TWEAG

Bayesian data analysis with TensorFlow Probability

www.tweag.io

Your hosts

Simeon will give the presentation

background in computational biology

Data Scientist at Tweag since 2019

Dorran will happily answer questions

previous positions in geophysics

Data Scientist at Tweag since 2019

Tweag I/O

TODO

Tweag I/O is a software innovation lab and consultancy based in Paris with employees all around the world.

We specialize in

- software engineering, with a focus on functional programming
- DevOps, with a focus on reproducible software systems and builds
- data science

What you're in for

This tutorial consists of alternating blocks of

- theory / example slides
- practical examples on either external websites or Google Colab notebooks. Links are provided at `https://github.com/tweag/tutorial-dsc-2020/`

Requirements:

- a Google account (for the practical exercises)
- elementary knowledge in probability theory and statistics

Bayesian vs frequentist probabilities

Example: fair coin flip with $p(\text{head}) = p(\text{tail}) = \frac{1}{2}$

Frequentist probability

$p(\text{"flip results in head"} | b = \frac{1}{2}) = \frac{1}{2}$:

$\frac{\text{\# of heads}}{\text{\# total flips}}$ for ∞ many fair coin flips

Bayesian probability

$p(\text{"flip results in head"} | b = \frac{1}{2}) = \frac{1}{2}$:

measure of *belief* in the statement "flip results in head" given single fair coin flip

TODO: this slide looks complicated and is aesthetically offputting

Prior beliefs

Assume: unknown bias b

Prior probability

Encodes prior belief in b *before* flipping the coin

What is known about b?

b is a probability: $0 \leq b \leq 1$

most coins are fair

$\rightarrow$ choose prior distribution defined between 0 and 1, with maximum at and symmetric around $b = \frac{1}{2}$

Example:

$$b \sim \text{Beta}(\alpha = 2, \beta = 2)$$

with

$$\text{Beta}(x; \alpha, \beta) \propto x^{\alpha-1}(1-x)^{\beta-1}$$

Posterior belief

Now: flip coin one time, result: head

Posterior belief

$p(b|\text{head})$: updated prior belief after obtaining new data

TODO: some illustration of shifting distribution

Update rule

Bayes' theorem

$$p(A|B) = \frac{p(B|A) \times p(A)}{p(B)}$$

(easily derived from rules for conditional probabilities)

In data analysis:

$$\underbrace{p(x|D,I)}_{\text{posterior}} = \underbrace{p(D|x,I)}_{\text{likelihood}} \times \underbrace{p(x|I)}_{\text{prior}} \; / \; \underbrace{p(D|I)}_{\text{evidence}}$$

x: model parameter

D: data

I: prior information (often not made explicit)

Likelihood

$p(D|x)$: probability of the data given fixed model parameters
$\rightarrow$ models data-generating process

In our case:

$$p(k|b) = b^k(1-b)^{k-1} = \mathrm{Ber}(k; b)$$

with

$$k = \begin{cases} 0: & \text{tail} \\ 1: & \text{head} \end{cases}$$

TODO: graph of likelihood function

Evidence

$p(D) = \int \mathrm{d}x\, p(D|x)p(x)$:

normalization constant (long story...)

In our case:

$$
\begin{aligned}
p(k=1) &= \int_0^1 \mathrm{d}b\, L(k=1|b)p(b) \\
&= \int_0^1 \mathrm{d}b\ \ \mathrm{Ber}(k;b) \times \mathrm{Beta}(k;\alpha=2,\beta=2)|_{k=1} \\
&= \int_0^1 \mathrm{d}b\ \ b^k(1-b)^{k-1}b(1-b)\Big|_{k=1} \\
&\ \ \vdots \\
&= \frac{1}{12}
\end{aligned}
$$

Evidence

$p(D) = \int \mathrm{d}x \, p(D|x)p(x)$:

normalization constant (long story...)

In our case:

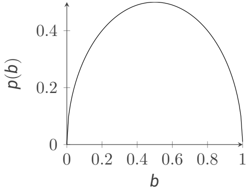
$$
\begin{aligned}
p(k=1) &= \int_0^1 \mathrm{d}b \, L(k=1|b)p(b) \\
&= \int_0^1 \mathrm{d}b \;\; \mathrm{Ber}(k;b) \times \mathrm{Beta}(k; \alpha=2, \beta=2)\big|_{k=1} \\
&= \int_0^1 \mathrm{d}b \;\; b^k(1-b)^{1-k} b(1-b)\Big|_{k=1} \\
&\;\vdots \\
&= \frac{1}{12}
\end{aligned}
$$

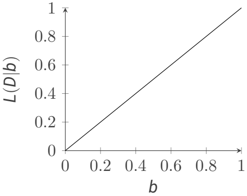
YIKES

Update rule

In our coin flip example:

prior
$$
\begin{aligned}
p(b) &= \mathrm{Beta}(b; \alpha = 2, \beta = 2) \\
&\propto b^{\frac{1}{2}}(1-b)^{\frac{1}{2}}
\end{aligned}
$$





likelihood
$$
\begin{aligned}
p(D|b) &= \mathrm{Bern}(k = 1; b) \\
&= b
\end{aligned}
$$