CHAPTER 1

Non-trivial Model Properties

The model we have created has been analysed with methods well-known in the field of statistics. However, using these methods to our model brings unique issues/decisions. In this chapter we will discuss these complications and explain the decisions we make going forward.

1.1 ISI Distribution

In this section we discuss the importance of the inter-spike interval (ISI) distribution. In Chapter 2 we gave an example whereby we created a time-dependent ISI distribution by extending the one dimensional Gamma(γ, γ) distribution. The extension is performed by applying a time-rescaling method via the intensity function $x(t)$ which accounts for the time-dependence of Ca^{2+} spiking. This approach can be used for all probability distributions defined on $(0, \infty)$, such as an Inverse Gaussian or Weibull probability distributions. We will demonstrate that the choice of the original distribution is crucial to how well the model fits the Ca^{2+} data. Furthermore, we will explain how to create a time-dependent ISI distribution from any probability

distribution on $(0, \infty)$. Care is required in constructing these inhomogeneous ISI distributions to avoid non-identifiability issues. We will show how to circumvent this problem whilst also providing biological meaning to the intensity function used in creating these models.

To begin, let us first consider the case of stationary Ca^{2+} spikes — spikes that depend only on the time since the last spike and not the time of the experiment. Hence if a Ca^{2+} spike occurs at time s then the probability that the next spike occurs at time $t > s$ only depends on the time between spikes $\tau = t - s$. Thus, we can convert a spike sequence $\{y_i\}_{i=1}^{N}$ into a sequence of ISIs $\{\tau_i\}_{i=1}^{M}$, where $\tau_i = y_{i+1} - y_i$ and $M = N - 1$. Note that the time until the first spike and the time from the last spike to the end of the experiment do not correspond to ISIs so we ignore them. Thus we can treat each τ_i as an observation of an unknown random variable Y which describes the stationary ISI distribution. Since the ISI times can only take positive values the sample space of Y is $(0, \infty)$.

Before we can use observed ISI times to estimate the ISI distribution we need to give structure to the ISI distribution. One method is to assume that Y comes from a known family of probability distributions, such as the Exponential or Gamma distributions. However, it is not obvious which family of distributions will best describe the ISI dynamics. If we assume that Y comes from, say the Exponential distribution, we call this the Exponential model for the ISI distribution. Thus, we shall test several different models to find which describes the Ca^{2+} ISIs the best. The models we consider are: the Exponential, Gamma, Weibull, Inverse Gaussian and Log Normal distributions.

We consider the Exponential distribution because is the simplest distribution of $(0, \infty)$, and it acts as a useful base case to compare more complex ISI models. Moreover, the Exponential distribution is the only distribution on $(0, \infty)$ with the memoryless property [1]. This means that if a spike occurs at time s the probability of the next spike occurring at time t is the same as the probability of the next spike occurring at time $t + r$ given that no spike occurs in (s, r). Thus, including this model allows us to test whether the time since the last spike influences the probability of

the next spike occurring. The Gamma distribution offers a natural extension to the Exponential distribution, since it contains the Exponential distribution as a special case. The Gamma distribution is an appealing candidate for the ISI distribution because it is a probability distribution for a combination of events occurring for the first time. Therefore, each event could be interpreted as a Ca^{2+} puff, and each spike occurs the first time that α puffs occur [2]. The Weibull distribution can also be viewed as an extension to the Exponential distribution. The Weibull distribution is often used to describe the "time to failure" of some component where the failure rate is proportional to a power of time [3]. For this reason, the Weibull distribution is justifiable if we believe Ca^{2+} spikes depend on a power of time, i.e. if a spike is more likely to occur as time goes on. Conceptually, Ca^{2+} spikes have been described as first passage events [4]. Thus, since the Inverse Gaussian has been used to model first passage problems with Brownian motion [5] it appears as a suitable candidate. Moreover, it has been shown that the ISI distribution of several neuronal systems under approximately stationary conditions are unimodal and right skewed [6]. This supports the Log Normal distribution since it is unimodal and right-skewed. Furthermore, this also gives support to the Gamma, Inverse Gaussian and Weibull distributions.

We now define each probability distribution, in terms of a general random variable Z. The random variable Z is said to have an Exponential distribution with rate $\alpha > 0$, shown as $Z \sim \text{Exponential}(\alpha)$, if its probability density function is given by

$$f_Z^{\text{ex}}(z;\alpha) = \alpha e^{-\alpha z}. \tag{1.1}$$

Z is said to have a Gamma distribution with shape $\alpha > 0$ and rate $\beta > 0$, shown as $Z \sim \text{Gamma}(\alpha, \beta)$, if its probability density function is given by

$$f_Z^{\text{G}}(z;\alpha,\beta) = \frac{\beta^\alpha z^{\alpha-1} e^{-\beta z}}{\Gamma(\alpha)}, \tag{1.2}$$

where $\Gamma(\alpha)$ is the gamma function $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1}e^{-x}\mathrm{d}x$. Z is said to have an Inverse Gaussian distribution with mean $\mu > 0$ and shape $\lambda > 0$, shown as $Z \sim \text{IG}(\mu, \lambda)$, if

its probability density function is given by

$$f_Z^{\mathrm{IG}}(z;\mu,\lambda) = \sqrt{\frac{\lambda}{2\pi z^3}} \exp\left\{-\frac{\lambda(z-\mu)^2}{2\mu^2 z}\right\}. \tag{1.3}$$

Z is said to have a Log Normal distribution with parameters $\mu \in (-\infty, \infty)$ and $\sigma^2 > 0$, shown as $Z \sim \mathrm{Lognormal}(\mu, \sigma^2)$, if its probability density function is given by

$$f_Z^{\mathrm{LN}}(z;\mu,\sigma^2) = \frac{1}{z\sigma\sqrt{2\pi}} \exp\left\{-\frac{(z-\mu)^2}{2\sigma^2}\right\}. \tag{1.4}$$

Z is said to have a Weibull distribution with scale $\lambda > 0$ and shape $k > 0$, shown as $Z \sim \mathrm{Weibull}(\lambda, k)$, if its probability density function is given by

$$f_Z^{\mathrm{W}}(z;\lambda,k) = \frac{k}{\lambda}\left(\frac{z}{\lambda}\right)^{k-1} e^{-(z/\lambda)^k}. \tag{1.5}$$

Now that we have defined our models for the ISI distribution we want to fit them to observed ISIs and also find which model describes the ISI distribution best. Suppose Y comes from the family of probability distributions parameterised by the set θ with probability distribution given by $f(z|\theta)$. To fit the model we need to estimate the value of θ that best describes the ISI times $\boldsymbol{\tau} = \{\tau_i\}_{i=1}^M$. One such method is maximum likelihood estimation (MLE). This involves maximising the likelihood function such that under the assumed model the ISI times are most probable. The likelihood function $L(\theta)$ is the joint probability density of observing all the ISI times. Since we are considering stationary Ca^{2+} spikes, the joint probability function reduces to the product of the probabilities of individual ISI times giving

$$L(\theta) = p(\tau_1, \tau_2, \ldots, \tau_M|\theta) = \prod_{i=1}^{M} f(\tau_i|\theta).$$

The maximal likelihood estimate $\hat{\theta}$ is the value of θ that maximises the likelihood

$$\hat{\theta} = \underset{\theta\in\Theta}{\mathrm{argmax}} \left\{\prod_{i=1}^{M} f(\tau_i|\theta)\right\},$$

where Θ is the space of all possible parameter values. In the case of the Expo nential, Inverse Gaussian and Log Normal analytical solutions for the MLE exist. Namely,

with observed ISI times $\boldsymbol{\tau}$, the MLE when $Y \sim \text{Exp}(\alpha)$ is $\hat{\alpha} = M/\sum_{i=1}^{M} \tau_i$, the inverse of the sample mean. If $Y \sim \text{IG}(\mu, \lambda)$ the MLE occurs at

$$\hat{\mu} = \frac{1}{M}\sum_{i=1}^{M} \tau_i \qquad \text{and} \qquad \frac{1}{\hat{\lambda}} = \frac{1}{M}\sum_{i=1}^{M}\left(\frac{1}{\tau_i} - \frac{1}{\hat{\mu}}\right).$$

Finally, if $Y \sim \text{Lognormal}(\mu, \sigma^2)$ the MLE occurs at

$$\hat{\mu} = \frac{1}{M}\sum_{i=1}^{M} \log \tau_i \qquad \text{and} \qquad \hat{\sigma}^2 = \frac{1}{M}\sum_{i=1}^{M} (\log \tau_i - \hat{\mu})^2$$

However, a closed form solution of the MLE does not exist for either the Gamma or Weibull distributions and thus numerical methods are required to calculate the MLE. We use the `optim` function from the `stats` package in the language R.

To visualise the difference between the five stationary ISI distribution models we fit each model to the same spike sequence generated by a HEK293 cell challenged with 10μM Carbachol. The cell exhibited 66 Ca^{2+} spikes and we convert the spikes into a sequence of ISI times — shown as the black ticks on the x-axis of Figure 1.1. Subsequently, we calculate the MLE for each of the distributions. The results are shown in the second column of Table 1.1. We compare the models by examining the log likelihood at the MLE, shown as the third column in Table 1.1. The log likelihood will give an indication into which models perform better, where superior models will have a higher log likelihood than other models. The method is used as a simple tool to illustrate the fact the some models better suit the Ca^{2+} spike data. Other more powerful tests could be done to compare the different models such as the Kolmogorov-Smirnov test [7] or Anderson-Darling test [8]. We find that the Exponential is the worst model having a log likelihood significantly lower than the other models. The Inverse Gaussian model has the largest log likelihood, implying it fits the spike sequence the best. However, the log likelihood of the Gamma is similar, thus either model could be used to describe the spike sequence.

To explain this difference in performance we plot the probability density function of the MLE for each model. This is shown in Figure 1.1 where dark green, orange, pink, light green and blue correspond to the Exponential, Gamma, Inverse Gaussian, Log Normal and Weibull distributions respectively. We see that the largest proportion

Distribution	MLE of ISI parameter/s	Log Likelihood
Exponential	$\alpha = 0.0221$	-312.75
Gamma	$(\alpha, \beta) = (0.229, 10.35)$	-261.86
Inverse Gaussian	$(\lambda, \mu) = (442.13, 45.22)$	-261.06
Log Normal	$(\mu, \sigma) = (3.762, 0.312)$	-268.78
Weibull	$(k, \lambda) = (3.226, 50.41)$	-266.48

Table 1.1: The MLE and log likelihood of the MLE of each distribution fit to Ca^{2+} ISIs from a HEK293 cell challenged with 10μM Carbachol. The colour of the distribution matches the colour of the probability densities shown in Figure 1.1.

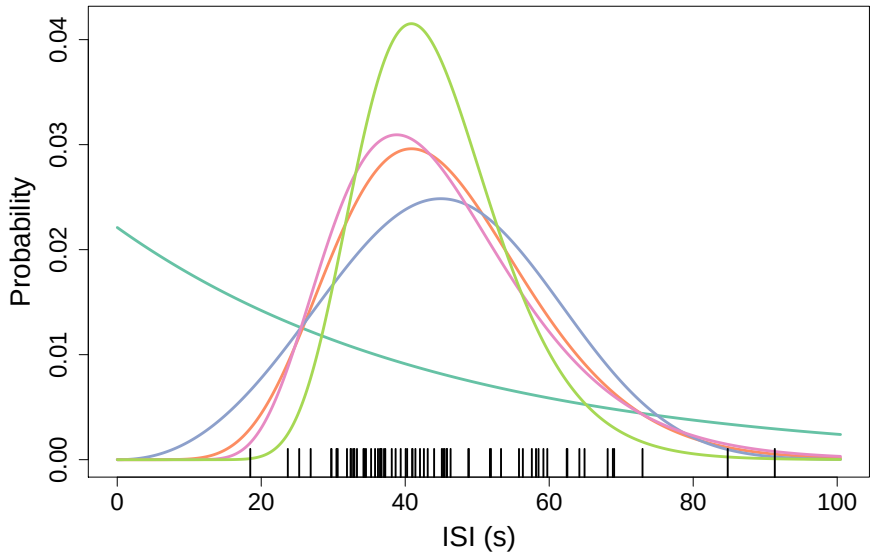

Figure 1.1: Probability density functions corresponding to the MLE of: an Exponential (dark green), a Gamma (orange), Inverse Gaussian (pink), Log Normal (light green) or Weibull Distribution (blue) fit to ISIs — shown as black ticks on the x-axis.

of spike times are centered around 40s, which is mirrored by the probability densities of all the distributions bar the Exponential. The Exponential model does not capture this due to the limitations of its shape, i.e. the mode of the exponential is zero. We see that the Gamma and Inverse Gaussian have similar probability density functions — the pink and orange lines — and hence similar likelihoods. We also find that the Log Normal and Weibull have similar centers as the Gamma and Inverse Gaussian but with narrower and wider tails respectively. Zooming in on the region between 0s and 20s we see a single ISI just before 20s. In this region we see that the Weibull's probability density is too high — we would expect more ISIs lower the 20s — and the Log Normal's probability density is too low — the density at the lowest ISI time is almost zero. This illustrates why the Log Normal and Weibull do not fit the ISIs as well as the Gamma and Inverse Gaussian models.

The above example demonstrates that the choice of ISI distribution is crucial, since we found that some distributions fit the Ca^{2+} ISIs better than other distributions. This is due to the fact that shape of the probability density function varies for each family of probability distributions. Thus, when considering the ISI distribution for non-stationary spike sequences we need to consider several distributions as some may describe the Ca^{2+} spikes better than others.

At this stage, one could argue to use a non-parametric method to estimate the ISI distribution, such as using splines or a Gaussian Process prior. This would indeed be a more flexible framework since the ISI distribution would not be limited to the shape of the chosen probability model. However, we have chosen to show this approach as our model for time-dependent Ca^{2+} spikes generalises these ISI distributions. If we began by using a non-parametric method to model the stationary spike sequences it would be unclear how to generalise to a time-dependent model. For example, as each ISI does not necessarily have the same statistics, it is not possible to find a 'common' ISI distribution that describes all ISIs. One way to visualise this is that in the stationary setting the 'true' ISI distribution is a fixed function, whereas generalising to time-dependent spikes the ISI distribution varies as the experiment progresses.

Now that we have shown the importance of the ISI distribution in the stationary case we want to generalise each distribution to allow for time-dependence. We use the approach of Barbieri et al [6] which uses a time rescaling transformation. This approach has was also used by Tilunaite et al [9] to model Ca^{2+} spikes. However, both only apply the method to one-parameter versions of the Gamma and Inverse Gaussian distributions. In particular, they generalise a Gamma(γ, γ) for $\gamma > 0$ and an IG$(\mu, 1)$ where $\mu > 0$. It is unclear why these particular one-parameter versions of the Gamma and Inverse Gaussian are taken as the initial stationary distribution. For example rather than rescaling IG$(\mu, 1)$ where $\mu > 0$ another option could be to rescale IG$(1, \lambda)$ where $\lambda > 0$. Moreover, it is not obvious why you cannot time rescale the full Inverse Gaussian IG(μ, λ).

We begin by deriving the time dependent ISI distribution from a general stationary ISI distribution. Recall that the random variable Y with probability density function $f(z|\theta)$ describes the stationary ISI distribution, where z represents a ISI time. Since we are now concerned with a time-dependent model of the ISI distribution we need to explicitly state which ISI we are modelling. Consider the kth spike y_k that occurs at time s. The probability density of the next spike occurring at time $t > s$ is $f(t-s|\theta)$, in other words we have $z = t - s$. Thus, the random variable Y_s describing the stationary ISI distribution of the next spike occurring after y_k has probability density function $f(z|\theta)$ on support $s < z < \infty$. To account for the time dependence we map the original time (s, ∞) to a new time via the one-to-one differentiable transformation $v(z) = \int_s^z x(u)du$, where x is a positive function that accounts for the time dependence of the ISI times. Therefore, by applying the change of variables formula we create a new random variable W describing the time-dependent ISI distribution with pdf $g(t|x, \theta)$ and support $0 < t < \infty$, where

$$\begin{aligned} g(t|x,\theta) =& |v'(t)| \times f(v(t)|\theta), \\ =& x(t) f(\int_s^t x(u)du|\theta), \\ =& x(t) f(X(s,t)|\theta), \end{aligned}$$

where $X(s,t) = \int_s^t x(u)du$. Note that the above result holds true for any time s of

the spike y_k. Thus the probability density of the next spike occurring at time t given the last spike occurred at time s is given by

$$g(t, s|x, \theta) = x(t) f(X(s,t)|\theta). \tag{1.6}$$

We can now apply the transformation to the stationary ISI distributions defined in (1.1) to (1.5) to construct inhomogeneous versions of the Exponential, Gamma, Inverse Gaussian, Log Normal and Weibull distributions.. For example by substituting (1.2) into (1.6) we construct the inhomogeneous Gamma, where the probability density that the next spike occurs at time t given the previous spike occurred at time t is given by

$$\begin{aligned} g_{\text{GAM}}(t, s|x, \alpha, \beta) &= x(t) f_{\text{GAM}}(X(s,t)|\alpha, \beta), \\ &= \frac{\beta x(t)}{\Gamma(\alpha)} \left[\beta X(s,t)\right]^{\alpha-1} \exp\left\{-\beta X(s,t)\right\}. \end{aligned} \tag{1.7}$$

Thus, we see that the time-dependent version of the ISI distribution consists of introducing the intensity function x as a new parameter, where the original ISI time — $z = t - s$ in equations (1.1) to (1.5) — is replaced with the integral of the intensity function between t and s. The intensity describes the time-dependent nature of the Ca^{2+} spiking, where the larger the intensity is over an interval the more spikes we expect in that interval. We also get a factor of $x(t)$ in the time-dependent ISI distribution from the time transformation.

At this stage, we have constructed a model for time-dependent Ca^{2+} spikes and if given the type of ISI distribution — Exponential, Gamma, etc — along with their parameters and the intensity function over the whole experiment time, we can simulate spike sequences. This is done by using the inverse transform method — see Chapter 2.

Next, we explore if we can infer the model parameters — the intensity function and ISI parameters — given a spike sequence. Let us begin by considering the inhomogeneous Gamma ISI distribution — shown in (1.7). Notice in the probability density that β is always multiplied by either $x(t)$ or $X(s,t)$. This could lead to difficulty disentangling the parameters β and x, since they may be non-identifiable.

To visualise this issue, imagine we want to estimate the area of a rectangle. Since we know the area of a rectangle is given by its base b multiplied by its height h we choose to model the area in terms of the product bh. However, suppose we are given sample areas $A_1, A_2, \ldots, A_N$, and we want to infer our parameters b and h. This is not possible because the area does not give us enough information about b and h, i.e if $b = 1$ and $h = 4$ this gives the same area as $b = 2$ and $h = 2$, thus we cannot identify b and h. This gives a picture of non-identifiability issues. Strictly speaking a model is non-identifiable if two unique parametisations lead to identical probability distributions. Conversely, a probability model is identifiable if we can learn the true parameter values if given an infinite set of observations.

Indeed, we find that the inhomogeneous Gamma ISI distribution is non-identifiable. Consider the inhomogeneous Gamma distribution with parameters $(ax, \alpha, \beta/a)$ with $a > 0$ a constant, where we have varied the values of β and x but maintained the value of their product. Rearranging we get

$$\begin{aligned} g_{\text{GAM}}(t, s|ax, \alpha, \beta/a) &= \frac{(\beta/a)ax(t)}{\Gamma(\alpha)}\left[(\beta/a)\int_s^t ax(u)du\right]^{\alpha-1}\exp\left\{-(\beta/a)\int_s^t ax(u)du\right\}, \\ &= \frac{\beta x(t)}{\Gamma(\alpha)}\left[(\beta/a)a\int_s^t x(u)du\right]^{\alpha-1}\exp\left\{-(\beta/a)a\int_s^t x(u)du\right\}, \\ &= \frac{\beta x(t)}{\Gamma(\alpha)}\left[\beta X(s,t)\right]^{\alpha-1}\exp\left\{-\beta X(s,t)\right\}, \\ &= g_{\text{GAM}}(t, s|x, \alpha, \beta). \end{aligned}$$

Thus, we found two unique parametisations with identical probability distributions and the Gamma ISI distribution is non-identifiable. Furthermore, we find that all the other considered time-dependent ISI distributions are also non-identifiable. This is shown in Table 1.2, where each distribution is given a pair of unique parameters sets that give rise to identical distributions — for proof these parameter sets lead to identical distributions see Appendix XX.

Therefore, this explains the use of one-parameter versions of the Gamma and Inverse Gaussian used in [6, 9], since restricting the stationary distributions to a single parameter alleviates the non-identifiability. To picture this, return to the

Distribution	Base parameters	Identical to ..
Exponential	(x, α)	$(ax, \alpha/x)$
Gamma	(x, α, β)	$(ax, \alpha, \beta/a)$
Inverse Gaussian	(x, λ, μ)	$(ax, a\lambda, a\mu)$
Log Normal	(x, μ, σ)	$(ax, \mu + \log a, \sigma)$
Weibull	(x, k, λ)	$(ax, k, a\lambda)$

Table 1.2: Parameterisations of the inhomogeneous ISI distributions which lead to non-identifiability, where $a > 0$ is a constant.

rectangle problem discussed previously. Constraining the stationary ISI distribution is similar to adding, the restriction that the rectangle must be a square. Then the base and height are equal and lead to a unique area $A = b^2 = h^2$. However, it is still unclear how to constrain the stationary probability distributions to resolve the non-identifiability. For example, we could restrict the stationary Gamma distribution from two parameters (α, β) to one parameter γ by setting $\alpha = \beta = \gamma$ or $\alpha = \gamma$ and $\beta = \gamma^2$. There are infinitely many possibilities for constraining the stationary ISI distribution. Thus we want to find a consistent approach to restrict these distributions whilst giving meaning to the remaining parameters.

Recall that Barbieri and Tilunaite generalised a Gamma(γ, γ) and an IG$(\mu, 1)$. Thus, the constrained Gamma distribution has mean equal to 1 and variance equal to $1/\gamma$, whereas the constrained Inverse Gaussian has mean equal to μ and variance equal to μ^3. From considering the mean and variance there is no clear pattern for how they constrained the stationary ISI distribution. Consider the probability density functions for the inhomogeneous ISI distributions constructed from an IG$(\mu, 1)$ and a Gamma(γ, γ)

$$g_{\mathrm{IG}^*}(t, s|x, \mu) = \frac{x(t)}{\sqrt{2\pi X(s,t)^3}} \exp\left\{-\frac{(X(s,t) - \mu)^2}{2\mu^2 X(s,t)}\right\} \tag{1.8}$$

and

$$g_{\mathrm{GAM}^*}(t, s|x, \gamma) = \frac{\gamma x(t)}{\Gamma(\gamma)} \left[\gamma X(s,t)\right]^{\gamma-1} \exp\left\{-\gamma X(s,t)\right\}, \tag{1.9}$$

respectively. Bear in mind that the corresponding stationary ISI distribution should be contained in the time-dependent model as a special case. Indeed, if we take the intensity function to be a constant, the time-dependent case reduces to its stationary counterpart. This is because the time-variability comes from the variability in the intensity function. Therefore, we set $s = 0$ and substitute $x(t) = a$ into (1.8) giving

$$\begin{aligned} g_{\mathrm{IG}^*}(t, s|a, \mu) =& \frac{a}{\sqrt{2\pi(at)^3}} \exp\left\{ -\frac{(at-\mu)^2}{2\mu^2 at} \right\}, \\ =& \sqrt{\frac{a^{-1}}{2\pi t^3}} \exp\left\{ -\frac{a^{-1}\left(t-\mu/a\right)^2}{2\left(\mu/a\right)^2 t} \right\}. \end{aligned} \tag{1.10}$$

Notice that (1.10) is the probability density function for a $\mathrm{IG}(\mu/a, 1/a)$. Similarly, it can be shown that the time-dependent Gamma ISI distribution, shown in (1.9), reduces to a $\mathrm{Gamma}(\gamma, \gamma a)$. Thus, comparing the mean and variance of these distributions we find that the Inverse Gaussian has mean μ/a and variance μ^3/a^2, whereas the Gamma distribution has mean $1/a$ and variance $1/\gamma a^2$. Therefore, for the Inverse Gaussian both the intensity and ISI parameter μ interact in the mean and variance, whereas in the Gamma the mean is controlled only by the intensity and not ISI parameter γ. In particular, the intensity can be viewed as the mean spiking rate and γ controls the variance of the ISI distribution.

In our modelling philosophy it would be advantageous to give meaning to the parameters that describe the ISI distribution, because this would lead to easier interpretation of the parameters. For example, in the case of the inhomogeneous Gamma distribution — Equation (1.9) — this is exactly what happens. In particular, the intensity function describes the mean spiking rate and γ describes the spiking variance. Thus, the intensity function has a biological meaning rather than just been a mathematical construct.

Therefore, we want to restrict all our stationary ISI distributions such that the mean ISI time corresponds to the inverse of the intensity. To do this we need to restrict the stationary ISI distributions to have mean one before applying the time rescaling. This is exactly what happens when you constrain a $\mathrm{Gamma}(\alpha, \beta)$ to a $\mathrm{Gamma}(\gamma, \gamma)$.

225 Therefore, when constructing time-dependent ISI models we begin with stationary
226 distributions with mean one, this is shown in Table 1.3.

Distribution	Original parameters	New parameter
Exponential	α	1
Gamma	(α, β)	(γ, γ), $\gamma > 0$
Inverse Gaussian	(μ, λ)	$(1, \lambda)$, $\lambda > 0$
Log Normal	(μ, σ^2)	$(-\mu, \sqrt{2\mu},)$, $\mu > 0$
Weibull	(k, λ)	$(k, \frac{1}{\Gamma(1+1/k)})$, $k > 0$

Table 1.3: One parameter version of the stationary ISI distributions with mean one.

227 To visualise the impact of setting the mean of the stationary ISI distribution
228 to one, we simulate spike sequences from a variety of inhomogeneous models. The
229 intensity function is the same for each model — the red line in Figure 1.2(A) and
230 (B) — and is drawn from a GP. The first model has a Gamma ISI distribution
231 with parameters $\alpha = 1.8$ and $\beta = 1.8$, and the second model also has Gamma ISI
232 distribution but with the parameters $\alpha = 3.6$ and $\beta = 1.8$. Recall that the mean of a
233 Gamma distribution is α/β, hence the mean for the first and second models are one
234 and two respectively. We simulate 1000 spike sequences from both models and bin
235 the spikes to calculate the mean spiking rate — this is shown as the light and dark
236 grey boxes in Figure 1.2(A) for the first and second model respectively. We find that
237 only the spikes from the first model have mean spiking rate equalling the intensity
238 function. In fact, the second model's mean spiking rate is equal to half the intensity
239 function, shown as the red dotted line. Consequently, we find that only the Gamma
240 ISI distribution with mean one has its mean spike rate coincide with the intensity
241 function.

242 Moreover, by constraining the stationary distributions in this way we can directly
243 compare intensity functions from models with different ISI distributions. For example
244 we could compare the the intensity function from a Gamma and Inverse Gaussian. If
245 the intensities were similar this would show consistency between the models.

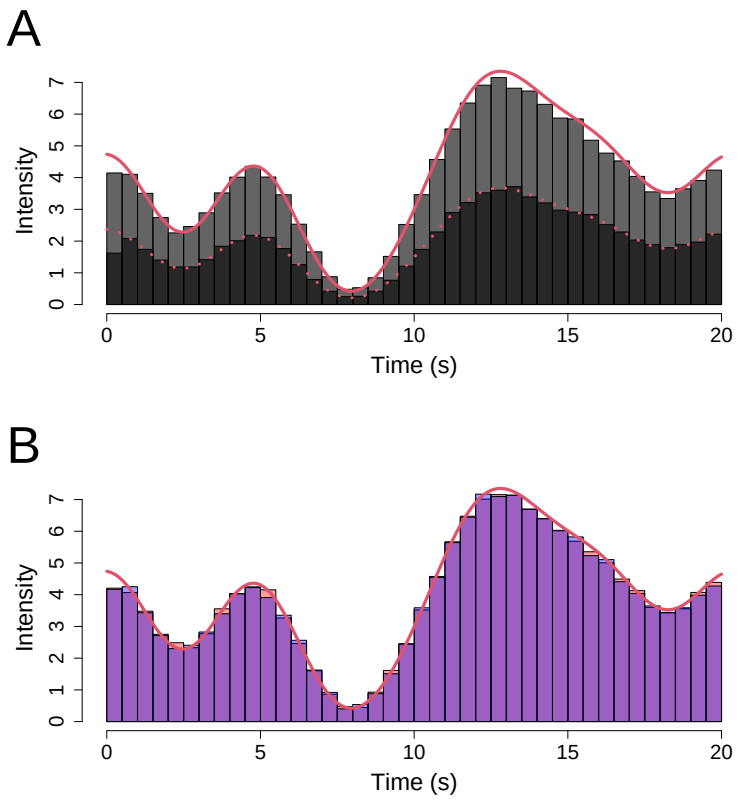

Figure 1.2: Comparing the intensity function — the red line in both plots — to histograms of 1000 simulated spike sequences. In (A) the light and dark boxes correspond to spikes generated from the model with Gamma ISI distributions with parameters Gamma(1.8,1.8) and Gamma(3.6,1.8) respectively. In (B) the spike sequences are simulated from a Gamma(1.8,1.8) and an Inverse Gaussian (1,1.8) resulting in the histograms in blue and red respectively, whose overlap is purple.

To show this, consider a third model with the same intensity function but whose ISI distribution follows an Inverse Gaussian with parameters $\mu = 1$ and $\lambda = 1.8$, which has mean one. We simulate 1000 spike sequences from this model. In figure 1.2(B), we compare the mean spike rate of this model with the first model — Gamma ISI distribution with mean one. The red and blue boxes correspond to binning the spike sequences from the Inverse Gaussian and Gamma models respectively. The overlap of the boxes is coloured purple. We see that the mean spike rate from both models are near identical. This shows that the intensity function of both models

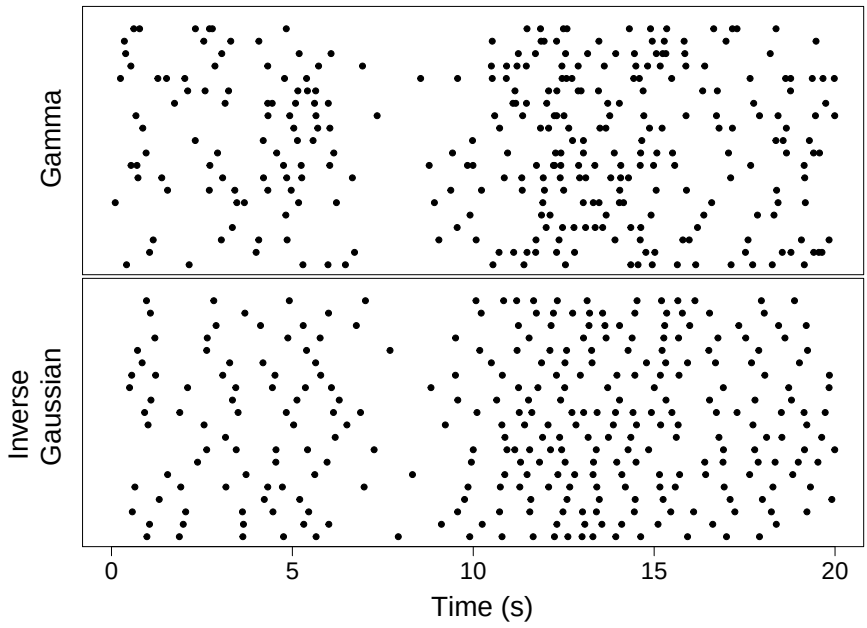

Figure 1.3: Rastor plots comparing 20 spike sequences drawn from the models in Figure 1.2(B). Both models share the same intensity function. The inhomogeneous Gamma has parameters $\alpha = 1.8$ and $\beta = 1.8$ and the inhomogeneous Inverse Gaussian has parameters $\mu = 1$ and $\lambda = 5$.

correspond to the mean spike rate, thus allowing us to compare intensity functions from models using different underlying probability distributions.

Note however, that even though the two models have the same intensity function this does not mean that the models are identical, the statistics of individual spike sequences could vary dramatically. For example, in Figure 1.3 we show rastor plots of 20 sequences corresponding to the Gamma and Inverse Gaussian models respectively, used in Figure 1.2(B). We see that the spikes from the Inverse Gaussian model has more regular spikes than those from the Gamma model.

Applying the time-rescaling transformation to the new restricted one-parameter version of the stationary distributions leads to the following inhomogeneous ISI distributions: the Exponential ISI distribution

$$g_{\text{EXP}}(t, s|x) = x(t)e^{-X(s,t)},$$

the Gamma ISI distribution

$$g_{\mathrm{GAM}}(t,s|x,\gamma) = \frac{\gamma x(t)}{\Gamma(\gamma)}\left[\gamma X(s,t)\right]^{\gamma-1}\exp\left\{-\gamma X(s,t)\right\},$$

the Inverse Gaussian ISI distribution

$$g_{\mathrm{IG}}(t,s|x,\lambda) = x(t)\left(\frac{\lambda}{2\pi X(s,t)^3}\right)^{0.5}\exp\left\{-\frac{\lambda(X(s,t)-1)^2}{2X(s,t)}\right\},$$

the Log Normal ISI distribution

$$g_{\mathrm{LN}}(t,s|x,\mu) = \frac{x(t)}{2X(s,t)\sqrt{\pi\mu}}\exp\left\{-\frac{(\log X(s,t)+\mu)^2}{4\mu}\right\},$$

and the Weibull ISI distribution

$$g_{\mathrm{W}}(t,s|x,k) = \frac{x(t)k}{\frac{1}{\Gamma(1+1/k)}}\left(\frac{X(s,t)}{\frac{1}{\Gamma(1+1/k)}}\right)^{k-1}\exp\left\{-\left(\frac{X(s,t)}{\frac{1}{\Gamma(1+1/k)}}\right)^{k}\right\}.$$

Thus, we have created five inhomogeneous ISI distributions based on the Exponential, Gamma, Inverse Gaussian, Log Normal and Weibull probability distributions. The time-dependence of each distribution comes from the intensity function x which represents the mean spiking rate. All the ISI distributions bar the Exponential also have a single ISI parameter γ, λ, μ and k for the Gamma, Inverse Gaussian, Log Normal and Weibull respectively. This parameter describes the shape/variance of the ISI distribution.

We will use the five ISI distributions — namely Exponential, Gamma, Inverse Gaussian, Log Normal and Weibull — in our analysis of Ca^{2+} spike sequences and explore which, if any, of the distributions best describe Ca^{2+} data and what we can learn from the results.

We now compare our inhomogeneous ISI distributions with those used in Tilunaite et al [9]. They used three ISI distributions in their work, the so-called inhomogeneous Poisson (IP), inhomogeneous Gamma (IG) and the inhomogeneous Inverse Gaussian (IIG). The IP and IG agree exactly with our Exponential ISI distribution and Gamma ISI distribution. However, their IIG ISI distribution has a different parameterisation to our Inverse Gaussian ISI distribution. Their model is generalised from the Inverse Gaussian with parameters $(1,\alpha)$ whereas our model is based on the parameters $(\lambda,1)$.

We do not use their parameterisation because the mean is not one, and as such the intensity function x does not agree with the mean spiking rate.

Thus the ISI distributions we consider builds upon previous work using the Exponential and Gamma ISI distributions and explores new distributions including the Log Normal and Weibull, which could better describe the Ca^{2+} spike sequences. Moreover, we have constructed the models in such a way that the intensity function coincides with the mean spiking rate. This allows for easier interpretation of the model parameters.

Bibliography

[1] Sheldon M. Ross. *A First Course in Probability*. Prentice Hall, Upper Saddle River, N.J., fifth edition, 1998.

[2] Jake Powell, Martin Falcke, Alexander Skupin, Tomas Bellamy, Theodore Kypraios, and Rüdiger Thul. *A Statistical View on Calcium Oscillations*, volume 1131, pages 799–826. 01 2020.

[3] Ayşe Kzlersü, Markus Kreer, and Anthony W. Thomas. The weibull distribution. *Significance*, 15(2):10–11, 2018.

[4] Thurley K, Smith IF, Tovey SC, Taylor CW, Parker I, and Falcke M. Timescales of IP_3-evoked Ca^{2+} spikes emerge from Ca^{2+} puffs only at the cellular level. *Biophysical Journal*, 101(11):2638–2644, Dec 2011.

[5] Raj S. Chhikara and J. Leroy Folks. Optimum test procedures for the mean of first passage time distribution in brownian motion with positive drift (inverse gaussian distribution). *Technometrics*, 18(2):189–193, 1976.

[6] Barbieri R, Quirk MC, Frank LM, Wilson MA, and Brown EN. Construction and analysis of non-Poisson stimulus-response models of neural activity. *The Journal of Neuroscience Methods*, 105(1):25–37, 2001.

[7] Frank J. Massey Jr. The kolmogorov-smirnov test for goodness of fit. *Journal of the American Statistical Association*, 46(253):68–78, 1951.

[8] T. W. Anderson and D. A. Darling. A test of goodness of fit. *Journal of the American Statistical Association*, 49(268):765–769, 1954.

[9] Tilunaite A, Croft W, Russell N, Bellamy TC, and Thul R. A Bayesian approach to modelling hetrogeneous calcium responses in cell populations. *PLoS Computational Biology*, 13(10):e1005794, Oct 17.