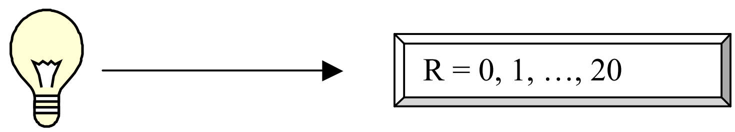


S= 0, 1, ……, 20

Problem: Estimation of Light Bulb Brightness. As shown above, you have a light bulb that you can adjust its brightness S, i.e. S can be any integer from 0 to 20. You are given a special type of meter that measures the brightness of the light it receives and gives you a reading, which is an integer from 0 to 20. However, the manufacturer did a poor job with this meter, and for a fixed S value, if you repeats your measure several times, it will reads differently, and vice versa. In addition, there seems no obvious simple relationship between S and R values. Nevertheless, since this is the only meter you have, you are determined to do the best you can to estimate the brightness level of the light bulb .

You start by repeatedly reading your meter for hundreds of times while you keep your light bulb at each brightness level S. The table below shows the results you got:

	S_1	S_2	S_3	……	S_{20}
R_1	$P[R_1 \mid S_1]$	$P[R_1 \mid S_2]$	$P[R_1 \mid S_3]$	……	$P[R_1 \mid S_{20}]$
R_2	$P[R_2 \mid S_1]$	$P[R_2 \mid S_2]$	$P[R_2 \mid S_3]$	……	$P[R_2 \mid S_{20}]$
R_3	$P[R_3 \mid S_1]$	$P[R_2 \mid S_2]$	$P[R_3 \mid S_3]$	……	$P[R_3 \mid S_{20}]$
……	……	……	……	……	……
R_{20}	$P[R_{20} \mid S_1]$	$P[R_{20} \mid S_2]$	$P[R_{20} \mid S_3]$	……	$P[R_{20} \mid S_{20}]$

Each entry in the table is the conditional probability of reading R when the light brightness is S, i.e. P[R|S]. Data within each column are collected from experiments in which brightness S is the same, while data within each row are from experiments in which the meter gives the same reading R.

The actual data look like:

	S_1	S_2	…
R_1	0.0814	0.1082	…
R_2	0.0198	0.0070	…
…	…	…	…

You need to go to the course website to download the actual data, which is saved in .mat format. You should be able to open it in Matlab. You will see 4 variables. P_RS1 (21x21) contains data when the probabilities of S taking on all possible values (0 to 20) are the same, i.e. P[S] is constant and independent of S. P_S1 (1x21) gives the actual P[S]. P_RS2 (21x21), contains data when P[S] increases linearly with S, i.e. S is more possible to be a bigger value than a smaller one, and P_S2 (1x21) gives the P[S] in this condition.

1) When P[S] is constant, and you read 3 from the meter.

a. For R = 3, in the same figure, plot both P[R|S] and P[S|R] as a function of S. What is the estimated S given by MAP? What is the estimated S given by ML? Mark them on your plot. Are they the same? Why?

b. The function that Bayesian inference minimizes is: F = ∑(L·P[S|R]) with R = 3. Let's define the loss function as $L(S, S_{bayes}) = (S – S_{bayes})^2$. Now, assume S_{bayes} equals 0, compute the value of F. Repeatedly, you can assume S_{bayes} equals 1 to 20 and compute the corresponding F values. Plot F as a function of S. Where does F take the minimum value? Now, let's define the loss function as $L(S, S_{bayes}) = |S – S_{bayes}|$, again, compute F and plot it as a function of S. Where does F take the minimum value now? Are they the same?

c. From the Dayan & Abbott book, we know that actually S_{bayes} will be either the mean or the median of S depending on the format of the loss function. Try to compute the mean and the median directly. Do they agree with the S value where F takes the minimum in question c?

2) When P[S] is constant, and you read 10 from the meter, repeat a, b, and c.

3) When P[S] increases with S, and you read 3, repeat a, b and c.

4) When P[S] increases with S, and you read 10, repeat a, b and c.

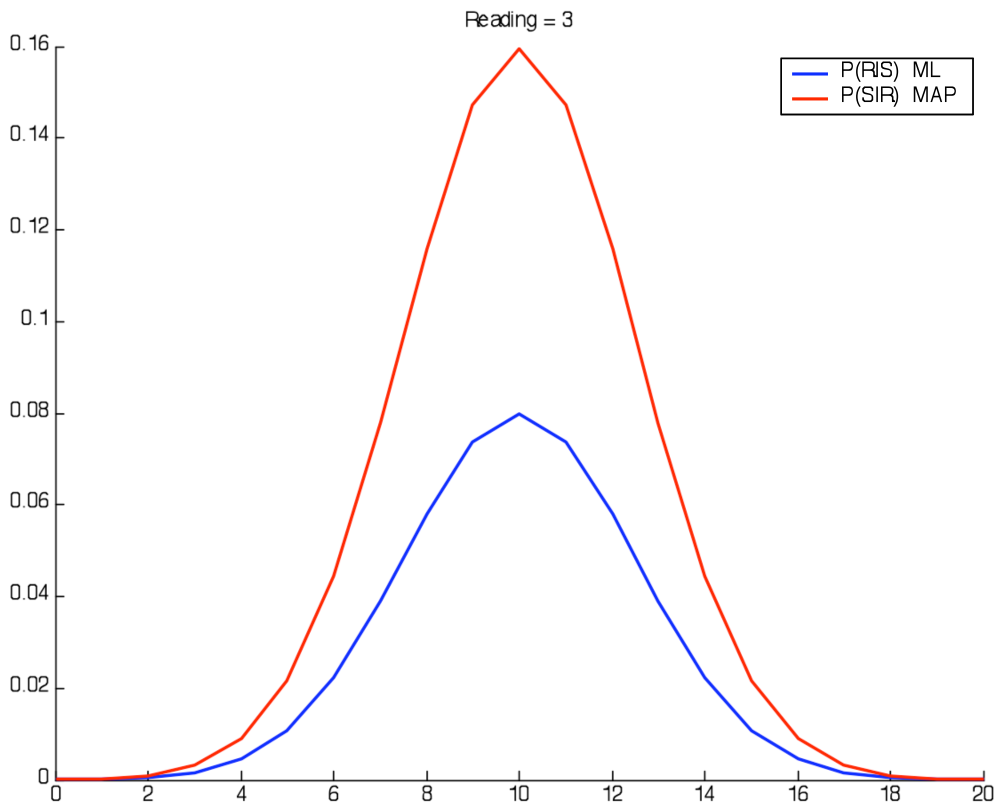
Some useful hints:

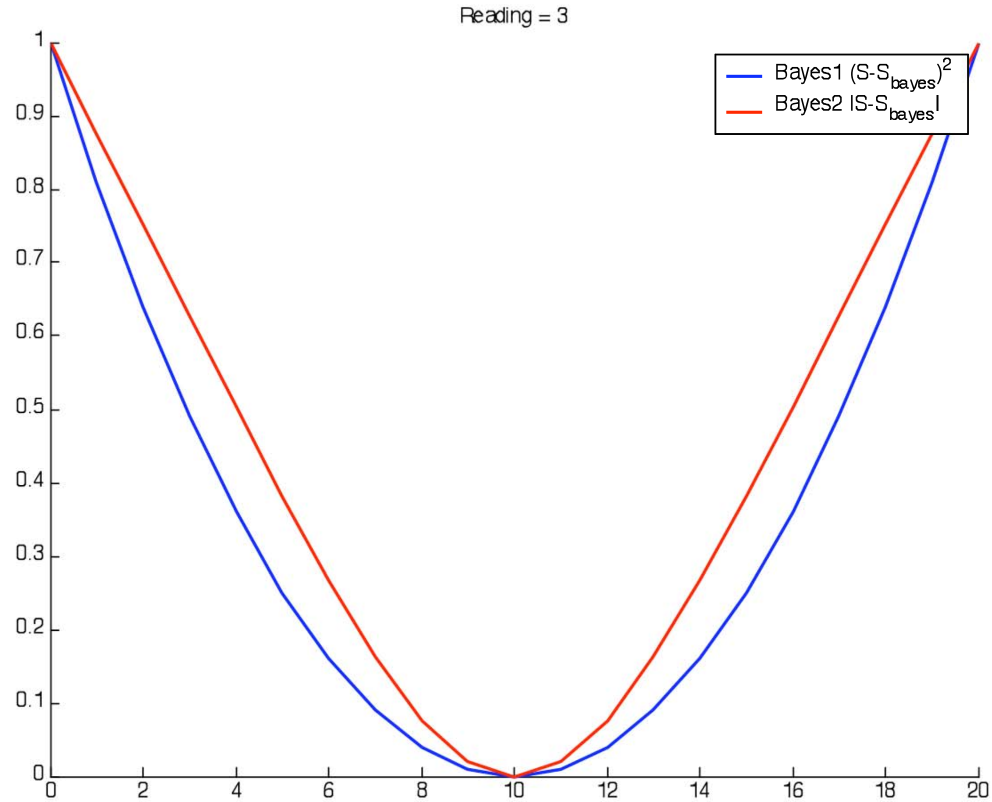
1. The table gives you P[R|S]. Write a simple matlab code that uses the Bayes theorem to compute the corresponding P[S|R].
2. You need to find P[R] before you use the Bayes theorem, and P[R]=∑(P[R|S]·S).
3. For a random variable S that takes different values with different probabilities, the mean is ∑(P[S]·S), and the median is the S value that satisfies: $P[S<S_{med}] = P[S>S_{med}]$. Easiest way to find the median is to try all S values and find out the one that gives the smallest difference between $P[S<S_{med}]$ and $P[S>S_{med}]$.

ANSWER: The idea here is to show the different estimations four methods will give. These four methods are Bayes1 (loss function is squared difference), Bayes2 (loss function is absolute value), MAP and ML.

The estimation will be carried out in 4 conditions.

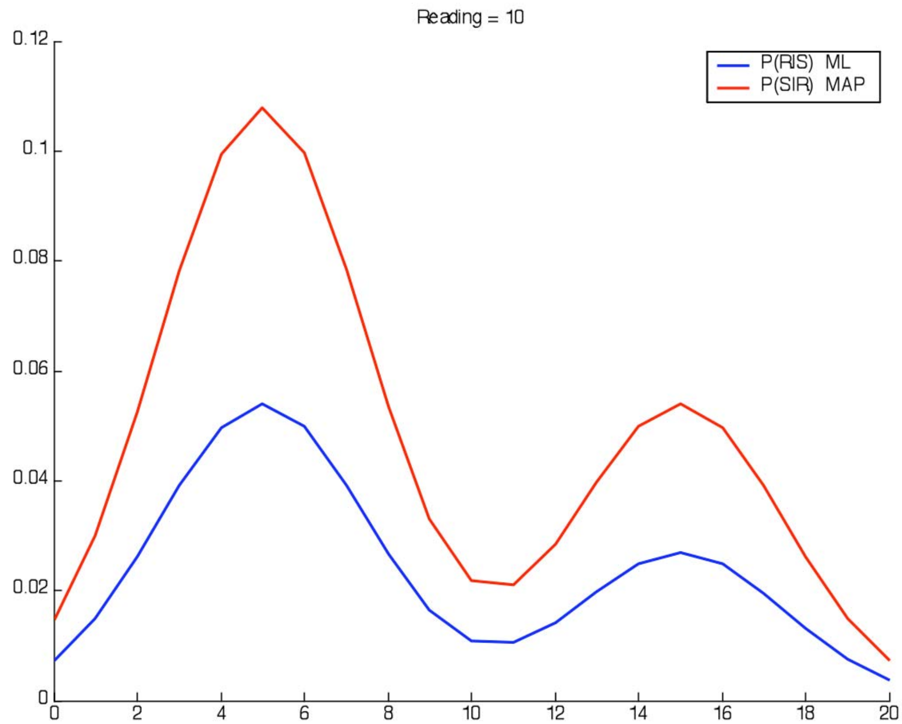
For constant P[S] and R = 3, P[R|S] is constructed to be symmetric Gaussian and all four methods will give the same answer (S = 10). The following figures are expected:

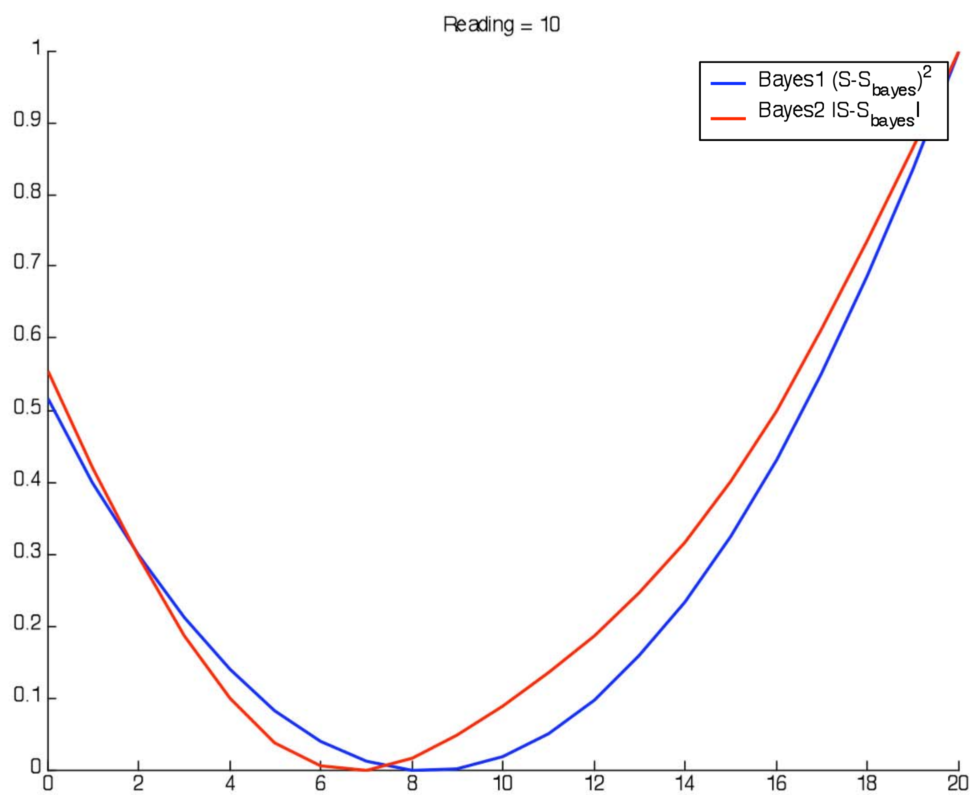




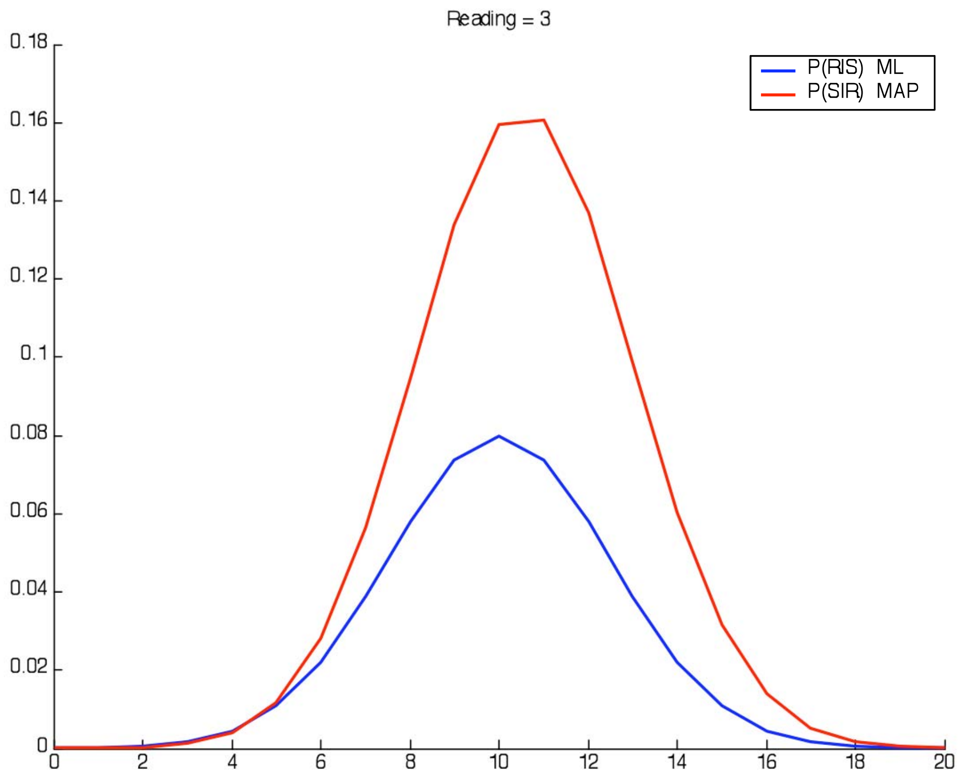
Here, we intentionally have people compute the function that Bayes inference will minimize (e.g, Question 1b) and this should give a very intuitive understanding. The S value obtained by computing mean/median directly should agree with the S value that you get by looking at the above plot.

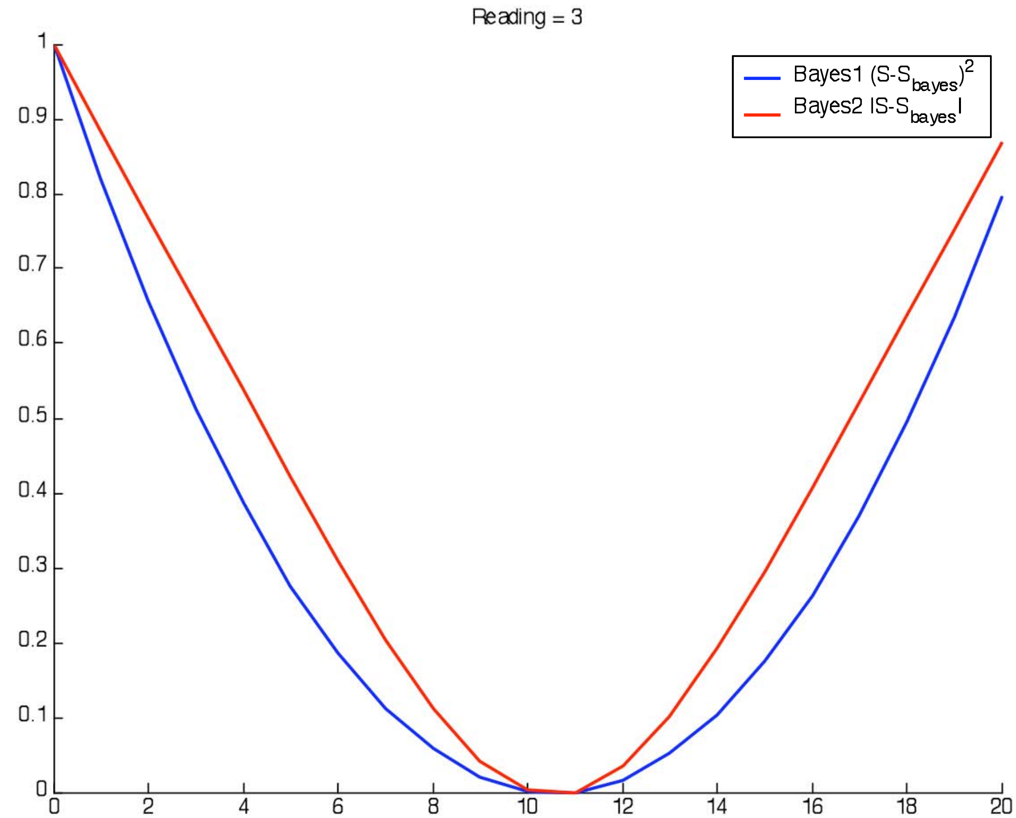
For constant P[S] and R = 10, P[R|S] is constructed to be a summation of two Gaussians (two humps). This time, MAP and ML gives the same answer (S = 5). Bayes1 (S = 8) and Bayes2 (S = 7) give different answers, since the mean and the median are different now.


Reading = 10
P(R|S) ML
P(S|R) MAP

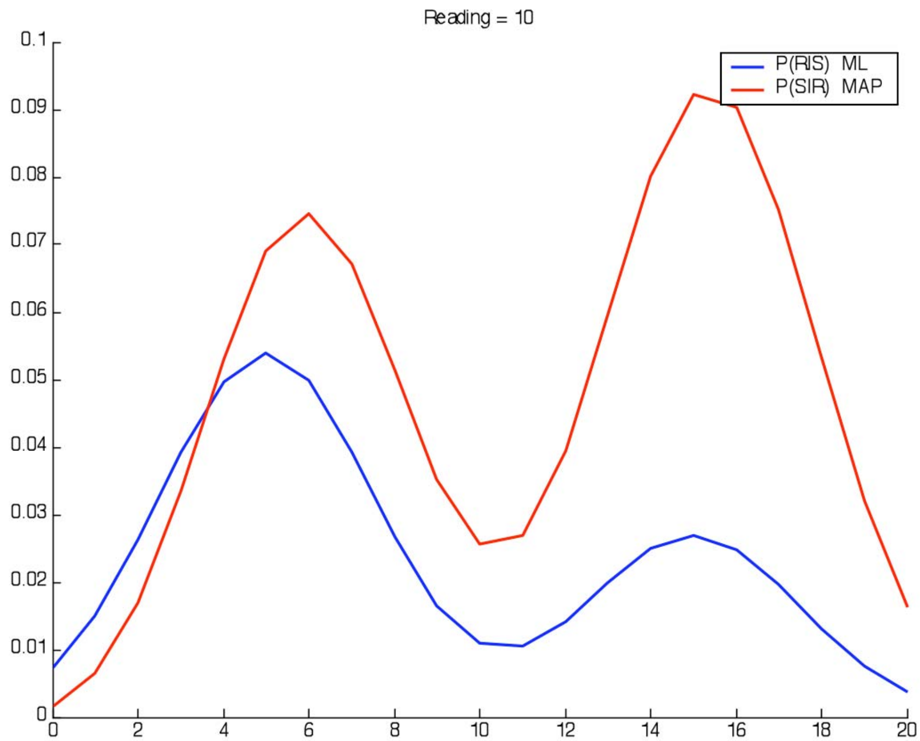

Reading = 10
Bayes1 (S-S_bayes)^2
Bayes2 |S-S_bayes|

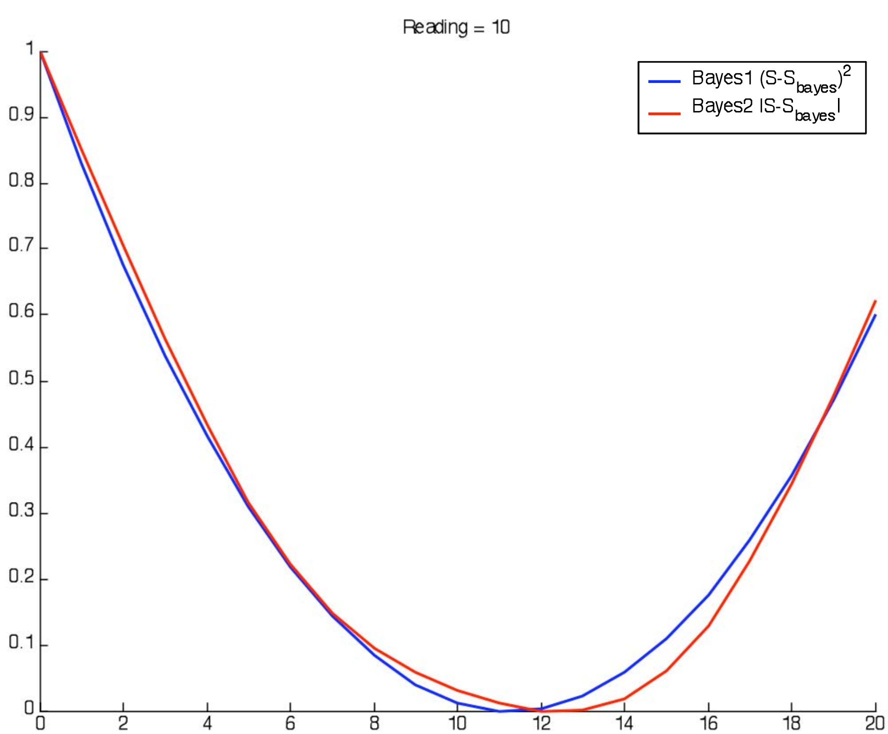
For increasing P[S] and R = 3, MAP (S = 11) is shifted toward the right due to the prior distribution P[S]. And ML stays the same. (S =10). Bayes1 and Bayes2 give the same estimate (S = 11).





For increasing P[S] and R = 10, all four methods gives different answers. (MAP: 15. ML: 5. Bayes1: 11. Bayes2: 12)


Reading = 10
P(R|S) ML
P(S|R) MAP
0.1
0.09
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
0
2
4
6
8
10
12
14
16
18
20


Reading = 10
Bayes1 $(S-S_{bayes})^2$
Bayes2 $|S-S_{bayes}|$
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
2
4
6
8
10
12
14
16
18
20