

CALCULATION OF THE KULLBACK-LEIBLER DISTANCE BETWEEN POINT PROCESS MODELS

*Charlotte M. Gruner**
Pronghorn Engineering
P.O. Box 1913
Casper, WY 82602-1913

Don H. Johnson
Rice University
Electrical & Computer Engineering Dept.
Houston, TX 77251-1892

ABSTRACT

We have developed a method for quantifying neural response changes in terms of the Kullback-Leibler distance between the intensity functions for each stimulus condition. We use empirical histogram estimates to characterize the intensity function of the neural response. A critical factor in determining the histogram estimates is selection of binwidth. In this work we analytically derive the Kullback-Leibler distance between two Poisson processes and two dead time modified Poisson processes in terms of the binwidth selected. Our results show that for constant intensity processes having the same number of expected counts, the distance between the dead time modified processes will be larger than between the Poisson processes.

1. INTRODUCTION

It is believed that neurons, the cells of the nervous system, code and process information in the times of occurrences of action potentials or “spikes,” the electrical signals they produce (see Figure 1). Researchers seek to understand the mechanisms of neural information coding. A key observation is that the pattern of action potentials elicited by a neuron (or group of neurons) to a given stimulus condition is stochastic, not deterministic. It is generally agreed that a good model of an action potential firing pattern is a point process with some unknown, probably complex, intensity description. A cornerstone of learning about neural coding is discovering stimulus properties that result in a “change” to a neural response. Quantifying a change in a neural response means showing a change in the intensity description of the neural response to different stimulus conditions.

We have developed methodology for quantifying neural response changes in terms of information-theoretic distances between the empirical intensity descriptions[1]. Our empirical intensity descriptions are empirical histogram estimates (or “types”) describing the probability of spike firing in a given time period or “bin” of the response. In this

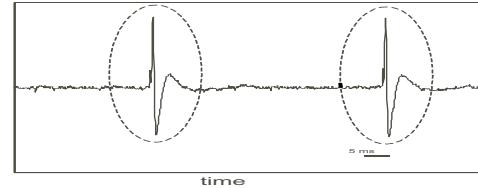


Fig. 1. This voltage trace is an example of an extracellular recording from a neuron in the auditory pathway of a cat. The dotted ovals show the occurrences of two action potentials.

work we present analytical results related to the selection of the binwidth for this procedure.

2. METHODS

We have developed a method for computing the Kullback-Leibler and Chernoff distances between empirically calculated intensity descriptions. Our method is outlined in Figure 2. Briefly, we use data gathered from an experimental paradigm in which two (or more) stimuli are repeatedly presented to a neuron. From the recorded spike patterns to each stimulus presentation an estimate of the probability of a spike happening in a given bin can be estimated. The Kullback-Leibler distance or Chernoff distance can be found between these empirical estimates. Since the probability of a spike in a given bin can depend on previous activity, we can also expand the histogram estimate to include dependency on previous bins (Figure 3). In expanding the “alphabet” of the histogram we obtain a more complete picture of the spike train’s statistical structure. This distance-based technique has several advantages:

- Quantifies how different responses are in terms of distances meaningful in classification theory.
- The technique can be expanded to measure distances between neural ensemble responses.
- Confidence bounds can be found using bootstrap resampling techniques.

*Work performed while at Rice University.

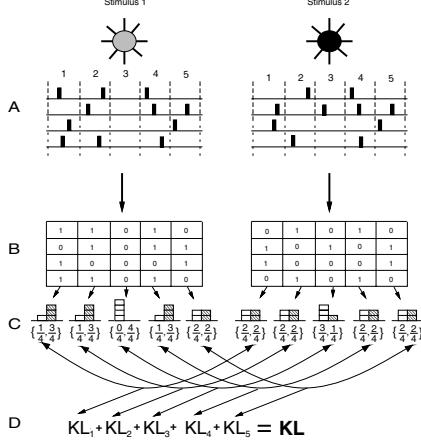


Fig. 2. Computing the Kullback-Leibler (KL) distance between neural responses. (A) A single neuron’s response to four presentations of each of two stimuli. Each vertical bar represents the occurrence of an action potential. The records are binned along the time axis as shown by the dashed vertical bars. (B) A table is formed by counting the number of spikes in each bin. (C) A histogram is formed for each bin and normalized by the number of stimulus repetitions. The KL distance between each empirical probability mass function is computed and the total KL distance for the time interval of interest is the sum of the individual estimates for each bin. The Kullback Leibler distance is defined as[2] $\sum_{x \in X} p(x) \log \frac{p(x)}{q(x)}$.

- Using a relationship between the Kullback-Leibler distance and the Fisher Information, it is possible in some cases, to measure how well information is coded about particular stimulus parameters in terms of the Cramér-Rao upper bound on expected mean square error.

In this work we analytically compute the Kullback-Leibler distance between responses characterized as Poisson processes and between responses characterized as dead time modified Poisson processes.

3. RESULTS

3.1. Kullback-Leibler distance between Poisson processes

The Kullback-Leibler distance between two Poisson processes depends strongly on the chosen evaluation binwidths. Consider two Poisson processes with time-varying intensities $\lambda_p(t)$ and $\lambda_q(t)$ compared over a time interval T . It is easily shown that the distance between the two processes evaluated over the entire interval T is less than the summed

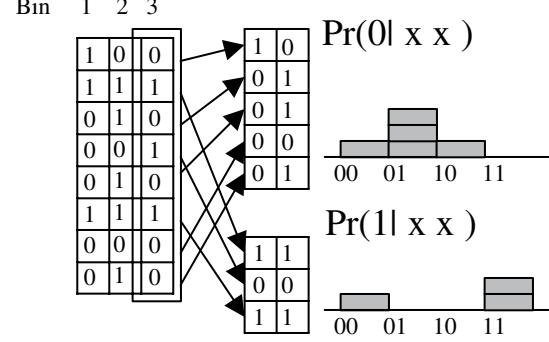


Fig. 3. Computing the probability of a spike in bin 3 given the activity patterns in bins 1 and 2.

distances computed over non overlapping subintervals of T , $\{\tau_i\}$, where $T = \bigcup_i \tau_i$, and $\bigcap_i \tau_i = \emptyset$:

$$D(p_T(x)||q_T(x)) \leq \sum_i D(p_{\tau_i}(x)||q_{\tau_i}(x)),$$

where

$$p_{\tau}(x) = \frac{\left(\int_{\tau} \lambda_p(t) dt\right)^x}{x!} \exp\left\{-\int_{\tau} \lambda_p(t) dt\right\},$$

with equality if and only if $\int_{\tau_i} \lambda_p(t) dt = c \int_{\tau_i} \lambda_q(t) dt$, for all $\{\tau_i\}$ where c is a constant. Thus, a smaller binwidth over an interval T increases the calculated Kullback-Leibler distance. However, if the equality condition holds, as in the case of constant rates, $\lambda_p(t) = \lambda_p$, $\lambda_q(t) = \lambda_q$, for a particular binwidth, then smaller binwidths will not further increase the Kullback-Leibler distance.

3.2. Kullback-Leibler distance between refractory processes

An interval between events of a Poisson process can be infinitely small. Because of the observed dead-time between action potentials, the Poisson process is not a realistic approximation of neural spiking behavior. The simplest refractory model adds a dead-time after each spike, during which the probability of another spike occurring is zero. After the dead-time, the process is essentially Poisson; the inter-event interval density is $p(\tau) = \lambda \exp(-\lambda(\tau - \Delta))$, $\tau > \Delta$, where Δ is the dead-time. Including a dead-time complicates analysis because time increments are no longer independent: the probability of an event during a particular time period is dependent on when the last event occurred. Thus, the probability of an event during a particular time period could depend on the entire process history. To examine how the addition of a refractory period affects the Kullback-Leibler distance be-

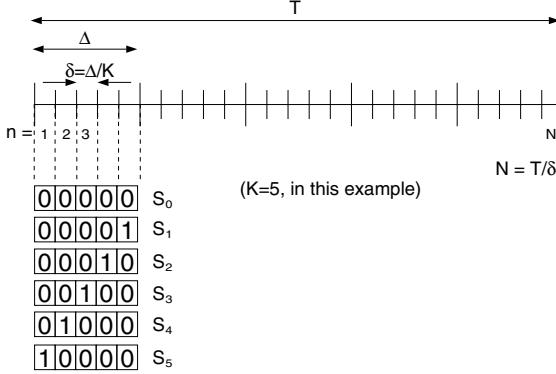


Fig. 4. An observation time T of a refractory process with dead-time Δ is broken up into bins of time duration $\delta = \frac{\Delta}{K}$, where $K = 5$ in this example. In each Δ -length time interval there are $K + 1$ possible states as shown.

tween responses this section develops an analysis framework for analytically computing the Kullback-Leibler distance between refractory processes.

If the refractory process has a dead-time Δ , during any time interval of length Δ no more than one event can occur. If the interval Δ is divided into K subintervals, a total of $K + 1$ possible “states” describe the possible occurrence of an event in the given time interval Δ . An example of the possible states, $\{s_j\}$, for each Δ increment are shown in the figure 4 for $K = 5$.

The transition matrix between states is given by

$$\mathbf{T}(n+1) = \begin{bmatrix} 1 - p(n+1) & 0 & 0 & \dots & 0 & 1 \\ p(n+1) & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ \vdots & & & & & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 \end{bmatrix},$$

where each matrix element is the conditional probability of moving from one state at time n to another in time interval $n+1$, $T_{ij}(n+1) = \Pr(s_i(n+1)|s_j(n))$. Most of the entries are either “1” or “0”. Because of the dead-time, state 1 at time n always implies state 2 at time $n+1$; state 2 at time n always implies state 3 at time $n+1$ etc. In fact, only if the state at time n is 0 is it possible for the state at time $n+1$ to be either 0 or 1, depending on the probability of an event in the time period $n+1$, $p(n+1)$. The transition matrix can be used to find the probability of the next state:

$$\mathbf{T}(n+1)\mathbf{S}(n) = \mathbf{S}(n+1),$$

where

$$\mathbf{S}(\mathbf{n}) = \begin{bmatrix} p(s_0(n)) \\ p(s_1(n)) \\ p(s_3(n)) \\ \vdots \\ p(s_K(n)) \end{bmatrix}.$$

Given initial state probabilities, the transition matrices can be used to find the probability of any state at any time by

$$\begin{aligned} \mathbf{S}(1) &= \mathbf{T}(1)\mathbf{S}(0) \\ \mathbf{S}(2) &= \mathbf{T}(2)\mathbf{S}(1) = \mathbf{T}(2)\mathbf{T}(1)\mathbf{S}(0) \\ &\vdots \\ \mathbf{S}(n) &= \left(\prod_{i=1}^n \mathbf{T}(i) \right) \mathbf{S}(0). \end{aligned}$$

Using the Markov state construction of the problem it is possible to calculate the Kullback-Leibler distance between two refractory processes. Let x_i be the value of the i th subinterval, $x_i \in \{0, 1\}$. The goal is to find the Kullback-Leibler distance $D(p_1(x_0, x_1, x_2, \dots, x_n) || p_2(x_0, x_1, x_2, \dots, x_n))$. Because of the refractory period lasting for K bins, the value of x_i only depends on the values of x at the previous K bins. The Kullback-Leibler distance between two refractory processes simplifies to

$$\begin{aligned} D(p_1(x_1, x_2, \dots, x_N) || p_2(x_1, x_2, \dots, x_N)) &= \\ &= D(p_1(\mathbf{s}(0)) || p_2(\mathbf{s}(0))) + \\ &\quad \sum_{i=1}^{N-k} D(p_1(x_i | \mathbf{s}(i-1)) || p_2(x_i | \mathbf{s}(i-1))). \end{aligned} \quad (1)$$

Additional simplifications to the Kullback-Leibler formula occur because all the transition probabilities are either 0 or 1 except those associated with state s_0 :

$$\begin{aligned} p_1(x_i = 1 | s_l(i-1), l \neq 0) &= \\ p_2(x_i = 1 | s_l(i-1), l \neq 0) &= 0 \\ p_1(x_i = 0 | s_l(i-1), l \neq 0) &= \\ p_2(x_i = 0 | s_l(i-1), l \neq 0) &= 1. \end{aligned}$$

Equation (1) simplifies to

$$\begin{aligned} D(p_1(x_1, x_2, \dots, x_N) || p_2(x_1, x_2, \dots, x_N)) &= \\ &= D(p_1(\mathbf{s}(0)) || p_2(\mathbf{s}(0))) + \\ &\quad \sum_{i=1}^{N-k} D(p_1(x_i | \mathbf{s}(i-1) = s_0) || p_2(x_i | \mathbf{s}(i-1) = s_0)) \end{aligned} \quad (2)$$

If the rate is assumed constant $\lambda(t) = \lambda$, then $p(n+1) \approx \lambda \frac{\Delta}{K}$ and the set of state probabilities $\mathbf{S}(n)$ will tend to limits π as $n \rightarrow \infty$ [3]. At equilibrium the following equation holds,

$$\begin{bmatrix} 1 - \lambda \frac{\Delta}{K} & 0 & 0 & \dots & 0 & 1 \\ \lambda \frac{\Delta}{K} & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ \vdots & & & & \vdots & \\ 0 & 0 & 0 & \dots & 1 & 0 \end{bmatrix} \begin{bmatrix} \pi_o \\ \pi_1 \\ \pi_2 \\ \vdots \\ \pi_K \end{bmatrix} = \begin{bmatrix} \pi_o \\ \pi_1 \\ \pi_2 \\ \vdots \\ \pi_K \end{bmatrix}$$

Solving for π :

$$\pi_0 = \frac{1}{(1 + \lambda \Delta)} \quad (3)$$

$$\pi_1 = \dots = \pi_K = \frac{\lambda \frac{\Delta}{K}}{(1 + \lambda \Delta)}. \quad (4)$$

Using the equilibrium state probabilities from equations (3) and (4) in the equation for the Kullback-Leibler distance in equation (2),

$$\begin{aligned} D(p_1(x_1, \dots, x_N) || p_2(x_1, \dots, x_N)) = & \log \frac{\lambda_2 \Delta + 1}{\lambda_1 \Delta + 1} + \frac{\lambda_1 \Delta}{\lambda_1 \Delta + 1} \log \frac{\lambda_1}{\lambda_2} \\ & + \frac{K(\frac{T}{\Delta} - 1)}{\lambda_1 \Delta + 1} \left[\lambda_1 \frac{\Delta}{K} \log \frac{\lambda_1}{\lambda_2} \right. \\ & \left. + (1 - \lambda_1 \frac{\Delta}{K}) \log \frac{(1 - \lambda_1 \frac{\Delta}{K})}{(1 - \lambda_2 \frac{\Delta}{K})} \right]. \quad (5) \end{aligned}$$

Taking the limit as $K \rightarrow \infty$ yields $D_{RP}(\lambda_1, \lambda_2, T)$,

$$\begin{aligned} \lim_{K \rightarrow \infty} D(p_1(x_1, \dots, x_N) || p_2(x_1, \dots, x_N)) & = \log \frac{\lambda_2 \Delta + 1}{\lambda_1 \Delta + 1} + \frac{\lambda_1 T}{\lambda_1 \Delta + 1} \log \frac{\lambda_1}{\lambda_2} + \\ & \quad \frac{(\frac{T}{\Delta} - 1)}{\lambda_1 \Delta + 1} (-\lambda_1 + \lambda_2) \Delta. \quad (6) \end{aligned}$$

3.3. Comparison between refractory Kullback-Leibler and Poisson Kullback-Leibler of the same virtual rates

Consider two Poisson processes with rates $\frac{\lambda_1}{\lambda_1 \Delta + 1}, \frac{\lambda_2}{\lambda_2 \Delta + 1}$. The Kullback-Leibler distance between these Poisson processes is

$$D_{VPP}(\lambda_1, \lambda_2, T) = \frac{T}{(1 + \lambda_1 \Delta)} \left[\lambda_1 \log \frac{\lambda_1}{\lambda_2} + \lambda_1 \log \frac{\lambda_2 \Delta + 1}{\lambda_1 \Delta + 1} + \frac{\lambda_2 - \lambda_1}{\lambda_2 \Delta + 1} \right]$$

These processes have the same expected number of events in an interval as the dead-time modified Poisson processes

with rates λ_1 and λ_2 with dead-time Δ considered previously. Subtracting D_{VPP} from D_{RP}

$$\begin{aligned} D_{RP} - D_{VPP} = & \frac{T}{(1 + \lambda_1 \Delta)} \left[(\lambda_2 - \lambda_1) - \lambda_1 \log \frac{\lambda_2 \Delta + 1}{\lambda_1 \Delta + 1} - \frac{(\lambda_2 - \lambda_1)}{(\lambda_2 \Delta + 1)} \right] \\ & + \log \frac{(1 + \lambda_2 \Delta)}{(1 + \lambda_1 \Delta)} - \frac{(\lambda_2 - \lambda_1) \Delta}{1 + \lambda_1 \Delta}. \quad (7) \end{aligned}$$

Approximating $\log(\lambda \Delta + 1) = \lambda \Delta$ (assuming $\lambda \Delta$ is small),

$$\begin{aligned} D_{RP} - D_{VPP} & \approx \frac{T}{(1 + \lambda_1 \Delta)} \left[\frac{(\lambda_2^2 \Delta - \lambda_1 \lambda_2 \Delta)}{(\lambda_2 \Delta + 1)} - \lambda_1 \lambda_2 \Delta + \lambda_1^2 \Delta \right] \\ & + \left[\lambda_2 \Delta - \lambda_1 \Delta - \frac{(\lambda_2 - \lambda_1) \Delta}{(1 + \lambda_1 \Delta)} \right]. \end{aligned}$$

Assuming that $\frac{1}{1 + \lambda \Delta} \approx 1$, the second bracketed term goes to 0 and the first term goes to

$$D_{RP} - D_{VPP} \approx T \frac{(\lambda_2 - \lambda_1)^2 \Delta}{(1 + \lambda_1 \Delta)},$$

which is always ≥ 0 , independent of the choice of λ_1, λ_2 , thus, the distance between two dead-time modified Poisson processes is greater than the distance between Poisson processes of the same virtual rates.

4. CONCLUSIONS

We have shown that the Kullback-Leibler distance between two Poisson processes is larger when the interval on which counts are compared is broken up into many subintervals. This analysis indicates that smaller binwidths increase the Kullback-Leibler distance. We have also shown that the Kullback-Leibler distance between dead-time modified Poisson processes is greater than the distance between comparable Poisson processes. This result indicates that the temporal correlation introduced by the refractory process increases the potential discriminability of the two responses, provided that the process data are compared with appropriate binwidth and order to “see” the temporal correlation.

5. REFERENCES

- [1] D. H. Johnson, C. M. Gruner, K. Baggerly, and C. Se-shagiri. Information-theoretic analysis of neural coding. To appear in *J. Comp. Neuroscience*, 2001.
- [2] T. M. Cover and J. A. Thomas. *Elements of Information Theory* John Wiley & Sons, New York, 1991.
- [3] D. R. Cox and H. D. Miller. *The Theory of Stochastic Processes*. Chapman & Hall, London, 1965.