

# **Workshop Bayesian Corpus Studies**

Christoph Finkensiep

Würzburg, Feb 2024

## **Session 2: Bayesian Inference**

 Benachrichtigung zu Ihrer Paketzustellung Nr. 34632900-371?



PAKETVERFOLGUNGSNUMMER:

58412233520000

VERFOLGEN



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?



PAKETVERFOLGUNGSNUMMER:

58412233520000

VERFOLGEN



# Spam or not?



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](#)

# 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

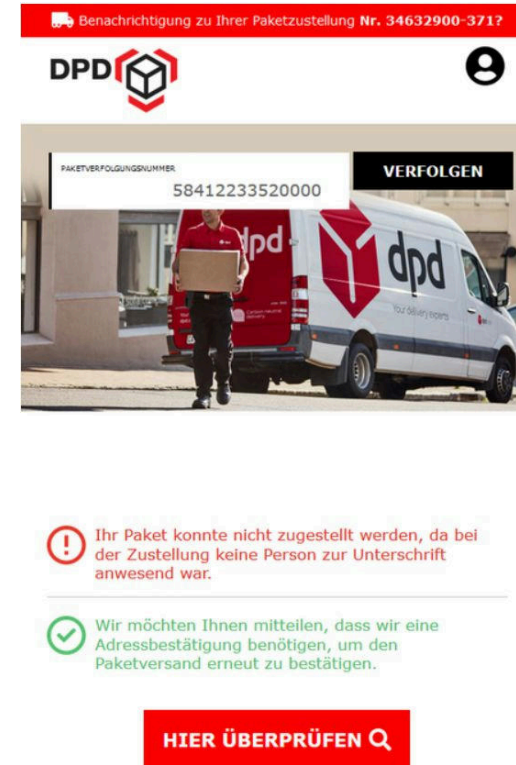
Generating emails:

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!



[Abbestellen...27](#)



# Operations on Distributions

**joint** distribution:

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

# Operations on Distributions

**joint** distribution:

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

from joint to **marginal** distribution:

$$p(x, y) = \int_z p(x, y, z)$$

$$p(y) = \int_{x,z} p(x, y, z)$$

# Operations on Distributions

**joint** distribution:

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

from joint to **marginal** distribution:

$$p(x, y) = \int_z p(x, y, z)$$

$$p(y) = \int_{x,z} p(x, y, z)$$

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

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

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

# The Rule of Bayes (Bayes' Theorem)

$x$ : observed variables

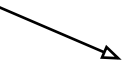
$z$ : latent variables

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

# The Rule of Bayes (Bayes' Theorem)

$x$ : observed variables

$z$ : latent variables

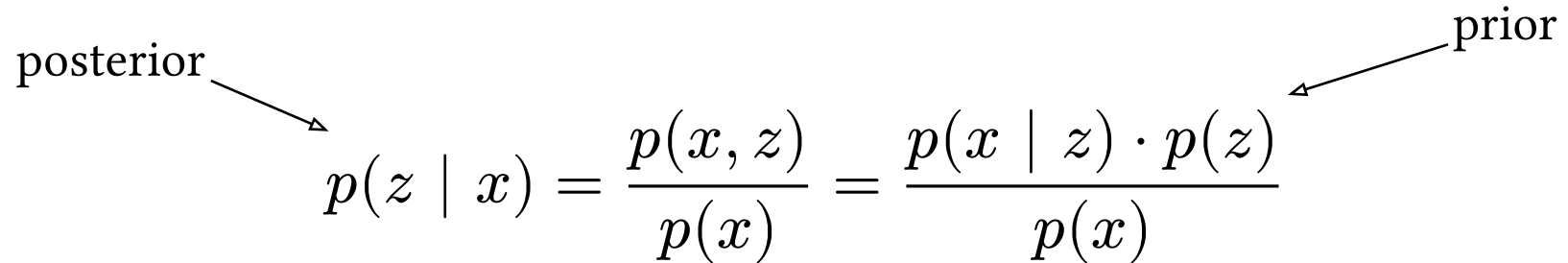
posterior 

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

# The Rule of Bayes (Bayes' Theorem)

$x$ : observed variables

$z$ : latent variables



The diagram shows the equation for Bayes' Theorem:  $p(z | x) = \frac{p(x, z)}{p(x)} = \frac{p(x | z) \cdot p(z)}{p(x)}$ . An arrow labeled 'posterior' points to the term  $p(z | x)$  on the left. Another arrow labeled 'prior' points to the term  $p(z)$  in the numerator of the right-hand fraction.

$$\text{posterior} \rightarrow p(z | x) = \frac{p(x, z)}{p(x)} = \frac{p(x | z) \cdot p(z)}{p(x)} \leftarrow \text{prior}$$

# The Rule of Bayes (Bayes' Theorem)

$x$ : observed variables

$z$ : latent variables

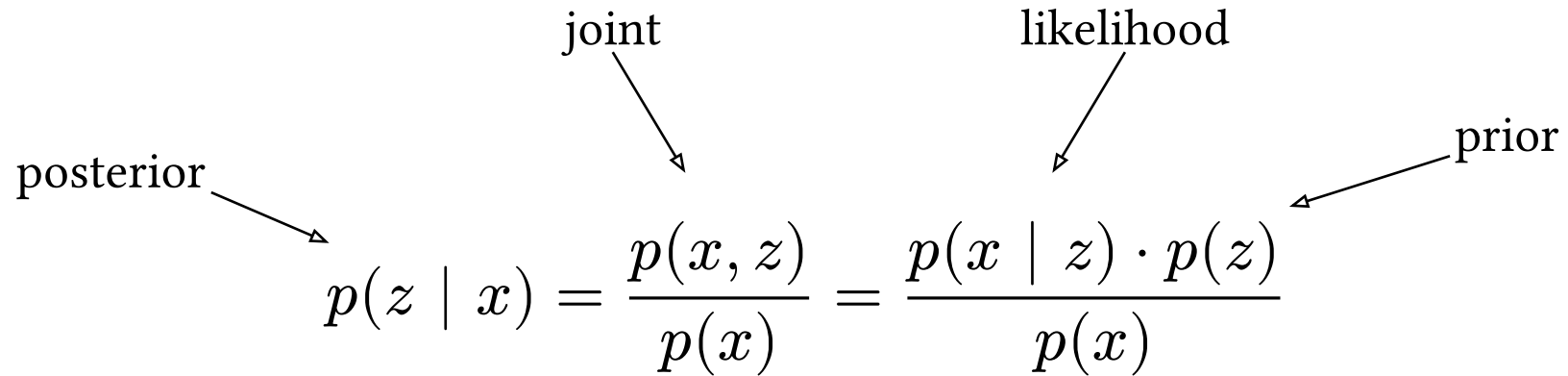
The diagram shows the equation for Bayes' Theorem with three labels and arrows: 'posterior' points to  $p(z | x)$ , 'likelihood' points to  $p(x | z)$ , and 'prior' points to  $p(z)$ .

$$\text{posterior} \rightarrow p(z | x) = \frac{p(x, z)}{p(x)} = \frac{\text{likelihood} \cdot \text{prior}}{p(x)}$$

# The Rule of Bayes (Bayes' Theorem)

$x$ : observed variables

$z$ : latent variables



The diagram shows the equation for Bayes' Theorem with four labels and arrows pointing to specific parts of the formula:

- posterior** points to  $p(z | x)$
- joint** points to  $p(x, z)$
- likelihood** points to  $p(x | z)$
- prior** points to  $p(z)$

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



# The Rule of Bayes (Bayes' Theorem)

$x$ : observed variables

$z$ : latent variables

The diagram shows the equation for Bayes' Theorem with arrows pointing from descriptive labels to the corresponding mathematical terms. The label 'posterior' points to  $p(z | x)$ . The label 'joint' points to the numerator  $p(x, z)$  of the first fraction. The label 'evidence' points to the denominator  $p(x)$  of the first fraction. The label 'likelihood' points to the numerator  $p(x | z) \cdot p(z)$  of the second fraction. The label 'prior' points to  $p(z)$  within the second fraction's numerator.

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

# 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?

	$s$	$\bar{s}$
$u$	9%	18%
$\bar{u}$	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?

	$s$	$\bar{s}$
$u$	9%	18%
$\bar{u}$	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

The diagram illustrates the formula for the posterior distribution  $p(z \mid x = D)$ . The formula is presented as an equality of two fractions. The left fraction is  $p(z \mid x = D) = \frac{p(x = D, z)}{p(x = D)}$ . The right fraction is  $= \frac{p(x = D \mid z) \cdot p(z)}{p(x = D)}$ . Arrows point from labels to specific parts of the formula: 'posterior' points to  $p(z \mid x = D)$ ; 'join' points to the joint probability  $p(x = D, z)$  in the numerator of the first fraction; 'evidence' points to the marginal likelihood  $p(x = D)$  in the denominator of the first fraction; 'likelihood' points to  $p(x = D \mid z)$  in the numerator of the second fraction; and 'prior' points to  $p(z)$  in the numerator of the second fraction.

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

constant  
distribution  
function



# Finding the Posterior Distribution

$x$ : observed variables

$z$ : latent variables

$D$ : data

The diagram illustrates the components of the posterior distribution formula. Arrows point from descriptive labels to terms in the equation:

- posterior** (blue) points to  $p(z \mid x = D)$ .
- joint** (orange) points to  $p(x = D, z)$ .
- evidence** (green) points to  $p(x = D)$ .
- likelihood** (orange) points to  $p(x = D \mid z)$ .
- prior** (blue) points to  $p(z)$ .

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

$Z \rightarrow \mathbb{P}$  is written above the posterior term, and  $Z \rightarrow \mathbb{R}$  is written above both the joint and likelihood terms.  $Z \rightarrow \mathbb{P}$  is also written above the prior term.  $\mathbb{P}$  is written below the evidence term.

**constant**  
**distribution**  
**function**

# The Problem

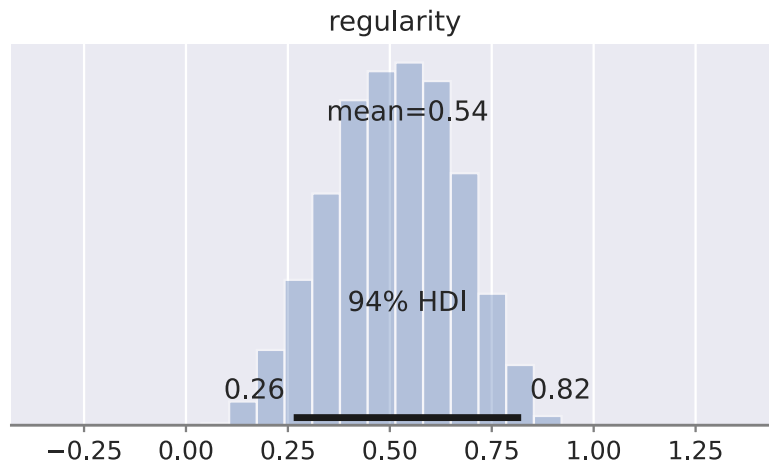
Computing  $p(x = D)$  is hard (often impossible)!

# The Problem

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

Solutions:

sample from  $p(z \mid x)$



approximate  $p(z \mid x)$

