Bayesian Inference

Mauricio A. Álvarez PhD,
H.F. Garcia C. Guarnizo (TA)

Universidad Tecnológica de Pereira, Pereira, Colombia

Outline

1. Bayes: conceptos básicos

Outline

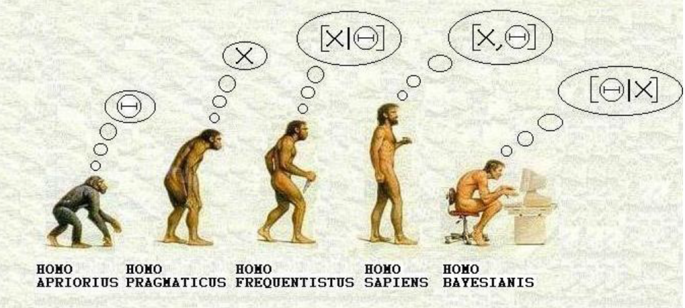
Bayes

El lenguaje de la incertidumbre.

- Teoría de probabilidad
- Específicamente *Bayesian probability theory* (1750!)

Cuando se aplica en ciencias de datos ...

- Bayesian modelling



- Construido sobre sólidos fundamentos matemáticos.
- Ortogonal al deep learning...

Que es el análisis estadístico Bayesiano?

- Un model Bayesiano se describe por parámetros. Su incertidumbre en los paráметros se describe utilizando distribuciones de probabilidad.
- Todos los aspectos derivados de procedimientos Bayesianos se realizan en términos de probabilidades.

> Practical methods for making inferences from data using probability models for quantities we observe and about which we wish to learn. – Gelman et al. 2013

$$p(\theta|y)$$

unknowns observations

Estadísticas bayesianas vs frecuentistas: ¿Cuál es la diferencia? I

Cualquier paradigma estadístico, bayesiano o de otro tipo, involucra al menos lo siguiente:

- Algunas cantidades desconocidas sobre las cuales estamos interesados en aprender o probar. Llamamos a estos parámetros.
- Algunos datos que se han observado, y con suerte contienen información sobre (1).
- Uno o más modelos que relacionan los datos con los parámetros, y es el instrumento que se utiliza para aprender.

The Frequentist World View

- Los datos que se han observado se consideran aleatorios, porque son realizaciones de procesos aleatorios y, por lo tanto, variarán cada vez que uno vaya a observar el sistema.
- Los parámetros del modelo se consideran fijos. Los valores de los parámetros son desconocidos, pero son fijos, por lo que los condicionamos.

Estadísticas bayesianas vs frecuentistas: ¿Cuál es la diferencia? III

The Bayesian World View

- Los datos se consideran **fijos**. Solían ser aleatorios, pero una vez se procesan no cambian.
- Los parámetros del modelo en sí pueden no ser aleatorios
- Se utilizan distribuciones de probabilidad para describir su incertidumbre en los valores de los parámetros (i.e. se consideran aleatorios).
- En algunos casos, es útil considerar los parámetros como muestras de distribuciones de probabilidad.

Probabilistic Modelling

- Un modelo describe datos que se podrían observar desde un sistema.
- Si usamos las matemáticas de la teoría de la probabilidad para expresar todas las formas de incertidumbre y ruido asociadas con nuestro modelo ...
- ... entonces la **probabilidad inversa** es decir, se *infiere de las observaciones a los parámetros*, o de los efectos a las causas.

Bayes' Formula

La fórmula de Bayes es el **único estimador** que los Bayesianos necesitan para obtener estimaciones de cantidades desconocidas que nos interesan.

$$\underbrace{\Pr(\theta|y)}_{\text{Posterior Probability}} = \frac{\overbrace{\Pr(y|\theta)}^{\text{Likelihood of Observations}}\,\overbrace{\Pr(\theta)}^{\text{Prior Probability}}}{\underbrace{\Pr(y)}_{\text{Normalizing Constant}}}$$

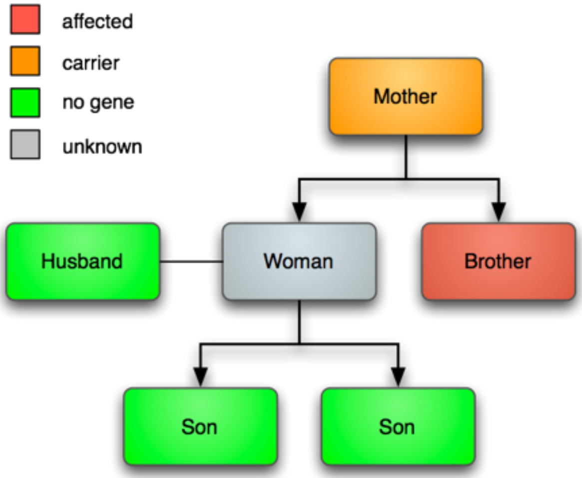
Ejemplo: Genetic probabilities I

- Pongamos la inferencia Bayesiana en acción usando un ejemplo muy simple. Mostraremos cómo se pueden utilizar los datos para actualizar nuestra creencia en hipótesis en competencia:
  - La hemofilia es un trastorno genético poco común que afecta la capacidad de los factores de coagulación del cuerpo para coagular la sangre en respuesta a vasos sanguíneos rotos.
  - La enfermedad es un rasgo recesivo ligado al gen X, lo que significa que solo hay una copia del gen en los hombres, pero dos en las mujeres, y el rasgo puede ser enmascarado por el alelo dominante del gen.
  - Esto implica que los hombres con 1 gen están afectados, mientras que las mujeres con 1 gen no se ven afectadas, pero sí son portadoras de la enfermedad.
  - Tener 2 copias de la enfermedad es fatal, por lo que este genotipo no existe en la población.

Ejemplo: Genetic probabilities II

- En este ejemplo, consideraremos a una mujer cuya madre es portadora (porque su hermano está afectado) y que se casa con un hombre no afectado.
- Observemos ahora algunos datos: la mujer tiene dos hijos consecutivos (no mellizos) que no se ven afectados. Estamos interesados en determinar **si la mujer es portadora**.

Ejemplo: Genetic probabilities III



Ejemplo: Genetic probabilities IV

Probability model

La cantidad desconocida de interés es simplemente una variable indicadora W que es igual a 1 si la mujer está afectada y cero si no lo está. Nos interesa la probabilidad de que la variable sea igual a uno, dado lo que hemos observado.

$$Pr(W = 1|s_1 = 0, s_2 = 0)$$

- Nuestra información previa (*prior*) se basa en lo que sabemos sobre la ascendencia de la mujer: su madre era portadora.
- Por lo tanto, el anterior es $Pr(W = 1) = 0.5$. Otra forma de expresar esto es en términos de probabilidades previas, o:

$$O(W = 1) = \frac{Pr(W = 1)}{Pr(W = 0)} = 1$$

Ejemplo: Genetic probabilities V

- Ahora para la *likelihood*: la forma de esta función es:

$$L(W|s_1 = 0, s_2 = 0)$$

- Podemos calcular la probabilidad de observar los datos para cualquier valor pasado para el parámetro.
- Aquí, la probabilidad toma solo dos valores posibles:

$$L(W = 1|s_1 = 0, s_2 = 0) = (0.5)(0.5) = 0.25$$
$$L(W = 0|s_1 = 0, s_2 = 0) = (1)(1) = 1$$

Ejemplo: Genetic probabilities VI

- Con todas las piezas en su lugar, ahora podemos aplicar la fórmula de Bayes para calcular la probabilidad posterior de que la mujer sea portadora:

$$
\begin{aligned}
Pr(W{=}1|s_1{=}0,s_2{=}0) &= \frac{L(W{=}1|s_1{=}0,s_2{=}0)Pr(W{=}1)}{L(W{=}1|s_1{=}0,s_2{=}0)Pr(W{=}1)+L(W{=}0|s_1{=}0,s_2{=}0)Pr(W{=}0)} \\
&= \frac{(0.25)(0.5)}{(0.25)(0.5)+(1)(0.5)} \\
&= 0.2
\end{aligned}
$$

- Por lo tanto, hay una probabilidad de 0.2 de que la mujer sea portadora.
- Ahora, ¿qué pasa si la mujer tiene un tercer hijo no afectado? ¿Cuál es nuestra estimación de su probabilidad de ser portadora entonces?

Ejemplo: Genetic probabilities VII

- Simplemente asignamos el análisis posterior del anterior al anterior prior para el nuevo análisis, y procedemos como antes:

$$= 0.1111111111111112$$

Por lo tanto, observar a un tercer niño no afectado ha reducido aún más nuestra creencia de que la madre es portadora

Bayesian Conjugate Priors

https://en.wikipedia.org/wiki/Conjugate_prior

LAB: Bayesian Inference