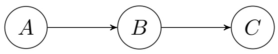
Dr. Peter Arndt
Dr. Konrad Völkel

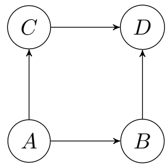
Machine Learning
Exercise Sheet 2
(4 Exercises, 120 Points)
Due: 25.10.2022, 10:00

Exercise 1: Bayes net calculations (30 Points)

Consider the following Bayes nets displaying conditional (in)dependencies between some random variables:



$\mathcal{G}_1$



$\mathcal{G}_2$

(a) Let A, B and C be binary random variables. Calculate $p(B)$ and $p(C)$ for the Bayes net defined by $\mathcal{G}_1$ and the (conditional) probabilities $p(A) = 0.3$, $p(B|A) = 0.2$, $p(B|\neg A) = 0.4$, $p(C|B) = 0.7$ and $p(C|\neg B) = 0.5$.

(b) Let A, B, C and D be binary random variables. Calculate $p(B)$ and $p(D)$ for the Bayes net defined by $\mathcal{G}_2$ and the (conditional) probabilities $p(A) = 0.3$, $p(B|A) = 0.2$, $p(B|\neg A) = 0.4$, $p(C|A) = 0.7$, $p(C|\neg A) = 0.6$, $p(D|B, C) = 0.9$, $p(D|B, \neg C) = 0.5$, $p(D|\neg B, C) = 0.3$ and $p(D|\neg B, \neg C) = 0.3$.

Exercise 2: Properties of expected value and variance (30 Points)

Let X and Y be two discrete, not necessarily independent, random variables and $a \in \mathbb{R}$ a real number.

(a) Show that the expectation is a linear operator, i.e. that $E(aX + bY) = aEX + bEY$.

(b) Show that $Var(aX) = a^2 Var(X)$

Exercise 3: Mode of the Gaussian distribution (30 Points)

The mode of a continuous probability distribution is the point at which the probability density function attains its maximum value. Prove that the mean μ is also the mode of the Gaussian normal distribution

$$\mathcal{N}(x|\mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-(x-\mu)^2}{2\sigma^2}\right).$$

Hint: Try maximizing $\log \mathcal{N}(x|\mu, \sigma^2)$ and argue why it has the same maximum.

Exercise 4: Plotting distributions (programming task) (30 Points)

The goal of this task is to make yourself familiar with some continuous distributions and plotting pdfs and histograms. First, add the following import statements:

```
import numpy as np
import matplotlib.pyplot as plt
from scipy.stats import beta, expon, gamma, laplace, norm
```

For each of the following distributions, plot its density function in the given interval and a histogram based on 1000 samples. Use `.pdf(x, *args)` to compute the probability densities for the given array of x values, e.g., `norm.pdf(x, a, b)`. Use `.rvs(*args, n)` to sample n values, e.g., `norm.rvs(a, b, n)`. For plotting you may either use `plotly` or `matblotlib`.

(a) Normal distribution: $\mathcal{N}(5, 1.5)$. Plot interval: $[0, 10]$

(b) Laplace distribution: $Laplace(5, 1.5)$. Plot interval: $[-5, 15]$

(c) Exponential distribution: $Exp(0, 1/1.5)$. Plot interval: $[-1, 8]$

(d) Gamma distribution: $Gamma(5, 0, 1)$. Plot interval: $[0, 15]$

(e) Beta distribution: $Beta(2, 1.5, 0, 1)$. Plot interval: $[0, 1]$

If you are interested, play around with the parameters of the distributions and see what influence they have on the probability densities.

Important: Make sure you also submit the output produced by your Python code.