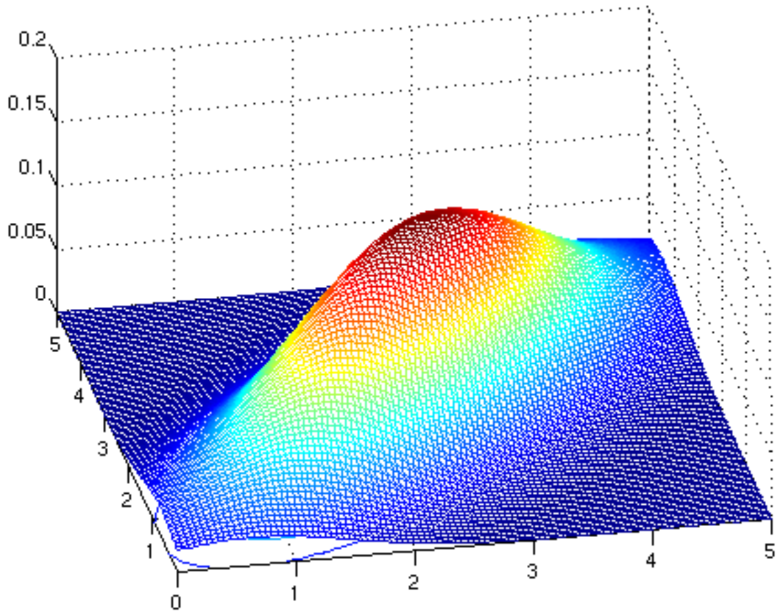
- Con variables multivariadas, muchos más números

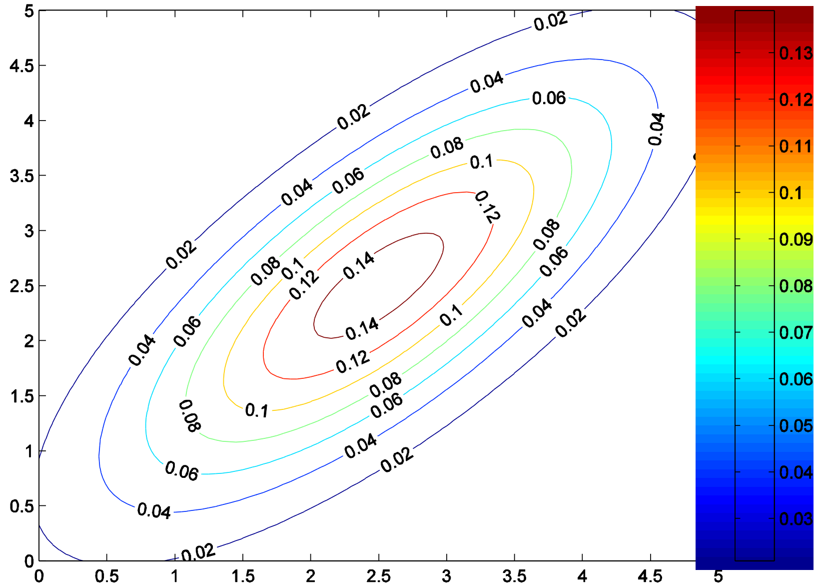
$$\left\{\mathbf{X}^{(1)},\mathbf{X}^{(2)},\ldots\right\}\rightarrow\left\{\mathbf{x}^{(1)},\mathbf{x}^{(2)},\ldots\right\}=\left\{\begin{bmatrix}-0,645325\\-1.461683\\-1.484642\end{bmatrix},\begin{bmatrix}-0,618750\\-1.377008\\-1.711383\end{bmatrix},\ldots\right\}$$

Número de Medida	**Log(Na/O)**	**Log(Mg/O)**	**Log(Al/O)**
1	-0.645325	-1.461683	-1.484642
2	-0.618750	-1.377008	-1.711383
3	-0.613208	-1.374241	-1.687958
4	-0.615041	-1.374233	-1.709916
…	…		

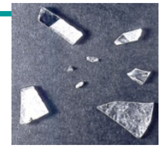
Modelado multivariable

- Para ello, hay que modelar distribuciones conjuntas
  - Si hay más de una variable, distribuciones multivariadas
- Ejemplo: gaussiana multivariada
  - Muestra la correlación entre las variables…





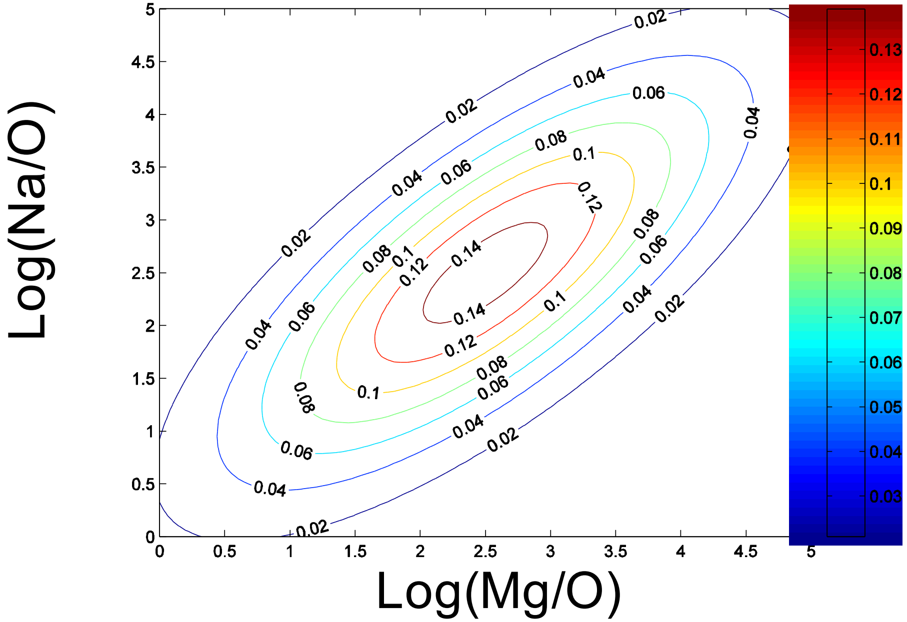
Ejemplo



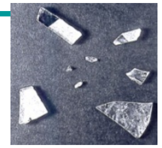
- Supongamos:
  - Fragmento de vidrio en la ropa de un sospechoso
    - Log(Mg/O): 2
    - Log(Na/O): 1,5

$$\mathbf{x}^{(1)} = \begin{bmatrix} 2 \\ 1,5 \end{bmatrix}$$

  - De una ventana en la escena del crimen se conoce la siguiente densidad:



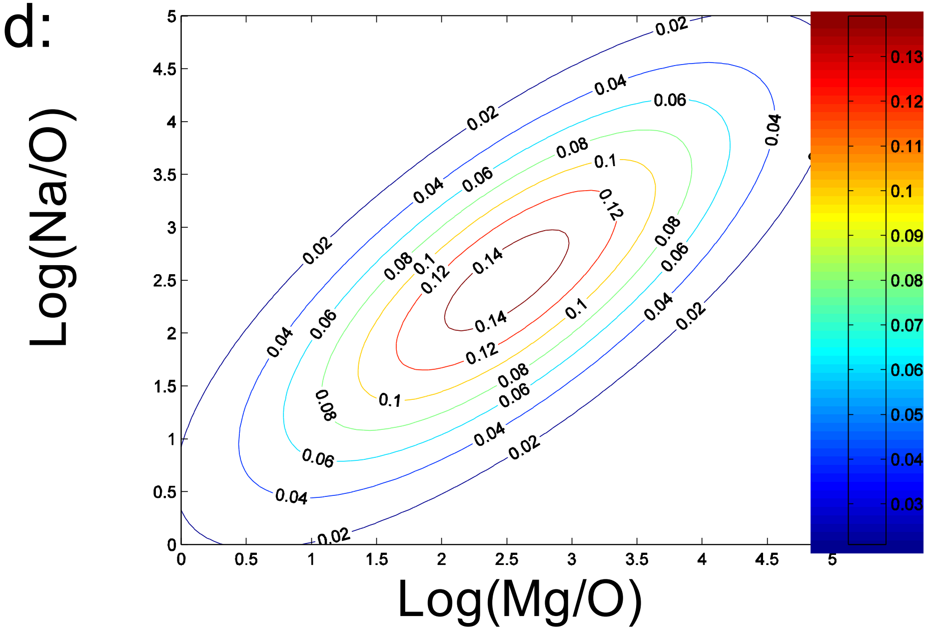
Ejemplo



- Supongamos:
  - Fragmento de vidrio en la ropa de un sospechoso
    - Log(Mg/O): 2
    - Log(Na/O): 1,5

$$\mathbf{x}^{(1)} = \begin{bmatrix} 2 \\ 1,5 \end{bmatrix}$$

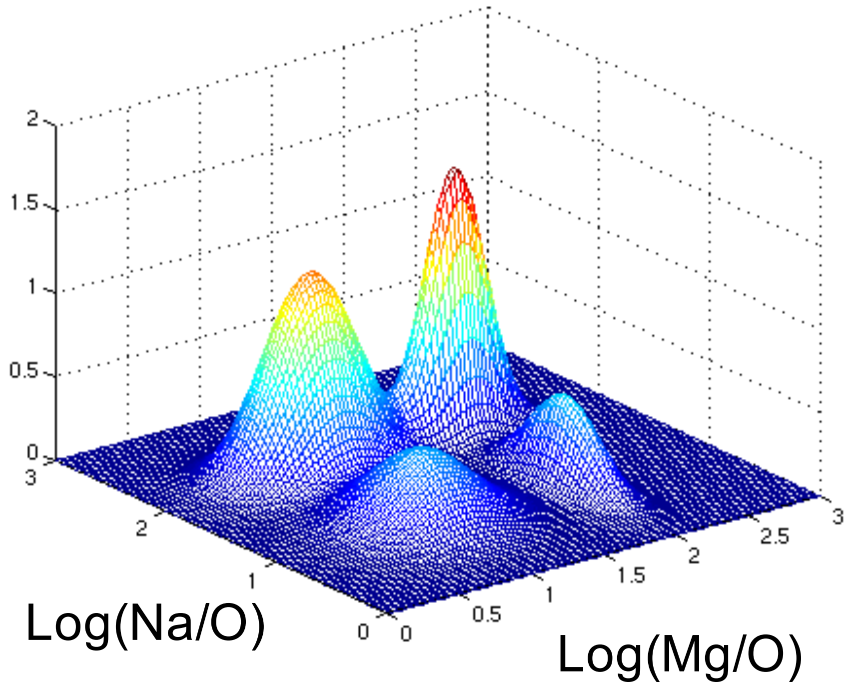
  - De una ventana en la escena del crimen se conoce la siguiente densidad:

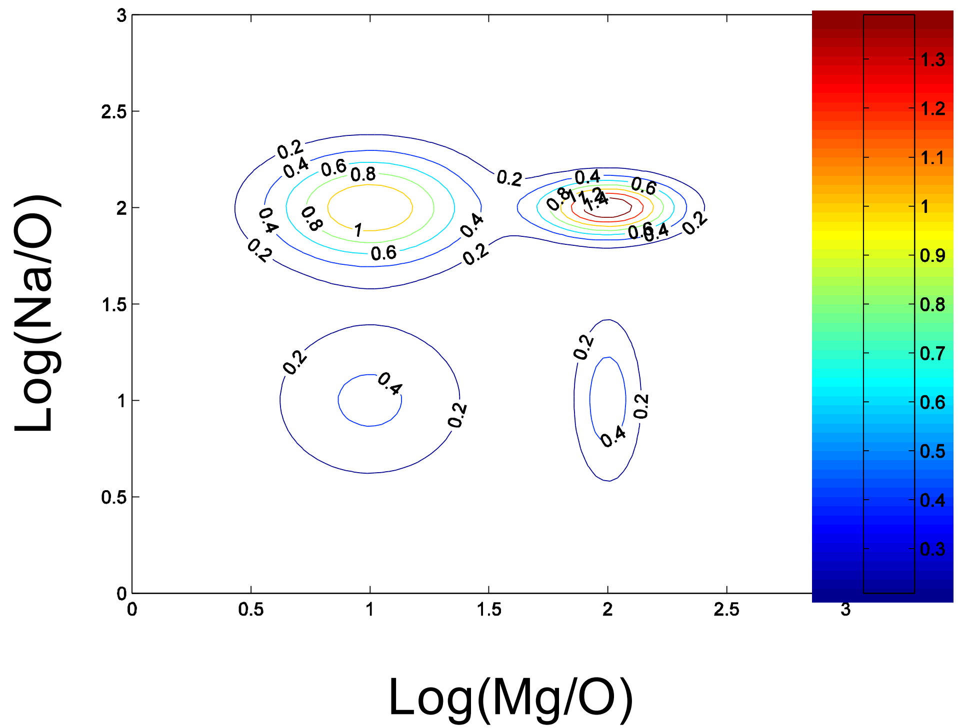


$$f_{\mathbf{X}}(\mathbf{x}) = 0,11$$

Ejemplo

- Imaginemos que la densidad de la ventana de control es en cambio así

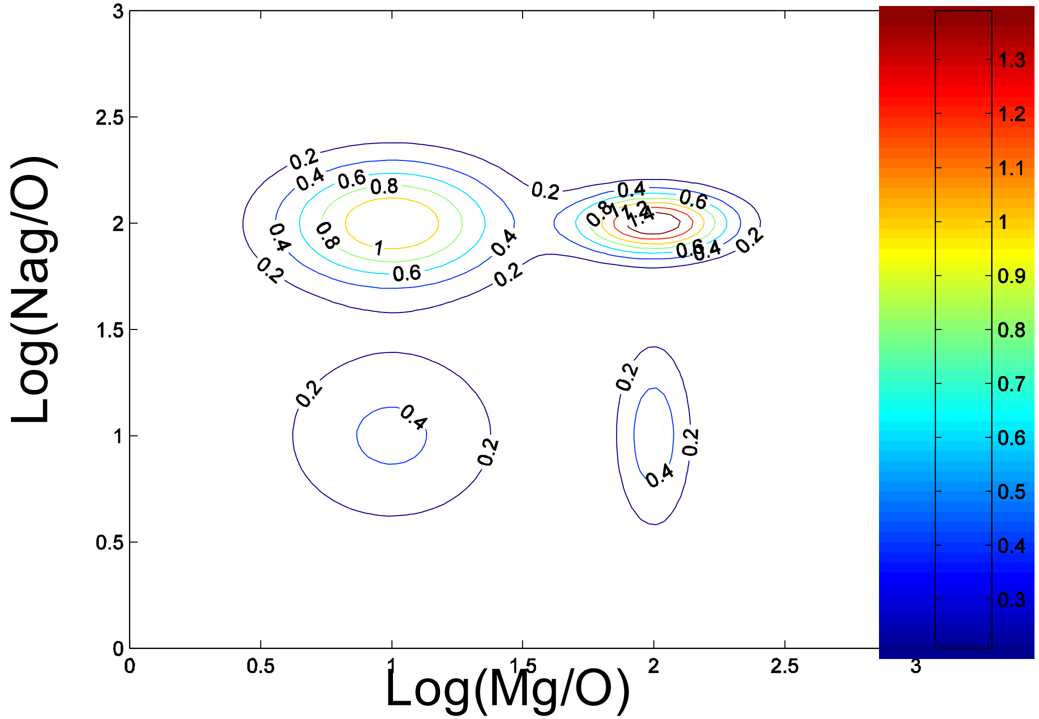




Ejemplo

- Vidrio en la ropa del sospechoso
  - Log(Mg/O): 2
  - Log(Na/O): 1,5

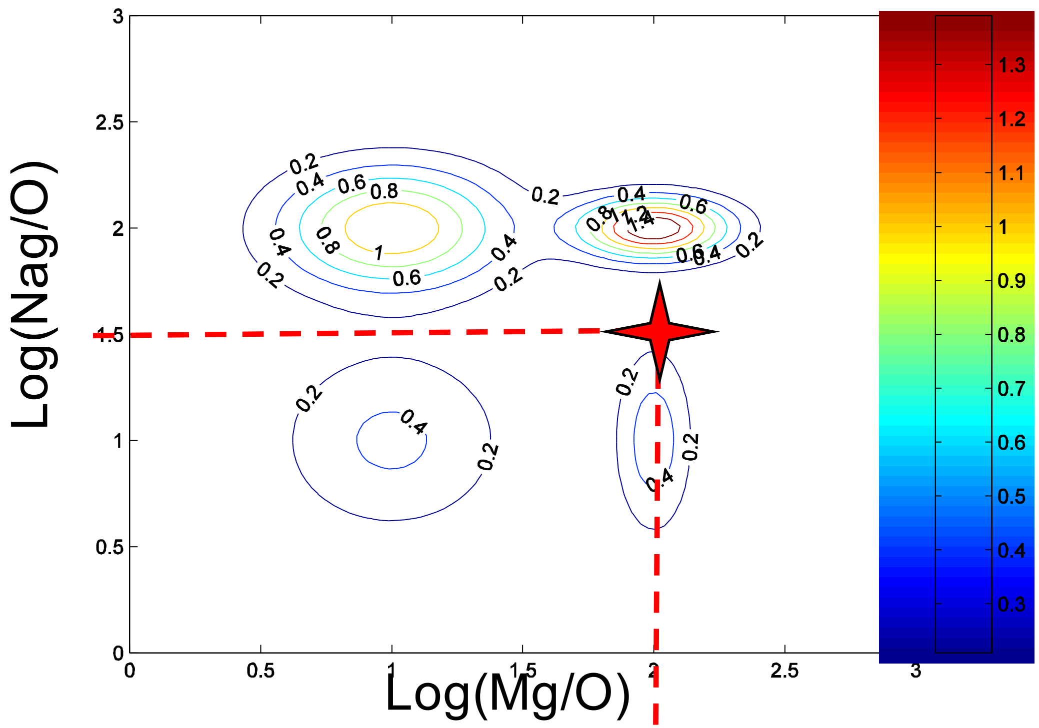
$$\mathbf{x}^{(1)} = \begin{bmatrix} 2 \\ 1,5 \end{bmatrix}$$



Ejemplo

- Vidrio en la ropa del sospechoso
  - Log(Mg/O): 2
  - Log(Na/O): 1,5

$$\mathbf{x}^{(1)} = \begin{bmatrix} 2 \\ 1,5 \end{bmatrix}$$



$$f_{\mathbf{X}}(\mathbf{x}) = 0,005$$

Referencias Adicionales

- D. Lindley (2006). "Understanding Uncertainty". John Wiley and Sons.
- C. Aitken, F. Taroni (2004). "Statistics and the Evaluation of the Evidence for Forensic Scientists". John Wiley and Sons.
- D. Lucy (2005). "Introduction to Statistics for Forensic Scientists". John Wiley and Sons.
- F. Taroni, S. Bozza, A. Biedermann, P. Garbolino, C. Aitken (2010). "Data Analysis in Forensic Science: A Bayesian Decision Perspective" John Wiley and Sons.
- G. Zadora, A. Martyna, D. Ramos, C. Aitken (2014). "Statistical Analysis in Forensic Science: Evidential Values of Multivariate Physicochemical Data". John Wiley and Sons.