

Chapter 1

A simple neuron model

Sparse coding has been observed in neurons in the visual tract [Olshausen and Field, 2004], a simple model is proposed for such neurons.

It is unlikely that a neuron is ever truly “off”, so it is assumed that a neuron has a base-firing state that will be referred to as the “off-state”. For the sake of simplicity, it is assumed that a neuron in such an “off-state” would have a constant firing rate, λ_d . Correspondingly, for sparse-coding a neuron must have a higher firing-rate when the feature that it codes for is present; this firing rate, λ_u , is also taken to be constant.

This idea is simplified to the extreme case, where a neuron is either in its up-state and has a high firing-rate, or it is “off”. The model is treated as a poisson process, for ease of calculation.

The data that is simulated has an “up rate” λ_u and a “down rate” λ_d . For the initial trials, $\lambda_d = 0$. When in the down-state, the state switches “up” with expected frequency u , and when in the up-state it switches “down” with expected frequency d . Thus, the average background rate, r , of the

inhomogeneous poisson process can be calculated as:

$$r = \frac{\frac{\lambda_u}{d}}{\frac{1}{u} + \frac{1}{d}} = \frac{u \lambda_u}{u + d} \quad (1.1)$$

since the expected time for being in the up-state is simply $1/d$ and the expected time in the down-state is $1/u$.

1.1 Estimating the firing rate $r(t)$

With the model for the firing rate as above, it is possible to explicitly calculate the probability of being in the up-state for the time following a spike. In the initial setting, where λ_d is set to zero, then it is known when there is a spike that the rate is in the up-state, so at the time of spiking t_0 the probability $p(t_0)$ of being in the up-state is equal to one. For the estimate of the rate, $\tilde{r}(t)$, this is reflected by setting $\tilde{r}(t_0) = \lambda_u$. Then, the probability is reset every time there is a spike, so it is only required to calculate the probability of being in the up-state given that there has been no spike since the time t_0 of the last spike.

It is possible, using Baye's Theorem, to calculate a first approximation of the probability of being in the up-state at a time $t + \Delta t$, given a spike at time $t = t_0$ for small Δt .

Let X = up at $t = t_0 + \Delta t$, Y = no spike since $t = t_0$ and Z = spike at $t = t_0$. Then,

$$P(X|Y|Z) = \frac{P(Y|X|Z)P(X|Z)}{P(Y|Z)} = \frac{(1 - \lambda_u \Delta t)(1 - d \Delta t)}{1 - r \Delta t} \quad (1.2)$$

Then, it is possible to calculate the first approximation to the probability of being up at time t , by letting $t_0 = 0$, $\Delta t = t/n$.

$$\begin{aligned}
P(\text{up at } t|Y|Z) &= \lim_{n \rightarrow \infty} \prod_{k=1}^n \frac{(1 - \lambda_u t/n)(1 - dt/n)}{1 - rt/n} \\
&= \lim_{n \rightarrow \infty} \prod_{k=1}^n \left(1 + t \frac{r - \lambda_u - d}{n} \right) \\
&= \lim_{n \rightarrow \infty} \left(1 + t \frac{r - \lambda_u - d}{n} \right)^n \\
&= e^{(r - \lambda_u - d)t}
\end{aligned} \tag{1.3}$$

Recalling from equation 1.1 that $\lambda_u = r(u + d)/u$, get:

$$p = e^{-\frac{d(r+u)}{u}(t-t_0)} \tag{1.4}$$

Comparing this to the exponential kernel in the Van Rossum metric [van Rossum, 2001], then:

$$\tau = \frac{u}{d(r+u)} \tag{1.5}$$

This is clearly just the first approximation, since it does not take into account the possibility of the model switching into the down-state and back up. To calculate the full probability, let $y(t) = P(\text{up at } t|Y|Z)$, with Y, Z as in equation 1.2 above. Then:

$$\begin{aligned}
y(t+h) &= y(t)P(\text{staying in up-state}|\text{up at } t|Y|Z) \\
&\quad + (1 - y(t))P(\text{switching up}|\text{down at } t|Y|Z) \\
&= y(t) \left(\frac{(1-dh)(1-\lambda_u h)}{1-rh} \right) + (1-y(t)) \left(\frac{(uh)(1)}{1-rh} \right) \\
&= y(t) (1 + h(r - d - \lambda_u - u)) + uh
\end{aligned} \tag{1.6}$$

This leads to the following ordinary differential equation:

$$y'(t) = (r - d - u - \lambda_u) y(t) + u \quad (1.7)$$

This ODE has general solution:

$$y(t) = C e^{(r-d-u-\lambda_u)(t-t_0)} + \frac{u}{u+d+\lambda_u-r} \quad (1.8)$$

Since $y(t_0) = 1$, get:

$$y(t) = \frac{d + \lambda_u - r}{d + u + \lambda_u - r} e^{(r-d-u-\lambda_u)(t-t_0)} + \frac{u}{u+d+\lambda_u-r} \quad (1.9)$$

Letting $\lambda_u = r(u+d)/u$, get:

$$y(t) = \frac{d(u+r)}{u^2 + d(u+r)} e^{-\frac{u^2+d(u+r)}{u}(t-t_0)} + \frac{u^2}{u^2 + d(u+r)} \quad (1.10)$$

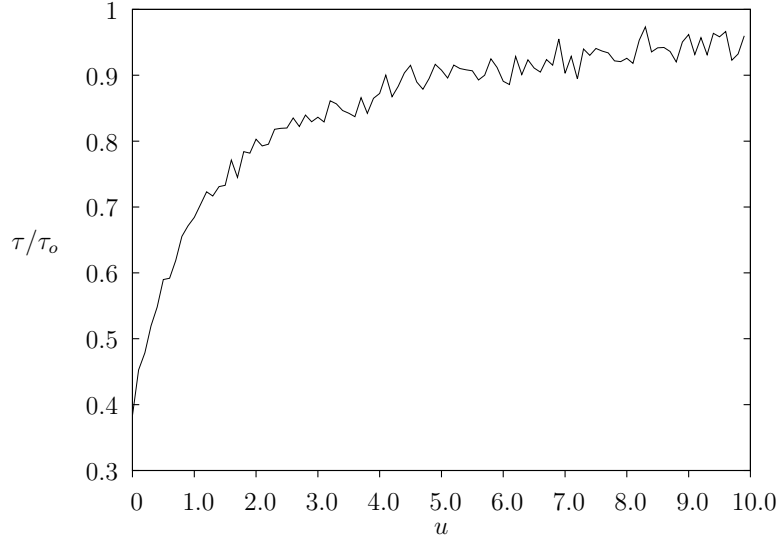


Figure 1.1: This figure shows the ratio of the predicted value for τ in an exponential model over the empirical optimised value τ_o .

This solution was tested against a standard exponential model, which was minimised over τ with a downhill simplex method. Figure 1.1 shows how the ratio of the predicted value for τ over the optimised value rises towards one as the value for u increases. This suggests that, while this method is clearly not completely correct, that it is at least a reasonable approximation.

1.2 Markov Process

The switching of the up and down states defines a Continuous-time Markov Chain (CTMC). That is, the future of the process does not depend on the past at all, merely whether the model is currently in the up-state or the down-state.

By the theory of Markov Chains, this process can be described by the simultaneous differential equations:

$$\begin{cases} p'_u(t) = -dp_u(t) + dp_d(t) \\ p'_d(t) = up_u(t) - up_d(t) \end{cases} \quad (1.11)$$

This is simply the linear ODE:

$$P'(t) = P(t)Q \quad (1.12)$$

where

$$Q = \begin{pmatrix} -d & d \\ u & -u \end{pmatrix} \quad (1.13)$$

Then the solution to the matrix $P(t)$ is simply the exponential, e^{tQ} of the transition matrix Q .

$$P(t) = e^{tQ} = \begin{pmatrix} \frac{u}{u+d} + \frac{d}{u+d}e^{-(u+d)t} & \frac{d}{u+d} - \frac{d}{u+d}e^{-(u+d)t} \\ \frac{u}{u+d} - \frac{u}{u+d}e^{-(u+d)t} & \frac{d}{u+d} + \frac{u}{u+d}e^{-(u+d)t} \end{pmatrix} \quad (1.14)$$

To calculate the probability of being in the up-state at a time t after a spike at $t = 0$, provided there has been no spike since $t = 0$, it suffices to calculate the exponential of the transition matrix Q_s below:

$$Q_s = \begin{pmatrix} -d - \lambda_u & d & \lambda_u \\ u & -u & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad (1.15)$$

The third state of the Markov chain described by the transition matrix Q_s is an “absorbing” state, the state of spiking, from which it is impossible to return to either the up or the down state.

The solution to the spiking Markov chain is:

$$P(t) = \begin{pmatrix} Ae^{-\alpha t} + (1-A)e^{-\beta t} & Be^{-\alpha t} - Be^{-\beta t} & 1 - (A+B)e^{-\alpha t} - (1-A-B)e^{-\beta t} \\ Ce^{-\alpha t} - Ce^{-\beta t} & (1-A)e^{-\alpha t} + Ae^{-\beta t} & 1 - (1-A+C)e^{-\alpha t} - (A+C)e^{-\beta t} \\ 0 & 0 & 1 \end{pmatrix} \quad (1.16)$$

where $-\alpha, -\beta$ are the roots to the characteristic polynomial of Q_s .

$$\begin{aligned} -\alpha &= \frac{-(u+d+\lambda_u) + \sqrt{(u-d-\lambda_u)^2 + 4ud}}{2} \\ -\beta &= \frac{-(u+d+\lambda_u) - \sqrt{(u-d-\lambda_u)^2 + 4ud}}{2} \end{aligned} \quad (1.17)$$

Since $u, d, \lambda_u > 0$, these probabilities are simply double exponentials, and there are no trigonometric terms.

Specifically:

$$\begin{aligned}
A &= \frac{\left(u - d - \lambda_u + \sqrt{(u - d - \lambda_u)^2 + 4ud}\right)}{2\sqrt{(u - d - \lambda_u)^2 + 4ud}} \\
B &= \frac{d}{\sqrt{(u - d - \lambda_u)^2 + 4ud}} \\
C &= \frac{u}{\sqrt{(u - d - \lambda_u)^2 + 4ud}}
\end{aligned} \tag{1.18}$$

It can be easily confirmed that $A, B, C > 0$, which is necessary to ensure that all the probabilities are between 0 and 1.

If the model is changed such that there is a non-zero spiking probability, $\lambda_d > 0$, in the down-state, then the transition matrix Q_s becomes:

$$Q_s = \begin{pmatrix} -d - \lambda_u & d & \lambda_u \\ u & -u - \lambda_d & \lambda_d \\ 0 & 0 & 0 \end{pmatrix} \tag{1.19}$$

This matrix has solution:

$$\begin{pmatrix} Ae^{-\alpha t} + (1 - A)e^{-\beta t} & B(e^{-\alpha t} - e^{-\beta t}) & 1 - e^{-\beta t} - (A + B)(e^{-\alpha t} - e^{-\beta t}) \\ C(e^{-\alpha t} - e^{-\beta t}) & (1 - A)e^{-\alpha t} + Ae^{-\beta t} & 1 - e^{-\alpha t} - (C - A)(e^{-\alpha t} - e^{-\beta t}) \\ 0 & 0 & 1 \end{pmatrix} \tag{1.20}$$

where $-\alpha, -\beta$ are once again the roots of the characteristic polynomial of Q_s .

$$\begin{aligned}
-\alpha &= \frac{-(u + d + \lambda_u + \lambda_d) + \sqrt{(u - d + \lambda_d - \lambda_u)^2 + 4ud}}{2} \\
-\beta &= \frac{-(u + d + \lambda_u + \lambda_d) - \sqrt{(u - d + \lambda_d - \lambda_u)^2 + 4ud}}{2}
\end{aligned} \tag{1.21}$$

again, $u, d > 0$ means that the probabilities above are all double exponentials, with no trigonometric terms.

$$\begin{aligned}
A &= \frac{u - d + \lambda_d - \lambda_u + \sqrt{(u - d + \lambda_d - \lambda_u)^2 + 4ud}}{2\sqrt{(u - d + \lambda_d - \lambda_u)^2 + 4ud}} \\
B &= \frac{d}{\sqrt{(u - d + \lambda_d - \lambda_u)^2 + 4ud}} \\
C &= \frac{u}{\sqrt{(u - d + \lambda_d - \lambda_u)^2 + 4ud}}
\end{aligned} \tag{1.22}$$

To simplify notation, set:

$$\begin{aligned}
\Delta &= \lambda_u - \lambda_d \\
\gamma &= \sqrt{(u - d - \Delta)^2 + 4ud}
\end{aligned} \tag{1.23}$$

Thus:

$$A = \frac{u - d - \Delta + \gamma}{2\gamma}, B = \frac{d}{\gamma}, C = \frac{u}{\gamma} \tag{1.24}$$

1.2.1 Estimating the rate function $r(t)$

These probabilities can now be used to calculate the estimated rate function, $r(t)$, based on the timing of spikes. Assuming that the probability of the last spike is known, then the rate function between any two spikes is estimated by this Markov model.

If the probability vector at the time of the last spike is $\mathbf{p}_0 = (p_0, 1 - p_0, 0)$, then the probabilities at a time t after the spike are:

$$\mathbf{p}(t) = \mathbf{p}_0 e^{Q_s t} = (p_u(t), p_d(t), p_s(t)) \tag{1.25}$$

where

$$\begin{aligned}
p_u(t) &= (p_0(A - C) + C)e^{-\alpha t} + (p_0(1 - A + C) - C)e^{-\beta t} \\
p_d(t) &= (p_0(A + B - 1) + 1 - A)e^{-\alpha t} + (p_0(-A - B) + A)e^{-\beta t} \\
p_s(t) &= 1 - p_u(t) - p_d(t)
\end{aligned} \tag{1.26}$$

The probabilities $p_u(t)$, $p_d(t)$ represent the probability of being in the up/down state and not spiking. The probability required to calculate the rate function, $r(t)$, between spikes is the probability of being in the up-state given that there has not been a spike. This is because it is known that there has been no spike since the last spike. Therefore, by Baye's Theorem, the probability $p(t)$ is just:

$$p(t) = \frac{p_u(t)}{p_u(t) + p_d(t)} \quad (1.27)$$

so, the rate function becomes:

$$r(t) = \Delta p(t) + \lambda_d = \frac{A_0 e^{-\alpha t} + B_0 e^{-\beta t}}{C_0 e^{-\alpha t} + D_0 e^{\beta t}} \quad (1.28)$$

where:

$$\begin{aligned} A_0 &= \Delta(p_0(-u - d - \Delta + \gamma) + 2u) + \lambda_d(p_0(-2\Delta) + u + d + \Delta + \gamma) \\ B_0 &= \Delta(p_0(u + d + \Delta + \gamma) - 2u) + \lambda_d(p_0(2\Delta) - u - d - \Delta + \gamma) \\ C_0 &= p_0(-2\Delta) + u + d + \Delta + \gamma \\ D_0 &= p_0(2\Delta) - u - d - \Delta + \gamma \end{aligned} \quad (1.29)$$

Since $-\beta < -\alpha$, this can be written as:

$$r(t) = \frac{A_0 + B_0 e^{-\gamma t}}{C_0 + D_0 e^{-\gamma t}} \quad (1.30)$$

so the asymptotic value of $r(t)$ is equal to the fraction A_0/C_0 , so this value should not depend on p_0 . Allowing $\delta = u + d + \Delta$, then $\delta^2 - 4u\Delta = \gamma^2$ which simplifies the calculation.

$A_0 = 2\Delta u + p_0\Delta(\gamma - \delta) + \lambda_d C_0$, and $C_0 = \gamma + \delta - 2\Delta p_0$:

$$\Delta p_0 = \frac{C_0 - \gamma - \delta}{2} \quad (1.31)$$

then

$$A_0 = 2\Delta u + \frac{1}{2}C_0(\gamma - \delta + 2\lambda_d) - \frac{1}{2}(\gamma + \delta)(\gamma - \delta) \quad (1.32)$$

Therefore:

$$\frac{A_0}{C_0} = \lambda_d + \frac{1}{2}(\gamma - \delta) \quad (1.33)$$

1.2.2 Change in probability when a spike arrives

In the previous section, the rate function, $r(t)$, and the pdf of the ISI distribution, $p_{ISI}(t)$, were calculated, given that the probability of being in the up-state at the time of the last spike, p_0 , is known. However, it is necessary to reevaluate the probability at the time a spike arrives.

If the probability of being in the up-state before the spike is $p_u(t)$, then the presence of a spike provides information on the state of the model. By Baye's Theorem:

$$\begin{aligned} P(\text{up-state}|\text{spike}) &= \frac{P(\text{spike}|\text{up-state})P(\text{up-state})}{P(\text{spike})} \\ &= \frac{\lambda_u p_u(t)}{\Delta p_u(t) + \lambda_d} \end{aligned} \quad (1.34)$$

Thus:

$$p_u(t) \rightarrow \frac{\lambda_u p_u(t)}{\Delta p(t) + \lambda_d} \quad (1.35)$$

equivalently:

$$\begin{aligned} r(t) &\rightarrow (\lambda_u - \lambda_d) \frac{\lambda_u p_u(t)}{\Delta p_u(t) + \lambda_d} + \lambda_d \\ &= \frac{\lambda_u^2 p_u(t) - \lambda_u \lambda_d p_u(t) + \lambda_u \lambda_d p_u(t) + \lambda_d^2 (1 - p_u(t))}{\lambda_u p_u(t) + \lambda_d (1 - p_u(t))} \\ &= \frac{\lambda_u^2 p_u(t) + \lambda_d^2 (1 - p_u(t))}{\lambda_u p_u(t) + \lambda_d (1 - p_u(t))} \end{aligned} \quad (1.36)$$

1.2.3 Calculating the ISI distribution from the estimated rate function

When dealing with spike-train data, it is very useful to know the inter-spike interval (ISI) distribution, as this can be observed much easier than the firing rate of a neuron.

The ISI distribution for an inhomogeneous Poisson process, with rate function $r(t)$, is:

$$p_{ISI}(t) = r(t)e^{-\int_0^t r(s) ds} \quad (1.37)$$

Now it is necessary to integrate the rate function from 0 to t :

$$\begin{aligned} \int_0^t r(s) ds &= \int_0^t \frac{A_0 e^{-\alpha s} + B_0 e^{-\beta s}}{C_0 e^{-\alpha s} + D_0 e^{\beta s}} ds \\ &= \frac{\alpha A_0 D_0 - \beta B_0 C_0}{(\alpha - \beta) C_0 D_0} t + \frac{B_0 C_0 - A_0 D_0}{(\alpha - \beta) C_0 D_0} \log \left(\frac{C_0 e^{\beta t} + D_0 e^{\alpha t}}{C_0 + D_0} \right) \end{aligned} \quad (1.38)$$

Substituting equations 1.29 and 1.23, and noting that $\beta - \alpha = \gamma$, the fractions in the above equation become:

$$\begin{aligned} \frac{\alpha A_0 D_0 - \beta B_0 C_0}{(\alpha - \beta) C_0 D_0} &= u + d + \lambda_u + \lambda_d = \alpha + \beta, \\ \frac{B_0 C_0 - A_0 D_0}{(\alpha - \beta) C_0 D_0} &= -1 \end{aligned} \quad (1.39)$$

Thus equation 1.38 becomes:

$$\begin{aligned} \int_0^t r(s) ds &= (\alpha + \beta)t - \log \left(\frac{C_0 e^{\beta t} + D_0 e^{\alpha t}}{C_0 + D_0} \right) \\ &= \log(e^{(\alpha + \beta)t}) + \log \left(\frac{C_0 + D_0}{C_0 e^{\beta t} + D_0 e^{\alpha t}} \right) \\ &= \log \left(\frac{C_0 e^{(\alpha + \beta)t} + D_0 e^{(\alpha + \beta)t}}{C_0 e^{\beta t} + D_0 e^{\alpha t}} \right) \\ &= \log \left(\frac{C_0 + D_0}{C_0 e^{-\alpha t} + D_0 e^{-\beta t}} \right) \end{aligned} \quad (1.40)$$

Now $p_{ISI}(t)$ becomes:

$$\begin{aligned}
p_{ISI}(t) &= \frac{A_0 e^{-\alpha t} + B_0 e^{-\beta t}}{C_0 e^{-\alpha t} + D_0 e^{-\beta t}} e^{-\log\left(\frac{C_0 + D_0}{C_0 e^{-\alpha t} + D_0 e^{-\beta t}}\right)} \\
&= \frac{A_0 e^{-\alpha t} + B_0 e^{-\beta t}}{C_0 e^{-\alpha t} + D_0 e^{-\beta t}} e^{\log\left(\frac{C_0 e^{-\alpha t} + D_0 e^{-\beta t}}{C_0 + D_0}\right)} \\
&= \frac{A_0 e^{-\alpha t} + B_0 e^{-\beta t}}{C_0 e^{-\alpha t} + D_0 e^{-\beta t}} \left(\frac{C_0 e^{-\alpha t} + D_0 e^{-\beta t}}{C_0 + D_0} \right) \\
&= \frac{A_0}{C_0 + D_0} e^{-\alpha t} + \frac{B_0}{C_0 + D_0} e^{-\beta t}
\end{aligned} \tag{1.41}$$

This is the probability density function (pdf) of a hyper-exponential distribution, H_2 .

1.3 Testing on data

The previous result allows for testing on real data. The data used in this report is the data used in [Sen et al., 2001]. An electrode was placed in L1 of anaesthetised zebra finches, and the birds were presented with 20 natural stimuli ten times each. As the model does not account for the refractory period of neurons, a well-documented feature of neurons **CITATION**, the ten presentations of each stimulus are overlaid to counter the impact of the refractory period.

The inter-spike intervals (ISIs) were then aggregated across 16 of the 20 stimuli, to give a representation of a typical ISI distribution for each neuron. The hyperexponential distribution, H_2 , was trained on the ISI distribution of the 16 stimuli by minimising the statistic D_n of the Kolmogorov-Smirnov test [Massey Jr, 1951] by downhill-simplex search. Then the remaining four stimuli were aggregated across the ten trials to form a test ISI distribu-

tion upon which the calculated minimum hyperexponential distribution was tested using the Kolmogorov-Smirnov test for goodness-of-fit.

The Kolmogorov-Smirnov (KS) test is a nonparametric goodness-of-fit test introduced in [Massey Jr, 1951]. It is used to compare a model probability distribution to a sample. The KS statistic is the d_∞ metric between the cumulative distribution function of the model distribution and the empirical distribution function of the data.

$$D_n = \sup_x \left| \hat{F}_n(x) - F(x) \right| \quad (1.42)$$

where $\hat{F}_n(x)$ is the empirical distribution function of the data set of n points and $F(x)$ is the cumulative distribution function of the proposed probability distribution.

Then the ISI distributions were tested against the hyperexponential distribution with both two and three modes (H_2 and H_3). As the exponential distribution is a special case of the hyperexponential distribution, and H_2 is a special case of H_3 , it is clear that there should be an improvement with each step. What was observed, as shown in figure 1.2, was that the H_2 distribution was a drastic improvement on the exponential distribution, but the the improvement of H_3 over H_2 was insignificant.

At the $p < 0.05$ significance level, only three out of 24 neurons were not rejected by the Kolmogorov-Smirnov test for both the H_2 and the H_3 distributions. This could be due to the number of data points, which would make the KS test extremely rigorous. It is also notable that this model is simply a first approximation to a neuron model; the KS statistic serves as a metric to determine goodness-of-fit rather than fitting the exact ISI distribution.

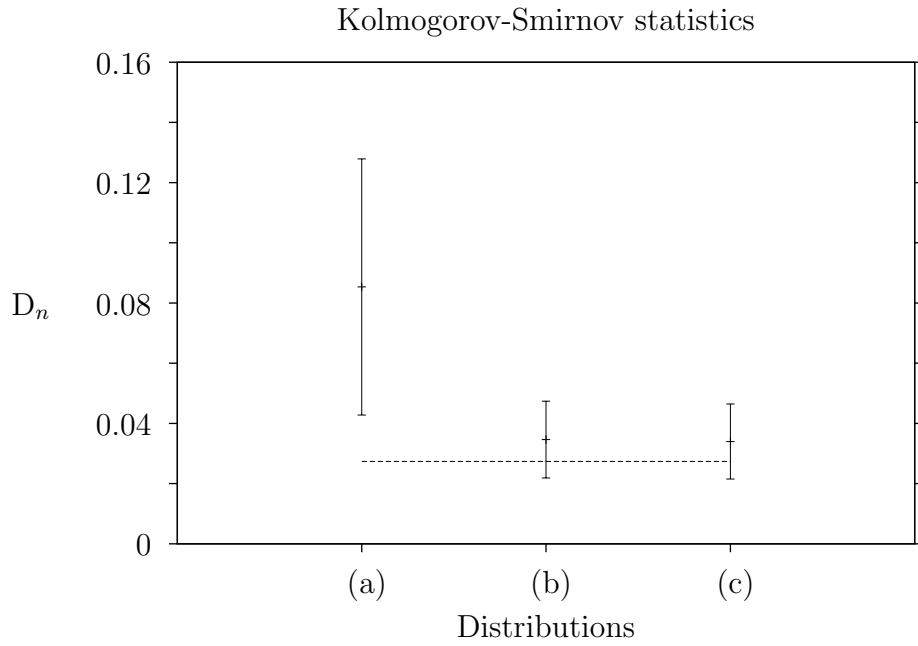


Figure 1.2: The Kolmogorov-Smirnov test statistic D_n for (a) the exponential distribution, (b) the hyperexponential distribution with two modes, and (c) the hyperexponential distribution with three modes. The mean of the $p < 0.05$ critical value is indicated by the broken line.

This seems to demonstrate that there is a bimodality of firing rates in neurons in the auditory forebrain of zebra finches. This may suggest a sparse coding of features in the auditory forebrain of songbirds.

Bibliography

- [Massey Jr, 1951] Massey Jr, F. J. (1951). The kolmogorov-smirnov test for goodness of fit. *Journal of the American statistical Association*, 46(253):68–78.
- [Olshausen and Field, 2004] Olshausen, B. A. and Field, D. J. (2004). Sparse coding of sensory inputs. *Current Opinion in Neurobiology*, 14:481–487.
- [Sen et al., 2001] Sen, K., Theunissen, F. E., and Doupe, A. J. (2001). Feature analysis of natural sounds in the songbird auditory forebrain. *Journal of Neurophysiology*, 86:1445–1458.
- [van Rossum, 2001] van Rossum, M. (2001). A novel spike distance. *Neural Computation*, 13(4):751–763.