Review of

PROBABILITY DISTRIBUTIONS

Probability distribution

Probability that a (continuous) **random variable** X is in (x,x+dx).

$$P\left(x < X < x + dx\right)$$

Probability density function (**PDF**).

$$f\left(x\right) = \lim_{dx \to +0} \frac{P\left(x < X < x + dx\right)}{dx}$$

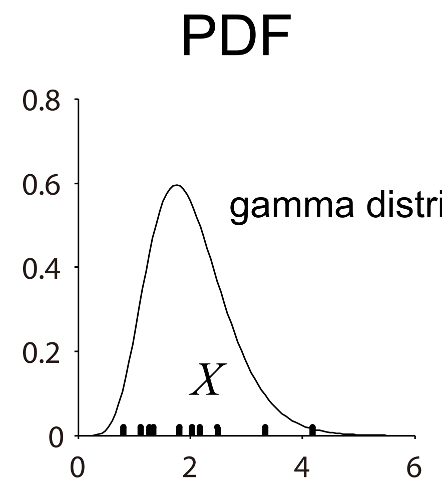
$$P\left(x < X < x + dx\right) \approx f\left(x\right) dx$$

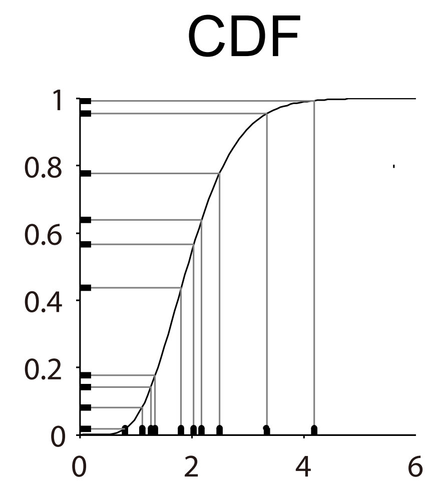
Cumulative distribution function (**CDF**)
= Life time distribution.

$$F\left(x\right) = P\left(X \le x\right) = \int_{-\infty}^{x} f\left(s\right) ds$$

Survival function
= Complementary CDF.

$$\bar{F}\left(x\right) = P\left(X > x\right) = \int_{x}^{\infty} f\left(s\right) ds = 1 - F\left(x\right)$$





The fundamental theorem of calculus

The relation between probability distribution and probability density function.

$$F(x) = \int_{-\infty}^{x} f(s)\,ds$$

$$\frac{d}{dx}F(x) = \frac{d}{dx}\left[\int_{-\infty}^{x} f(s)\,ds\right] = f(x)$$

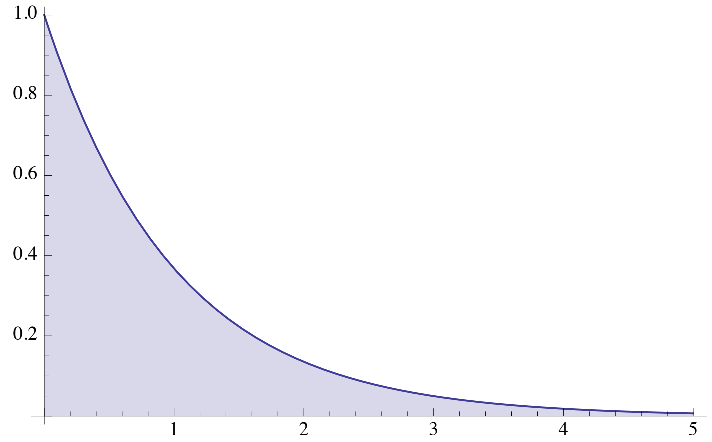
Probability distribution: Example

Exponential distribution

Probability density function (PDF). $f(x) = \lambda e^{-\lambda x}$

Cumulative distribution function (CDF). $F(x) = P(X \le x) = 1 - e^{-\lambda x}$

Survival function. $\bar{F}(x) = P(X > x) = 1 - F(x) = e^{-\lambda x}$



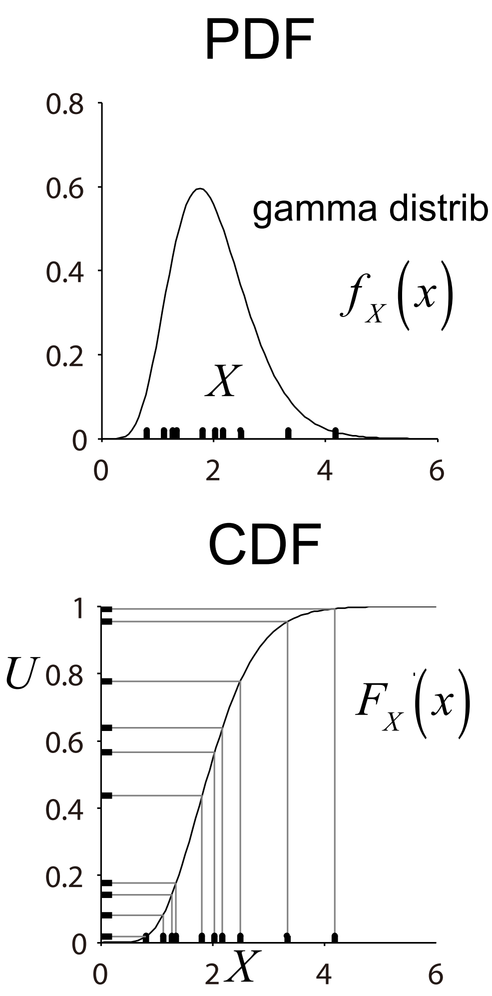
Expectation $E[X] = \lambda^{-1}$ Variance $\mathrm{var}[X] = \lambda^{-2}$

Samples from a distribution

If a random variable X follows a cdf F_X, then the r.v. U given by

$$U = F_X(X) = \int_{-\infty}^{X} f_X(s)\,ds$$

follows an uniform distribution in [0,1].



Homework 1-1

Problem

Confirm that $U = F_X\left(X\right)$ follows an uniform distribution in [0,1].
when X is a random variable from $F_X\left(x\right) = \int_0^x f_X\left(s\right)ds$

Hint

Suppose a random variable (r.v.) X given by a pdf $f_X\left(x\right)$

Consider a r.v. Y that is transformed from X
using a monotonic function g. $Y = g\left(X\right)$

The distribution of Y is given as $f_Y\left(y\right) = f_X\left(x\right)\left|\frac{dx}{dy}\right| = f_X\left(x\right)\left|g'\left(x\right)\right|^{-1}$

Consider a particular relation given by CDF, $g\left(x\right) = F_X\left(x\right) = \int_{-\infty}^x f_X\left(s\right)ds$

Then we obtain $f_Y\left(y\right) = f_X\left(x\right)\left|g'\left(x\right)\right|^{-1} = f_X\left(x\right)\left|f_X\left(x\right)\right|^{-1} = 1$

Samples from a distribution

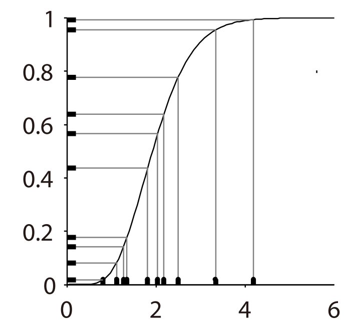
Inverse function method

Let U be an uniform r.v..

X that satisfies $U = \int_{-\infty}^{X} f_X(s)\,ds$ follows the density $f_X(x)$.

The r.v. $X = F^{-1}(U)$ follows the CDF given by F.

CDF



Exponential distribution: $f_X(x) = \lambda e^{-\lambda x}$

$$U = F_X(X) = \int_{-\infty}^{X} \lambda e^{-\lambda s}\,ds = 1 - e^{-\lambda X}$$

A r.v. that follows an exponential distribution with rate λ is generated using an uniform r.v. X as

$$X = F_X^{-1}(U) = -\lambda^{-1}\log(U)$$

Memo: The distribution that can be inverted analytically is limited, e.g., Exp., Weibull, Pareto, etc. For other distributions, one can numerically compute CDF to obtain *X*.

Introduction to

POISSON POINT PROCESS

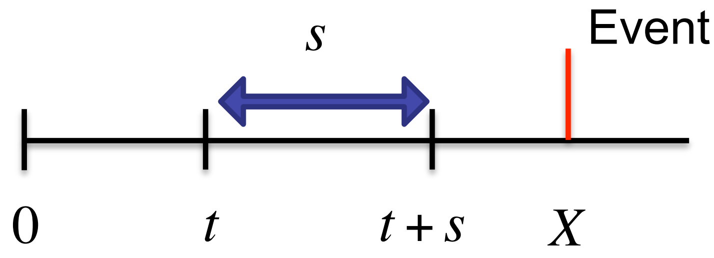
Memoryless process

A point process is memoryless if intervals X satisfy

$$P(X > s + t \mid X > t) = P(X > s)$$

Question.

Suppose that you bought a refrigerator *t* years ago. It still works well.
How long will it continue to work from now?



Answer.

We wish to know the probability that the refrigerator survives another s years given that it survived *t* years. This is given as $P(X > s + t \mid X > t)$.

From the memoryless property, it is equivalent to $P(X > s)$.

Thus, the fact that the refrigerator survived *t* years did not provide any information to predict a failure of the refrigerator in future.

Exponential distribution

Q. What distribution possesses the memoryless property?

Ans. The exponential distribution: $f\left(x\right)=\lambda e^{-\lambda x}$

The survival function can be written in a conditional form as

$$P\left(X>s+t\right)=P\left(X>s+t\,|\,X>t\right)P\left(X>t\right)$$

Using the memoryless property, we obtain

$$P\left(X>s+t\right)=P\left(X>s\right)P\left(X>t\right)$$

The exponential distribution satisfies the memoryless property.

If $P\left(X>x\right)=e^{-\lambda x}$, then $P\left(X>s\right)P\left(X>t\right)=e^{-\lambda s}e^{-\lambda t}$

$$=e^{-\lambda(s+t)}=P\left(X>s+t\right)$$

Density is given as $f\left(x\right)=-\dfrac{d}{dx}P\left(X>x\right)=\lambda e^{-\lambda x}$

Poisson point process

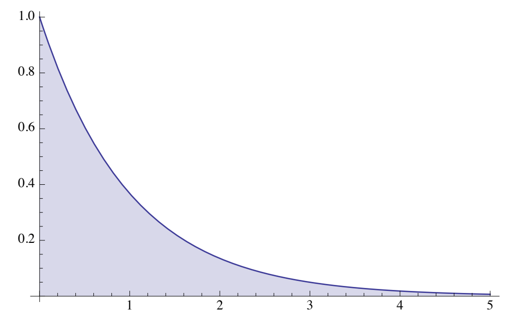
If intervals of events are samples from the exponential distribution, and the intervals are independent each other, then the process is memoryless.

This process is called a (homogeneous) **Poisson point process**.

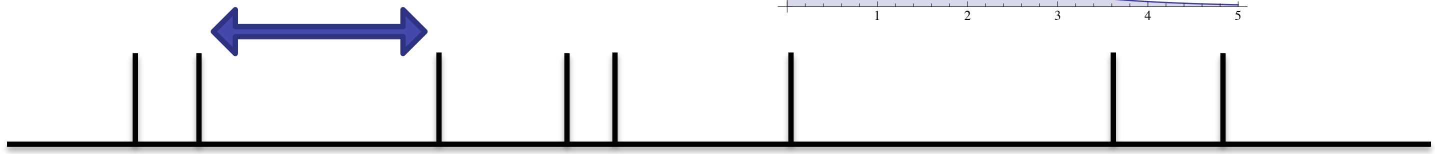
Simulating a Poisson process

Question. How can we simulate a Poisson point process?

Ans. Use the inverse function method to generate an inter-spike interval (ISIs) that follows an exponential distribution.



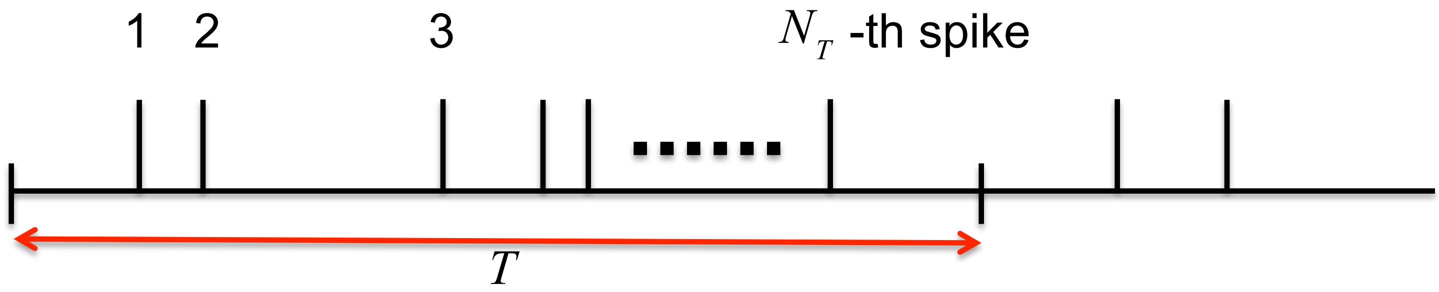
ISI $X = -\lambda^{-1} \log U$



The Poisson distribution

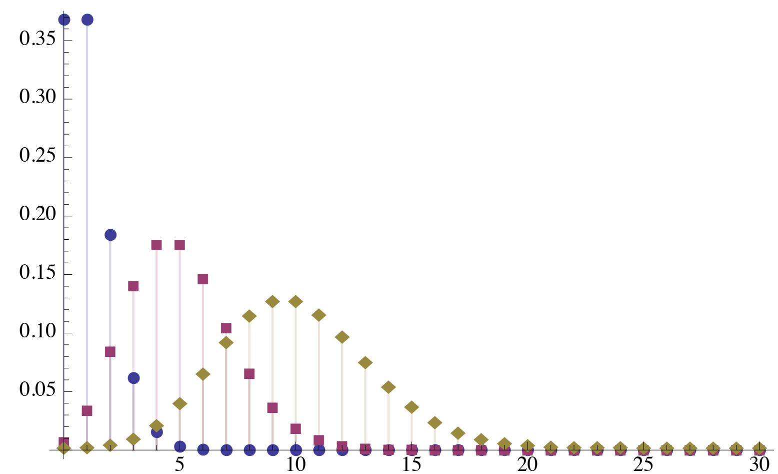
Let's consider a distribution of the number of spikes that occurred in T[s].

N_T : the number of spikes that happens in T [s]



A count distribution of a Poisson point process is given by a Poisson distribution.

$$P(N_T = n) = \frac{(\lambda T)^n}{n!} e^{-\lambda T}$$

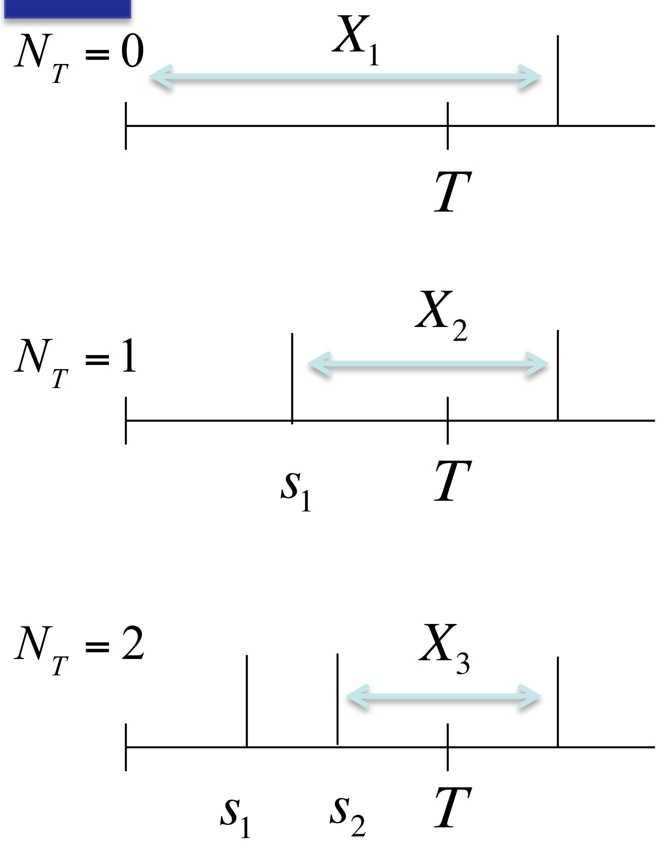


Homework 1-2

Problem

Derive the Poisson distribution for the first few spike counts, $N_T = 0, 1, 2.$

Hint



If there is no spike in T, then the first spike occurred after T. This is given by survival function of an exponential distribution.

$$P\left(N_T = 0\right) = P\left(X_1 > T\right) = \bar{F}\left(T\right)$$

Let the first spike occurred at s_1, the first ISI follows an exponential density, the second ISI should be greater than T-s_1.

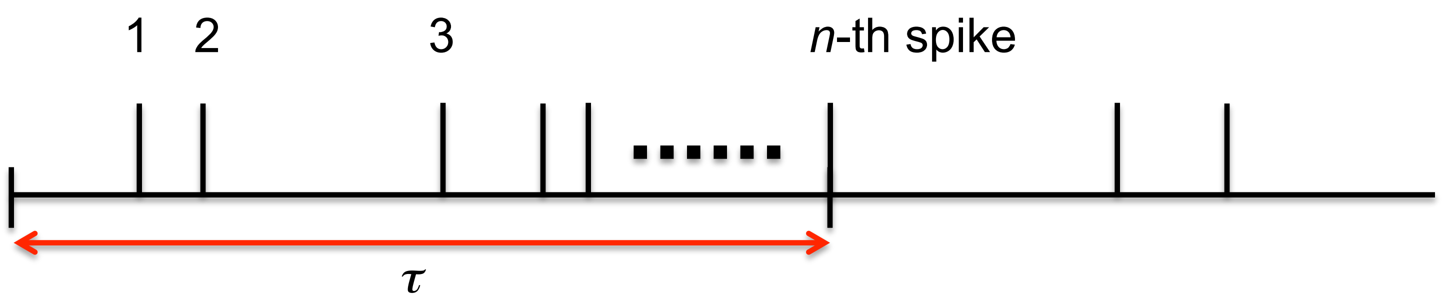
$$P\left(N_T = 1\right) = \int_0^T f\left(s_1\right)\bar{F}\left(T - s_1\right) ds_1$$

Similarly if you have two spikes in T,

$$P\left(N_T = 2\right) = \int_0^T \int_{s_1}^T f\left(s_1\right) f\left(s_2 - s_1\right)\bar{F}\left(T - s_2\right) ds_1\, ds_2$$

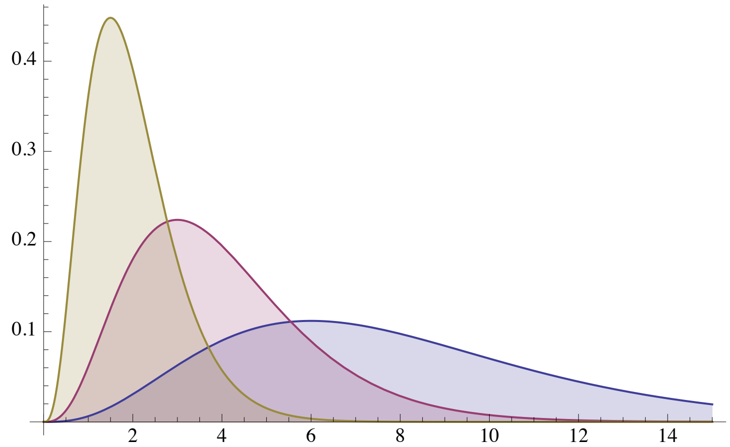
The Erlang distribution

Let's consider a distribution of a waiting time τ until spikes occurred n times.



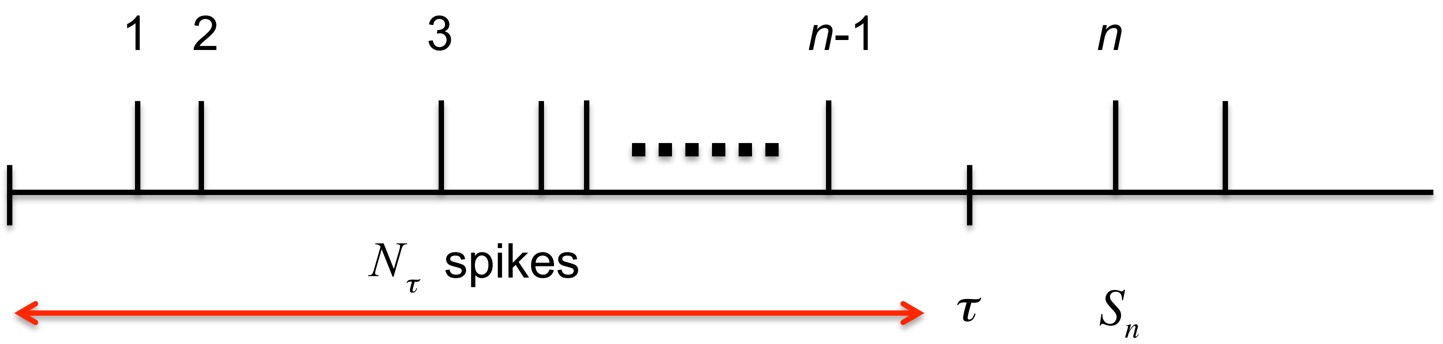
Waiting time of multiple spikes from a Poisson process is given by an Erlang distribution.

$$P\left(\tau < X < \tau + d\tau\right) = \frac{\lambda^n \tau^{n-1}}{\left(n-1\right)!} e^{-\lambda\tau} d\tau$$



Spike count and ISI

The relation between a Poisson and Erlang distribution.



If $S_n > \tau$, the number of spikes in is at most n-1.

$$P(N_\tau < n) = P(S_n > \tau)$$

Rhs is a survival function of an Erlang distribution.

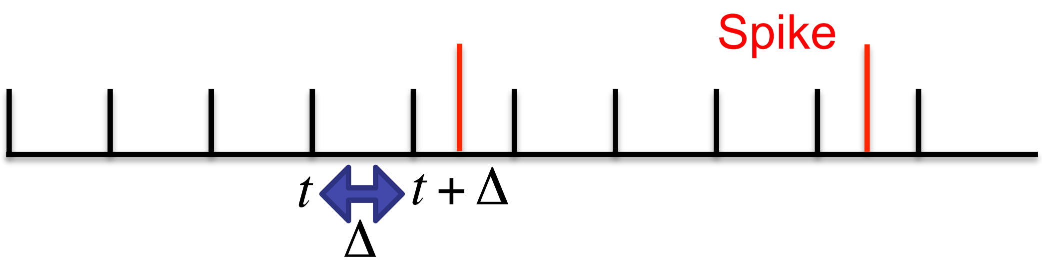
$$P(S_n > \tau) = \int_\tau^\infty \lambda^n t^{n-1} e^{-\lambda t}\, dt = \sum_{k=0}^{n-1} \frac{(\lambda\tau)^k}{k!} e^{-\lambda\tau}$$

$$P(N_\tau = n) = P(N_\tau < n+1) - P(N_\tau < n) = \frac{(\lambda\tau)^n}{n!} e^{-\lambda\tau}$$

Instantaneous spike rate

Alternative definition of a Poisson point process

Divide T into small N bins of a small width Δ (T=NΔ).



Instantaneous spike rate

$$P\left(\text{a spike in } [t, t+\Delta)\right) = \lambda\Delta + o(\Delta)$$

$$P\left(>1 \text{ spikes in } [t, t+\Delta)\right) = o(\Delta)$$

$$P\left(\text{no spikes in } [t, t+\Delta)\right) = 1 - \lambda\Delta + o(\Delta)$$

$o(\Delta)$ is a function that approaches to zero faster than Δ : $\lim_{\Delta \to +0} \frac{o(\Delta)}{\Delta} = 0$

Relation to the Poisson distribution

The probability of a spike/no spike is an approximation of the Poisson distribution for a small time bin.

The spike count in a small bin is approximated as

$$P\left(N_{\Delta}=n\right)=\frac{\left(\lambda\Delta\right)^{n}}{n!}e^{-\lambda\Delta}=\frac{\left(\lambda\Delta\right)^{n}}{n!}\left[1-\lambda\Delta+\frac{1}{2}\left(\lambda\Delta\right)^{2}+\cdots\right]$$

In particular,

$$P\left(N_{\Delta}=0\right)=1\left[1-\lambda\Delta+\frac{1}{2}\left(\lambda\Delta\right)^{2}+\cdots\right]=1-\lambda\Delta+o\left(\Delta\right)$$

$$P\left(N_{\Delta}=1\right)=\lambda\Delta\left[1-\lambda\Delta+\frac{1}{2}\left(\lambda\Delta\right)^{2}+\cdots\right]=\lambda\Delta+o\left(\Delta\right)$$

$$P\left(N_{\Delta}=2\right)=\left(\lambda\Delta\right)^{2}\left[1-\lambda\Delta+\frac{1}{2}\left(\lambda\Delta\right)^{2}+\cdots\right]=o\left(\Delta\right)$$

Exponential / Poisson distribution revisited

The exponential distribution from an instantaneous spike rate.

The probability that no spikes happened in N-1 bins and a spike happened in the last bin (geometric distribution). ($x=N\Delta$)

$$P(x < X < x+\Delta) = (1-\lambda\Delta)^{N-1}\lambda\Delta + o(\Delta)$$

ISI distribution: $$f(x) = \lim_{\Delta\to 0}\frac{P(x < X < x+\Delta)}{\Delta} = \lambda e^{-\lambda x}$$

The Poisson distribution from an instantaneous spike rate.

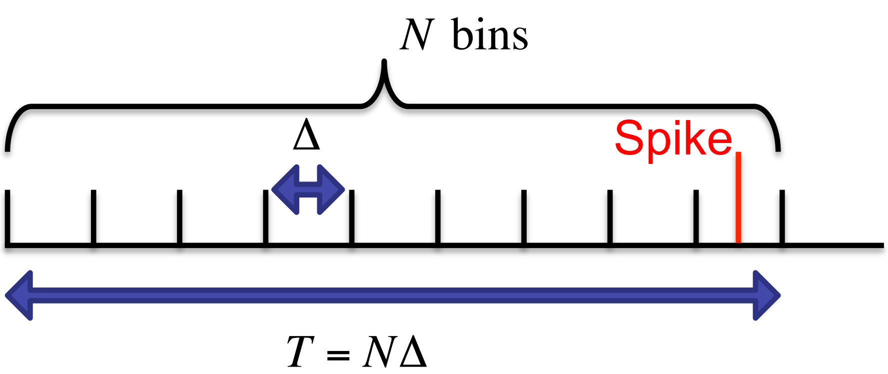
The probability that n spikes happened in N bins ($T=N\Delta$).

$$P(N_T = n) = \binom{N}{n}(1-\lambda\Delta)^{N-n}(\lambda\Delta)^n \to \frac{(\lambda T)^n}{n!}e^{-\lambda T} \text{ as } \Delta \to 0$$

Exponential distribution revisited

We can derive the exponential distribution from an instantaneous spike rate, λ [spike/sec]

Divide T into small N bins of a width Δ.



Probability

a spike in a bin: $\lambda\Delta$

>2 spikes in a bin: $o(\Delta)$

no spikes in a bin: $1-\lambda\Delta+o(\Delta)$

The probability that no spikes happened in N-1 bins and a spike happened in the last bin.

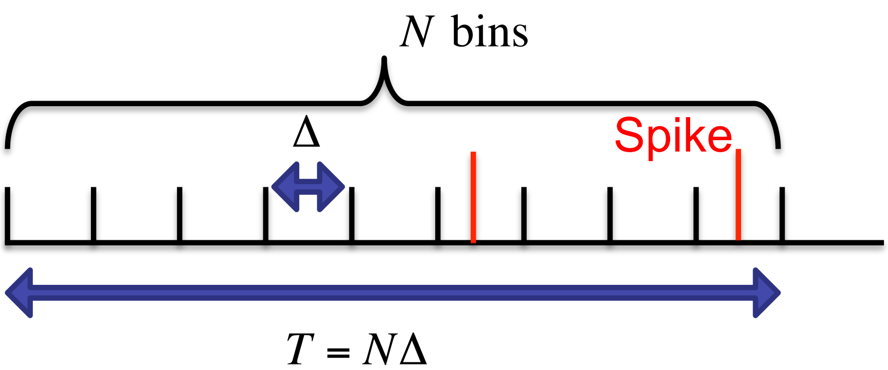
$$P(x<T<x+\Delta) = (1-\lambda\Delta)^{N-1}\lambda\Delta + o(\Delta)$$

ISI distribution: $f(x) = \lim_{\Delta\to 0}\frac{P(x<T<x+\Delta)}{\Delta} = \lambda e^{-\lambda x}$

Poisson distribution revisted

We can derive the Poisson distribution from an instantaneous spike rate, λ [spike/sec]

Divide T into small N bins of a width Δ.



Probability

a spike in a bin: $\lambda\Delta$

>2 spikes in a bin: $o(\Delta)$

no spikes in a bin: $1-\lambda\Delta+o(\Delta)$

The probability that n spikes happened in T [s].

$$P(N_T = n) = \begin{pmatrix} N \\ n \end{pmatrix} (1-\lambda\Delta)^{N-n} (\lambda\Delta)^n \rightarrow \frac{(\lambda T)^n}{n!} e^{-\lambda T} \text{ as } \Delta \rightarrow 0$$

Homework 1-3

Problem

The probability that no spikes happened in N-1 bins and a spike happened in the last bin is given as

$$P(x < X < x+\Delta) = (1-\lambda\Delta)^{N-1}\lambda\Delta + o(\Delta) = \frac{\lambda\Delta}{1-\lambda\Delta}(1-\lambda\Delta)^{N} + o(\Delta)$$

Show that the ISI distribution becomes an exponential distribution.

Hint

Approximate the terms by the Taylor expansions.

$$\frac{1}{1-x} = 1 + x + x^2 + x^3 \cdots$$

$$\frac{\lambda\Delta}{1-\lambda\Delta} = \lambda\Delta\left\{1 + \lambda\Delta + (\lambda\Delta)^2 + \cdots\right\} = \lambda\Delta + o(\Delta)$$

$$\log(1-x) = -x - \frac{1}{2}x^2 - \frac{1}{3}x^3 \cdots$$

$$\begin{aligned}(1-\lambda\Delta)^N &= \exp\left[N\log(1-\lambda\Delta)\right] \\ &= \exp\left[N\left\{-\lambda\Delta - \frac{1}{2}(\lambda\Delta)^2 - \cdots\right\}\right] \\ &= e^{-\lambda x}\left(1 - \frac{1}{2}\lambda^2 x\Delta + \cdots\right)\end{aligned}$$

Hence we obtain

$P(x < X < x+\Delta) = \lambda\Delta\exp[-\lambda x] + o(\Delta)$, and $f(x) = \lim_{\Delta\to 0}\frac{P(x < X < x+\Delta)}{\Delta} = \lambda e^{-\lambda x}$

Homework 1-4

Problem Probability that n spikes occur in N bins (T [s]) is given as

$$P(N_T = n) = \binom{N}{n} (\lambda\Delta)^n (1-\lambda\Delta)^{N-n}$$

Show that the probability becomes the Poisson distribution for large N and small Δ with a relation given by a constant T=NΔ.

Hint

Using an approximation given by the Taylor expansions (See previous slides).

$$P(N_T = n) = \frac{N!}{(N-n)!n!}\left(\frac{\lambda\Delta}{1-\lambda\Delta}\right)^n (1-\lambda\Delta)^N \approx \frac{N!}{(N-n)!n!}(\lambda\Delta)^n \exp[-\lambda\Delta N]$$

Use the relation, T=NΔ, to obtain

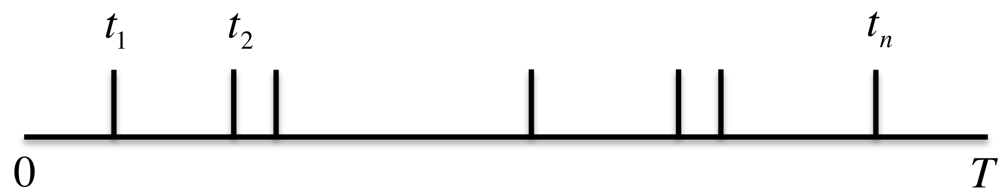
$$P(N_T = n) = \frac{N!}{(N-n)!N^n}\frac{1}{n!}(\lambda T)^n \exp[-\lambda T]$$

Use the Stirling's formula $\ln N! \sim N \ln N - N$

and prove that $\ln \frac{N!}{(N-n)!N^n} = \left(1-\frac{n}{N}\right)\ln\left(1-\frac{n}{N}\right)^N - n = 1 \cdot \ln e^{-n} - n \to 0$

Likelihood function

Likelihood function



$$p\left(t_1,t_2,\ldots,t_n \cap N_T = n\right) = \lambda^n \exp\left[-\lambda T\right]$$

Proof

$$p\left(t_1,t_2,\ldots,t_n \cap N_T = n\right)\Delta^n$$

$$= f\left(t_1\right)\Delta\prod_{i=2}^{n} f\left(t_i - t_{i-1}\right)\Delta \cdot P\left(t_{n+1} > T \mid t_n\right)$$

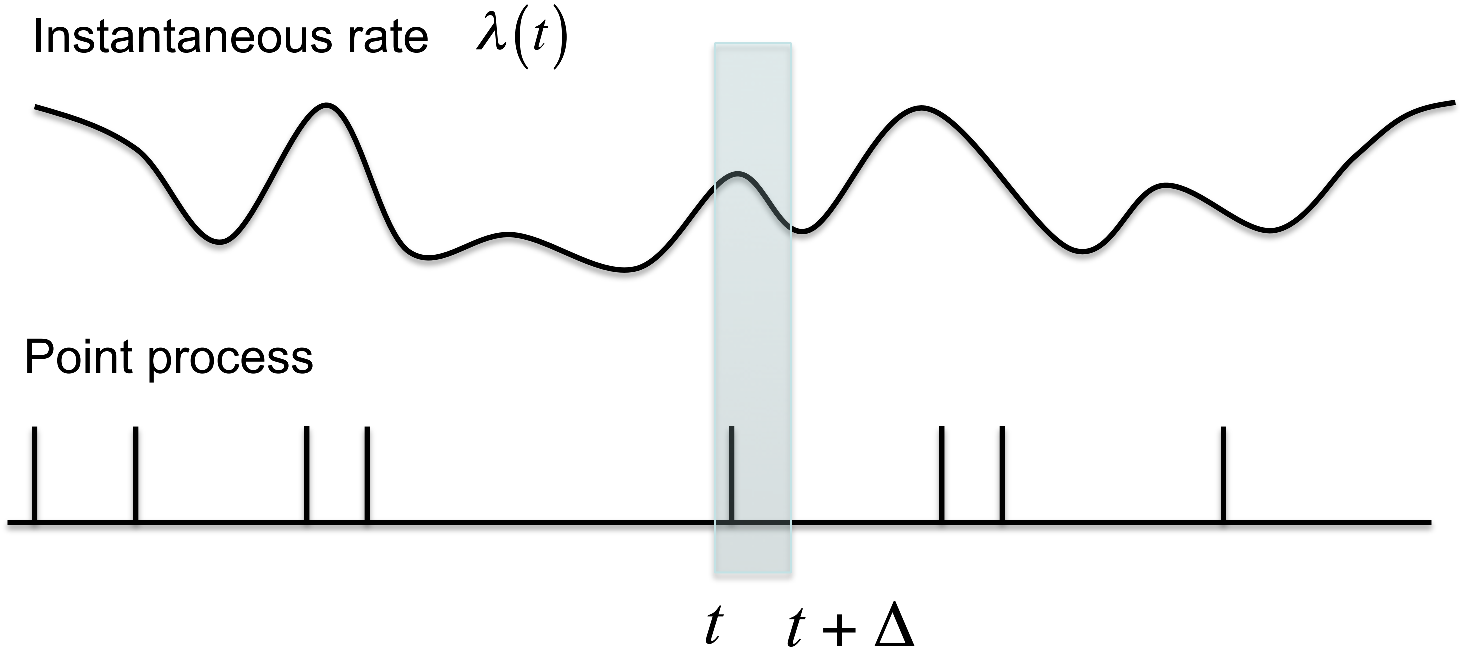
$$= \lambda\exp\left(-\lambda t_1\right)\Delta\prod_{i=2}^{n}\lambda\exp\left[-\lambda\left(t_i - t_{i-1}\right)\right]\Delta \cdot \exp\left[-\lambda\left(T - t_n\right)\right]$$

$$= \lambda^n \Delta \exp\left[-\lambda T\right]$$

Here the ISI distribution is $f\left(x\right) = \lambda\exp\left[-\lambda x\right]$

INHOMOGENEOUS POISSON PROCESS

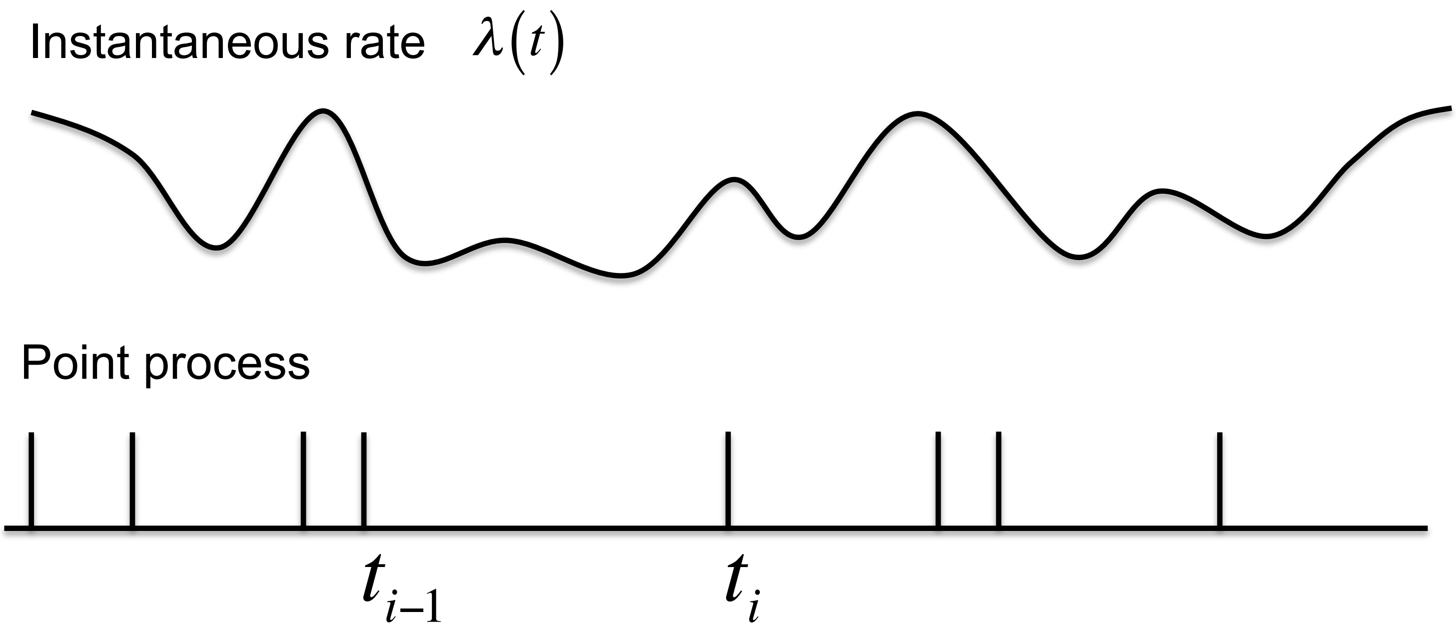
Inhomogeneous Poisson process



$$P\left(\text{a spike in } [t, t+\Delta)\right) = \lambda(t)\Delta + o(\Delta)$$

↑ Time-dependent rate

Inhomogeneous Poisson process



$$f(t \mid t_{i-1}) = \lambda(t)\exp\left[-\int_{t_{i-1}}^{t} \lambda(u)\,du\right] \quad \text{for} \quad t > t_{i-1}$$

Note that the above formula extends the exponential ISI of homogeneous Poisson process.

Homework 1-5

Problem Derive the ISI density of the inhomogeneous Poisson process.

Hint Let the last spike occurred at time 0.

The probability of a spike occurrence in [t, t+Δ) is given as

$$P(t < X < t+\Delta) = \prod_{k=1}^{N-1}\left[1-\lambda(k\Delta)\Delta\right]\cdot\left[\lambda(N\Delta)\Delta\right] + o(\Delta)$$

$$= \frac{\lambda(N\Delta)\Delta}{1-\lambda(N\Delta)\Delta}\cdot\prod_{k=1}^{N}\left[1-\lambda(k\Delta)\Delta\right] + o(\Delta)$$

Using the following approximation,

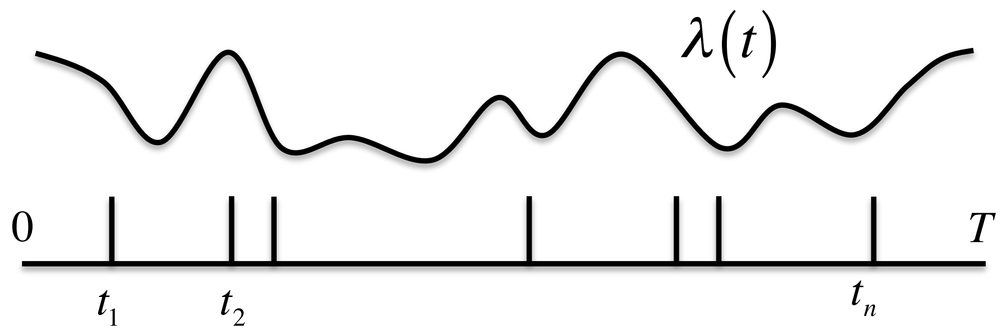
$$\prod_{k=1}^{N}\left[1-\lambda(k\Delta)\Delta\right] = \exp\left[\sum_{k=1}^{N}\log\left\{1-\lambda(k\Delta)\Delta\right\}\right] = \exp\left[\sum_{k=1}^{N}\left\{-\lambda(k\Delta)\Delta + o(\Delta)\right\}\right]$$

$$\frac{\lambda(N\Delta)\Delta}{1-\lambda(N\Delta)\Delta} = \lambda(N\Delta)\Delta + o(\Delta)$$

The ISI density is given as $f(t) = \lim_{\Delta\to 0}\frac{P(t < X < t+\Delta)}{\Delta} = \lambda(t)\exp\left[-\int_0^t \lambda(u)\,du\right]$

Likelihood function

Likelihood function



$$p\left(t_1,t_2,\ldots,t_n \cap N_T = n\right) = \prod_{i=1}^{n} \lambda(t_i) \exp\left[-\int_0^T \lambda(u)\,du\right]$$

Proof

$$p\left(t_1,t_2,\ldots,t_n \cap N_T = n\right)\Delta^n = f(t_1)\Delta \prod_{i=2}^{n} f(t_i \mid t_{i-1})\Delta \cdot P(t_{n+1} > T \mid t_n)$$

$$= \prod_{i=1}^{n} \lambda(t_i)\Delta \exp\left[-\int_0^T \lambda(u)\,du\right]$$

Here the ISI distribution is given as $f(t_i \mid t_{i-1}) = \lambda(t_i)\exp\left[-\int_{t_{i-1}}^{t_i} \lambda(u)\,du\right]$

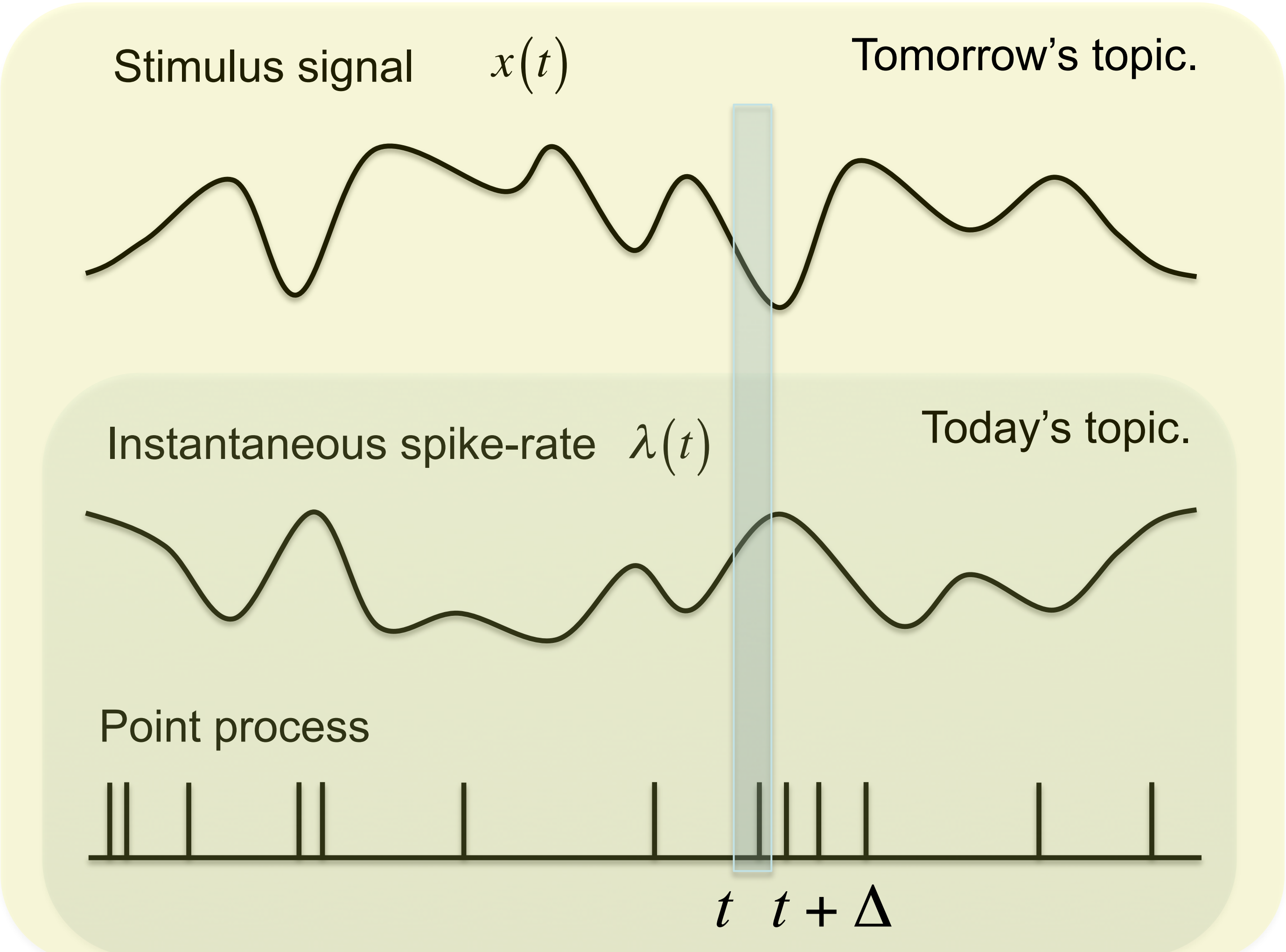
What we learned

1. **Memoryless property** of a process.
2. **Exponential** distribution = **Poisson process.**
3. Poisson count distribution, Waiting time distribution (Erlang).
4. Definition of a Poisson process using **instantaneous firing rate**.
5. Exponential and a Poisson count distribution revisited.
6. **Inhomogeneous Poisson process**: ISI distribution and likelihood

Tomorrow, we will learn

1. **Renewal process: Conditional intensity function** and **ISI density**.
2. **non-Poisson process** (Point process written by the conditional intensity function)
3. **Time-rescaling theorem**
4. **How to simulate a point process via the time-rescaling theorem.**
5. **How to assess a point process model** via the time-rescaling theorem. (Q-Q plot and K-S test)

Follow-up from the lecture feedback



$$P(\text{a spike in } [t, t+\Delta)) = \lambda(t)\Delta + o(\Delta)$$

Likelihood function

Density function $f(x) = \lambda e^{-\lambda x}$

An observed sample X

Likelihood function $L(\lambda) = f(X) = \lambda e^{-\lambda X}$

Multiple observation $X_1, X_2, \cdots, X_n$

Likelihood function $L(\lambda) = \prod_{i=1}^{n} p(X_i \mid \lambda)$

Maximum likelihood estimation (MLE)

$$\lambda_{\mathrm{MLE}} = \underset{\lambda}{\mathrm{argmax}} \prod_{i=1}^{n} p(X_i \mid \lambda)$$

Likelihood of a Poisson process

Likelihood function of a homogeneous Poisson process

Inter-spike Intervals (ISIs)

$$p\left(t_1,t_2,\ldots,t_n \cap N_T = n \mid \lambda\right) = p\left(t_1,t_2-t_1,\ldots,t_n-t_{n-1} \cap N_T = n\right)$$

$$= \prod_{i=1}^{n} f\left(t_i - t_{i-1};\lambda\right)\cdot \bar{F}\left(T-t_n;\lambda\right)$$

$$f\left(t_i - t_{i-1}\right) = \lambda e^{-\lambda\left(t_i - t_{i-1}\right)}$$

$$= \lambda^n \exp\left[-\lambda T\right]$$

Likelihood function of a inhomogeneous Poisson process

$$p\left(t_1,t_2,\ldots,t_n \cap N_T = n \mid \lambda_{0:T}\right) = p\left(t_1,t_2-t_1,\ldots,t_n-t_{n-1} \cap N_T = n\right)$$

$$= \prod_{i=1}^{n} f\left(t_i - t_{i-1};\lambda_{t_i:t_{i-1}}\right)\cdot F\left(T-t_i;\lambda_{t_n:T}\right)$$

$$f\left(t_i - t_{i-1}\right) = \lambda\left(t_i\right)\exp\left[-\int_{t_{i-1}}^{t_i}\lambda\left(u\right)du\right]$$

$$= \prod_{i=1}^{n}\lambda\left(t_i\right)\exp\left[-\int_0^T \lambda\left(u\right)du\right]$$