Modeling Waiting Times

There are many situations in which we would like to model the amount of time you have to wait for an event to occur where we know the average frequency of occurrence, but the exact timing is unknown. Determining the amount of time you have to wait for the next customer to enter a store. We may know that on average how frequently a customer enters, but the exact amount of time to wait is unknown. Another situation to model would be the time between neuronal spikes. Additionally, we could use a model to determine how many hours of smart phone use before an unfortunate accident results in the screen breaking. A common model for determining the waiting time is to use an exponential model of the form

$$p(t) = \begin{cases} Ce^{-\lambda t} & \text{if } t \geq 0 \\ 0 & \text{if } t < 0 \end{cases}$$

where $\lambda > 0$ is the average rate of occurrence and C is a constant. The function $p(t)$ is called a probability density and can be used to compute the probability of the wait time T being within a given interval $[t_1, t_2]$ by

$$\mathbb{P}(T \in [t_1, t_2]) = \int_{t_1}^{t_2} p(t)dt.$$

1. For a probability distribution, we require $\displaystyle\int_{-\infty}^{\infty} p(t)dt = 1$. Use this to show that $C = \lambda$.
2. The expected value for a probability distribution is given by $\mathbb{E}(T) = \displaystyle\int_{-\infty}^{\infty} tp(t)dt$. Use this to show that $\mathbb{E}(T) = \frac{1}{\lambda}$. Does the term "expected rate" for λ make sense?
3. Plot the distribution function for different values of λ and compare the different distribution functions.
4. One useful property about the exponential distribution is that it is "memoryless." This means that the probability of waiting does not depend on how long you have already waited (i.e. if you have already waited a time of t_1, the probability of waiting an additional amount t_2 would be the same as if you just started waiting). Mathematically, we write this as $\mathbb{P}(T \geq t_1 + t_2 | T \geq t_1) = \mathbb{P}(T \geq t_2)$. Use the fact that $P(A|B) = \frac{\mathbb{P}(A \text{ and } B)}{\mathbb{P}(B)}$ to show that $\mathbb{P}(T \geq t_1 + t_2 | T \geq t_1) = \mathbb{P}(T \geq t_2)$.

   Hint: $\mathbb{P}(\{T \geq t_1 + t_2\} \text{ and } \{T \geq t_1\}) = \mathbb{P}(T \geq t_1 + t_2) = \displaystyle\int_{t_1+t_2}^{\infty} p(t)dt$.
5. For the exponential distribution, we rely on the assumptions that events are independent, occur at a known rate on average, and do not depend on the amount of time that has already passed. While these assumptions make for a relatively convenient model to analyze, they may not hold in every situation. For example, suppose that instead of trying to determine the amount of time before a phone screen breaks, you want to model the time at which your phone wears out. In this case, the time that has already passed will likely impact the amount of time before the components fail. Thus, the exponential model would not be suitable. Furthermore, in a neural firing model, on certain time scales, there might be a cool-down period after an action potential which prevents the following action potential from occurring too soon. For such a situation, it may be better to model the waiting time with a density of

   $$p(t) = \begin{cases} Cte^{-\frac{2t}{\theta}} & \text{if } t \geq 0 \\ 0 & \text{if } t < 0 \end{cases}$$

   where $\theta > 0$ is the expected lifetime and C is again an unknown constant.

   (a) Show that $C = \frac{4}{\theta^2}$.

   (b) Show that $\mathbb{E}(T) = \theta$. Does the term "expected lifetime" make sense for this parameter?

   (c) Plot this distribution for different values of θ and compare the different distributions.
6. Can you think of any other situations that can be modeled by either of the previous distributions? Alternatively, can you think of any other situations in which neither of these distributions seems adequate? For your scenario, can you come up with a distribution that would be better suited for modeling the wait time?