Quiz 3

Daniel Salgado Rojo

Computational Neuroscience - Coursera, January 2018.

1) **Likelihood ratio test with asymmetric costs**.

Suppose we have a stimulus defined by a single variable called s. s can take one of two values, which we will call s_1 and s_2. You could think of these as lights flashing in the eyes at one of two possible frequencies. Or perhaps listening to punk rock vs. listening to Dvorak.

Let's call the firing rate response of a neuron to this stimulus r.

Suppose that under stimulus s_1 the response rate of the neuron can be roughly approximated with a Gaussian distribution with the following parameters:

- μ (mean): 5,
- σ (standard deviation): 0.5

And likewise for s_2:

- μ: 7
- σ: 1

Lets say that both stimuli are equally likely and we are given no other prior information.

Now let's throw in another twist. Let's say that we receive a measurement of the neuron's response and want to guess which stimulus was presented, but that to us, it is twice as bad to mistakenly think it is s_2 than to mistakenly think it is s_1.

Which of these firing rates would make the best decision threshold for us in determining the value of s given a neuron's firing rate?

- 5.667
- 5.830
- 2.69
- 5.978

Solution. From the first part of the problem statement we have that

$$P(r|s_1) = \mathcal{N}(5, 0.5^2), \qquad P(r|s_2) = \mathcal{N}(7, 1^2)$$

where $\mathcal{N}(\mu, \sigma^2)$ denotes a Gaussian distribution of mean μ and standard deviation σ. Moreover, we have that $P(s_1) = P(s_2) = 0.5$.

Now following the notation from lecture 3.1, but now labelling using sub-indices 1 and 2 instead of - and +, respectively, we can write the following loss balance condition:

$$\mathrm{LOSS}_1 = L_1 P(s_2|r) < L_2 P(s_1|r) = \mathrm{LOSS}_2 \tag{1}$$

where L_1 is the penalty factor when we say s_1 but it was s_2 (and L_2 is the penalty factor if we say s_2 but it was s_1). The information of the problem statement gives us that

$$2L_1 = L_2 \tag{2}$$

By combining equations (1) and (2), and applying Bayes' Theorem, we arrive to:

$$L_1 \frac{P(r|s_2)P(s_2)}{P(r)} < L_2 \frac{P(r|s_1)P(s_1)}{P(r)} \iff L_1 P(r|s_2) < L_2 P(r|s_1) \iff \frac{1}{2} < \frac{P(r|s_1)}{P(r|s_2)} \tag{3}$$

and where $P(r|s_i) = \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left(-\frac{(r-\mu_i)^2}{2\sigma_i^2}\right)$.

The best decision threshold will be given by the solution of the equation:

$$\frac{1}{2} = \frac{P(r|s_1)}{P(r|s_2)} \tag{4}$$

which can be easily solved numerically. In Figures 1 we can see the two Gaussian distributions displayed together. In Figure 2 we have plotted $P(r|s_1)$ in blue v.s. $\frac{1}{2}P(r|s_2)$ in red; the point where both distributions coincide is the threshold we are looking for.

Either graphically or numerically we can conclude that the threshold has to be equal to 0.978 (see `quiz3.R` for more details on the code).

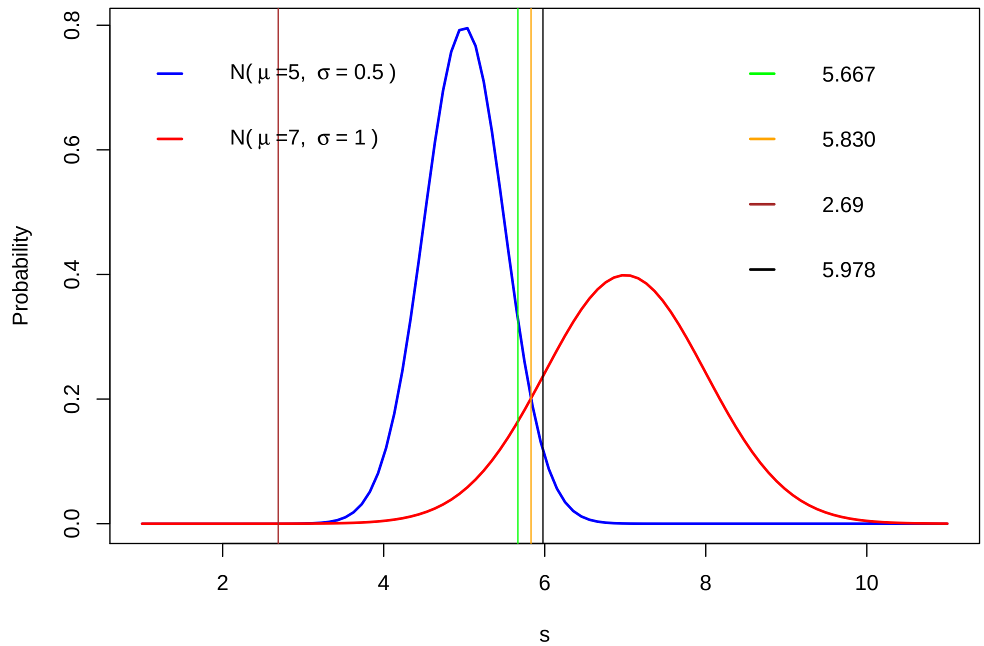

Figure 1

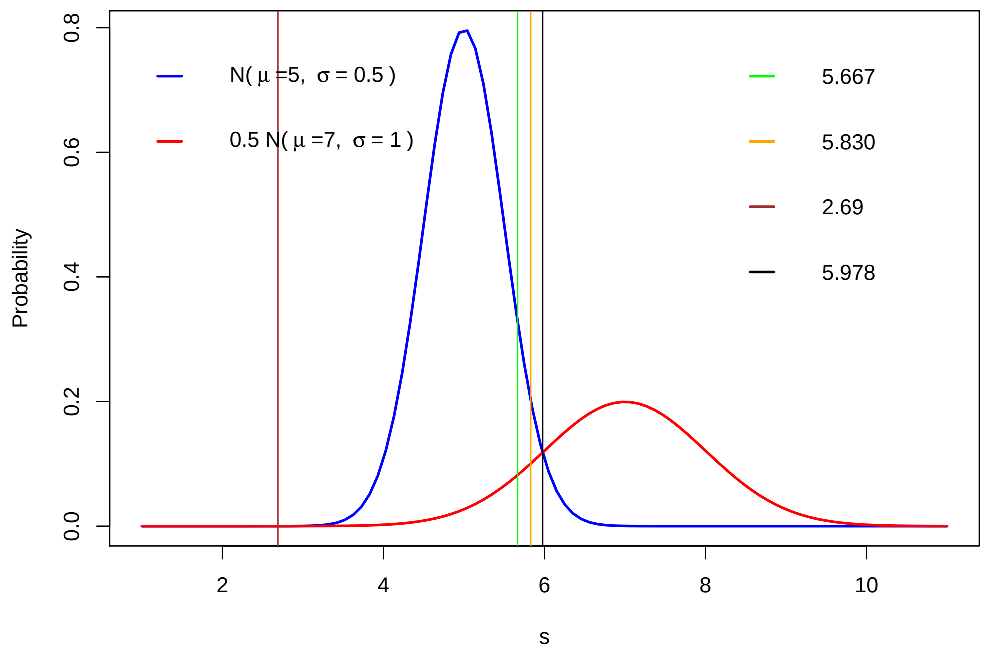

Figure 2

◀

2) Suppose we are diagnosing a very rare illness, which happens only once in 100 million people on average. Luckily, we have a test for this illness but it is not perfectly accurate. If somebody has the disease, it will report positive 99% of the time. If somebody does not have the disease, it will report positive 2% of the time.

   Suppose a patient walks in and tests positive for the disease. Using the maximum likelihood (ML) criterion, would we diagnose them positive?

   Solution. `Answer`: YES.

   We have the following information (D = disease, T = test)

   - $P(D=+) = 1/10^8$,
   - $P(T=+|D=+) = 0.99$,
   - $P(T=+|D=-) = 0.02$.

   Using the Bayes' Theorem we have that

$$
\begin{aligned}
P(D=+|T=+) &= \frac{P(T=+|D=+)P(D=+)}{P(T=+)} \\
&= \frac{P(T=+|D=+)P(D=+)}{P(T=+|D=+)P(D=+) + P(T=+|D=-)(1-P(D=+))} \\
&= 0.99 \times 1/10^8 + 0.02 \times (1 - 1/10^8) \approx 4.95 \times 10^{-7}
\end{aligned}
$$

   Thus,

   $P(D=-|T=+) \approx 0.9999995.$ ◀

3) Continued from Question 2:

   What if we used the maximum a posteriori (MAP) criterion?

   Solution. `Answer`: NO.

   To be continued. ◀

4) Continued from Question 2:

   Why do we see a difference between the two criteria, if there is one?

   a) The role of prior probability is different between the two.
   b) Unlike MAP, ML assumes the same model for all people.
   c) There is no difference between the two, because in this case they are equivalent.
   d) Since ML assumes a Gaussian distribution, unlike MAP, it oversimplifies the world.

   Solution. `Answer`: a). ◀