Workshop Bayesian Corpus Studies

Christoph Finkensiep

Würzburg, Feb 2024

Session 2: Bayesian Inference

Benachrichtigung zu Ihrer Paketzustellung **Nr. 34632900-371?**

DPD

Ihr Paket konnte nicht zugestellt werden, da bei der Zustellung keine Person zur Unterschrift anwesend war.

Wir möchten Ihnen mitteilen, dass wir eine Adressbestätigung benötigen, um den Paketversand erneut zu bestätigen.

HIER ÜBERPRÜFEN

[Abbestellen..27]


Benachrichtigung zu Ihrer Paketzustellung Nr. 34632900-371?
DPD
PAKETVERFOLGUNGSNUMMER
58412233520000
VERFOLGEN
dpd
Ihr Paket konnte nicht zugestellt werden, da bei der Zustellung keine Person zur Unterschrift anwesend war.
Wir möchten Ihnen mitteilen, dass wir eine Adressbestätigung benötigen, um den Paketversand erneut zu bestätigen.
HIER ÜBERPRÜFEN
Abbestellen..27

Spam or not?

The Old Question

Generating emails:

The Old Question

Generating emails:

1. decide intention (spam or not)

The Old Question

Generating emails:

1. decide intention (spam or not)
2. write email
   - if spam:
     - generic information (made-up)
     - makes reader click on a link
   - if not spam:
     - specific information for reader
     - legitimate request

The Old Question

Generating emails:

1. decide intention (spam or not)
2. write email
   - if spam:
     - generic information (made-up)
     - makes reader click on a link
   - if not spam:
     - specific information for reader
     - legitimate request
3. send!

The Old Question

Generating emails:

1. decide intention (spam or not)
2. write email
   - if spam:
     - generic information (made-up)
     - makes reader click on a link
   - if not spam:
     - specific information for reader
     - legitimate request
3. send!

Operations on Distributions

joint distribution:

$$p(x, y, z)$$

Operations on Distributions

joint distribution:

$$p(x, y, z)$$

from joint to **marginal** distribution:

$$p(x, y) = \sum\!\!\!\!\!\!\!\int_z p(x, y, z) \qquad\qquad p(y) = \sum\!\!\!\!\!\!\!\int_{x,z} p(x, y, z)$$

Operations on Distributions

joint distribution:

$$p(x, y, z)$$

from joint to **marginal** distribution:

$$p(x, y) = \sum\!\!\!\!\!\!\!\int_z p(x, y, z) \qquad p(y) = \sum\!\!\!\!\!\!\!\int_{x,z} p(x, y, z)$$

from joint to **conditional** distribution (“chain rule”):

$$p(x, z \mid y) = \frac{p(x, y, z)}{p(y)} \qquad p(y \mid x, z) = \frac{p(x, y, z)}{p(x, z)}$$

The Rule of Bayes (Bayes' Theorem)

x: observed variables

z: latent variables

$$p(z \mid x) = \frac{p(x, z)}{p(x)} = \frac{p(x \mid z) \cdot p(z)}{p(x)}$$

The Rule of Bayes (Bayes' Theorem)

x: observed variables

z: latent variables

posterior

$$p(z \mid x) = \frac{p(x, z)}{p(x)} = \frac{p(x \mid z) \cdot p(z)}{p(x)}$$

The Rule of Bayes (Bayes' Theorem)

x: observed variables

z: latent variables

prior

posterior

$$p(z \mid x) = \frac{p(x, z)}{p(x)} = \frac{p(x \mid z) \cdot p(z)}{p(x)}$$

The Rule of Bayes (Bayes' Theorem)

x: observed variables

z: latent variables

likelihood

prior

posterior

$$p(z \mid x) = \frac{p(x, z)}{p(x)} = \frac{p(x \mid z) \cdot p(z)}{p(x)}$$

The Rule of Bayes (Bayes' Theorem)

x: observed variables

z: latent variables

joint likelihood

prior

posterior

$$p(z \mid x) = \frac{p(x, z)}{p(x)} = \frac{p(x \mid z) \cdot p(z)}{p(x)}$$

The Rule of Bayes (Bayes' Theorem)

x: observed variables

z: latent variables

joint likelihood

posterior prior

$$p(z \mid x) = \frac{p(x, z)}{p(x)} = \frac{p(x \mid z) \cdot p(z)}{p(x)}$$

evidence

Finding Spam

- 10% of my email is spam
- 90% of all spam emails contain the word "urgent"
- 20% of all non-spam emails contain the word "urgent"

Finding Spam

- 10% of my email is spam
- 90% of all spam emails contain the word “urgent”
- 20% of all non-spam emails contain the word “urgent”

I get an email containing the word “urgent”. Is it spam or not?

Finding Spam

- 10% of my email is spam
- 90% of all spam emails contain the word "urgent"
- 20% of all non-spam emails contain the word "urgent"

I get an email containing the word "urgent". Is it spam or not?

	$\boldsymbol{s}$	$\bar{\boldsymbol{s}}$
$\boldsymbol{u}$	9%	18%
$\bar{\boldsymbol{u}}$	1%	72%

Finding Spam

- 10% of my email is spam
- 90% of all spam emails contain the word "urgent"
- 20% of all non-spam emails contain the word "urgent"

I get an email containing the word "urgent". Is it spam or not?

	$\boldsymbol{s}$	$\bar{\boldsymbol{s}}$
$\boldsymbol{u}$	9%	18%
$\bar{\boldsymbol{u}}$	1%	72%

$$p(s|u) = \frac{p(u,s)}{p(u)} = \frac{p(u,s)}{p(u,s) + p(u,\bar{s})} = \frac{0.09}{0.09 + 0.18} = \frac{0.09}{0.27} = \frac{1}{3}$$

Finding the Posterior Distribution

x: observed variables

z: latent variables

$$p(z \mid x) = \frac{p(x, z)}{p(x)} = \frac{p(x \mid z) \cdot p(z)}{p(x)}$$

Finding the Posterior Distribution

x: observed variables

z: latent variables

D: data

$$p(z \mid x = D) = \frac{p(x = D, z)}{p(x = D)} = \frac{p(x = D \mid z) \cdot p(z)}{p(x = D)}$$

Finding the Posterior Distribution

x: observed variables

z: latent variables

D: data

join likelihood prior

posterior

$$p(z \mid x = D) = \frac{p(x = D, z)}{p(x = D)} = \frac{p(x = D \mid z) \cdot p(z)}{p(x = D)}$$

evidence

constant

distribution

function

Finding the Posterior Distribution

x: observed variables

z: latent variables

D: data

$$\underbrace{p(z \mid x = D)}_{\substack{\text{posterior} \\ Z \to \mathbb{P}}} = \frac{\overbrace{p(x = D, z)}^{\substack{Z \to \mathbb{R} \\ \text{joint}}}}{\underbrace{p(x = D)}_{\substack{\text{evidence} \\ \mathbb{P}}}} = \frac{\overbrace{p(x = D \mid z)}^{\substack{Z \to \mathbb{R} \\ \text{likelihood}}} \cdot \overbrace{p(z)}^{\substack{Z \to \mathbb{P} \\ \text{prior}}}}{\underbrace{p(x = D)}_{\text{evidence}}}$$

constant

distribution

function

The Problem

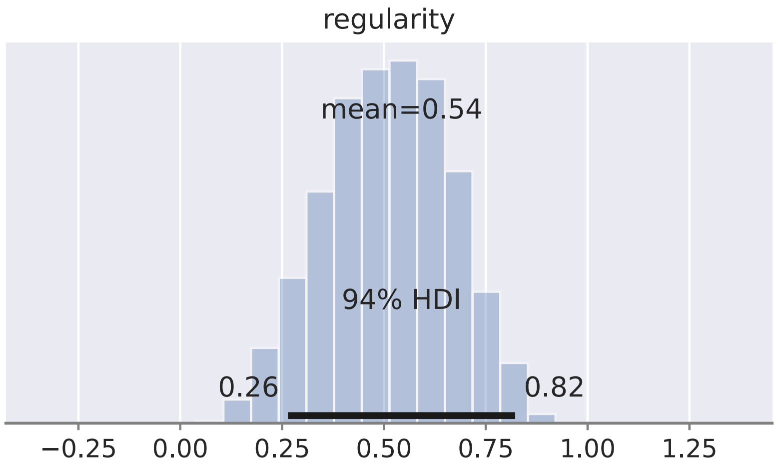
Computing $p(x = D)$ is hard (often impossible)!

The Problem

Computing $p(x = D)$ is hard (often impossible)!

Solutions:

sample from $p(z \mid x)$



approximate $p(z \mid x)$