Columbia University – Fall 2013

Machine Learning Homework #3

Joe Ellis - jge2105

November 6, 2013

1 Problem 1

1.1 Problem1.1 - Discrete

We assume that we have a coin that can land on heads or tails, and we have 3 possible values for μ, which are $\frac{1}{2}, \frac{1}{4}, \frac{3}{4}$. The first problem that we address is to find the minimum number of tosses we'd need to see in order to conclude that $p(\mu = \frac{1}{2}) > \frac{1}{2}$. We will refer to this series of tosses as, D. Using Bayes rule we know that $p(\mu|D) = p(D|\mu)p(\mu)/p(D)$. For this example, we can assume that the prior on each choice of mu is equal. We know that the flipping of a coin is given by the bernoulli distribution, whereN_H and N_T are the number of heads and tails in our set respectively. We will ignore $p(\mu)$ in the equations since the priors are uniform. The quickest time that we can be sure that the $p(\mu = \frac{1}{2}|D) > \frac{1}{2}$ is $D = (H, T, H, T, H, T)$

$$p(\mu = \frac{1}{2}|D) = \frac{(\mu)^{N_H}(1-\mu)^{N_T}}{\sum\limits_{\mu}(\mu)^{N_H}(1-\mu)^{N_T}} \tag{1}$$

$$= .5424 \tag{2}$$

The second test that we will do on this is to find the minimal set D such that $p(\mu = \frac{3}{4}|D) > \frac{1}{2}$. For this the smallest dataset that makes this a reality is $D = (H, H)$, by the same logic.

$$p(\mu = \frac{3}{4}|D) = \frac{(\mu)^{N_H}(1-\mu)^{N_T}}{\sum\limits_{\mu}(\mu)^{N_H}(1-\mu)^{N_T}} \tag{3}$$

$$= .6429 \tag{4}$$

1.2 Problem 1.2 - Continuous

Consider 2 possible distributions: (A) $\mu\ uniform[0,1]$; (B) we have some reason to think the coin is likely to be fair, so we have a parabola that is peaked at $\frac{1}{2}$ and then goes to 0 at 0 and 1.

Use the two datasets, and answer a variety of questions for each, $D_1 = (H,T)$; $D_2 = (T,T,T)$.

1.2.1 A − 1

2 Problem 2

We are assuming that we have a good test for swine flu that is highly accurate, with the probabilities given below. We want to find the likelihood if I (Joe) take the test and it outputs true, what is the probability that I have swine flu.

$$p(test = true|flu = True) = 0.99 \tag{5}$$

$$p(test = false|flu = false) = 0.98 \tag{6}$$

$$p(flu = true) = .0001 \tag{7}$$

$$p(flu = false) = .999 \tag{8}$$

Let's use Bayes Rule to solve this problem.

$$p(flu = true|test = True) = \frac{p(test = true|flu = True)p(flu = true)}{p(test = true)} \tag{9}$$

$$= \frac{p(test = true|flu = True)p(flu = true)}{p(test = true|flu = True)p(flu = true) + p(test = true|flu = False)p(flu =} \tag{10}$$

$$= \frac{0.99 * 0.0001}{0.99 * 0.0001 + (1 - 0.98) * 0.999} \tag{11}$$

So still not that likely, shows the power of the prior.

3 Problem 3

4 Problem 4

The weibull distribution is a probablility distribution over non-negative scalar values. The distribution is as follows, $p(x|\lambda) = \frac{3}{\lambda}(\frac{x}{\lambda})^2 exp(-(\frac{x}{\lambda})^3)$. Given a dataset $D = (x_0, x_1, ..., x_N)$, what is the ML estimate of λ. Let's use the log-likelihood.

$$
\begin{aligned}
l(D|\lambda) &= \sum log(p(x_i|\lambda) & (12)\\
&= \sum log(\frac{3}{\lambda}) + 2log(\frac{x_i}{\lambda}) - (\frac{x_i}{\lambda})^3 & (13)\\
&= \sum log(3) - log(\lambda) + 2log(x_i) - 2log(\lambda) - (\frac{x_i}{\lambda})^3 & (14)\\
& & (15)
\end{aligned}
$$

Now let's take the derivative of the log-likelihood, and set to 0.

$$
\begin{aligned}
&\frac{\partial l}{\partial \lambda} = \frac{\partial}{\partial \lambda} \sum log(3) - log(\lambda) + 2log(x_i) - 2log(\lambda) - (\frac{x_i}{\lambda})^3 & (16)\\
&\sum -\frac{1}{\lambda} - \frac{2}{\lambda} + 3\lambda^{-4}x_i^3 = 0 & (17)\\
&\frac{-3N}{\lambda} + \sum 3\lambda^{-4}x_i^3 = 0 & (18)\\
&\frac{3N}{\lambda} = \sum 3\lambda^{-4}x_i^3 & (19)\\
&\frac{3N\lambda^4}{3\lambda} = \frac{\lambda^4}{x_i^3} \sum 3\lambda^{-4}x_i^3 & (20)\\
&\lambda^3 = \sum \frac{x_i^3}{N} & (21)\\
&\lambda = (\sum \frac{x_i^3}{N})^{1/3} & (22)
\end{aligned}
$$

5 Problem 5