STA 360: Assignment 2

Michael Lin

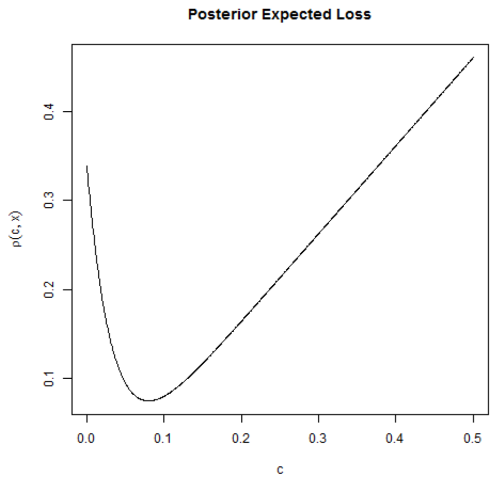
January 19, 2015

4. The 0-1 loss function is given by
$$l(S,a) = \mathbb{1}(S \neq a)$$
Therefore, the posterior expected loss is
$$\mathbb{E}(l(S,a)|x_{1:n}) = 1 \cdot \mathbb{P}(S \neq a|x_{1:n}) + 0 \cdot \mathbb{P}(S = a|x_{1:n})$$
To minimize the posterior expected loss, we would need to find a that would minimize $\mathbb{P}(S \neq a|x_{1:n})$. Since the events $S \neq a$ and $S = a$ are mutually exclusive, minimizing the probability of the former event is equivalent to maximizing the probability of the latter. In particular, we would need the action a that maximizes $\mathbb{P}(S = a|x_{1:n})$, i.e.
$$\delta(x) = \text{argmax}_a \mathbb{P}(S = a|x_{1:n})$$

5. Since $x_1, ..., x_n$ are Bernoulli random variables, the intuitive prediction of x_{n+1} would be the mode of $x_1, ..., x_n$, i.e., the value (either 1 or 0) that occurs more often. Based on the previous problem, the Bayes procedure for making this prediction would be choosing the a that would maximize $\mathbb{P}(x_{n+1} = a|x_{1:n})$. Let $a_n = a + \sum x_i$ and $b_n = b + n - \sum x_i$. Thus, given the posterior distribution
$$p(\theta|x_{1:n}) = \text{Beta}(\theta|a_n, b_n)$$
the posterior predictive, from section 3.2 of notes, would be
$$p(x_{n+1}|x_{1:n}) = \begin{cases} \frac{a_n}{a_n+b_n} & \text{if } x_{n+1} = 1 \\ \frac{b_n}{a_n+b_n} & \text{if } x_{n+1} = 0 \end{cases}$$
Therefore, the Bayes procedure is
$$\delta(x) = \text{argmin}_a \rho(a,x) = \begin{cases} 1 & \text{if } \frac{a_n}{a_n+b_n} \geq \frac{b_n}{a_n+b_n} \\ 0 & \text{otherwise} \end{cases}$$

6. The hyperparameters would be a, b such that
$$\frac{a + \sum x_i}{a+b+n} \geq \frac{b + n - \sum x_i}{a+b+n}$$
But notice this is true for $a = b = 0$. Thus this would be one such setting of the hyperparameters that agrees with the above Bayes procedure. The values a and b determine the posterior predictive and shift the intuitive choice. A larger a shifts the prediction towards 1; a larger b shifts the prediction towards 0.

8. The following Riemann sum with equal partition $\{x_0 = 0, x_1, x_2 ..., x_N = 1\}$ was used to approximate an integral:
$$\int_0^1 f(x)dx \approx \frac{1}{N}\sum_{i=1}^{N} f(x_{i-1})$$
Below is the plot of the posterior expected loss.


Posterior Expected Loss
ρ(c, x)
0.1
0.2
0.3
0.4
0.0
0.1
0.2
0.3
0.4
0.5
c

See below for R code used to plot posterior expected loss.

```
## Precision parameters
N.graph = 1000
c.val = seq(0,0.5,length=N.graph);
N = 1000; # number of partitions


## Distribution parameters
a = 1.05;
b = 30;


## Initialize vectors
pel.val = rep(0,N.graph);
theta.val = seq(0,1,length=N); # all partitions
loss.val = rep(0,N);


## Calculate posterior expected loss values
for(k in 1:N.graph) {

  for(i in 1:N) {
    if(c.val[k] < theta.val[i]) {
      loss.val[i] = 10*abs(theta.val[i]-c.val[k]);
    }
    else loss.val[i] = abs(theta.val[i]-c.val[k]);
  }

  fun.val = loss.val*dbeta(theta.val, a, b);

  int.val = rep(0,N-1);

  for(j in 1:N-1) {
    int.val[j] = fun.val[j]/(N-1);
  }

  pel.val[k] = sum(int.val);
}


## Graph
png("pel.png")
plot(c.val, pel.val, type="l", lty=1, xlab="c",
     ylab=expression(rho(c,x)), main="Posterior Expected Loss");
dev.off()
```