HW4 Solutions

Gabe Hope
March 2017

1 Problem 1

1.1 Part a

Recall that the maximum likelihood estimate for a Bernoulli distribution is:

\[ \hat{\mu}_{MLE} = \frac{k}{n} \]

Where \( k \) is the number of 1’s (heads) and \( n \) is the total number of samples.

Therefore our maximum likelihood estimate for \( \mu \) given the results of our coin tosses is:

\[ \hat{\mu}_{MLE} = \frac{4}{4+9} = \frac{4}{13} \]

1.2 Part b

Code to generate posterior plots:

```matlab
% Compute values between 0 and 1
x = [0:0.01:1];
y1 = betapdf(x, 4+1, 9+1);
y2 = betapdf(x, 4+2, 9+2);
y3 = betapdf(x, 4+10, 9+10);
y4 = betapdf(x, 4+100, 9+100);

% Create the plot with a legend
figure
hold on
plot(x, y1);
plot(x, y2);
plot(x, y3);
plot(x, y4);
legend('a=b=1', 'a=b=2', 'a=b=10', 'a=b=100')
```
1.3 Part c

Recall that the mean ($\bar{\mu}$) of the posterior predictive distribution given a set of Bernoulli samples $D$ and a beta prior parametrized by $\alpha, \beta$ is:

$$\bar{\mu} = P(x = 1|D) = \frac{m + \alpha}{m + \alpha + l + \beta}$$

Where $m$ is the number of 1’s (heads) in $D$ and $l$ is the number of 0’s (tails) in $D$. We see that this is equivalent to the probability that the next toss will be a heads. Applying this formula to each of our hyperparameter settings we get:

$$P(x = 1|D) = \frac{4 + 1}{4 + 1 + 9 + 1} = \frac{5}{15} = \frac{1}{3} \quad (1)$$

$$P(x = 1|D) = \frac{4 + 2}{4 + 2 + 9 + 2} = \frac{6}{17} \quad (2)$$

$$P(x = 1|D) = \frac{4 + 10}{4 + 10 + 9 + 10} = \frac{14}{33} \quad (3)$$

$$P(x = 1|D) = \frac{4 + 100}{4 + 100 + 9 + 100} = \frac{104}{213} \quad (4)$$

Note that a stronger prior brings our posterior beliefs closer to that of a fair coin.
1.4 Part d
Recall that the MAP estimate for \( \mu \) was derived to be:

\[
\mu_{MAP} = \mu = \frac{m + (\alpha - 1)}{m + l + (\alpha - 1) + (\beta - 1)}
\]

Plugging in our different hyperparameter settings we get:

\[
\mu_{MAP} = \frac{4 + (1 - 1)}{4 + 9 + (1 - 1) + (1 - 1)} = \frac{4}{13} \quad (1)
\]

\[
\mu_{MAP} = \frac{4 + (2 - 1)}{4 + 9 + (2 - 1) + (2 - 1)} = \frac{5}{15} \quad (2)
\]

\[
\mu_{MAP} = \frac{4 + (10 - 1)}{4 + 9 + (10 - 1) + (10 - 1)} = \frac{13}{31} \quad (3)
\]

\[
\mu_{MAP} = \frac{4 + (100 - 1)}{4 + 9 + (100 - 1) + (100 - 1)} = \frac{103}{211} \quad (4)
\]

2 Problem 2
We can first write the expected loss of each choice in terms of \( \rho = p(y = 1|x) \):

\[
E[L(y, r(x))] = \begin{cases} 
\rho & r(x) = 0 \\
1 - \rho & r(x) = 1 \\
\lambda & r(x) = reject 
\end{cases}
\]

For any given input \( x \) we want our rule to choose the labels with the minimal expected loss. We see that given these three values our rule should reject if:

\[
\lambda < \rho < 1 - \lambda
\]

Otherwise we should assign a label of 1 if:

\[
\rho > \frac{1}{2}
\]