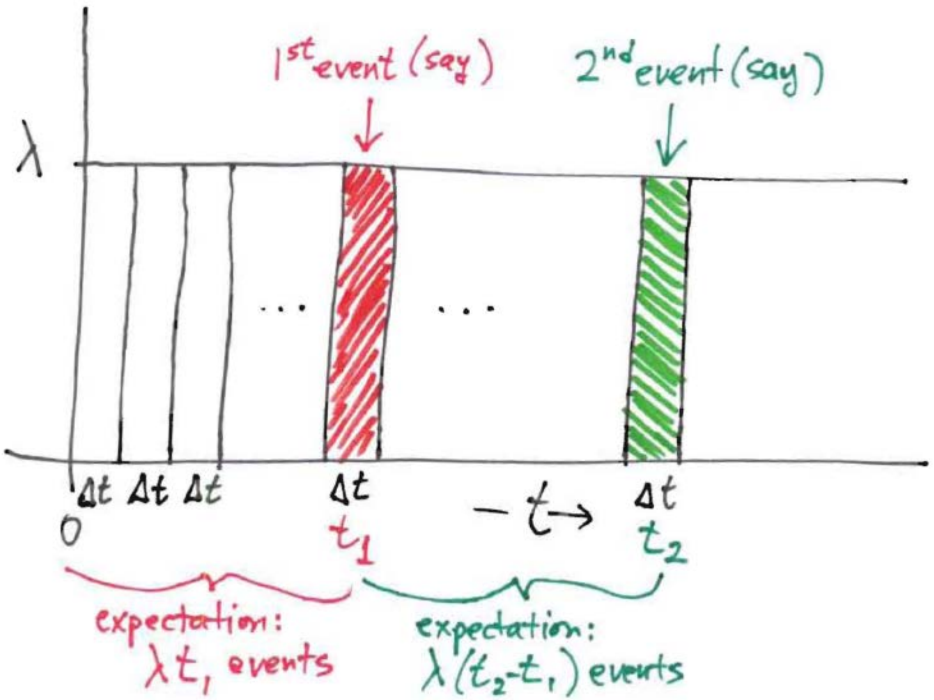
Opinionated Lessons in Statistics

by Bill Press

#15.5 Poisson Processes and Order Statistics

Professor William H. Press, Department of Computer Science, the University of Texas at Austin

In a "constant rate Poisson process", independent events occur with a constant probability per unit time



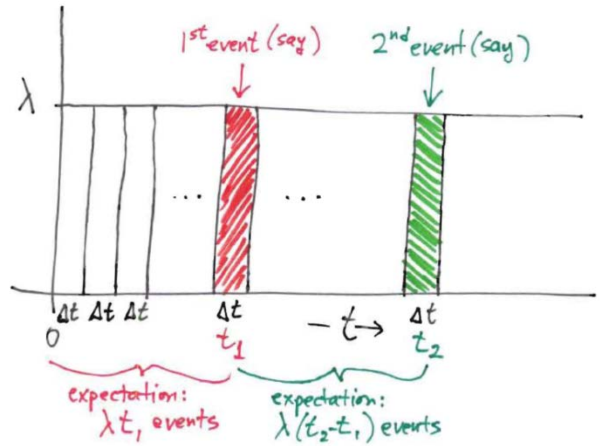
In any small interval Δt, the probability of an event is $\lambda\Delta t$

In any finite interval T, the mean (expected) number of events is λT

What is the probability distribution of the waiting time to the 1st event, or between events?

It's the product of $t/\Delta t$ "didn't occur" probabilities, times one "did occur" probability.

$$
\begin{aligned}
p_{T_1}(t)\Delta t &= [1-\lambda\,\Delta t]^{t/\Delta t}\,\lambda\Delta t \\
&= e^{\frac{t}{\Delta t}\,\ln[1-\lambda\,\Delta t]}\lambda\Delta t \\
&\approx \lambda e^{-\lambda t}\Delta t
\end{aligned}
$$



(random variable T_1 with value t)

$$p_{T_1}(t) = \lambda e^{-\lambda t} = \lambda P_{\text{Poisson}}(0|\lambda t)$$

$$t_1 \sim \text{Exponential}(\lambda)$$

Get it? It's the probability that 0 events occurred in a Poisson distribution with λt mean events up to time t, times the probability (density) of one event occurring at time t.

Once we understand the relation to Poisson, we immediately know the waiting time to the k^{th} event

$$p_{T_k}(t) = \lambda P_{\text{Poisson}}(k-1|\lambda t)$$
$$= \lambda \frac{(\lambda t)^{k-1}}{(k-1)!} e^{-\lambda t}$$
$$= \frac{\lambda^k t^{k-1}}{\Gamma(k)} e^{-\lambda t}$$

$$t_k \sim \text{Gamma}(k, \lambda)$$

It's the probability that k-1 events occurred in a Poisson distribution with λt mean events up to time *t*, times the probability (density) of one event occurring at time t.

We could also prove this with characteristic functions (sum of k independent waiting times):

```
pexpon = lam Exp[-lam x]

pexponCF = Integrate[pexpon Exp[I t x],
  {x, 0, Infinity},
  GenerateConditions → False]
```

$$\frac{\text{lam}}{\text{lam} - i\,t}$$

```
pgamm = lam^k x^(k - 1) Exp[-lam x] / Gamma[k]

pgammCF = Integrate[pgamm Exp[I t x],
  {x, 0, Infinity},
  GenerateConditions → False]
```

$\text{lam}^k (\text{lam} - i\,t)^{-k}$ k^{th} power of above !

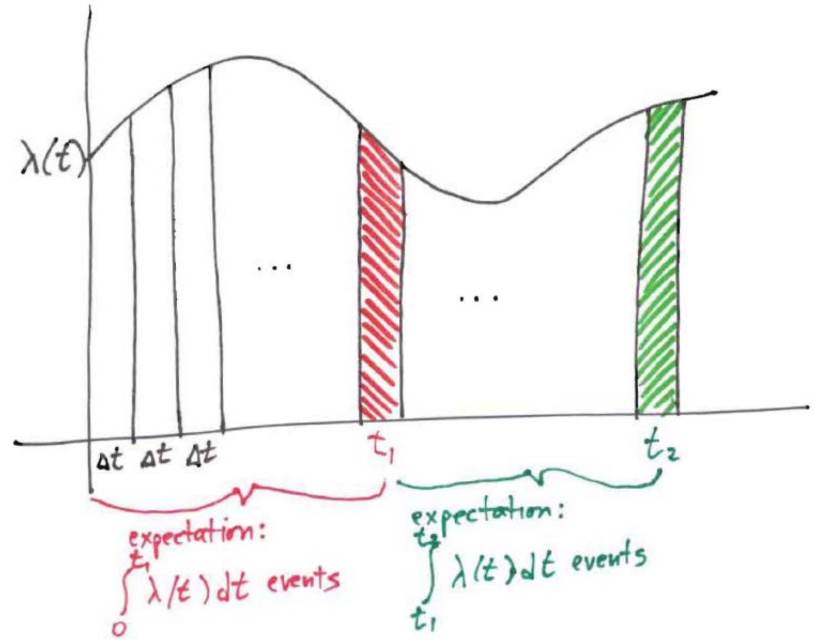
"Waiting time to the k^{th} event in a Poisson process is Gamma distributed with degree k."

Same ideas go through for a "variable rate Poisson process"

Waiting time to first event or between events:

$$
\begin{aligned}
p_{T_1}(t) &= \lambda(t) \prod_j [1 - \lambda(t_j)\,\Delta t] \\
&= \lambda(t)\, e^{\sum_j \ln[1-\lambda(t_j)\Delta t]} \lambda(t) \\
&\approx \lambda(t)\, e^{-\sum_j \lambda(t_j)\Delta t} \\
&= \lambda(t)\, e^{-\int_0^t \lambda(t) dt} \\
&\equiv \lambda(t)\, e^{-\Lambda(t)}
\end{aligned}
$$

where $\Lambda(t) \equiv \int_0^t \lambda(t) dt$



so basically the area Λ(t) replaces the area λt

Thus, waiting time for the kth event in a variable rate Poisson process is…

$$p_{T_k}(t) = \lambda(t)\ P_{\text{Poisson}}[k-1|\Lambda(t)]$$
$$= \lambda(t)\ \frac{[\Lambda(t)]^{k-1}}{(k-1)!} e^{-\Lambda(t)}$$

$$\Lambda(t) \equiv \int_0^t \lambda(t)dt$$

Notice the $\lambda(t)$ – not $\Lambda(t)$ – in front. So, in this form it is not Gamma distributed.
We can recover the Gamma if we compute the probability density of $\Lambda(t)$

$$p_{\Lambda_k}(\Lambda) = p_{T_k}(t)\frac{dt}{d\Lambda}$$

← $1/\lambda(t)$

$$= \frac{[\Lambda(t)]^{k-1}}{(k-1)!} e^{-\Lambda(t)}$$

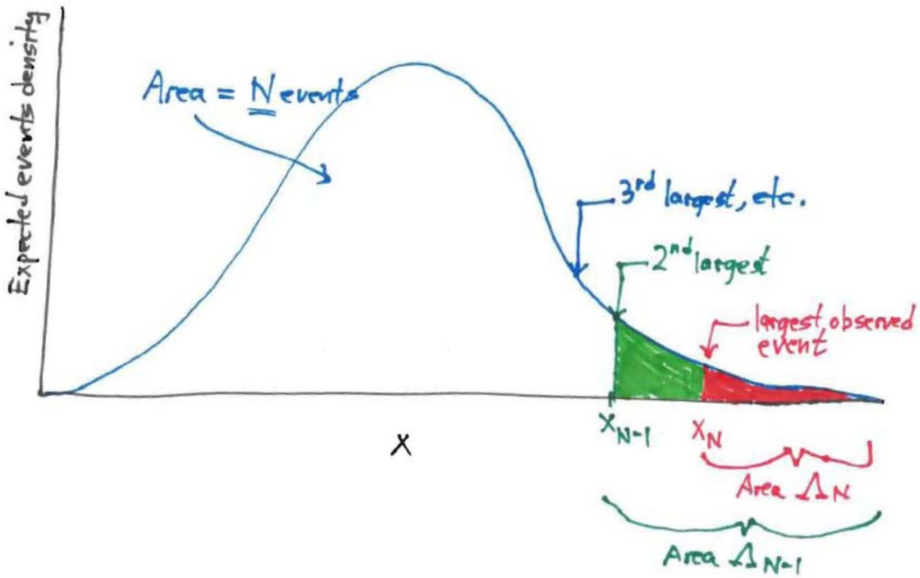
So the "area (mean events) up to the kth event" is Gamma distributed,

$$\Lambda_k \sim \text{Gamma}(k,1)$$

How to simulate variable rate Poisson: Draw from Gamma(1,1) [or Exponential(1) which is the same thing] and then advance through that much area under $\lambda(t)$. That gives the next event.

What does this have to do with "order statistics"?

If *N* i.i.d. numbers are drawn from a univariate distribution, the kth order statistic is the probability distribution of the kth largest number.



Near the ends, with $k \ll N$ or $N - k \ll N$, this is just like variable-rate Poisson.

How to simulate order statistics (approximation near extremes, large N): Draw from Exponential(1) and then advance through that much area under the distribution of expected number of events (total area N). That gives the next event.

Order statistics: the exact result

The previous approximation is just the approximation of Binomial by Poisson (for large N and small k). Instead of what we had before,

$$p_{T_k}(t) = \lambda(t)\; P_{\text{Poisson}}[k-1|\Lambda(t)]$$

a moment of thought gives the exact result,

$$p_{T_k}(t) = \lambda(t)\; P_{\text{Binomial}}[k-1\,|\,N, \Lambda(t)/N]$$

- $\lambda(t)$: probability density of the event
- $k-1$: number of events in the tail
- N: total number of events
- $\Lambda(t)/N$: binomial probability p for each event being in the tail

Beta has the same relation to Binomial as Gamma (or Exponential) has to Poisson:

$$\lim_{N\to\infty} N\text{Beta(1,N)} = \text{Exponential}(1)$$
$$\lim_{N\to\infty} N\text{Beta(k,N)} = \text{Gamma}(k,1)$$

How to simulate order statistics (exact):
Draw from Beta(1,N), giving a value between 0 and 1. Multiply by N. Advance through that much area under the distribution of expected number of events (total area N). That gives the next event. Decrement N by 1. Repeat.