Title

- Text

Exercise 13: Bayes Inference

MAD

bacdavid@student.ethz.ch

Outline

1. Information
2. Goals
3. Theory/ Recap
4. Exercises
5. Further proceeding wrt. Exam

ETH zürich

Information

ETH zürich

General

- Lecture material & problem sets available here
- Tutorial material available here

Goals

Goals of Today

- Understand probability basics (again)
- Understand Bayes Rule (again)
- Understand Change of Variables (again)
- Understand Maximum Likelihood (“Parameter Estimation”)

ETH zürich

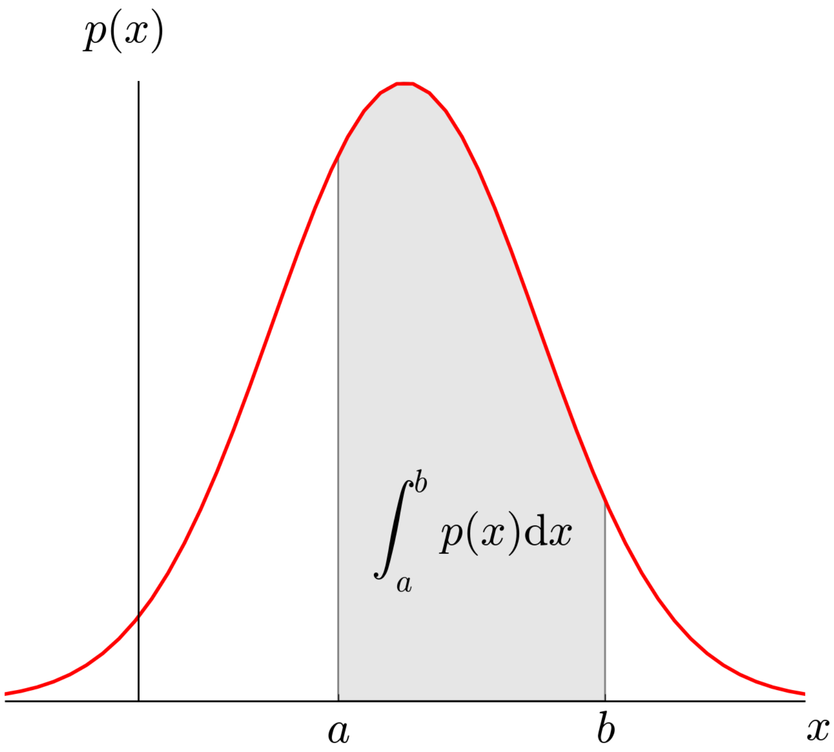
Theory / Recap

ETH zürich

Probability Review

- Discrete Random Variables
  - $p: \Omega \to \mathbb{R}$, where eg. Ω={0, 1, 2, …}
  - $\sum_{\Omega} p = 1,\ p \geq 0$

- Continuous Random Variables
  - $p: \Omega \to \mathbb{R}$, where eg. Ω=[0, 1]
  - $\int_{\Omega} p = 1,\ p \geq 0$



ETH zürich

Marginalization, Conditional Probability, and Bayes Rule (again)

- Marginalization:

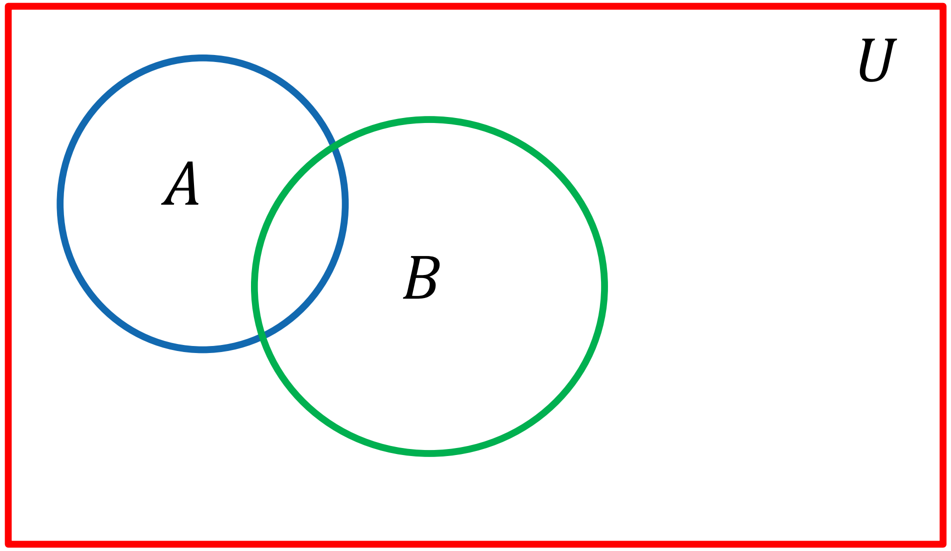
$$p(x) = \int_y p(x,y)dy$$

- Conditional Probability

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

- Bayes Rule:

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



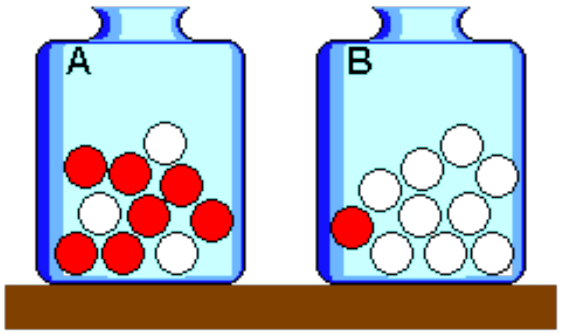
Venn Diagram of Bayes Rule: Have a look at
https://oscarbonilla.com/2009/05/visualizing-bayes-theorem/ for more info

Example 1: Marbles

- We have two jars (A, B) with red (R) and white (W) marbles.
- Drawing a marble from A or B is equally likely: $P(A) = P(B) = \frac{1}{2}$

- Write out Bayes Rule with the appropriate random variables for $P(A \mid R)$
- Compute the normalization factor $P(R)$
- Compute $P(A \mid R)$

- Tipps:
  - $p(x \mid y) = \frac{p(y|x)p(x)}{p(y)}$

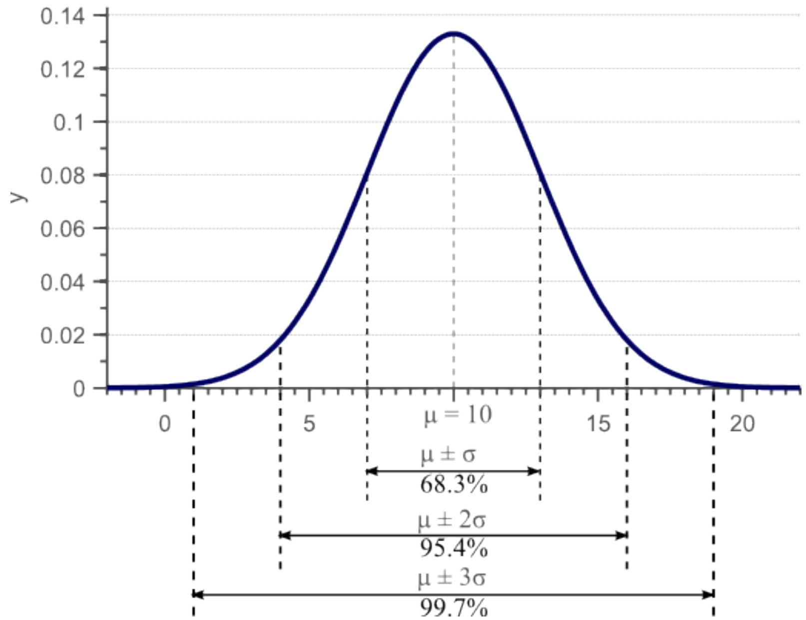


ETH zürich

Expectation value and variance (again)

- Expectation value:

$$\mathbb{E}[f(x)] = \int_{-\infty}^{\infty} f(x)p(x)dx$$



- Variance:

$$Var\big(f(x)\big) = \sigma^2 = \mathbb{E}[f^2(x)] - \mathbb{E}^2[f(x)]$$

Change of random variables (again)

- We can infere the probability distribution of on random varibales from another:

$$p(x) = \frac{p\big(f(x)\big)}{\left|\frac{f^{-1}(y)}{dy}\right|}$$

- Where the variables are related by $y = f(x)$

Example 2: Mean and variance of Gaussian distribution

- Given $x = \mu + \sigma y$, where $y \sim \mathcal{N}(0, 1)$

- Show that $x \sim \mathcal{N}(\mu, \sigma^2)$

- Tipps:
  - $p(x) = \frac{p\left(f(x)\right)}{\left|\frac{f^{-1}(y)}{dy}\right|}$
  - $\mathcal{N}(\mu, \sigma^2) = 1/\sqrt{2\pi\sigma^2} \cdot \exp - \frac{(x-\mu)^2}{2\sigma^2}$

Maximum Likelihood

- Assume we have some data D and a parametrized model, then

$$p(D \mid \theta)$$

- We want our model to describe the data as well as possible

$$\theta^* = \arg\max p(D \mid \theta)$$

Maximum Log-Likelihood

- We are interested in finding

$$\theta^* = \arg\max p(D \mid \theta)$$

- We assume that the sample are independent

$$p(D \mid \theta) = p(x_1 \mid \theta) \cdot p(x_2 \mid \theta) \cdot \ldots$$

- For convenience solve the following

$$\theta^* = \arg\max \underbrace{\log p(D \mid \theta)}_{Log-likelihood\ function} = \arg\max \log p(x_1 \mid \theta) + \log p(x_2 \mid \theta) + \cdots$$

Example 3: Maximum log-likelihood for Gauss

- Given some data $D = \{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$ and a model

$$p(\, y \mid x;\ \theta\,) = 1/\sqrt{2\pi} \cdot \exp - \frac{(y - \theta x)^2}{2}$$

- Write down the log-likelihood function $l(\theta) = \log p(D \mid \theta)$
- Find $\kappa(\theta)$ such $\arg\max \kappa(\theta)$ contains no unnecessary terms, in other words: Strap $l(\theta)$ from all terms which don't affect the $\arg\max \cdot$.
- What does the resulting term compare to? (seen in one of the first lectures)

Exercises

Q1

- Compute a few things from the Gaussian distribution (similar to what was done in examples)

ETH zürich

Q2

- Use Bayes Rule (similar to examples)

Q3

- Compute Maximum Log Likelihood
- Compute the variance of the data (see lecture notes)

ETH zürich

Further Proceedings

Exam Information (*disclaimer: Not official, “my experience”*)

- PVK (2 x halfday) by me
- Release of all material on AMIV (probably ~1 week before the PVK)
- Questions: Email or (better) Piazza

- Tipps :
  - Able to manually use the algorithms (“you have this and this data, compute one iteration…”)
  - Know about attributes of methods (“to what order accurate, etc …”)
  - Solve all hand-exercises, all in-class examples, & examples from TA slides
  - Study lecture notes (some things are not covered in exercises – might still be on exam – but focus on main concepts)

“Future” Information (*disclaimer: Not official, “my experience”*)

- Computational Methods for Engineering Application (HS19)
- Stochastik (HS19)
- Introduction to Machine Learning (master)

- At least one course of coding per semester! (Highly beneficial for almost all masters)

Questions?