Introducción al pensamiento bayesiano

Ezequiel Koile

1. Introducción

Hay dstintas formas de definir el término “bayesiano”, pero mayormente denota una interpretación particular de la probabilidad. Básicamente, la inferencia bayesiana es simplemente contar la cantidad de formas en que pueden pasar las cosas, según nuestras suposiciones. Las cosas que pueden pasar de más formas son más plausibles. Dado un conjunto de suposiciones, la inferencia bayesiana fuerza a un procedimiento lógico de procesamiento para producir una inferencia.

El enfoque frecuentista define todas sus probabilidades conectándolas con eventos contables y sus frecuencias en muestras muy grandes. Esto explica las incertezas frecuentistas como el resultado de un muestreo imaginario de los datos que se repitieran muchas veces. También, implica que los parámetros y los modelos no pueden tener distribuciones de probabilidad: solo los datos (mediciones) pueden tenerlos. Este remuestreo infinito no tiene sentido en muchos casos, y puede considerarse un artefacto del método.

En los procedimientos estadísticos más comunes, como regresiones lineales, esta diferencia entre conceptos de probabilidad no es tan relevante en el sentido de que los resultados finales serán casi iguales. Sin embargo, mientras el modelo frecuentista imagina una repetición de muestreo, el modelo bayesiano trata la aleatoriedad como una proppiedad de la información, no del mundo (ver ejemplo sobre la foto de Saturno tomada por Galileo en [1]).

La gran ventaja de los métodos bayesianos es la siguiente: Dados unos datos de entrada y un modelo, los métodos bayesianos encontrarán la solución óptima. Por supuesto, no puede garantizar la mejor solución en el mundo “real”, sino dentro del marco del modelo matemático en consideración. .

2. El jardín de los datos que se bifurcan a.k.a. Juguemos con pelotitas

Juguemos un juego

PPT

Conjecture	Ways to produce ●○●
[○○○○]	$0 \times 4 \times 0 = 0$
[●○○○]	$1 \times 3 \times 1 = 3$
[●●○○]	$2 \times 2 \times 2 = 8$
[●●●○]	$3 \times 1 \times 3 = 9$
[●●●●]	$4 \times 0 \times 4 = 0$

Figura 1: Actualización bayesiana. Fuente: [1].

3. Construcción de un modelo

Para construir un modelo, necesitamos tres pasos:

- Historia de los datos: motivar el modelo narrando cómo surgen los datos,
- Actualización: educar el modelo incorporándole datos.
- Evaluación: supervisar y modificar el modelo de ser necesario.

3.1. En nuestro ejemplo

- Historia de los datos:
  - hay un número entero n de piedras negras y b de piedras blancas. Ambos son positivos y cumplen $n + b = 4$.
  - n y b son constantes a través de las distintas extracciones
  - en cada extracción, la probabilidad de obtener un dado color de piedra es proporcional a la proporción de estas en la bolsa (no es más fácil ni más difícil sacar una piedra blanca)
  - n y b en una extracción no dependen de extracciones anteriores (¡válido también para la ruleta!)
- Actualización: usar los datos de observaciones previas para mejorar nuestro modelo
- Evaluación: el modelo bayesiano aprende de manera óptima, asumiendo que el mundo real está descripto de manera exacta por el modelo.

3.2. Otro ejemplo: golobo terráqueo - Probabilidad continua

- Historia de los datos:
  - La verdadera proporción de agua cubriendo el globo es p
  - En cada tirada del globo hay una probabilidad p de observar agua W y una probabilidad $1 - p$ de observar tierra L

- n y b cada tirada es independiente de las otras

- Actualización bayesiana: figura 2

- Evaluación

4. Componentes del modelo

Likelihood

La *likelihood* (lit. verosimilitud) especifica la plausibilidad de los datos. Nos da la cantidad de formas en que puede producirse una observación *dado un modelo* (conjunto de parámetros).

Parámetros

Cuantifica lo que queremos responder de los datos. A partir de ellos, construimos las conjeturas o configuraciones posible. En la estadística bayesiana, no es trivial la distinción entre parámetros y datos: estos son parte de un mismo tipo de entidad.

Prior

Para cada parámetro que se quiere estimar, hace falta una asignación de plausibilidad inicial para cada uno de sus valores. Estas asignaciones iniciales se llaman *priors*, que pueden ser más o menos informativos (regularizantes, relajados, estrictos, etc.).

Posterior

Para cada combinación de datos, *likelihood*, parámetros y *prior*, hay un único conjunto de estimadores. Estos tendrán una plausibilidad relativa para el valor de cada parámetro, condicional a los datos. Esto se llama distribución *posterior*. El *posterior* de un paso puede tomarse como el *prior* del siguiente, al agregar nuevos datos.

En nuestro ejemplo

- *Likelihood*: cantidad de formas de generar los datos (color observado) en una sacada, dada una configuración particular.

- Parámetros: Uno solo, el número de piedras negras n, (b queda fijado por $b = 4 - n$) . Este define la configuración de piedras dentro de la bolsa.

- *Prior*: Probabilidad *a priori* de cada configuración. Homogéneo en el primer caso, y agregando la información del precio de las piedras en el segundo.
- *Posterior*: Probabilidad de cada configuración, dados tanto el *prior* como la *likelihood*. En este caso, proporcional al conteo de configuraciones luego de sacar cada piedra.

Recordemos el teorema de Bayes:

$$\begin{aligned} P(A \mid B) &= P(B \mid A)\,\frac{P(A)}{P(B)} \\ &= \frac{P(B \mid A) \times P(A)}{P(B)}. \end{aligned}$$

En este caso,

$$P(M \mid D) = \frac{P(D \mid M) \times P(M)}{P(D)}.$$

o bien,

$$\textit{Posterior} = \frac{\textit{Likelihood} \times \textit{Prior}}{\textit{Average likelihood}}.$$

Donde

- *Likelihood*: $P(D \mid M)$
- Parámetros: M
- *Prior*: $P(M)$
- *Posterior*: $P(M \mid D)$

5. Implementación

Ver sección 2.4 de [1]

Referencias

[1] Richard McElreath, 2015. *Statistical Rethinking: A Bayesian course with examples in R and Stan*, Chapman and Hall CRC. Capítulo II.

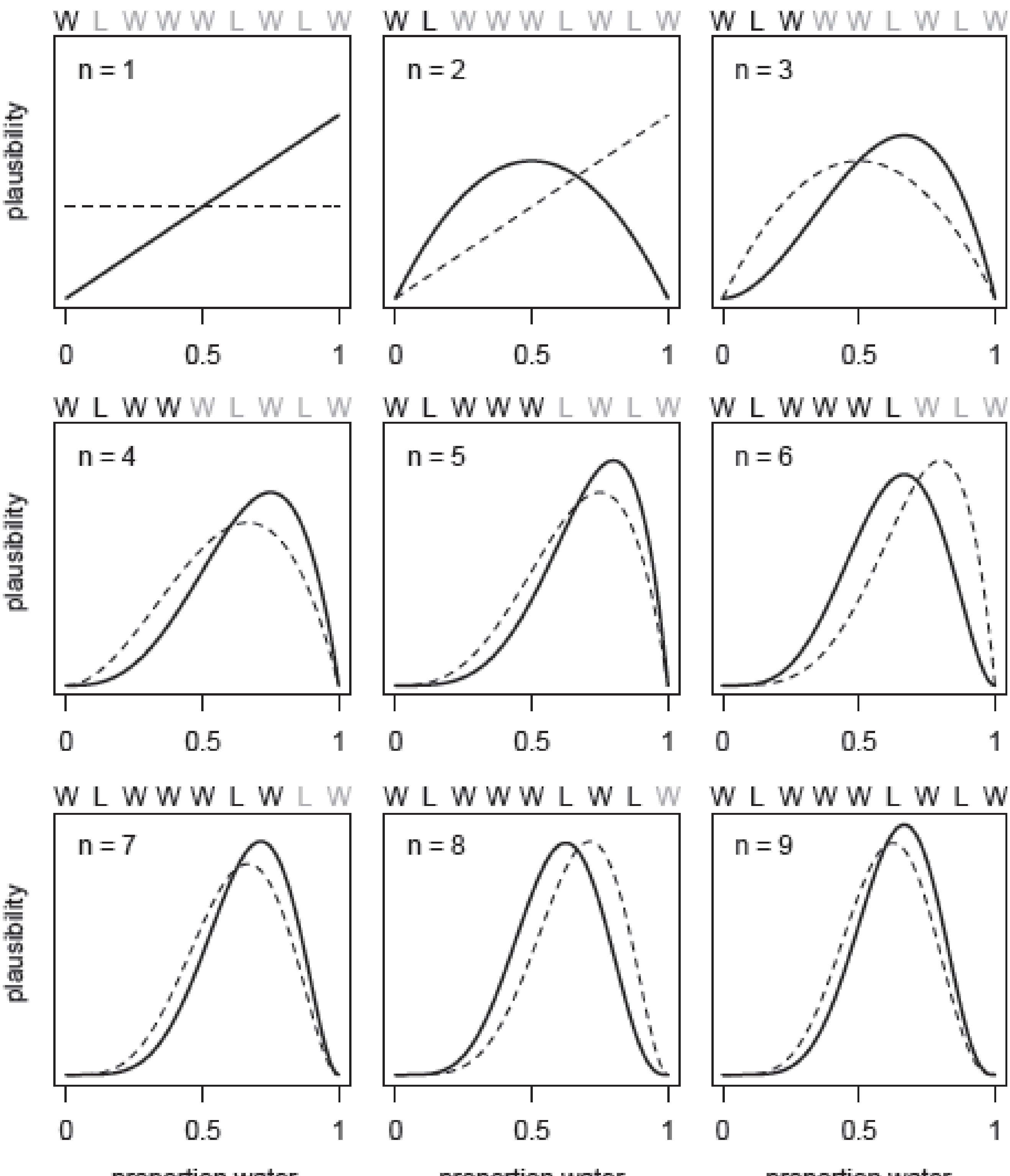

FIGURE 2.5. How a Bayesian model learns. Each toss of the globe produces an observation of water (W) or land (L). The model's estimate of the proportion of water on the globe is a plausibility for every possible value. The