GEORGE MASON UNIVERSITY
Systems Engineering and Operations Research
OR750/610: Deep Learning, Fall Semester 2019: Homework Assignment 1. Due: Sep 18 (before class)

1. As a result of medical examination, one of the tests revealed a serious illness in a person. This test has a high precision of 99% (the probability of a positive response in the presence of the disease is 99%, the probability of a negative response in the absence of the disease is also 99%). However, the detected disease is quite rare and occurs only in one person per 10,000. Calculate the probability that the person being examined does have an identified disease.

2. Consider the following probabilistic model. The student does poorly in a class ($c = 1$) or well ($c = 0$) depending on the presence / absence of depression ($d = 1$ or $d = 0$) and weather he/she partied last night ($v = 1$ or $v = 0$) . Participation in the party can also lead to the fact that the student has a headache ($h = 1$). As a result of poor student's performance, the teacher gets upset ($t = 1$). The probabilities are given by:

$p(c = 1\|d, v)$	v	d
0.999	1	1
0.9	1	0
0.9	0	1
0.01	0	0

$p(h = 1\|v)$	v
0.9	1
0.1	0

$p(t = 1\|c)$	c
0.95	1
0.05	0

$p(v = 1) = 0.2$, and $p(d = 1) = 0.4$.

Draw the causal relationships in the model. Calculate $p(v = 1|h = 1)$, $p(v = 1|t = 1)$, $p(v = 1|t = 1, h = 1)$.

3. Let $x_i \sim \text{Poss}(\lambda)$, $1 = 1, \ldots, N$. Find λ using MLE.

4. Let $x_i \sim \text{Poss}(\lambda)$, $1 = 1, \ldots, N$. Find $p(\lambda|x_1, \ldots x_N)$, assuming the Gamma prior $\lambda \sim \Gamma(\lambda|a, b)$.

5. Let $x_1, x_2, \ldots, x_N$ be an independent sample from the exponential distribution with density $p(x|\lambda) = \lambda \exp(-\lambda x)$, $x \geq 0$, $\lambda > 0$. Find the maximum likelihood estimate λ_{ML}. Choose the conjugate prior distribution $p(\lambda)$, and find the posterior distribution $p(\lambda|x_1, \ldots, x_N)$ and calculate the Bayesian estimate for λ as the expectation over the posterior.

6. SGD for Ridge Regression Solve the ℓ_2 regularized logistic regression. The objective function is

$$f(\theta) = \frac{1}{m} \sum_{i=1}^{m} \left(-y^{(i)} \log(h_\theta(x^{(i)})) - (1 - y^{(i)}) \log(1 - h_\theta(x^{(i)})) \right) + \frac{\lambda}{2m} ||\theta||_2^2$$

where

$$h_\theta(x) = g(\theta^T x), \qquad g(z) = \frac{1}{1 + e^{-z}}$$

There are three parameters in the model $\theta = (\theta_1, \theta_2, \theta_3)$, where θ_1 corresponds to intersect, so you need to add column of ones to the data. Fit the logistic regression to the `reg-lr-data` dataset. Write down the derivative for the objective function.

(a) Solve using gradient descent method

(b) Solve using stochastic gradient descent.

Run the algorithms with different step sizes and different values of λ. Plot data and the regression line as well as convergence plot (iteration vs sopping criteria) for several runs.