Machine Learning Assignment # 3
Universität Bern

Due date: 11/10/2017

Late submissions will incur a penalty. Submit your answers in ILIAS (as a pdf or as a picture of your written notes if the handwriting is very clear). Submission instructions will be provided via email. You are not allowed to work with others.

Probability theory review **[Total 100 points]**

Solve each of the following problems and show all the steps of your working.

1. Prove the following statements regarding the covariance:
   (a) Show that if X and Y are independent then $\text{Cov}[X, Y] = 0$.
   Give an example that shows that the opposite is not true. **[10 points]**
   (b) Show that the covariance matrix is always symmetric and positive semidefinite. **[10 points]**

2. $X \in R^n$ and $Y \in R^m$ are independent random variables. Their expectations and covariances are $E[X] = 0$, $\text{Cov}[X] = I$, $E[Y] = \mu$ and $\text{Cov}[Y] = \sigma I$, where I is the identity matrix of the appropriate size and σ is scalar. What is the expectation and covariance of the random variable $Z = AX + Y$, where $A \in R^{m \times n}$? **[25 points]**

3. Thomas and Viktor are friends. It is friday night and Thomas does not have phone. Viktor knows that there is a $2/3$ probability that Thomas goes to the party in downtown. There are 5 pubs in downtown and there is an equal probability of Thomas going to any of them if he goes to the party. Viktor already looked for Thomas in 4 of the bars.
   What is the probability of Viktor finding Thomas in the last bar? **[10 points]**

4. Let $A \in \mathbb{R}^{n \times n}$ be a positive definite square matrix, $b \in \mathbb{R}^n$, and c be a scalar. Prove that **[25 points]**

$$\int_{x \in \mathbb{R}^n} e^{-\frac{1}{2}x^T A x - x^T b - c} dx = \frac{(2\pi)^{n/2} |A|^{-1/2}}{e^{c - \frac{1}{2} b^T A^{-1} b}}.$$

   Hint: Recall the Gaussian distribution.

5. Let $X_1, X_2, ..., X_n$ be i.i.d. Poisson random variables, with $P(X = x) = \frac{\lambda^x e^{-\lambda}}{x!}$.
   Find the λ that maximizes the likelihood. **[10 points]**

6. From the definition of conditional probabilities for multiple random variables, show that **[10 points]**

$$f(x_1, x_2, ..., x_n) = f(x_1) \prod_{i=2}^{n} f(x_i | x_1, ..., x_{i-1})$$