Calculations

CAPEL Alexandre

11 April, 2024

Contents

1 Introduction and notation

1.1 Termination time T

Let the random variable $T > 0$ represent the *termination* time at which an individual ceases its activity. Either by death or by capture in a trap.

We can describe the distribution of T using its:

- Probability Density Function (PDF): $f(t) = \lim_{\mathrm{d}t \to 0} \frac{\mathbb{P}(t < T \leq t + \mathrm{d}t)}{\mathrm{d}t} = F'(t)$

- Cumulative Distribution Function (CDF): $F(t) = \mathbb{P}(T \leq t) = \int_0^t f(u)\,\mathrm{d}u$,

- Survival function: $S(t) = \mathbb{P}(T > t) = 1 - F(t)$

- Hazard Rate function: $h(t) = \lim_{\mathrm{d}t \to 0} \frac{\mathbb{P}(t < T \leq t + \mathrm{d}t | T > t)}{\mathrm{d}t}$

The hazard function can be interpreted as the instantaneous rate of events (i.e. the expected number of terminations per unit of time). It is often more convenient to express hypotheses in terms of the hazard function and then derive the other expressions from it. The following relationships can be useful.

Since $\mathbb{P}(t < T \leq t + \mathrm{d}t \mid T > t) = \frac{\mathbb{P}(t < T \leq t + \mathrm{d}t, T > t)}{\mathbb{P}(T > t)} = \frac{\mathbb{P}(t < T \leq t + \mathrm{d}t)}{S(t)}$, it follows from the definition that,

$$h(t) = \frac{f(t)}{S(t)} = -\frac{S'(t)}{S(t)} = -\frac{\mathrm{d}}{\mathrm{d}t} \log S(t) \tag{1}$$

Reciprocally, given a hazard function, survival can be obtained using

$$S(t) = \exp\left(-\int_0^t h(u)\,\mathrm{d}u\right), \tag{2}$$

assuming that $S(0) = 1$.

1.2 Termination cause C

Let the categorical random variable $C \in \{0, 1, \ldots, I\}$ be an indicator of the termination **cause**. Either *death* ($C = 0$) or *capture* in trap i ($C = i$).

Its Probability Mass Function is:

$$\mathbb{P}(C = i) = \omega_i, \tag{3}$$

with $0 \leq \omega_i \leq 1$ and $\sum \omega_i = 1$.

Due to the dispersal process, the relative risks of being caught in either trap will shift over time. Thus, it will be easier to specify these risks conditional on time and then average over the PDF of T:

$$\omega_i = \int_0^\infty \mathbb{P}(C = i \mid T = t) f(t) \, \mathrm{d}t. \tag{4}$$

1.3 Specific termination times T_i

Let T_i be the termination time in a scenario where the **only possible termination cause** is i.

T and T_i are different random variables, with different PDFs, CDFs, survivals and hazard rates, for all of which we define the corresponding **specific** versions f_i, F_i, S_i and h_i.

The relationship between these variables is that

$$T = \min\{T_i : \; i = 0, \dots, I\} \tag{5}$$

We can think of it as if at every instant, all specific termination causes were tested independently. The first that occurs determines both the termination time T and cause C.

1.4 Position X

Finally, let $X(t) \in \mathbb{R}^2$ be the random location of an individual at a given time $0 \leq t \leq T$. Without loss of generality, we assume that $X(0) = (0, 0)$.

2 Preliminary results

2.1 Fixed and known position

2.2 Different constant conditional hazard rates imply non-constant overall hazard rate

Assume a fixed position (x_0) of an individual, and no mortality.

$$
\begin{aligned}
\mathbb{P}(t < T \leq t + \mathrm{d}t \mid T > t) &= \frac{\mathbb{P}(t < T \leq t + \mathrm{d}t)}{\mathbb{P}(T > t)} \\
&= \frac{\sum \mathbb{P}(t < T \leq t + \mathrm{d}t \mid C = i) \mathbb{P}(C = i)}{\mathbb{P}(T > t \mid C = i) \mathbb{P}(C = i)}
\end{aligned}
\tag{6}
$$

Letting $\omega_i = \mathbb{P}(C = i)$, assuming constant conditional hazard rates λ_i implies $f_i(t \mid C = i) = \lambda_i e^{-\lambda_i t}$, thus:

$$
\begin{aligned}
h(t) &= \lim_{dt \to 0} \frac{\mathbb{P}(t < T \le t + dt \mid T > t)}{dt} = \lim_{dt \to 0} \frac{\mathbb{P}(t < T \le t + dt)}{S(t)\,dt} \\
&= \frac{\sum f_i(t \mid C = i)\omega_i}{\sum S(t \mid C = i)\omega_i} = \frac{\sum \omega_i \lambda_i e^{-\lambda_i t}}{\sum \omega_i e^{-\lambda_i t}} \\
&= \sum \lambda_i \theta_i(t),
\end{aligned}
\tag{7}
$$

with $\theta_i(t) = \frac{\omega_i e^{-\lambda_i t}}{\sum \omega_j e^{-\lambda_j t}}$.

The overall hazard rate is thus a weighted average of conditional hazard rates, with coefficients that evolve in time.

The coefficients $\theta_i(t)$ are a weighted version of the *softmax*[1] transformation of $\{\lambda_i\}$ with *temprature t*. They verify $\sum \theta_i(t) = 1$, $\forall t > 0$, $\theta_i(0) = \omega_i / \sum w_j$ and $\lim_{dt \to \infty} \theta_i(t) = 0$ except for i such that λ_i is smallest.

Thus, the coefficients start by reproducing the distribution of the termination causes ω_i and progressively concentrate on the trap with lower hazard rate.

2.3 The conditional hazard rates must be all equal to the overall hazard rate

The conditional hazard rates are the hazard rates of capture events in a specific trap, conditional to capture in that trap.

$$
\begin{aligned}
h(t \mid C = i) &= \lim_{dt \to 0} \frac{\mathbb{P}(t < T \le t + dt \mid C = i, T > t)}{dt} \\
&= \lim_{dt \to 0} \frac{\mathbb{P}(t < T \le t + dt, C = i \mid T > t)}{P(C = i \mid T > t)\,dt}
\end{aligned}
$$

The key observation is that since the position of the individual is fixed, the scenario is constantly the same and thus $P(C = i \mid T > t) = P(C = i)$, $\forall t$. So, the termination cause C is independent from the termination time T, and we have,

$$
\begin{aligned}
\mathbb{P}(t < T \le t + dt, C = i \mid T > t) &= \mathbb{P}(t < T \le t + dt \mid T > t)\mathbb{P}(C = i \mid T > t) \\
&= \mathbb{P}(t < T \le t + dt \mid T > t)\mathbb{P}(C = i) \\
&= \omega_i\,\mathbb{P}(t < T \le t + dt \mid T > t)
\end{aligned}
$$

---

[1] https://en.wikipedia.org/wiki/Softmax_function

Thus,

$$
\begin{aligned}
h(t \mid C = i) &= \lim_{\mathrm{d}t \to 0} \frac{\omega_i \mathbb{P}(t < T \leq t + \mathrm{d}t \mid T > t)}{\omega_i \, \mathrm{d}t} \\
&= \lim_{\mathrm{d}t \to 0} \frac{\mathbb{P}(t < T \leq t + \mathrm{d}t \mid T > t)}{\mathrm{d}t} \\
&= h(t), \quad \forall i
\end{aligned} \tag{8}
$$

Alternatively, one could also argue that the overall hazard rate should be constant since the situation with fixed position remains unchanged.

Let $h(t) = \lambda_0$. From (7), we have

$$
\sum \lambda_i \theta_i(t) = \lambda_0. \tag{9}
$$

Which is only possible if $\lambda_i = \lambda_0, \ \forall i$.

This is confirmed with an empirical simulation with 2 traps and fixed positions as displayed in figure 1).
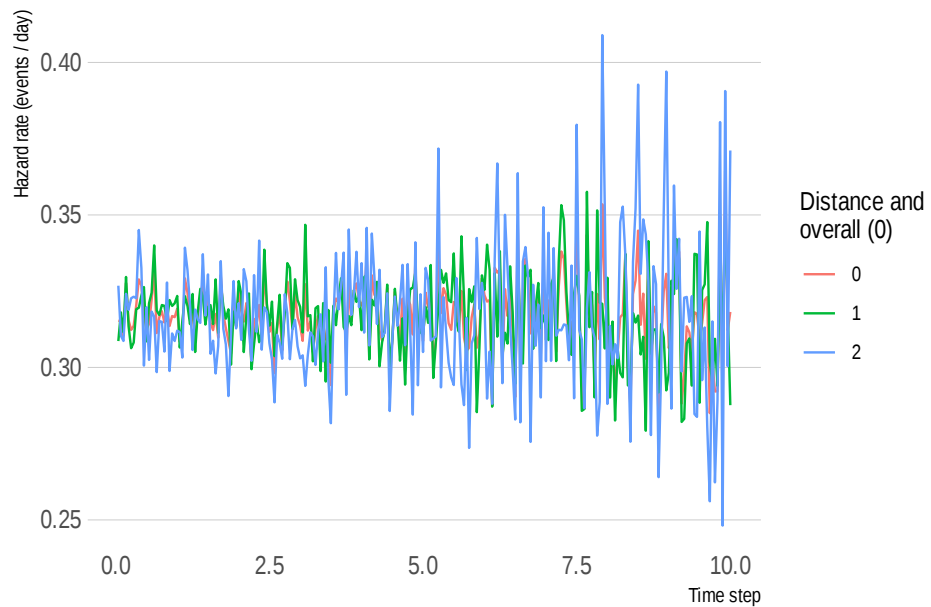

Figure 1: Conditional and overall empirical hazard rates for a scenario with fixed position of individuals and 2 traps at different distances.

2.4 The overall hazard rate is the sum of specific hazard rates

Let $\lambda_i = h_i(t)$, $i = 0, \dots, I$ the specific hazard rates associated to death ($i = 0$) and capture in trap i. This is, the (constant) hazard functions of the corresponding events if the specific cause was the only one at play. In other words, if the individual was active in its fixed position until it was eventually terminated by i.

Since $T = \min\{T_i : i = 0, \dots, I\}$, in an instant dt,

$$
\begin{aligned}
\mathbb{P}(t < T \leq t + dt \mid T > t) &= 1 - \prod_{i=0}^{I}(1 - \mathbb{P}(t < T_i \leq t + dt \mid T > t)) \\
&= 1 - \prod_{i=0}^{I}(1 - \lambda_i\, dt) \\
&= 1 - [1 - \lambda_0\, dt - \cdots - \lambda_I\, dt + O(dt^k)], \quad k \geq 2 \\
&= dt \sum_{i=0}^{I} \lambda_i + O(dt^k)
\end{aligned}
\tag{10}
$$

Thus, the overall hazard rate is:

$$
\begin{aligned}
h(t) &= \lim_{dt \to 0} \frac{\mathbb{P}(t < T \leq t + dt \mid T > t)}{dt} \\
&= \lim_{dt \to 0} \sum_{i=0}^{I} \lambda_i + O(dt^{k-1}), \quad k \geq 2 \\
&= \sum_{i=0}^{I} \lambda_i
\end{aligned}
\tag{11}
$$

3 Probability of capture with fixed location of the individual

Let's first assume that $X(t) = x_0$, $\quad 0 \leq t \leq T$.

3.1 With only one trap

Only one trap means that $I = 1$ and C can only take the value 0 in case of death prior to capture, or the value 1 in the converse.

We assume that, if an individual stays at the same location, the hazard function of the capture time $T \mid C = 1, X(t) = x_0$ is constant in time, and decreases exponentially with the distance $r = \|x_0\| > 0$ from the trap:

$$
h_{1r}(t) = h(t \mid C = 1, X = x_0) = \lambda_r = \alpha e^{-\beta r}, \quad \alpha, \beta > 0.
$$

From equation (2), it follows that:

$$S(t \mid C = 1, X = x_0) = e^{-\lambda_r t}$$

Hence, $f(t \mid C = 1, X = x_0) = -S'(t \mid C = 1, X = x_0) = \lambda_r e^{-\lambda_r t}$, which is the PDF of the exponential distribution of rate λ_r.

$$T \mid C = 1, \|X\| = r \sim \mathrm{Ex}(\alpha e^{-\beta r}) \tag{12}$$

We can also compute the daily capture probability:

$$\mathbb{P}(0 \leq T \leq 1 \mid C = 1, X = x_0) = \int_0^1 \lambda_r e^{-\lambda_r t} \, \mathrm{d}t = 1 - e^{-\lambda_r}$$

3.2 With several traps

Consider now the case where we have $I > 1$ traps in different locations. The aim is to compute the probability of capture where these traps are active together and to model the capture time of each trap.

Here, the difficulty comes from the dependence between these variables because if an individual is trapped by the trap i, then the capture time of the other ones will be equal to infinity.

This is why we consider the global *termination* time T jointly with the termination *cause* C described initially.

3.2.1 Multinomial capture probability

First, let K a categorical variable which takes the value 0 if the individual is not trapped before the next day and i if it is trapped by the trap i.

This is: $K = 0$ if $T > 1$ and $K = i$ if $T \leq 1, C = i > 0$.

We can easily compute the probability of non-capture which means that individually, the traps didn't work, i.e. :

$$\mathbb{P}(K = 0) = \mathbb{P}(K \neq 1, ..., K \neq I) = \prod_{i=1}^{I} \mathbb{P}(K \neq i) = \prod_{i=1}^{I} (1 - \omega_i) = \omega_0$$

where $\omega_i = 1 - e^{-\lambda_k}$.

Now, we would like to build a distribution conditionally to the event $K \neq 0$ using the weights ω_k. One way to do this, is to assume that the distribution is :

$$\mathbb{P}(K = k | K \neq 0) = \frac{\omega_k}{\sum_{i=1}^{I} \omega_i}$$

which is of course a probability distribution. Thus, we can calculate the value for the global distribution which is :

$$\begin{aligned} \mathbb{P}(K = k) &= \mathbb{P}(K = k, K \neq 0) \\ &= \mathbb{P}(K = k | K \neq 0)\mathbb{P}(K \neq 0) \\ &= (1 - \omega_0)\frac{\omega_k}{\sum_{j=1}^{I} \omega_i} \end{aligned}$$

3.2.2 Capture time

Let $T_1, \ldots, T_I$ the capture time for each trap. We are going to build a discrete process helping us to compute the survival function of each time.

Daily capture process

Let $(K_m)_{m \in \mathbb{N}}$ a categorical process describing the state of the individual at the day m. We can see this as a markov chain with the transition matrix :

$$\mathbb{P}(K_{m+1} = j | K_m = i) = \begin{cases} 0 & \text{si } i \neq 0 \text{ et } j \neq i \\ 1 & \text{si } i \neq 0 \text{ et } j = i \\ \omega_0 & \text{si } i = j = 0 \\ (1 - \omega_0)\frac{\omega_k}{\sum_{j=1}^{I} \omega_i} & \text{si } i = 0 \text{ et } j \neq 0 \end{cases}$$

Hence, this process can take only two values along time : 0 at the beginning and another one after a certain amount of time. Indeed, if the individual is trapped by one trap, it can not escape from it.

Then, let's define the variable $M_k = \inf\{m \in \mathbb{N}, K_m = k\}$ for $k \neq 0$.

For now, we can just say that :

$$(T_k < \infty) \Leftrightarrow (M_k < \infty)$$

To compute this probability, we will calculate the mass function of the variable M_k. In order to do that, we can express the event $M_k = m$ using the process K_m and also its markov property. Therefore, we have :

$$\mathbb{P}(M_k = m) = \mathbb{P}(K_1 = 0, \ldots, K_m = k)$$
$$= \mathbb{P}(K_m = k | K_{m-1} = 0)\mathbb{P}(K_1 = 0, \ldots, K_{m-1} = 0)$$
$$= \omega_0^{m-1}(1 - \omega_0)\frac{\omega_k}{W}$$

Thus, the probability is :

$$\mathbb{P}(M_k < \infty) = \sum_{m=1}^{\infty} \mathbb{P}(M_k = m) = \frac{\omega_k}{W}$$

We can also deduce that $\mathbb{P}(M_k = \infty) = 1 - \frac{\omega_k}{W}$.

Survival function and density

Using the previous results, we are able to compute the survival function for each capture time. Indeed, we have :

$$\mathbb{P}(T_k > t) = \mathbb{P}(T_k > t | M_k < \infty)\mathbb{P}(M_k < \infty) + \mathbb{P}(T_k > t | M_k = \infty)\mathbb{P}(M_k = \infty)$$
$$= \mathbb{P}(T_k > t | M_k < \infty)\frac{\omega_k}{W} + 1 - \frac{\omega_k}{W}$$
$$= 1 - \frac{\omega_k}{W}(1 - \mathbb{P}(T_k > t | M_k < \infty))$$

However, we can consider that knowing $(M_k < \infty)$, the probability that $T_k > t$ is the same as when there is only the trap k which is active. Then, we have $\mathbb{P}(T_k > t | M_k < \infty) = e^{-\lambda_k t}$ and :

$$S_k(t) = 1 - \frac{\omega_k}{W}(1 - e^{-\lambda_k t})$$

It is possible to compute the density function :

$$f_k(t) = -S_k'(t) = \lambda_k \frac{\omega_k}{W}e^{-\lambda_k t}$$

We observe that the integral of this function is not equal to 1, but $\frac{\omega_k}{W}$. It is actually due to the fact that the variable T_k can take the value ∞ with probality $1 - \frac{\omega_k}{W}$.

4 Probability of capture with unknown location

We assume that the individual is no longer immobile and that its position along time can be modeled as a Brownian motion (interpretation ?) that means the density g satisfies the diffusion equation :

$$\frac{\partial g}{\partial t} = D\nabla^2 g$$

where D is a diffuse constant which represents the rate of spread of the individuals and ∇^2 is the divergence of the gradient.

Hence, we can express the density g as follow :

$$f(x|a, D) = \mathcal{N}(x|a, (4Dt)I_2)$$

Let's note X_t the Brownian motion at time t and f_{X_t} its density.

4.1 With only one trap

All the previous section was about the computation of the probability of capture for a constant location. Here, we would like to extend the expression assuming the location is random.

Consider the hazard function, we have, by the definition that :

$$h(t) = \lim_{dt \to 0} \frac{\mathbb{P}(t \leq T \leq dt|T \geq t)}{dt} = \lim_{dt \to 0} \int_{\mathbb{R}^2} \frac{\mathbb{P}(t \leq T \leq dt|T \geq t, X_t = x)}{dt} f_{X_t}(x)\,dx$$

As everything is positive, we can switch between limit and integral :

$$h(t) = \int_{\mathbb{R}^2} \lim_{dt \to 0} \frac{\mathbb{P}(t \leq T \leq t + dt|T \geq t, X_t = x)}{dt} f_{X_t}(x)\,dx$$

More, the limit allows us to consider this quantity as the hazard function of the capture time with a constant location. So we can write this :

$$h(t) = \int_{\mathbb{R}^2} \lambda(x) f_{X_t}(x)\,dx$$

where $\lambda(x) = \alpha e^{-\beta\|x - x_0\|^2}$ where x_0 is the location of the trap.

Computation of the integral

We would like to calculate :

$$\int_{\mathbb{R}^2} \alpha e^{-\beta||x-x_0||^2} f_{X_t}(x) \, \mathrm{d}x = \alpha \mathbb{E}[e^{-\beta||X_t-x_0||^2}]$$

We have :

$$\mathbb{E}[e^{-\beta||X_t-x_0||^2}] = \int_{\mathbb{R}^2} e^{-\beta||x-x_0||^2} \frac{1}{8\pi Dt} e^{-\frac{||x||^2}{8Dt}} \, \mathrm{d}x$$

$$= \frac{1}{8\pi Dt} \int_{\mathbb{R}^2} e^{-(\frac{||x||^2}{8Dt}+\beta||x-x_0||^2)} \, \mathrm{d}x$$

But the term in the exponential can be written as :

$$\frac{||x||^2}{8Dt} + \beta||x-x_0||^2 = \frac{||x||^2}{8Dt} + \beta(||x||^2 - 2 <x|x_0> + ||x_0||^2)$$

$$= \gamma_t(||x||^2 - 2 <x|\frac{\beta}{\gamma_t}x_0>) + \beta||x_0||^2$$

$$= \gamma_t(||x - \frac{\beta}{\gamma_t}x_0||^2 - ||\frac{\beta}{\gamma_t}x_0||^2) + \beta||x_0||^2$$

$$= \gamma_t||x - \frac{\beta}{\gamma_t}x_0||^2 + \beta(1 - \frac{\beta}{\gamma_t})||x_0||^2$$

where $\gamma_t = \frac{8D\beta t+1}{8Dt}$.

We can use this equality to have :

$$\mathbb{E}[e^{-\beta||X_t-x_0||^2}] = \frac{e^{-\beta(1-\frac{\beta}{\gamma_t})||x_0||^2}}{8\pi Dt} \int_{\mathbb{R}^2} e^{-\gamma_t||x-\frac{\beta}{\gamma_t}x_0||^2} \, \mathrm{d}x$$

But, it is well known that :

$$\int_{\mathbb{R}^2} e^{-\gamma_t||x-\frac{\beta}{\gamma_t}x_0||^2} \, \mathrm{d}x = \frac{\pi}{\gamma_t}$$

Thus we finally have:

$$h(t) = \mathbb{E}[\alpha e^{-\beta||X_t-x_0||^2}] = \frac{\alpha}{1+8D\beta t} e^{-\frac{\beta||x_0||^2}{1+8D\beta t}}$$
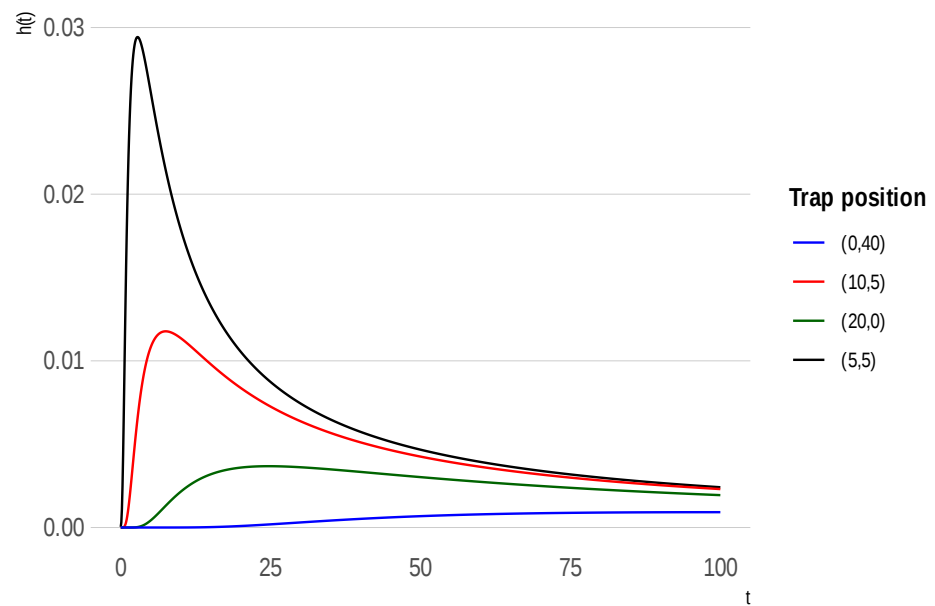
Figure 2: Hazard function of the capture time for different location of the trap.

These are the corresponding curve for different location of the trap :

Integral modification for survival function approximation

We know that :

$$S(t) = \exp(-\int_0^t h(s)ds)$$

but this integral is not calculable for our hazard function.

Thus, we have to approximate this value. However, the fact that the bound of the interval goes to infinity makes the computation really difficult. But, we can rewrite the integral (by using the transformation $u = \frac{1}{1+8\beta Ds}$) :

$$
\begin{aligned}
\int_0^t h(s)ds &= \int_0^t \frac{\alpha}{1+8D\beta s} e^{-\frac{\beta||x_0||^2}{1+8D\beta s}} ds \\
&= \int_1^{\frac{1}{1+8D\beta t}} \alpha u e^{-\beta||x_0||^2 u}(-\frac{1}{8D\beta u^2})du \\
&= \frac{\alpha}{8D\beta} \int_{\frac{1}{1+8D\beta t}}^1 \frac{1}{u} e^{-\beta||x_0||^2 u} du
\end{aligned}
\tag{13}
$$

which will easier to compute for each terms (see function `cumhazard_one` in the file `functions.R`.

4.2 With several traps

Now, we want to compute this in the case that there are I trap located at position x_k during the experiment.

FM: What is "this"? where does this come from?

$$h_k(t|x) = \frac{\omega_k(x)}{W(x)}\lambda_k(x)$$

Thus, the previous reasoning from the one trap case gives :

$$h_k(t) = \int_{\mathbb{R}^2} \frac{\omega_k(x)}{W(x)}\lambda_k(x)f_{X_t}(x)\,\mathrm{d}x$$

We can rewrite this formula :

$$h_k(t) = \int_{\mathbb{R}^2} \frac{\omega_k(x)}{W(x)} \lambda_k(x) f_{X_t}(x) \, \mathrm{d}x$$

$$= \alpha \int_{\mathbb{R}^2} \frac{\omega_k(x)}{W(x)} e^{-\beta||x-x_k||^2} f_{X_t}(x) \, \mathrm{d}x$$

$$= \frac{\alpha}{8\pi Dt} \int_{\mathbb{R}^2} \frac{\omega_k(x)}{W(x)} e^{-\left(\frac{||x||^2}{8Dt} + \beta||x-x_k||^2\right)} \, \mathrm{d}x$$

Using what we have done in the previous section, we can deduce that :

$$h_k(t) = \frac{\alpha e^{-\frac{\beta||x_k||^2}{1+8D\beta t}}}{8\pi Dt} \int_{\mathbb{R}^2} \frac{\omega_k(x)}{W(x)} e^{-\gamma_t||x-\frac{\beta}{\gamma_t}x_k||^2} \, \mathrm{d}x = \frac{\alpha e^{-\frac{\beta||x_k||^2}{1+8D\beta t}}}{1+8D\beta t} \int_{\mathbb{R}^2} \frac{\omega_k(y)}{W(y)} f_{Y_t}(y) dy$$

where $Y_t \sim \mathcal{N}(\frac{\beta}{\gamma_t} x_k, \frac{\gamma_t}{2} I_2)$. Thus we can compute the hazard function seeing that :

$$h_k(t) = \frac{\alpha e^{-\frac{\beta||x_k||^2}{1+8D\beta t}}}{1+8D\beta t} \mathbb{E}\left[\frac{\omega_k(Y_t)}{W(Y_t)}\right]$$

It's funny because we can see Y_t as a path to the trap k.

5 General model including survival of the individual

To end the calculation, we would like to add the death risk of each individual.

Lets consider the model as right censored data that is to say as $T = \min(X, C)$ where X is the capture time of a trap, and C the censure time which can be seen as the death time of the individual.

The censure time C is a random variable following an exponential distribution and conditionally to the capture at the trap k, X has the hazard function as written as upper.

Thus, the density of the couple can be expressed as :

$$f_{(X,C)}(x,c) = f_X(x) f_C(c)$$

But the collected data will be only from those such that $X < C$. We can try to compute the corresponding survival function :

$$\mathbb{P}(T > t, X \leq C) = \mathbb{P}(X > t, X \leq C) = \int_{\mathbb{R}_+} \int_{\mathbb{R}_+} 1_{x<c} 1_{t<x} f_X(x) f_C(c) \, \mathrm{d}x dc = \int_t^\infty \left(\int_x^\infty f_C(c) dc \right),$$

If we look at the time conditionally at the event $X \leq C$, we get a new density which is proportional to $f_X(x) S_C(x)$.

6 Simulation verification

Now we would like to see if our calculation are available. Thus, we simulate data with mortal individuals. We chose :

- a dispersion parameters $D = 150$.
- a probability of capture for an individual located at the trap of 0.8
- a probability of capture for an individual located at 5 meters from the trap of 0.1 (useful to get β parameters).
- a probability of dailey survival of 0.8

We get this with only one trap :

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

and we get this for a trap with others :

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

The curves fit well with the histograms. So, we can continue our study.

7 Cumulative likelihood with censored data

Now, we would like to use the data to estimate the unknown parameters of the model, but there is a problem arising from the type of data we have.

During the experiment, we do not see exactly the capture time of an individual.
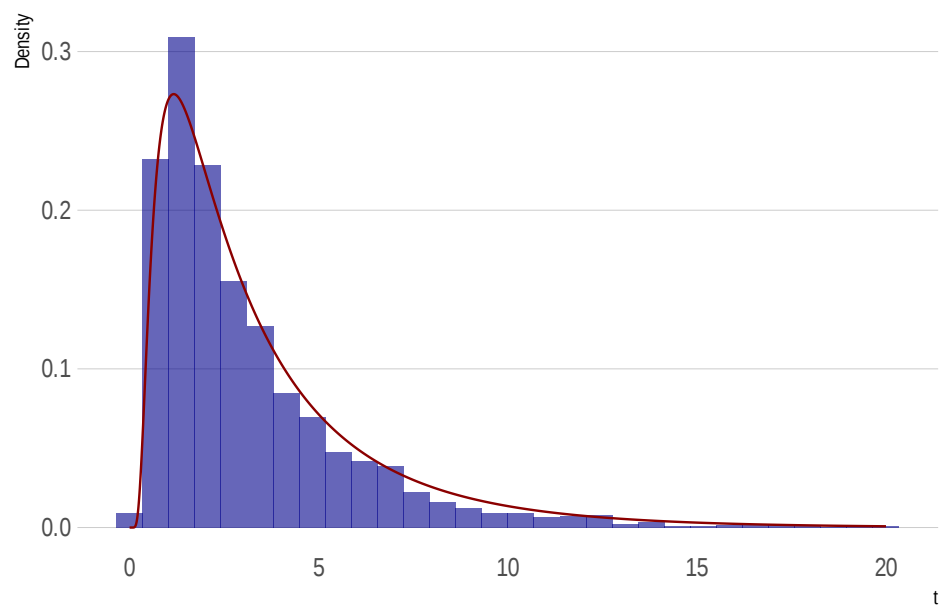
Figure 3: Histogram of the simulation with the pdf of the capture time at a one trap setting
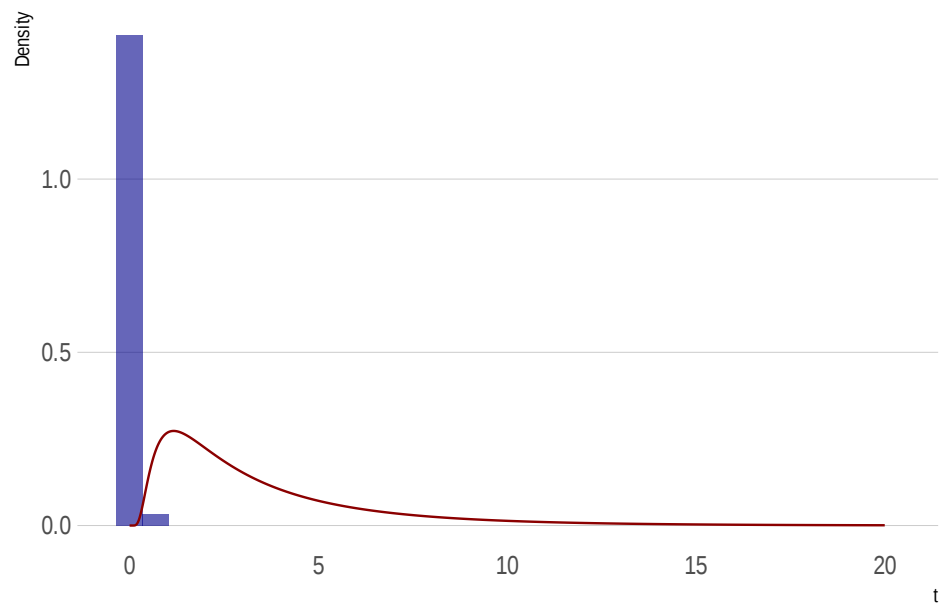
Figure 4: Histogram of the simulation with the pdf of the capture time at the trap 4 in a several traps setting

7.1 Type of censor

There is a lot of different type of censor. In our context, the capture times are censored randomly and at the end, we will get an interval-censored data.

Formally, the model of this type of data is defined as a triplet of random variable (L, T, R) where R is the value of the right censored bound, L is the left censored bound and T the real capture time.

We also get another random variable, that we call δ which can take three values :

- $\delta = 1$ if we know the true value of T

- $\delta = 2$ if $R = \infty$

- $\delta = 3$ else

The last case is the casual interval-censored case.

We can compute an estimator of the survival function with this type of data using the Turnbull's estimator.

7.2 A different likelihood

With these paradigm, we have to change the structure of the likelihood and take in consideration this variation of the data. Consider a sample of censored data $(t_i, [l_i, r_i], \delta_i)_{i=1,\dots,n}$, we write :

$$\mathcal{L} = \prod_{i:\delta_i=1} f(t_i) \prod_{i:\delta_i=2} S(l_i) \prod_{i:\delta_i=3} (S(l_i) - S(r_i))$$

But, in our context, we will have $\delta_i = 3, \forall i$, so the likelihood can be written as :

$$\mathcal{L} = \prod_{i=1}^{n} (S(l_i) - S(r_i))$$

Now, we suppose that we have a sequence of time $(t_0 = 0, t_1, \dots, t_n)$ which correspond at the time when the trap are checked. For every check at t_j, we get a list of numbers $(n_{1j}, \dots, n_{Kj})$ which are the number of capture between t_{j_1} and t_j for every trap.

With this setting, we can write our likelihood as :

$$\mathcal{L} = \prod_{j=1}^{n} \prod_{k=1}^{K} (S_k(t_j) - S_k(t_{j-1}))^{n_{jk}}$$

So we get the following log-likelihood :

$$l(N; D, \beta, \alpha, \lambda) = \sum_{j=1}^{n} \sum_{k=1}^{K} n_{jk} \log \left(S_k(t_j; D, \beta, \alpha, \lambda) - S_k(t_{j-1}; D, \beta, \alpha, \lambda) \right)$$

where $N = (n_{jk})$ is the matrix of capture count.

8 Estimation of the parameters

8.1 Maximum likelihood estimator

8.2 Bayesian estimator by MCMC

TODO : - end this section with survival times of the individuals - plot the histogram with the corresponding pdf (for one and several traps cases) - write the cumulative likelihood with specific data - introduce two methods : likelihood maximization, and Bayesian inference with priors defined by $\alpha \sim \mathcal{B}eta(a, 1)$ $\beta \sim \mathcal{B}eta(1, b)$ + research a good candidate for D's prior and (a, b) parameters. - estimation of the parameters with these two methods and see results with non parametric estimator of the survival function. - application with real data.