
Poisson process problems 07-12-17

N. T. Gladd

Initialization: Be sure the files *NTGStylesheet2.nb* and *NTGUtilityFunctions.m* is are in the same directory as that from which this notebook was loaded. Then execute the cell immediately below by mousing left on the cell bar to the right of that cell and then typing “shift” + “enter”. Respond “Yes” in response to the query to evaluate initialization cells.

```
SetDirectory[NotebookDirectory[]];
(* set directory where source files are located *)
SetOptions[EvaluationNotebook[], (* load the StyleSheet *)
  StyleDefinitions -> Get["NTGStylesheet2.nb"]];
Get["NTGUtilityFunctions.m"]; (* Load utilities package *)
```

Purpose

Some recent issues involving option trading got me interested in Poisson processes and problems. To gain perspective and become more familiar with the topic I worked through some basic derivations for Poisson processes and solve some problems I found on the web.

Basics

- Visualization of Poisson process
- Derivation of Poisson Process from recursive differential equations
- Derivation of Poisson Process from Bernoulli Trials with a low probability of success
- Derivation of Poisson Process from fact that interarrival times are exponentially distributed

Problems

- 1 Waiting for the doctor
- 2 Covariance, correlation
- 3 Conditional Poisson distribution
- 4 Thinning of Poisson process
- 5 Memoryless property of exponential distribution
- 6 Distribution of $\min[X, Y]$, $\max[X, Y]$ for exponential distribution
- 7 Fire Brigade
- 8 Conditional Poisson arrivals
- 9 Total waiting time
- 10 Poisson shocks
- 11 Hen crossing road
- 12 Waiting for a reward

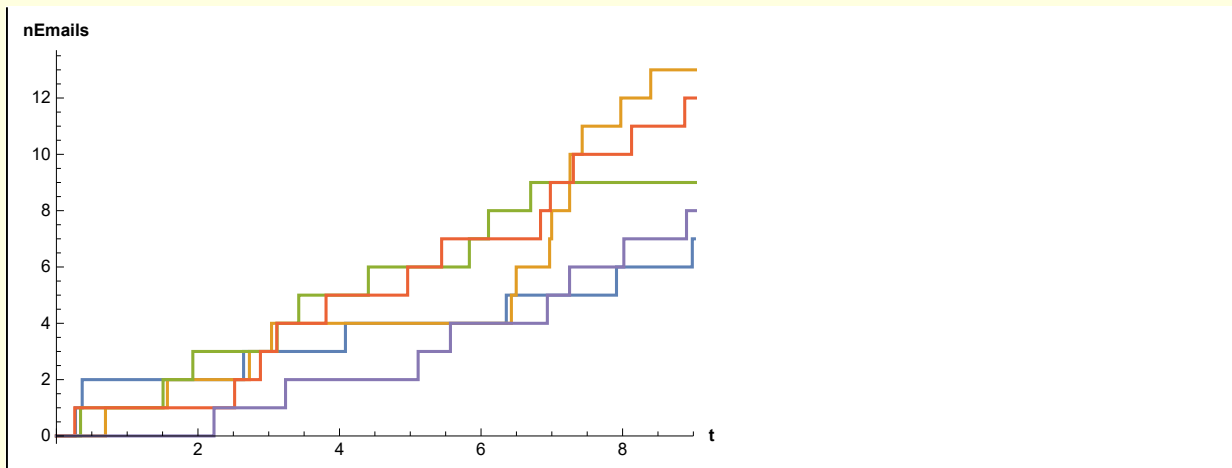
- 13 Trucks and cars
- 14 Light bulb failures
- 15 Three actors
- 16 Used car
- 17 Comparing two Poisson processes
- 18 Arriving students
- 19 Distribution of relative arrivals of two Poisson processes
- 20 Probability of reaching a point in space
- 21 London Bombs
- 22 Store Credit (compound Poisson process)
- 23 Poisson and the Law

Basics: Visualization and three different derivations of Poisson processes

Poisson processes are widely used as stochastic models for random arrivals. They have been extensively studied and a wealth of material related to them is readily available on the web. I will not attempt to provide a comprehensive or even a systematic presentation of the topic but focus here on calculations, making use of Mathematica capabilities whenever convenient (see Mathematica documentation for *PoissonProcess* and *PoissonDistribution*).

Visualization of Poisson process

To set the stage, suppose that emails arrive in your in-box at a rate of approximately one per hour. A quick simulation shows what the record of one week's arrivals during a 9 hour work day might look like.

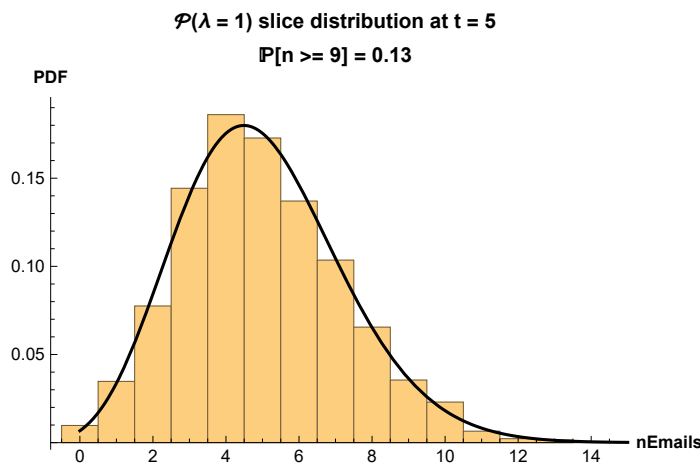


One day you arrive at the office just after lunch (5 hours after start of work). What is the distribution of emails you would expect and the probability that 8 or more emails are waiting for you?

```

Module[{λ = 1, tMax = 8, nSamples = 4000, tSlice = 5, ϕ, path, sliceData, lab, g},
  ϕ = PoissonProcess[λ];
  path = RandomFunction[ϕ, {0, tMax}, nSamples];
  sliceData = path["SliceData", tSlice];
  lab = Stl@StringForm["P(λ = ``) slice distribution at t = ``\nP[n >= 9] = ``",
    λ, tSlice, NF2@N@Probability[n ≥ 8, n ≈ ϕ[tSlice]]];
  g[1] = Histogram[sliceData, {-1/2, Max[sliceData], 1}, "PDF",
    PlotLabel → lab, AxesLabel → {Stl["nEmails"], Stl["PDF"]};
  g[2] = Plot[Evaluate@PDF[ϕ[tSlice], x], {x, Min[sliceData], Max[sliceData]},
    PlotRange → All, PlotStyle → Black];
  Show[{g[1], g[2]}]

```



Derivations of Poisson Processes

I Derivation of Poisson Process from recursive differential equations

A key assumption of the Poisson Process is that the probability of an arrival event is linearly proportional to the elapsed time since the last event

$$\mathbb{P}(\mathcal{N}(\Delta t) = 1) = \Delta t \lambda + O(\Delta t)$$

The second term represents the probability that two or more events arrived during the interval Δt . For a standard Poisson process, that probability is assumed to be negligibly small.

To derive differential equations for the probabilities of Poisson events, I start with the probability that no events occur during an interval Δt .

$$\mathbb{P}[\mathcal{N}(t + \Delta t) = 0] = \mathbb{P}[\mathcal{N}(t) = 0 \cap \mathcal{N}(t + \Delta t) - \mathcal{N}(t) = 0]$$

$$= \mathbb{P}[\mathcal{N}(t) = 0] \mathbb{P}[\mathcal{N}(t + \Delta t) - \mathcal{N}(t) = 0]$$

$$= \mathbb{P}[\mathcal{N}(t) = 0] \mathbb{P}[\mathcal{N}(\Delta t) = 0] \tag{1}$$

$$= \mathbb{P}[\mathcal{N}(t) = 0] (1 - \mathbb{P}[\mathcal{N}(\Delta t) \leq 0])$$

$$= \mathcal{P}_0(t) (1 - \lambda \Delta t + O(\Delta t))$$

where the short hand notation $\mathcal{P}_0(t) = \mathbb{P}[\mathcal{N}(t) = 0]$ has been introduced. The first and last expressions can be rearranged

$$\mathcal{P}_0(t + \Delta t) = \mathcal{P}_0(t) (1 - \lambda \Delta t + O(\Delta t))$$

implies

$$\frac{\mathcal{P}_0(t + \Delta t) - \mathcal{P}_0(t)}{\Delta t} = -\lambda \quad (2)$$

where the term associated with two or more events is neglected. In the limit $\Delta t \rightarrow 0$

$$\mathcal{P}_0(t + \Delta t) = \mathcal{P}_0(t) (1 - \lambda \Delta t + O(\Delta t))$$

implies

$$\lim_{\Delta t \rightarrow 0} \frac{\mathcal{P}_0(t + \Delta t) - \mathcal{P}_0(t)}{\Delta t} = \frac{d\mathcal{P}_0}{dt} = -\lambda \quad (3)$$

The initial condition for this differential equation is that the probability of no events at time $t = 0$ is 1.

That is $\mathcal{P}_0(0) = 1$.

By analogous logic and reasoning an equation can be derived for the probability of n events.

$$\mathbb{P}[\mathcal{N}(t + \Delta t) = n] =$$

$$\mathbb{P}[\mathcal{N}(t) = 0] \mathbb{P}[\mathcal{N}(t + \Delta t) - \mathcal{N}(t) = 0] +$$

$$\mathbb{P}[\mathcal{N}(t) = 1] \mathbb{P}[\mathcal{N}(t + \Delta t) - \mathcal{N}(t) = 1] + \mathbb{P}[\mathcal{N}(t) \geq 2] \mathbb{P}[\mathcal{N}(t + \Delta t) - \mathcal{N}(t) \leq 2]$$

$$= \mathcal{P}_n(t) (1 - \lambda \Delta t + O(\Delta t)) + \mathcal{P}_{n-1}(t) (\lambda \Delta t) + \mathcal{P}_{n-2}(t) O(\Delta t) \quad (4)$$

or

$$\frac{d\mathcal{P}_n(t)}{dt} = -\lambda \mathcal{P}_n(t) + \lambda \mathcal{P}_{n-1}(t)$$

I use Mathematica to solve these odes.

```
w1[1] = DSolve[{D[P0[t], t] == -λ P0[t], P0[0] == 1}, P0[t], t][[1, 1]]
```

```
P0[t] -> e^{-t λ}
```

```
w1[2] =
```

```
DSolve[{D[Pn[t], t] == -λ Pn[t] + λ P_{n-1}[t], Pn[0] == 0}, Pn[t], t][[1, 1]] /. K[1] -> u
```

```
Pn[t] -> -e^{-t λ} \left( \int_1^0 e^{u λ} λ P_{-1+n}[u] du - \int_1^t e^{u λ} λ P_{-1+n}[u] du \right)
```

```
w1[3] = w1[2] /. n -> 1 /. (w1[1] /. t -> u) // Simplify
```

$$\mathcal{P}_1[t] \rightarrow e^{-t\lambda} t \lambda$$

```
w1[4] = w1[2] /. n -> 2 /. (w1[3] /. t -> u) // Simplify
```

$$\mathcal{P}_2[t] \rightarrow \frac{1}{2} e^{-t\lambda} t^2 \lambda^2$$

```
w1[5] = w1[2] /. n -> 3 /. (w1[4] /. t -> u) // Simplify
```

$$\mathcal{P}_3[t] \rightarrow \frac{1}{6} e^{-t\lambda} t^3 \lambda^3$$

The pattern is clear

$$\mathcal{P}_n(t) = e^{-\lambda t} \frac{(\lambda t)^n}{n!}$$

This probability distribution is built-in to Mathematica

```
{PDF[PoissonProcess[\lambda][t], n], PDF[PoissonDistribution[\lambda t], n]}
```

$$\left\{ \left\{ \begin{array}{ll} \frac{e^{-t\lambda} (t\lambda)^n}{n!} & n \geq 0 \\ 0 & \text{True} \end{array} \right\}, \left\{ \begin{array}{ll} \frac{e^{-t\lambda} (t\lambda)^n}{n!} & n \geq 0 \\ 0 & \text{True} \end{array} \right\} \right\}$$

2 Derivation of Poisson Process from Bernoulli Trials with a low probability of success

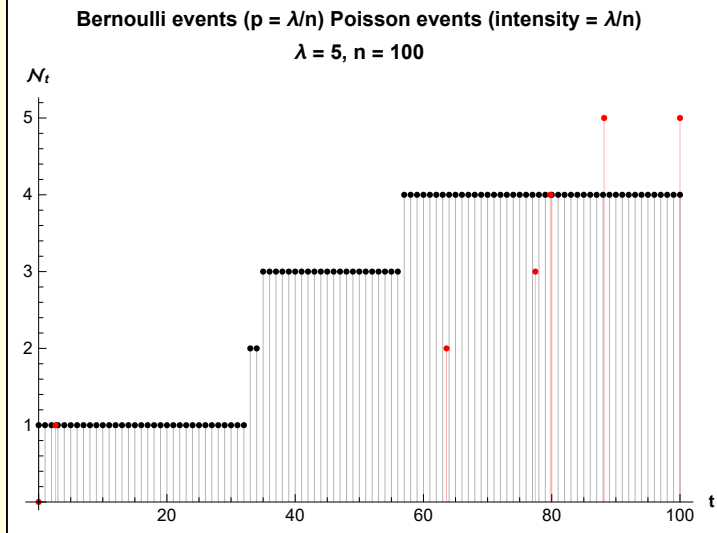
If successful Bernoulli events are counted and interpreted as arrival events, then the distribution of such Bernoulli arrivals approximates Poisson arrivals in the case that an individual Bernoulli success has low probability. A simulation illustrates this idea

```

Module[{n = 100, λ = 5, t = 100, p, B, P, pathB, cpathB, ppath, lab},
  p = λ / n;
  B = BernoulliProcess[p];
  P = PoissonProcess[p];
  pathB = RandomFunction[B, {0, t}]["Paths"];
  (* accumulate the successful Bernoulli events as arrivals *)
  cpathB = FoldList[Plus, 0, pathB][[1]] [[All, 2]] [[2 ;; -1]];
  cpathB = Transpose[{pathB[[1]] [[All, 1]], cpathB}];

  ppath = RandomFunction[P, {0, t}]["Paths"][[1]];
  lab = Stl@StringForm[
    "Bernoulli events (p = λ/n) Poisson events (intensity = λ/n)\nλ = ``, n = ``,
    λ, n];
  ListPlot[{cpathB, ppath}, PlotStyle → {Black, Red}, Filling → Axis,
    AxesLabel → {Stl["t"], Stl["N_t"]}, PlotLabel → lab]

```

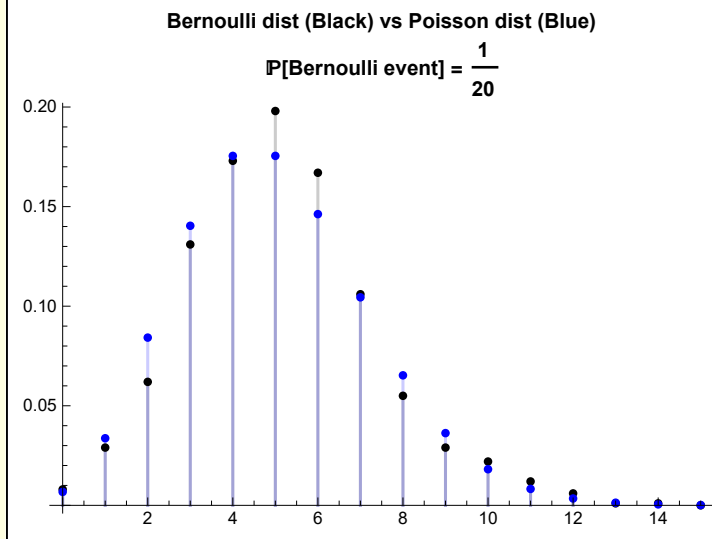


The similarity of the number of events becomes clear when the distributions are compared.

```

Module[{n = 100, λ = 5, t = 100, nSamples = 1000, p, B, H, events, lab},
  p = λ / n;
  B = BernoulliProcess[p];
  (* Simulate some Bernoulli events *)
  events = Table[Total@RandomFunction[B, {0, t}] ["Paths"] [[1]] [[All, 2]] , {nSamples}];
  H = HistogramDistribution[events];
  lab = Stl@StringForm[
    "Bernoulli dist (Black) vs Poisson dist (Blue)\nP[Bernoulli event] = ``", p];
  DiscretePlot[{PDF[H, k], PDF[PoissonProcess[p][t], k]}, {k, 0, 15},
    PlotStyle → {Black, Blue}, PlotLabel → lab]

```



The Poisson distribution can be derived from the sum of Bernoulli events. By simple counting arguments, the distribution of Bernoulli event is binomial.

$$w2[1] = \text{Binomial}[n, k] (\lambda \delta)^k (1 - \lambda \delta)^{n-k}$$

$$(\delta \lambda)^k (1 - \delta \lambda)^{-k+n} \text{Binomial}[n, k]$$

Consider an interval of time τ and subdivide it in n intervals of size $\delta = \tau/n$

$$w2[2] = w2[1] /. \delta \rightarrow 1/n$$

$$\left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{-k+n} \text{Binomial}[n, k]$$

The idea is to let $n \rightarrow \infty$ subject to the constraint $\lambda \tau = n p$. Introduce the explicit form for Binomial

$$w2[3] = w2[2] /. \text{Binomial}[n, k] \rightarrow \frac{n!}{k! (n-k)!}$$

$$\frac{\left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{-k+n} n!}{k! (-k+n)!}$$

Rewrite this as

$$w2[4] = w2[3] /. \left(\frac{\lambda}{n}\right)^k \rightarrow \lambda^k \left(\frac{1}{n}\right)^k$$

$$\frac{\left(\frac{1}{n}\right)^k \lambda^k \left(1 - \frac{\lambda}{n}\right)^{-k+n} n!}{k! (-k+n)!}$$

Note

$$\frac{n!}{(n-k)!} = n(n-1) \dots (n-(k+1)) \quad k \text{ terms}$$

and so

$$\left(\frac{1}{n}\right)^k \frac{n!}{(n-k)!} = \frac{n(n-1) \dots (n-(k+1))}{n n n \dots} \quad \text{both numerator and denominator have } k \text{ terms}$$

or, letting Mathematica do the work

$$w2[5] = \text{Limit}\left[\left(\frac{1}{n}\right)^k \frac{n!}{(n-k)!}, n \rightarrow \infty\right]$$

1

Also note

$$w2[6] = \text{Limit}\left[\left(1 - \frac{\lambda}{n}\right)^{-k+n}, n \rightarrow \infty\right]$$

$e^{-\lambda}$

Thus, in the limit $n \rightarrow \infty$

$$w2[7] = w2[4] /. \left(1 - \frac{\lambda}{n}\right)^{-k+n} \rightarrow \text{Exp}[-\lambda] /. n! \rightarrow (n-k)! n^k // \text{PowerExpand}$$

$$\frac{e^{-\lambda} \lambda^k}{k!}$$

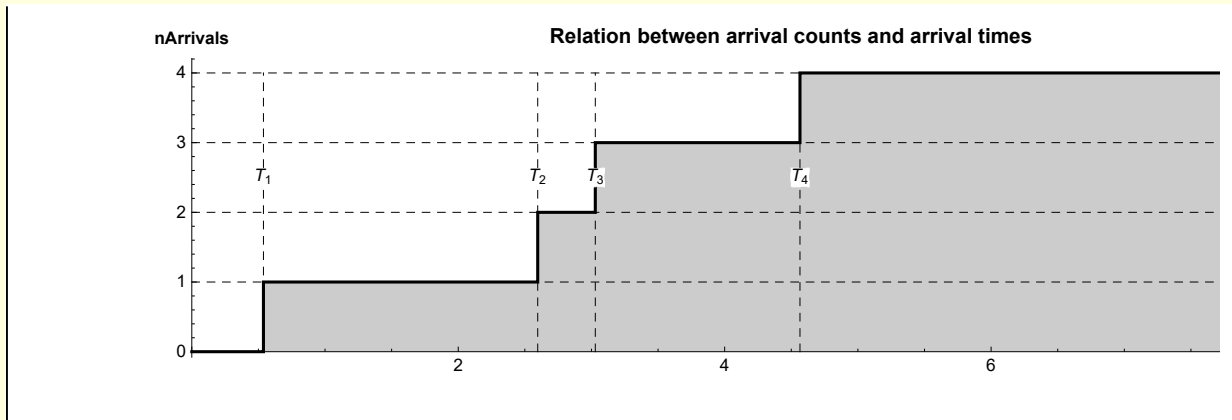
which should be compared with

$$\text{PDF}[\text{PoissonDistribution}[\lambda], k]$$

$$\begin{cases} \frac{e^{-\lambda} \lambda^k}{k!} & k \geq 0 \\ 0 & \text{True} \end{cases}$$

3 Derivation of Poisson Process from fact that interarrival times are exponentially distributed

As time increases, Poisson arrival events occur as illustrated here.



There are useful logical relationships between the arrival times and increases in event counts

$$\text{Event}(T_1 > t) \iff \text{Event}(\mathcal{N}(t) = 0)$$

So we know

$$\mathbb{P}[T_1 > t] = \mathbb{P}[\mathcal{N}_1 = 0] = e^{-\lambda t}$$

where, from the previous section, I have used the expression $e^{-\lambda t}$ for the explicit probability that no event has occurred.

Similarly, consider the second arrival event. The probability that the second arrival event has not occurred conditional on the first arrival event occurring at $T_1 = u$

$$\mathbb{P}[T_2 > t \mid T_1 = u] = \mathbb{P}[\text{no events between } u \text{ and } u + t \mid T_1 = u]$$

But, events in the interval $\{u, u+t\}$ are not influenced by events in the interval $\{0, u\}$, so

$$\mathbb{P}[T_2 > t \mid T_1 = u] = \mathbb{P}[\text{no events between } u \text{ and } u + t \mid T_1 = u] = \mathbb{P}[\mathcal{N}(t) = 0] = e^{-\lambda t}$$

This argument can be extended to conclude that all interarrival times are distributed like $e^{-\lambda t}$. Introducing W_n to be the time waited for the n^{th} arrival to occur. This is just the sum of the interarrival times

$$W_n = T_1 + T_2 + \dots + T_n$$

Notice how this sum is distributed

```
With[{ε = ExponentialDistribution[λ]},
{Probability[T1 > t, T1 ≈ ε, Assumptions → {t > 0}],
Probability[T1 + T2 > t, {T1 ≈ ε, T2 ≈ ε}, Assumptions → {t > 0}],
Probability[T1 + T2 + T3 > t, {T1 ≈ ε, T2 ≈ ε, T3 ≈ ε}, Assumptions → {t > 0}]}]
```

$$\left\{ e^{-t\lambda}, e^{-t\lambda} (1 + t\lambda), \frac{1}{2} e^{-t\lambda} (2 + 2t\lambda + t^2\lambda^2) \right\}$$

In general

$$\mathbb{P}[W_n > t] \approx e^{-\lambda t} \sum_{k=1}^{n-1} \frac{(\lambda t)^k}{k!}$$

Note that this probability is equivalent to the probability of the n^{th} increase in the count having not occurred

$$\mathbb{P}[W_n > t] = \mathbb{P}[N(t) < n] = \text{CDF}[N(t)]$$

The rhs is the CDF of the random variable $N(t)$. This sum involves the Gamma function

?? Gamma

Gamma[z] is the Euler gamma function $\Gamma(z)$.

Gamma[a, z] is the incomplete gamma function $\Gamma(a, z)$.

Gamma[a, z0, z1] is the generalized incomplete gamma function $\Gamma(a, z_0) - \Gamma(a, z_1)$. >>

Attributes[Gamma] = {Listable, NumericFunction, Protected, ReadProtected}

$$\text{w3}[1] = \text{Sum}[\text{Exp}[-\lambda t] \frac{(\lambda t)^k}{k!}, \{k, 0, n-1\}]$$

$$\frac{\text{Gamma}[n, t \lambda]}{\text{Gamma}[n]}$$

We can differentiate to calculate the PDF

$$\text{w3}[2] = -\text{D}[\text{w3}[1] /. t \lambda \rightarrow u, u] /. u \rightarrow t \lambda$$

$$\frac{e^{-t \lambda} (t \lambda)^{-1+n}}{\text{Gamma}[n]}$$

Simplifying

$$\text{w3}[3] = \text{w3}[2] /. \text{Gamma}[n] \rightarrow (n-1)!$$

$$\frac{e^{-t \lambda} (t \lambda)^{-1+n}}{(-1+n)!}$$

Independently of this calculation, the Poisson PDF for $\mathbb{P}[N(t) < n]$ is

$$\text{w3}[4] = \text{Refine}[\text{PDF}[\text{PoissonProcess}[\lambda][t], n-1], n-1 \geq 0]$$

$$\frac{e^{-t \lambda} (t \lambda)^{-1+n}}{(-1+n)!}$$

which is the same as w3[3]

```
w3[3] - w3[4] // Simplify
```

```
0
```

and the Poisson process PDF has been derived starting with the fact that interarrival times are exponentially distributed.

Problems involving Poisson processes

I found a nice collection of problems on the web and work through the details of the solution procedures.

[\[PDF\] solved problems in random processes - Valjhun](#) ✓

valjhun.fmf.uni-lj.si/~raicm/Vaje/SPI/SPI_ex.pdf ▼

by M Raič - 2015 - [Related articles](#)

M. RAIČ & A. TOMAN: SOLVED PROBLEMS IN RANDOM PROCESSES. 4 A homogeneous Poisson process with intensity $\lambda > 0$ can be defined as a certain ...

I Waiting for the doctor

Raič, Toman: Homogeneous Poisson Process - 1

1. Patients arrive in a surgery according to a homogeneous Poisson process with intensity 6 patients an hour. The doctor starts to examine the patients only when the third patient arrives.
 - a) Compute the expected time from the opening of the surgery until the first patient starts to be examined.
 - b) Compute the probability that in the first opening hour, the doctor does not start examining at all.

a) This can be immediately solved using Mathematica probability capabilities

The expected arrival time of the 3rd patient is

```
Module[{ε, result},
  ε = ExponentialDistribution[λ];
  result = Expectation[T1 + T2 + T3, {T1 ≈ ε, T2 ≈ ε, T3 ≈ ε}];
  {result, result /. λ → 6}]
```

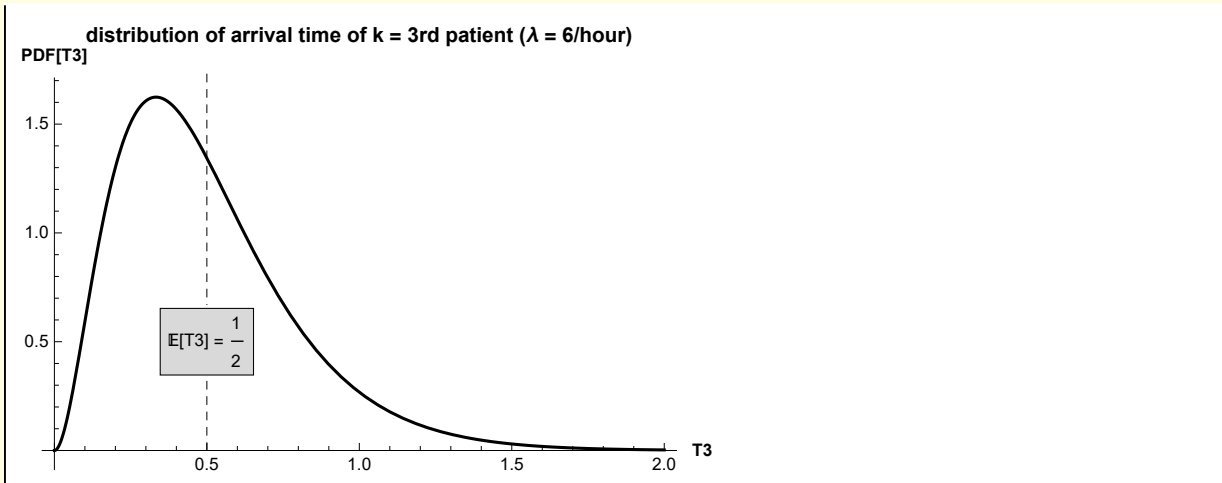
```
{ $\frac{3}{\lambda}$ ,  $\frac{1}{2}$ }
```

Alternatively, the sum of exponential random variables is distributed according to an Erlang distribution.

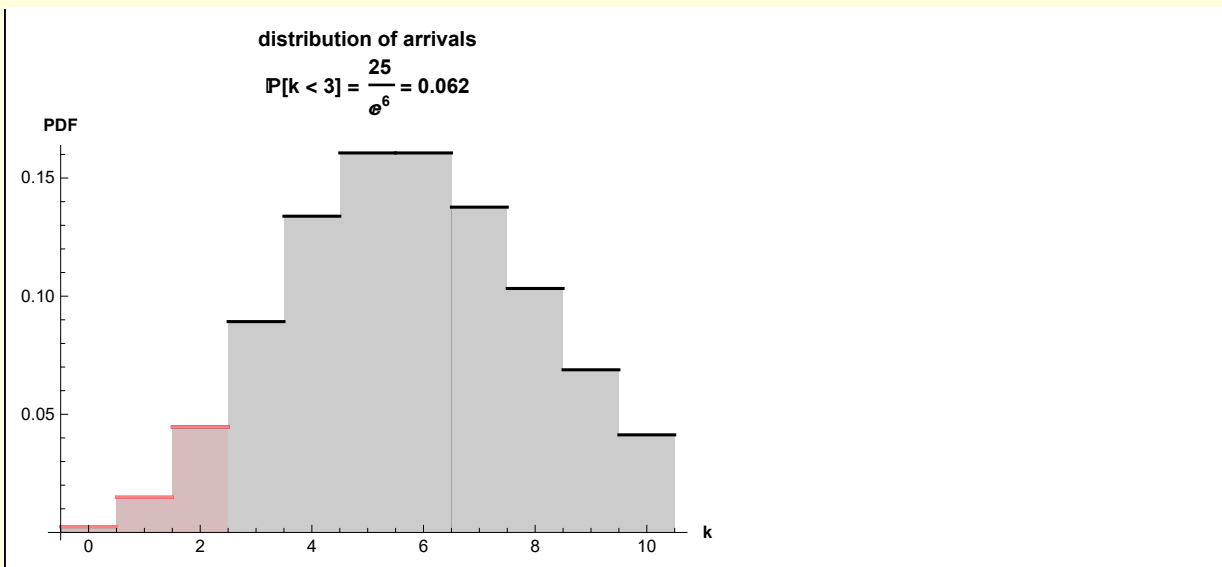
```
Module[{ε = ErlangDistribution[k, λ], results},
  results =
  {{Refine[PDF[ε, t], t > 0], Refine[CDF[ε, t], t > 0], Mean[ε], Variance[ε]}};
  PrependTo[results, {"PDF", "CDF", "mean", "variance"}];
  LGrid[results, "Properties of Erlang distribution"]]
```

Properties of Erlang distribution

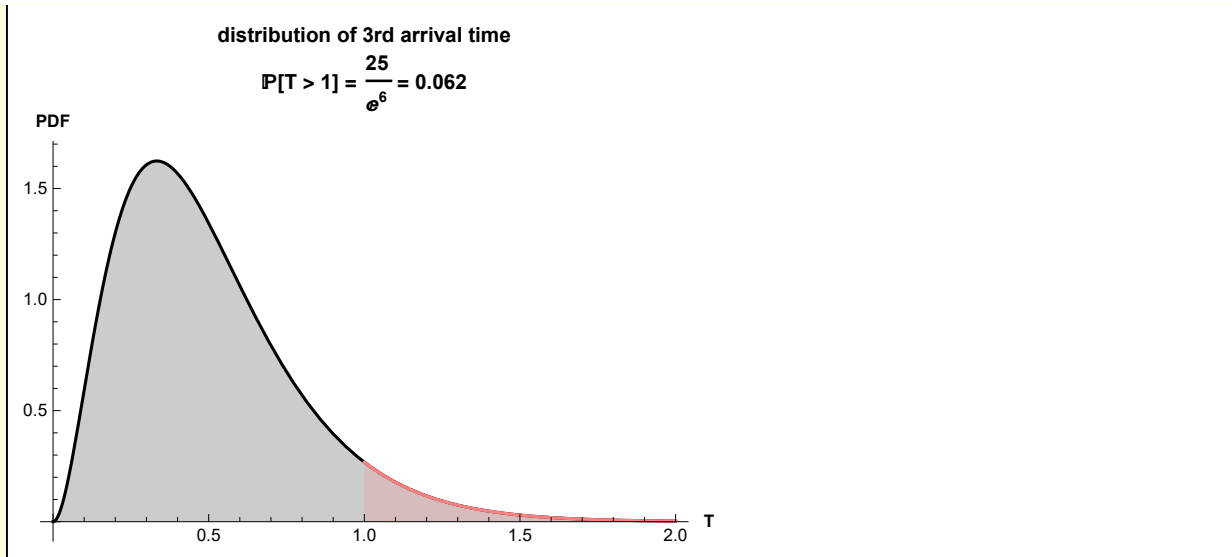
PDF	CDF	mean	variance
$\frac{e^{-t\lambda} t^{k-1} \lambda^k}{\text{Gamma}[k]}$	$\text{GammaRegularized}[k, 0, t\lambda]$	$\frac{k}{\lambda}$	$\frac{k}{\lambda^2}$



b) Probability that less than three patients arrive during the first hour



or that the third arrival time occurs after one hour



2 Covariance, correlation

Raic Homogeneous Poisson Process - 2

2. Consider a homogeneous Poisson process with intensity λ .

- a) Compute the auto-covariance function of the family N_t , i. e., all covariances $\text{cov}(N_t, N_s)$.
- b) Compute the auto-correlation function of the family N_t , i. e., all correlations $\text{cov}(N_t, N_s)$.
- c) For $t_1 \leq t_2 \leq \dots \leq t_n$, find the covariance matrix of the random vector $(N_{t_1}, N_{t_2}, \dots, N_{t_n})$.

```
Module[{P},
  P = PoissonProcess[λ];
  {CovarianceFunction[P, s, t], CorrelationFunction[P, s, t]}
```

$$\left\{ \lambda \text{Min}[s, t], \frac{\text{Min}[s, t]}{\sqrt{s t}} \right\}$$

Alternatively,

```
Module[{P, meant},
  P = PoissonProcess[λ];
  Print["mean = ", Mean[P[t]]];
  Expectation[(x[t] - λ t) (x[s] - λ s), x ≈ P]]
```

mean = $t \lambda$

```
ConditionalExpression[s λ, s < t]
```

```

Module[{λ = 1, tMax = 3, G},
  G[1] = Plot3D[λ Min[s, t], {s, 0, tMax}, {t, 0, tMax}, Mesh → False,
    AxesLabel → {St1["s"], St1["t"], ""}, PlotLabel → St1["Covariance"]];
  G[2] = Plot3D[λ  $\frac{\text{Min}[s, t]}{\sqrt{s t}}$ , {s, 0, tMax}, {t, 0, tMax}, Mesh → False,
    AxesLabel → {St1["s"], St1["t"], ""}, PlotLabel → St1["Correlation"]];
  Grid[{{G[1], G[2]}}]]

```

3 Conditional Poisson distribution

Raic Homogeneous Poisson Process - 4

4. Let $X \sim \text{Pois}(\lambda)$ and $Y \sim \text{Pois}(\mu)$ be independent random variables. Find the conditional distribution of the random variable X given $Z := X + Y$.

Note

$$\begin{aligned}
 \mathbb{P}[X = k \mid X + Y = n] &= \frac{\mathbb{P}[X = k \cap X + Y = n]}{\mathbb{P}[X + Y = n]} && \text{conditional prob definition} \\
 &= \frac{\mathbb{P}[X = k \cap k + Y = n]}{\mathbb{P}[X + Y = n]} && \text{meaning of intersection} \\
 &= \frac{\mathbb{P}[X = k] \mathbb{P}[Y = n - k]}{\mathbb{P}[X + Y = n]} && X \text{ and } Y \text{ are independent}
 \end{aligned}$$

Each of these three expressions can be evaluated. The two terms in the numerator are

```
w3[1] = Refine[Probability[X == k, X ≈ PoissonDistribution[λx]], k == Floor[k] && k ≥ 0]
```

$$\frac{e^{-\lambda x} \lambda_x^k}{k!}$$

$$\begin{aligned} w3[2] &= \text{Refine}[\text{Probability}[Y = n - k, Y \approx \text{PoissonDistribution}[\lambda_Y]], k + \text{Floor}[-k + n] == n \&\& k \leq n] \\ &= \frac{e^{-\lambda_Y} \lambda_Y^{k+n}}{(n-k)!} \end{aligned}$$

The term in the denominator can also be calculated because Poisson processes are additive.

$$\begin{aligned} w3[3] &= \text{Refine}[\text{Probability}[Z = n, \{Z \approx \text{PoissonDistribution}[\lambda_X + \lambda_Y]\}], n == \text{Floor}[n] \&\& n \geq 0] \\ &= \frac{e^{-\lambda_X - \lambda_Y} (\lambda_X + \lambda_Y)^n}{n!} \end{aligned}$$

Thus

$$\begin{aligned} w3[4] &= \frac{w3[1] w3[2]}{w3[3]} \\ &= \frac{n! \lambda_X^k \lambda_Y^{k+n} (\lambda_X + \lambda_Y)^{-n}}{k! (n-k)!} \end{aligned}$$

This suggests a binomial distribution. Note

$$\begin{aligned} w3[5] &= \text{Refine}[\text{PDF}[\text{BinomialDistribution}[n, p], k], 0 \leq k \leq n] \\ &= (1-p)^{-k+n} p^k \text{Binomial}[n, k] \end{aligned}$$

Using this result

$$\begin{aligned} w3[6] &= w3[4] /. \text{Sol}[\text{Binomial}[n, k] == \frac{n!}{k! (n-k)!}, n!] \\ &= \text{Binomial}[n, k] \lambda_X^k \lambda_Y^{k+n} (\lambda_X + \lambda_Y)^{-n} \end{aligned}$$

I conclude

$$\mathbb{P}[X = k \mid X + Y = n] \approx \text{Binomial}\left[n, \frac{\lambda_X}{\lambda_X + \lambda_Y}\right]$$

To establish that the sum of Poisson stochastic variables is Poisson distributed I can also use

$$\begin{aligned} &\text{CharacteristicFunction}[\text{PoissonDistribution}[\lambda], k] \\ &= e^{(-1 + e^{ik}) \lambda} \end{aligned}$$

The sum of Poisson variables $X_1 + X_2$ is the product of the characteristic functions

```

CharacteristicFunction[PoissonDistribution[λ1], k]
CharacteristicFunction[PoissonDistribution[λ2], k] // Simplify

$$e^{(-1+e^{ik})(\lambda_1+\lambda_2)}$$


```

4 Thinning of Poisson process

Raic Homogeneous Poisson Process - 5

5. Let N be a random variable denoting the number of arrivals, distributed by Poisson $\text{Pois}(\lambda)$. Each arrival is *successful* with probability p , independently of other arrivals, as well as of the number of arrivals. Denote by S the number of successful and by T the number of unsuccessful arrivals, that is, $T = N - S$.

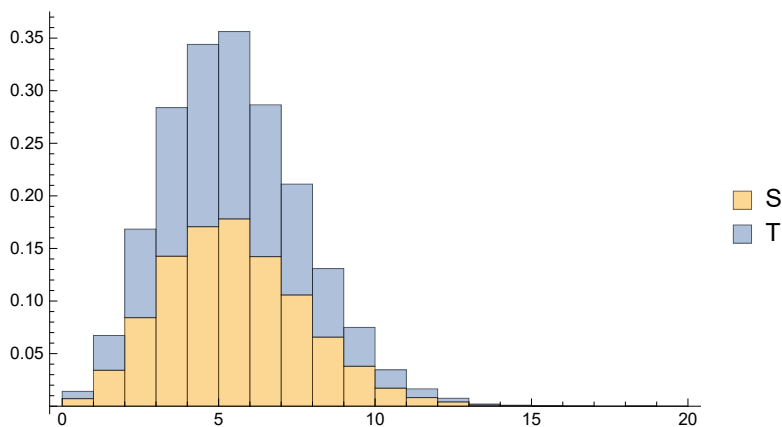
- Find the distribution of S and T .
- Show that the random variables S and T are independent.
- Show that under some other choice of the distribution of N , S and T are no longer necessarily independent.

To gain some perspective, I start with a simulation

```

Module[{λ = 5, nSamples = 100000, p = 1/4, ϕ, path, results, resultsS, resultsT},
ϕ = PoissonDistribution[λ];
path = RandomVariate[ϕ, nSamples];
results = ({#, If[RandomReal[] < p, "S", "T"]}) & /@ path;
resultsS = Select[results, (#[[2]] == "S") &];
resultsT = Select[results, (#[[2]] == "T") &];
Histogram[{resultsS[[All, 1]], resultsT[[All, 1]]}, Automatic,
"PDF", ChartLayout -> "Stacked", ChartLegends -> {"S", "T"} ]

```



These distributions appear similar.

a) The process S is successful with probability p — thus

$$P[S = k \mid N = n] = \binom{n}{k} p^k (1-p)^{n-k} \text{ Binomial}[n, k]$$

where $k \leq n$. The total probability is

$$(4-2) \quad \mathbb{P}[S = k] = \sum_{n=k}^{\infty} \mathbb{P}[S = k \mid N = n] \mathbb{P}[N = n]$$

Since the arrivals are Poisson distributed with intensity λ

$$(4-3) \quad \mathbb{P}[N = n] = \frac{e^{-\lambda} \lambda^n}{n!}$$

$$\mathbf{w4[1]} = \text{Refine}[\text{PDF}[\text{BinomialDistribution}[n, p], k], 0 \leq k \leq n]$$

$$(1-p)^{-k+n} p^k \text{Binomial}[n, k]$$

$$\mathbf{w4[2]} = \text{Refine}[\text{PDF}[\text{PoissonDistribution}[\lambda], n], n \geq 0]$$

$$\frac{e^{-\lambda} \lambda^n}{n!}$$

$$\mathbf{w4[3]} = \text{Sum}[\mathbf{w4[1]} \mathbf{w4[2]}, \{n, k, \infty\}]$$

$$\frac{e^{-p\lambda} p^k \lambda^k}{k!}$$

Thus, we see

$$\mathbb{P}[S = k] = \frac{e^{-\lambda p} (\lambda p)^k}{k!}$$

By analogy

$$\mathbb{P}[T = k] = \frac{e^{-\lambda(1-p)} (\lambda(1-p))^k}{k!}$$

b) We need to calculate

$$\mathbb{P}[S \cap T] = \mathbb{P}[S] \mathbb{P}[T]$$

Note

$$\mathbf{w4b[1]} = \mathbb{P}\text{Intersection}[S == j, T == k]$$

$$\mathbb{P}\text{Intersection}[S == j, T == k]$$

Since $n = j + k$

$$\mathbf{w4b[2]} = \mathbf{w4b[1]} /. T == k \rightarrow N == j + k$$

$$\mathbb{P}\text{Intersection}[S == j, N == j + k]$$

$$w4b[3] = w4b[2] /. \mathbb{P}\text{Intersection}[A_, B_] \rightarrow \mathbb{P}\text{conditional}[A, B] \mathbb{P}[B]$$

$$\mathbb{P}[N = j + k] \mathbb{P}\text{conditional}[S = j, N = j + k]$$

$$w4b[4] = w4b[3] /. \mathbb{P}\text{conditional}[S = j, N = j + k] \rightarrow \text{PDFBin}[j + k, p, j] /.$$

$$\mathbb{P}[N = j + k] \rightarrow \text{PDFPois}[\lambda, j + k]$$

$$\text{PDFBin}[j + k, p, j] \text{PDFPois}[\lambda, j + k]$$

$$w4b[5] = w4b[4] /.$$

$$\text{PDFBin}[j + k, p, j] \rightarrow \text{Refine}[\text{PDF}[\text{BinomialDistribution}[j + k, p], j], 0 \leq j \leq j + k]$$

$$(1 - p)^k p^j \text{Binomial}[j + k, j] \text{PDFPois}[\lambda, j + k]$$

$$w4b[6] = w4b[5] /.$$

$$\text{PDFPois}[\lambda, j + k] \rightarrow \text{Refine}[\text{PDF}[\text{PoissonDistribution}[\lambda], j + k], j + k \geq 0] // \text{Simplify}$$

$$\frac{e^{-\lambda} (1 - p)^k p^j \lambda^{j+k} \text{Binomial}[j + k, j]}{(j + k)!}$$

This can be factored. Let $q \rightarrow 1 - p$

$$\text{Binomial}[a_, b_] \rightarrow \frac{a!}{b! (a - b)!}$$

$$w4b[7] = w4b[6] /. 1 - p \rightarrow q /. \text{Exp}[-\lambda] \rightarrow \text{Exp}[-\lambda p] \text{Exp}[-\lambda q] /.$$

$$\text{Binomial}[a_, b_] \rightarrow \frac{a!}{b! (a - b)!}$$

$$\frac{e^{-p\lambda - q\lambda} p^j q^k \lambda^{j+k}}{j! k!}$$

Note that the independence condition $\mathbb{P}[S = j] \mathbb{P}[T = k]$. I calculate

$$\mathbb{P}S = \text{Refine}[\text{Probability}[S == j, S \approx \text{PoissonDistribution}[\lambda p]], j == \text{Floor}[j] \&\& j \geq 0]$$

$$\frac{e^{-p\lambda} (p\lambda)^j}{j!}$$

$$\mathbb{P}T = \text{Refine}[\text{Probability}[T == k, T \approx \text{PoissonDistribution}[\lambda q]], k == \text{Floor}[k] \&\& k \geq 0]$$

$$\frac{e^{-q\lambda} (q\lambda)^k}{k!}$$

The expression $w4b[7]$ is the same as the independence condition

```
w4b[8] = w4b[7] - PSPT // PowerExpand
0
```

5 Memoryless property of exponential distribution

Raic Homogeneous Poisson Process - 6

6. Show that the exponential distribution is also memoryless: for $T \sim \text{Exp}(\lambda)$ we have for sufficiently large $t, s \in \mathbb{N}$:

$$\mathbb{P}(T \leq t + s \mid T > t) = \mathbb{P}(T \leq s). \quad (*)$$

Moreover, show that the exponential distribution is the only memoryless distribution which is continuous and with density being continuous on $(0, \infty)$ and zero elsewhere. We assume that the conditional distribution in $(*)$ makes sense, i. e., that $\mathbb{P}(T > t) > 0$ for all $t \geq 0$.

The calculations corresponding to the lhs is

```
Refine[Probability[T ≤ s + t | T > t, T ≈ ExponentialDistribution[λ]],
  t ≥ 0 && s > 0] // Expand
1 - e-sλ
```

and the rhs

```
Refine[Probability[T ≤ s, T ≈ ExponentialDistribution[λ]], s > 0]
1 - e-sλ
```

The two probabilities are identical. The fact that these probabilities do not depend on the time parameter t means that the exponential process is “memoryless.”

6 Distribution of $\min[X, Y]$, $\max[X, Y]$ for exponential distribution

Raic, Toman: Homogeneous Poisson Process - 7

7. For independent random variables $X \sim \text{Exp}(\lambda)$ and $Y \sim \text{Exp}(\mu)$, find the distribution of $U := \min\{X, Y\}$ and $V := \max\{X, Y\}$.

M cannot calculate this directly.

```
Probability[Min[X, Y],
  {X ≈ ExponentialDistribution[λ], Y ≈ ExponentialDistribution[μ]}]
Probability[Min[X, Y],
  {X ≈ ExponentialDistribution[λ], Y ≈ ExponentialDistribution[μ]}]
```

However, the two conditions can be calculated separately

When $X < Y$

$$\text{Probability}[X \leq Y, \{X \approx \text{ExponentialDistribution}[\lambda X], Y \approx \text{ExponentialDistribution}[\lambda Y]\}]$$

$$\frac{\lambda X}{\lambda X + \lambda Y}$$

When $Y < X$

$$\text{Probability}[Y \leq X, \{X \approx \text{ExponentialDistribution}[\lambda X], Y \approx \text{ExponentialDistribution}[\lambda Y]\}]$$

$$\frac{\lambda Y}{\lambda X + \lambda Y}$$

Thus,

$$\mathbb{P}[U] = \mathbb{P}[\min(X, Y)] = \frac{\min[\lambda X, \lambda Y]}{\lambda X + \lambda Y}$$

The problem asks for the distribution.

For the minimum $U = \min[X, Y]$, the CDF is

$$\text{w6[1]} = \text{Probability}[X > u \ \&\& \ Y > u, \{X \approx \text{ExponentialDistribution}[\lambda X], Y \approx \text{ExponentialDistribution}[\lambda Y]\}, \text{Assumptions} \rightarrow \{u \geq 0\}]$$

$$e^{-u \lambda X - u \lambda Y}$$

The PDF is

$$\text{w6[2]} = -D[\#, u] \ \&\& \ \text{w6[1]}$$

$$-e^{-u \lambda X - u \lambda Y} (-\lambda X - \lambda Y)$$

$$\text{w6[2]} = -D[\text{w6[1]}, u] \ // \ \text{Simplify}$$

$$e^{-u (\lambda X + \lambda Y)} (\lambda X + \lambda Y)$$

$$\text{Clear}[\text{PDFU}];$$

$$\text{PDFU}[u_, \lambda X_, \lambda Y_] := e^{-u (\lambda X + \lambda Y)} (\lambda X + \lambda Y)$$

For the maximum $V = \max[X, Y]$

$$\text{w6[3]} = \text{Probability}[X < v \ \&\& \ Y < v, \{X \approx \text{ExponentialDistribution}[\lambda X], Y \approx \text{ExponentialDistribution}[\lambda Y]\}, \text{Assumptions} \rightarrow \{v \geq 0\}]$$

$$(1 - e^{-v \lambda X}) (1 - e^{-v \lambda Y})$$

In this case, the PDF is

```
w6[4] = D[#, v] &@w6[3]
```

$$e^{-v\lambda X} (1 - e^{-v\lambda Y}) \lambda X + e^{-v\lambda Y} (1 - e^{-v\lambda X}) \lambda Y$$

```
Clear[PDFV];
```

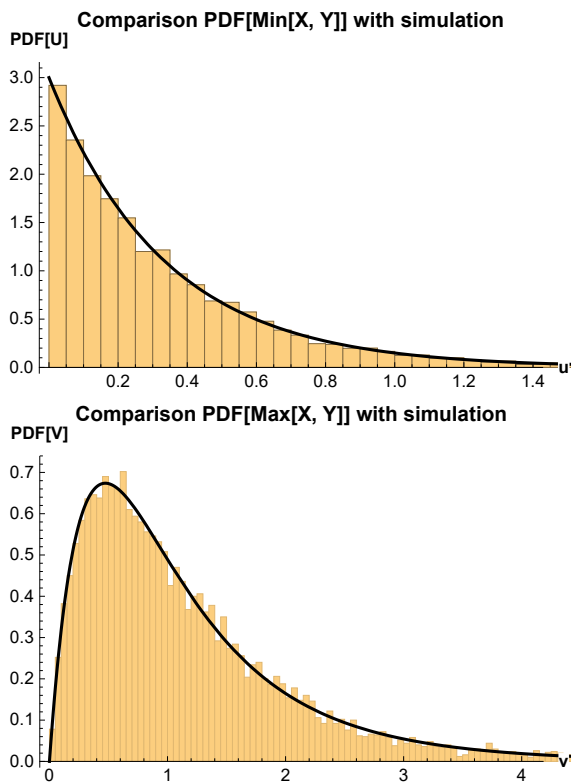
```
PDFV[u_, λX_, λY_] := e-vλX (1 - e-vλY) λX + e-vλY (1 - e-vλX) λY
```

I verify these derived PDFs against simulation

```

Module[{λX = 1, λY = 2, nSamples = 10000, εX, εY, pairsXY, U, V, lab, g},
  {εX, εY} = {ExponentialDistribution[λX], ExponentialDistribution[λY]};
  pairsXY = Transpose[{RandomVariate[εX, nSamples], RandomVariate[εY, nSamples]}];
  U = Min[#] & /@ pairsXY;
  V = Max[#] & /@ pairsXY;
  lab = Stl@StringForm["Comparison PDF[Min[X, Y]] with simulation"];
  g[1] = Histogram[U, {0.05}, "PDF", AxesLabel → {Stl["u"], Stl["PDF[U]"]},
    PlotLabel → lab, ImageSize → {500, 200}];
  g[2] = Plot[PDFU[u, λX, λY], {u, 0, Max[U]}, PlotStyle → Black, PlotRange → All];
  lab = Stl@StringForm["Comparison PDF[Max[X, Y]] with simulation"];
  g[3] = Histogram[V, {0.05}, "PDF", AxesLabel → {Stl["v"], Stl["PDF[V]"]},
    PlotLabel → lab, ImageSize → {500, 200}];
  g[4] = Plot[PDFV[v, λX, λY], {v, 0, Max[V]}, PlotStyle → Black];
  Grid[{{Show[{g[1], g[2]}]}, {Show[{g[3], g[4]}]}}]]

```



7 Fire Brigade

Raic, Toman: Homogeneous Poisson Process - 8

8. A fire station receives emergency calls according to a homogeneous Poisson process with intensity half a call per hour. Each time, the fire brigade needs certain time to process the call and prepare for further calls (this total period will be called *intervention*). During an intervention, the calls are redirected to other fire stations. Suppose that the intervention times are distributed uniformly over the interval from half an hour to one hour and independent of each other as well as of the emergency calls.

Suppose that the fire brigade is able to respond to the calls at the moment (i. e., there is no intervention). Find the distribution of the number of the calls to which the fire brigade responds before any call is redirected.

The first call can be taken, then an intervention period begins. At first cut I consider the intervention time ΔT to be constant. So, for the second call to not be diverted, it must arrive at least ΔT after the first call. Interarrival times are distributed exponentially with parameter λ . The probability p that a call will be diverted is

```
pCallDiverted =
  Probability[T ≤ ΔT, T ≈ ExponentialDistribution[λ], Assumptions → {ΔT ≥ 0}]
1 - e-ΔT λ
```

A sequence of calls arrive each of which is diverted with probability p . Such sequences of Bernoulli trials are modeled by a Geometric distribution — defined as the distribution of the number of failures in a sequence of Bernoulli trials before a success with probability p occurs.

In this context, failure corresponds to accepting a call, while diverting a call corresponds to success. For constant ΔT .

```
Module[{G},
  G = GeometricDistribution[p];
  Refine[PDF[G, k], k ≥ 0]]
(1 - p)k p
```

In the stated problem, ΔT is not constant but uniformly distributed between 1/2 and 1 hour. In this case, the probability $\mathbb{P}[T \leq \Delta T]$ in that case.

```
Probability[T ≤ ΔT,
  {T ≈ ExponentialDistribution[λ], ΔT ≈ UniformDistribution[{1/2, 1}]]]

$$\frac{2 e^{-\lambda} - 2 e^{-\lambda/2} + \lambda}{\lambda}$$

```

For the stated arrival rate of calls $\lambda = 1/2$

$$p = \frac{2e^{-\lambda} - 2e^{-\lambda/2} + \lambda}{\lambda} \quad /. \quad \lambda \rightarrow 1/2 \quad // \quad N$$

$$p = 0.31092$$

and the distribution of calls accepted before a diversion is

$$(1-p)^k p \quad /. \quad p \rightarrow \frac{2e^{-\lambda} - 2e^{-\lambda/2} + \lambda}{\lambda}$$

$$\frac{(2e^{-\lambda} - 2e^{-\lambda/2} + \lambda) \left(1 - \frac{2e^{-\lambda} - 2e^{-\lambda/2} + \lambda}{\lambda}\right)^k}{\lambda}$$

```
Clear[TheoreticalDistributionOfCallsAccepted];
TheoreticalDistributionOfCallsAccepted[λ_, kMax_] :=
Module[{p, vals},
  p =  $\frac{2e^{-\lambda} - 2e^{-\lambda/2} + \lambda}{\lambda}$  // N;
  vals = Table[(1-p)^k p, {k, 0, kMax}];
  Transpose[{Range[1, kMax + 1], vals}]]
```

```
Module[{λ = 0.5, kMax = 10, dat, info},
  dat = TheoreticalDistributionOfCallsAccepted[λ, kMax];
  info = ({#[[1]], NF3@#[[2]]) & /@ dat;
  PrependTo[info, {"nCalls accepted before diversion", "P[nCalls accepted]"}];
  LGrid[info, "test"]]
```

test

nCalls accepted before diversion	P[nCalls accepted]
1	0.311
2	0.214
3	0.148
4	0.102
5	0.070
6	0.048
7	0.033
8	0.023
9	0.016
10	0.011
11	0.008

Checking against simulation

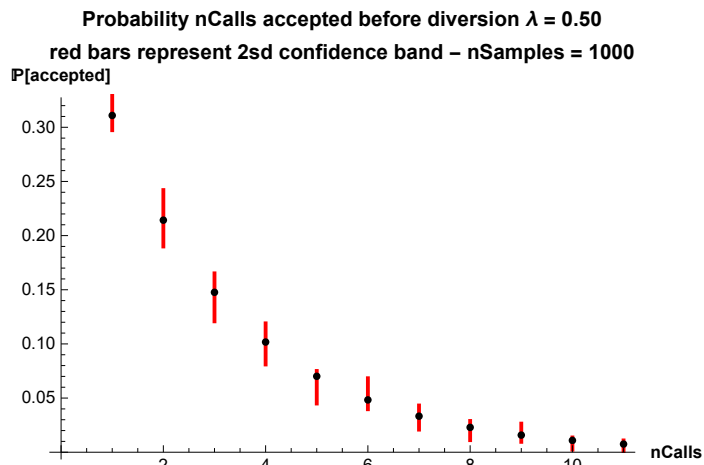

```

Module[{λ = 1/2, α = 1/2, β = 1, nTrials = 20,
  nSamples = 1000, kMax = 10, analyticalResults, ε, u, Tvals,
  ΔTvals, resultsSim, resultTrial, estimates, MCEstimate},
  analyticalResults = TheoreticalDistributionOfCallsAccepted[1/2, kMax];

  ε = ExponentialDistribution[λ];
  u = UniformDistribution[{α, β}];
  resultsSim = Table[
    (* Crude - I just make sure nTrials is large enough for a diversion to occur *)
    Tvals = RandomVariate[ε, nTrials];
    ΔTvals = RandomVariate[u, nTrials];
    resultTrial =
      (Boole[#[[1]] < #[[2]]) & /@ Transpose[{Tvals, ΔTvals}], {nSamples}];

  MCEstimate[pos_] :=
    Module[{d}, d = Boole[FirstPosition[#, 1][[1]] == pos] & /@ resultsSim;
    {pos, Mean[d], StandardDeviation[d] / Sqrt[Length[d]]} // N];
  estimates = (MCEstimate[#]) & /@ Range[1, kMax + 1];
  Module[{Bar, lab},
    lab =
      Stl@StringForm["Probability nCalls accepted before diversion λ = ``\nred bars
        represent 2sd confidence band - nSamples = ``", NF2@N@λ, nSamples];
    Bar[{{x_, y_, dy_}}] := {Red, Thick, Line[{{x, y - 2 dy}, {x, y + 2 dy}}]};
    ListPlot[analyticalResults, Prolog → Bar /@ estimates, PlotStyle → Black,
      AxesLabel → {Stl["nCalls"], Stl["P[accepted]"]}, PlotLabel → lab]]]

```



8 Conditional Poisson arrivals

Raic, Toman: Homogeneous Poisson Process - 9

9. Consider a homogeneous Poisson process with intensity λ .

- Suppose that up to time t , exactly one arrival occurred. Given this information, find the conditional distribution of the arrival time.
- Suppose that exactly two arrivals occurred. Compute the conditional expectations of both arrival times.

a) For $s < t$, what is $\mathbb{P}[S \leq s | \mathcal{N}(t)=1]$?

$$\mathbb{P}[S \leq s | \mathcal{N}(t) = 1] = \mathbb{P}[\mathcal{N}(s) \geq 1 \cap \mathcal{N}(t) = 1]$$

This probability involves a time related event and a counting related event. To make progress, one must reexpress the time event as a counting event. For the first arrival event occurring at time T .

$$\mathbb{P}[T \leq t] = \mathbb{P}[\mathcal{N}[t] \geq 1]$$

Using this result

$$\mathbb{P}[S \leq s | \mathcal{N}(t) = 1] = \mathbb{P}[\mathcal{N}(s) \geq 1 | \mathcal{N}(t) = 1]$$

and then

$$\begin{aligned} \mathbb{P}[S \leq s | \mathcal{N}(t) = 1] &= \mathbb{P}[\mathcal{N}(s) \geq 1 | \mathcal{N}(t) = 1] \\ &= \mathbb{P}[\mathcal{N}(s) = 1 | \mathcal{N}(t) = 1] \end{aligned}$$

The last statement is a bit mysterious. Since we know that $\mathcal{N}[s] \geq 1$ means a counting event has occurred and we are given that that event occurred at $t = 1$, then the event $\mathcal{N}(s) \geq s$ is equivalent to the event $\mathcal{N}[s] = 1$.

Proceeding

$$\begin{aligned} \mathbb{P}[S \leq s | \mathcal{N}(t) = 1] &= \mathbb{P}[\mathcal{N}(s) \geq 1 | \mathcal{N}(t) = 1] = \mathbb{P}[\mathcal{N}(s) = 1 | \mathcal{N}(t) = 1] = \frac{\mathbb{P}[\mathcal{N}(s) = 1 \cap \mathcal{N}(t) = 1]}{\mathbb{P}[\mathcal{N}(t) = 1]} \\ &= \frac{\mathbb{P}[\mathcal{N}(s) = 1 \cap \mathcal{N}(t) - \mathcal{N}(s) = 0]}{\mathbb{P}[\mathcal{N}(t) = 1]} \end{aligned}$$

This last step is often used in this type of calculation. It allows the independence of time intervals for Poisson processes to be evoked

$$\begin{aligned} \mathbb{P}[S \leq s | \mathcal{N}(t) = 1] &= \mathbb{P}[\mathcal{N}(s) \geq 1 | \mathcal{N}(t) = 1] = \mathbb{P}[\mathcal{N}(s) = 1 | \mathcal{N}(t) = 1] = \frac{\mathbb{P}[\mathcal{N}(s) = 1 \cap \mathcal{N}(t) = 1]}{\mathbb{P}[\mathcal{N}(t) = 1]} \\ &= \frac{\mathbb{P}[\mathcal{N}(s) = 1 \cap \mathcal{N}(t) - \mathcal{N}(s) = 0]}{\mathbb{P}[\mathcal{N}(t) = 1]} \\ &= \frac{\mathbb{P}[\mathcal{N}(s) = 1] \mathbb{P}[\mathcal{N}(t) - \mathcal{N}(s) = 0]}{\mathbb{P}[\mathcal{N}(t) = 1]} \end{aligned}$$

Note

$$\{\text{Probability}[x[t] == 1, x \approx \text{PoissonProcess}[\lambda]],$$

$$\text{Probability}[x[t - s] == 0, x \approx \text{PoissonProcess}[\lambda]]\}$$

$$\{e^{-t\lambda} t \lambda, e^{(s-t)\lambda}\}$$

Using these probabilities

$$\begin{aligned} \mathbb{P}[S \leq s | \mathcal{N}(t) = 1] &= \mathbb{P}[\mathcal{N}(s) \geq s | \mathcal{N}(t) = 1] = \mathbb{P}[\mathcal{N}(s) = 1 | \mathcal{N}(t) = 1] = \frac{\mathbb{P}[\mathcal{N}(s) = 1 \cap \mathcal{N}(t) = 1]}{\mathbb{P}[\mathcal{N}(t) = 1]} \\ &= \frac{\mathbb{P}[\mathcal{N}(s) = 1 \cap \mathcal{N}(t) - \mathcal{N}(s) = 0]}{\mathbb{P}[\mathcal{N}(t) = 1]} \\ &= \frac{\mathbb{P}[\mathcal{N}(s) = 1] \mathbb{P}[\mathcal{N}(t) - \mathcal{N}(s) = 0]}{\mathbb{P}[\mathcal{N}(t) = 1]} \\ &= \frac{(e^{-s\lambda} s \lambda) (e^{-(t-s)\lambda})}{(e^{-t\lambda} t \lambda)} = \frac{s}{t} \end{aligned}$$

b) There are two events S_1 and S_2 . What are $\mathbb{P}[S_1 \leq s | \mathcal{N}(t) = 2]$ and $\mathbb{P}[S_2 \leq s | \mathcal{N}(t) = 2]$?

The calculation of the second term is analogous to the calculation in part a)

$$\mathbb{P}[S_2 \leq s | \mathcal{N}(t) = 2] = \frac{\mathbb{P}[\mathcal{N}(s) = 2 \cap \mathcal{N}(t) = 2]}{\mathbb{P}[\mathcal{N}(t) = 2]} = \frac{\mathbb{P}[\mathcal{N}(s) = 2] \mathbb{P}[\mathcal{N}(t) - \mathcal{N}(s) = 0]}{\mathbb{P}[\mathcal{N}(t) = 2]}$$

$$\{\text{Probability}[x[t] == 2, x \approx \text{PoissonProcess}[\lambda]],$$

$$\text{Probability}[x[t - s] == 0, x \approx \text{PoissonProcess}[\lambda]]\}$$

$$\left\{ \frac{1}{2} e^{-t\lambda} t^2 \lambda^2, e^{(s-t)\lambda} \right\}$$

$$\begin{aligned} \mathbb{P}[S_2 \leq s | \mathcal{N}(t) = 2] &= \\ \mathbb{P}[\mathcal{N}(s) \geq 2 | \mathcal{N}(t) = 2] &= \frac{\mathbb{P}[\mathcal{N}(s) = 2 \cap \mathcal{N}(t) = 2]}{\mathbb{P}[\mathcal{N}(t) = 2]} = \frac{\mathbb{P}[\mathcal{N}(s) = 2] \mathbb{P}[\mathcal{N}(t) - \mathcal{N}(s) = 0]}{\mathbb{P}[\mathcal{N}(t) = 2]} \\ &= \frac{\left(\frac{1}{2} e^{-s\lambda} s^2 \lambda^2\right) (e^{(s-t)\lambda})}{\left(\frac{1}{2} e^{-t\lambda} t^2 \lambda^2\right)} = \frac{s^2}{t^2} \end{aligned}$$

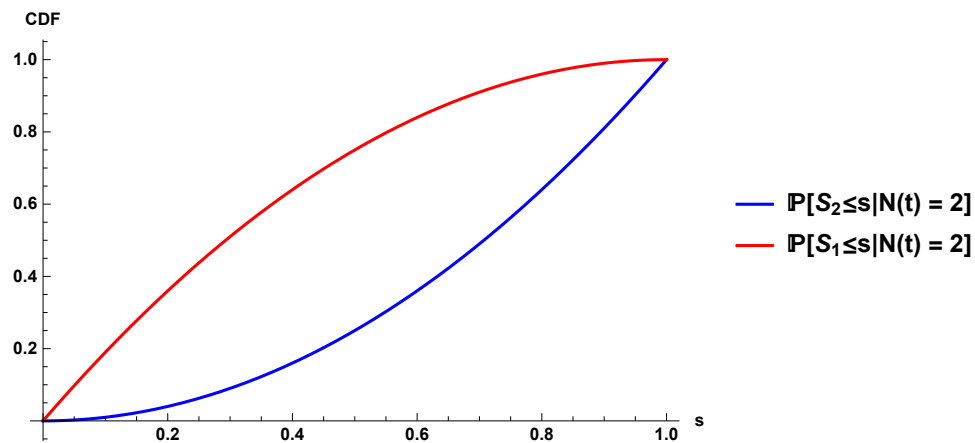
On the other hand, the first term has two parts that must be calculated.

$$\begin{aligned}
 \mathbb{P}[S_1 \leq s \mid \mathcal{N}(t) = 2] &= \mathbb{P}[\mathcal{N}(s) \geq 1 \mid \mathcal{N}(t) = 2] = \frac{\mathbb{P}[\mathcal{N}(s) \geq 1 \cap \mathcal{N}(t) = 2]}{\mathbb{P}[\mathcal{N}(t) = 2]} \\
 &= (\mathbb{P}[\mathcal{N}(s) = 1] \mathbb{P}[\mathcal{N}(t) - \mathcal{N}(s) = 1] + \mathbb{P}[\mathcal{N}(s) = 2] \mathbb{P}[\mathcal{N}(t) - \mathcal{N}(s) = 0]) / \mathbb{P}[\mathcal{N}(t) = 2] \\
 &= \frac{(e^{-s\lambda} s \lambda) (e^{-(t-s)\lambda} \lambda (t-s))}{\left(\frac{1}{2} e^{-t\lambda} t^2 \lambda^2\right)} + \frac{\left(\frac{1}{2} e^{-s\lambda} s^2 \lambda^2\right) (e^{-(t-s)\lambda})}{\left(\frac{1}{2} e^{-t\lambda} t^2 \lambda^2\right)} \\
 &= 2 \frac{s(t-s)}{t^2} + \frac{s^2}{t^2} \\
 &= 2 \frac{s}{t} - \frac{s^2}{t^2}
 \end{aligned}$$

```

Module[{t = 1},
  Plot[{ $\frac{s^2}{t^2}$ ,  $\frac{2st - s^2}{t^2}$ }, {s, 0, t}, AxesLabel -> {St1["s"], St1["CDF"]},
  PlotStyle -> {Blue, Red}, PlotLegends -> {" $\mathbb{P}[S_2 \leq s \mid \mathcal{N}(t) = 2]$ ", " $\mathbb{P}[S_1 \leq s \mid \mathcal{N}(t) = 2]$ "}]]

```

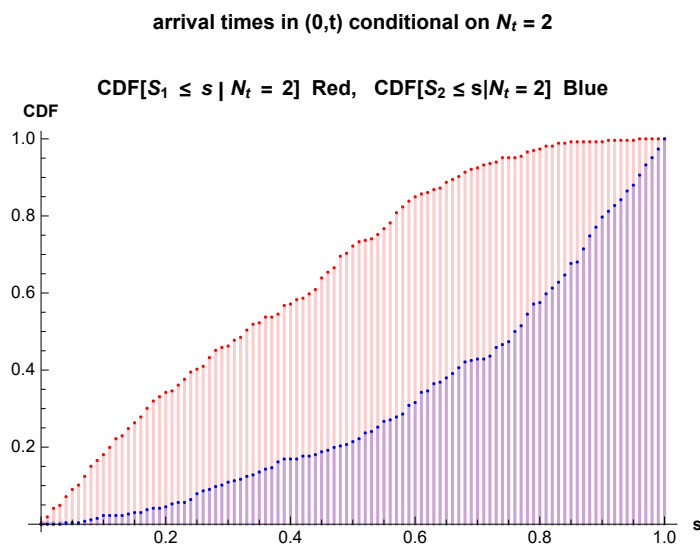


Check these results against simulation

```

Module[{λ = 2, t = 1, nSamples = 1000, ϕ, results,
  resultsConditional, S1vals, S2vals, arrivalTimes, H, lab},
  ϕ = PoissonProcess[λ];
  results = RandomFunction[ϕ, {0, t}, nSamples];
  (* Select those results that are conditional on Nt = 2 *)
  resultsConditional = Select[results["Paths"], (#[-1, 2] == 2) &];
  (* list of first arrival times *)
  S1vals = (#[[2]] & /@ resultsConditional);
  S1vals = S1vals [[All, 1]];
  (* list of second arrival times *)
  S2vals = (#[[3]] & /@ resultsConditional);
  S2vals = S2vals [[All, 1]];
  H[1] = EmpiricalDistribution[S1vals];
  H[2] = EmpiricalDistribution[S2vals];
  lab = St1["arrival times in (0,t) conditional on Nt = 2\n
CDF[S1 ≤ s | Nt = 2] Red, CDF[S2 ≤ s | Nt = 2] Blue"];
  DiscretePlot[{CDF[H[1], s], CDF[H[2], s]}, {s, 0, t, 0.01},
    PlotStyle → {Red, Blue}, AxesLabel → {St1["s"], St1["CDF"]}, PlotLabel → lab]

```



9 Total waiting time

Raic, Toman: Homogeneous Poisson Process - 10

10. Passengers arrive at a railway station according to a homogeneous Poisson process with intensity λ . At the beginning (time 0), there are no passengers at the station. The train departs at time t . Denote by W be the total waiting time of all passengers arrived up to the departure of the train. Compute $\mathbb{E}(W)$.

The expected time for passenger to arrive is

```
Tarrival == Expectation[x[s], x ≈ PoissonProcess[λ]]
```

```
Tarrival == s λ
```

The waiting time is

$$W = \int_0^t T_{\text{arrival}}(s) ds$$

So

```
W == Integrate[λ s, {s, 0, t}]
```

$$W == \frac{t^2 \lambda}{2}$$

10 Poisson shocks

Raic, Toman: Homogeneous Poisson Process - 11

11. *Poisson shocks*. Each arrival in a homogeneous Poisson process with intensity λ causes a *shock*. Its effect s time units later equals $e^{-\theta s}$. Denote by $X(t)$ the total effect of all the shocks from the interval $[0, t]$ at time t . Compute the expectation $\mathbb{E}[X(t)]$.

This is similar to the previous problem. The expected arrival time is

```
Tarrival == Expectation[x[s], x ≈ PoissonProcess[λ]]
```

```
Tarrival == s λ
```

In this case we want the cumulative shock at time t .

$$\text{shock} = \int_0^t \text{Exp}[-\theta(t - T_{\text{arrival}}(s))] ds$$

or

```
shock == Integrate[Exp[-θ (t - λ s)], {s, 0, t}] // Simplify
```

$$\text{shock} == \frac{e^{-t\theta} (-1 + e^{t\theta\lambda})}{\theta\lambda}$$

I confirm with a simulation. I use the following function to generate arrival times.

```
(* Simulation of Poisson arrival times *)
Clear[SimulateArrivalTimes];
SimulateArrivalTimes[λ_, tMax_, nSamples_] :=
Module[{result},
  result = RandomFunction[PoissonProcess[λ], {0, tMax}, nSamples] ["Paths"];
  result = #[[2 ;; -2, 1]] & /@ result ]
```

For example,

```
Module[{λ = 1, tMax = 10, nSamples = 1},
  SimulateArrivalTimes[λ, tMax, nSamples]]
{ {4.84707, 7.08518, 9.32553, 9.39195, 9.88382, 9.94405} }
```

Note: I could also have simulated the arrival times as the sum of exponential random variables

```
Module[{λ = 1, tMax = 10, nSamples = 10, variates, arrivalTimes},
  variates = RandomVariate[ExponentialDistribution[λ], nSamples];
  arrivalTimes = FoldList[Plus, 0, variates][[2 ;; -1]]
{1.10836, 2.96512, 3.3785, 6.0416, 7.48959, 7.60273, 8.00323, 11.0246, 11.1447, 11.7115}
```

I compare theory and simulation for nominal parameter values.

```
Module[{λ = 1 (* per hour *), tMax = 12 (* hours *), nSamples = 1000, θ = 1/10,
  times, shockEffects, cumulativeShockEffects, theory, info, lab, ShockEffect},
  ShockEffect[startTimes_] := Exp[-θ (tMax - #)] & /@ startTimes;
  times = SimulateArrivalTimes[λ, tMax, nSamples];
  shockEffects = ShockEffect /@ times;
  cumulativeShockEffects = Total /@ shockEffects;
  theory = 
$$\frac{e^{-tMax \theta} (-1 + e^{tMax \theta})}{\theta \lambda}$$
;
  info = {{N@theory, Mean[cumulativeShockEffects],
    StandardDeviation[cumulativeShockEffects] / Sqrt[nSamples]}};
  PrependTo[info, {"theory", "mean", "sim error"}];
  lab =
  StringForm["expected total effect of simulated shocks\nλ = `` tMax = `` θ = ``\n
    nSamples = ``", λ, tMax, θ, nSamples];
  LGrid[
  info,
  lab]
```

expected total effect of simulated shocks

$\lambda = 1$ $tMax = 12$ $\theta = \frac{1}{10}$ $nSamples = 1000$

theory	mean	sim error
6.98806	6.96522	0.0691489

II Hen crossing road

Raic, Toman: Homogeneous Poisson Process - 12

12. A hen wants to cross a one-way road, where cars drive according to a homogeneous Poisson process with intensity λ cars a time unit, all with the same speed. It takes c time units for the hen to cross the road. Assume that the hen starts to cross the road immediately when there is a chance to do it without being run over by a car. Compute the expected total waiting and crossing time for the hen.

The description of how the process starts seems ambiguous. I interpret it as follows.

- The chicken arrives at the edge of the roads and observes a car coming. The time it takes the car to reach the point where the chicken wants to cross is T_1 .
- If $T_1 > T_c$ the chicken crosses immediately so the waiting time for this case is $T_{\text{wait}} = T_c$
- If $T_1 \leq T_c$, the chick waits for the first car to pass, then crosses when possible.

Another way of interpreting this problem is that the distribution of cars along the road is determined by a Poisson process. The chicken, however, can only see the next car and make his decision to cross based on its time to arrival at his location.

Solution method based on counting

The probability that the chicken can cross is easily calculated

```
w12A[1] = pCrossSafely ==
  Refine[Probability[T1 > TC, T1 ≈ ExponentialDistribution[λ]], TC ≥ 0]
pCrossSafely == e-TC λ
```

and, if it is safe to cross, the crossing time is just T_c

In general, the chicken may have to wait for several cars to pass before it is safe to cross. Assume that it is the appearance of the \mathcal{N}^{th} car that allows the chicken to cross. Thus, the waiting time is

$$T_{\text{wait}} = T_1 + T_2 + \dots + T_{\mathcal{N}-1} + T_c$$

where the T_1 , etc are random times and \mathcal{N} is also a random variable. The expected time of crossing in this set up is

$$\begin{aligned} \mathbb{E}[T_{\text{wait}} | \mathcal{N} = n] &= \mathbb{E}[T_1 | T_1 \leq T_c] + \dots + \mathbb{E}[T_{n-1} | T_{n-1} \leq T_c] + T_c \\ &= (n-1) \mathbb{E}[T_i | T_i \leq T_c] + T_c \end{aligned}$$

Now, \mathcal{N} is a random variable distributed geometrically

$$\begin{aligned} \mathbb{E}_{\mathcal{N}}[\mathbb{E}[T_{\text{wait}} | \mathcal{N} = n]] &= \\ \mathbb{E}[T_{\text{wait}}] &= \mathbb{E}_{\mathcal{N}}[(\mathcal{N}-1) \mathbb{E}[T_i | T_i \leq T_c] + T_c] \end{aligned}$$

The individual waiting times are

w12A[2] = E["T_i | T_i ≤ T_c"] ==
Refine[Expectation[TI | TI ≤ TC, TI ≈ ExponentialDistribution[λ]], TC > 0]

$$\mathbb{E}[T_i | T_i \leq T_c] = \frac{-1 + e^{T_c \lambda} - T_c \lambda}{(-1 + e^{T_c \lambda}) \lambda}$$

Since the geometric distribution is the number of failures before success with probability pCrossSafely

w12A[3] =
E \mathcal{N} ["N = n-1"] == Expectation[N = n, N ≈ GeometricDistribution[pCrossSafely]]

$$\mathbb{E}\mathcal{N}[N = n-1] = \frac{1 - p_{\text{CrossSafely}}}{p_{\text{CrossSafely}}}$$

Calculating the waiting time

w12A[4] = E \mathcal{N} ["N = n-1"] E["T_i | T_i ≤ T_c"] + TC /. (w12A[2] // ER) /. (w12A[3] // ER)

$$TC + \frac{(1 - p_{\text{CrossSafely}}) (-1 + e^{T_c \lambda} - T_c \lambda)}{(-1 + e^{T_c \lambda}) p_{\text{CrossSafely}} \lambda}$$

w12A[5] = Twait == w12A[4] /. (w12A[1] // ER) // Simplify

$$T_{\text{wait}} = \frac{-1 + e^{T_c \lambda}}{\lambda}$$

Solution method based on renewal

After the first car passes, the chicken is faced with exactly the same situation.

$$\mathbb{E}[T_{\text{wait}} | T_1] = T_c \mathbf{1}\{T_1 \geq T_c\} + (T_1 + \mathbb{E}[T_{\text{wait}}]) \mathbf{1}\{T_1 < T_c\}$$

Take the expectation with respect T₁

$$\mathbb{E}_{T_1}[\mathbb{E}[T_{\text{wait}} | T_1]] = \mathbb{E}_{T_1}[T_c \mathbf{1}\{T_1 \geq T_c\}] + \mathbb{E}_{T_1}[(T_1 + \mathbb{E}[T_{\text{wait}}]) \mathbf{1}\{T_1 < T_c\}]$$

$$\mathbb{E}[T_{\text{wait}}] = T_c \mathbb{E}_{T_1}[\mathbf{1}\{T_1 \geq T_c\}] + \mathbb{E}_{T_1}[T_1 \mathbf{1}\{T_1 < T_c\}] + \mathbb{E}_{T_1}[\mathbb{E}[T_{\text{wait}}] \mathbf{1}\{T_1 < T_c\}]$$

$$\mathbb{E}[T_{\text{wait}}] = T_c \mathbb{P}[T_1 \geq T_c] + \mathbb{E}_{T_1}[T_1 \mathbf{1}\{T_1 < T_c\}] + \mathbb{E}[T_{\text{wait}}] \mathbb{P}[T_1 < T_c]$$

Note $\mathbb{P}[T_1 \geq T_c]$ and $\mathbb{P}[T_1 < T_c]$ are given by

{Refine[Probability[T1 ≥ TC, T1 ≈ ExponentialDistribution[λ]], TC ≥ 0],
Refine[Probability[T1 < TC, T1 ≈ ExponentialDistribution[λ]], TC ≥ 0]}

$$\{e^{-T_c \lambda}, 1 - e^{-T_c \lambda}\}$$

Also

```
Refine[Expectation[T1 Boole[T1 < TC], T1 ≈ ExponentialDistribution[λ]], TC > 0] //
Simplify
```

$$\frac{e^{-TC\lambda} (-1 + e^{TC\lambda} - TC\lambda)}{\lambda}$$

$$w12B[1] = \mathbb{E}[Twait] == TC e^{-TC\lambda} + \frac{e^{-TC\lambda} (-1 + e^{TC\lambda} - TC\lambda)}{\lambda} + \mathbb{E}[Twait] (1 - e^{-TC\lambda})$$

$$\mathbb{E}[Twait] == e^{-TC\lambda} TC + \frac{e^{-TC\lambda} (-1 + e^{TC\lambda} - TC\lambda)}{\lambda} + (1 - e^{-TC\lambda}) \mathbb{E}[Twait]$$

```
w12B[2] = Solve[w12B[1], E[Twait]] [[1, 1]]
```

$$\mathbb{E}[Twait] \rightarrow \frac{-1 + e^{TC\lambda}}{\lambda}$$

which is the same result as obtained with the counting method.

12 Waiting for a reward

Raic, Toman : Homogeneous Poisson Process - 13

13. In a certain place at a fair, prizes are shared every now and then. Everyone being in that place when the prizes are being shared gets a prize. The sharing times form a homogeneous Poisson process with intensity λ . At time zero, Tony observes that a sharing is in progress. He rushes towards the sharing place, but is too late. Then he waits for the next sharing, but at most for time δ : after that time, he gets bored and moves elsewhere. As soon as another sharing starts, he rushes towards the place, but is again too late and starts waiting for the next sharing, again at most for time δ . So he repeats until he gets the prize. Denote by T the time when Tony eventually gets his prize. Compute $\mathbb{E}(T)$, assuming that the fair is open infinitely long.

This is actually just a restatement of the chicken crossing the road problem.

The probability a prize is won on any given drawing is $p = \mathbb{P}[T < \delta]$.

```
Refine[Probability[T < δ, T ≈ ExponentialDistribution[λ]], δ ≥ 0]
```

$$1 - e^{-\delta\lambda}$$

Suppose the drawing is won on the n^{th} attempt. Then

$$\mathbb{E}[T | \mathcal{N} = n] = (n - 1)\mathbb{E}[\Delta T | \Delta T \geq \delta] + \mathbb{E}[\Delta T | \Delta T < \delta]$$

We have

```
Refine[Expectation[T | T ≥ δ, T ≈ ExponentialDistribution[λ]], δ > 0] // Expand
```

$$\delta + \frac{1}{\lambda}$$

Also,

```
Refine[Expectation[T | T < δ, T ≈ ExponentialDistribution[λ]], δ > 0] // Simplify
```

$$\frac{1 - e^{-\delta\lambda} + \delta\lambda}{\lambda - e^{-\delta\lambda}\lambda}$$

Take the expectation with respect to \mathcal{N}

$$\mathbb{E}_{\mathcal{N}}[\mathbb{E}[T | \mathcal{N} = n]] = \mathbb{E}_{\mathcal{N}}[(n - 1) \mathbb{E}[\Delta T \mathbf{1}\{\Delta T \geq \delta\}] + \mathbb{E}[\Delta T \mathbf{1}\{\Delta T < \delta\}]]$$

$$\mathbb{E}[T] = \mathbb{E}_{\mathcal{N}}[(n - 1) \mathbb{E}[\Delta T \mathbf{1}\{\Delta T \geq \delta\}] + \mathbb{E}[\Delta T \mathbf{1}\{\Delta T < \delta\}]]$$

```
Expectation[N, N ≈ GeometricDistribution[p]]
```

$$\frac{1 - p}{p}$$

$$w13[1] = \frac{1 - p}{p} \left(\delta + \frac{1}{\lambda} \right) + \frac{1 - e^{-\delta\lambda} + \delta\lambda}{\lambda - e^{-\delta\lambda}\lambda}$$

$$\frac{(1 - p) \left(\delta + \frac{1}{\lambda} \right)}{p} + \frac{1 - e^{-\delta\lambda} + \delta\lambda}{\lambda - e^{-\delta\lambda}\lambda}$$

Finally,

```
w13[2] = w13[1] /. p → 1 - e^{-δλ} // Simplify
```

$$\frac{e^{-\delta\lambda}}{(-1 + e^{-\delta\lambda})\lambda}$$

which is Raic's result.

13 Trucks and cars

Raic, Toman : Marking, Thinning, Superposition - 1

1. Suppose that the night traffic on the Jadranska street can be modelled by a homogeneous Poisson process with intensity 40 vehicles per hour, 10% thereof being lorries and 90% being cars. Suppose that the types of particular vehicles are independent.
 - a) Find the probability that in one hour, at least one lorry passes the Faculty of Mathematics and Physics.
 - b) Suppose that during the first hour of observation, exactly 10 lorries have passed by. Find the expected number of the cars passing the FMF during the same period.
 - c) Suppose that in the first hour of observation, exactly 50 vehicles have passed by. Find the probability that among these vehicles, there were exactly 5 lorries and 45 cars.
 - d) Compute the expected number of cars until the first lorry has passed by.

These are straightforward calculations

Part a)

Probability

```
Module[{λT = 0.1 × 40, PT, result},
  PT = PoissonProcess[λT];
  result = Probability[x[1] ≥ 1, x ≈ PT]
```

0.981684

Part b) The expected number of cars is independent of how many trucks have passed

```
Module[{λT = 0.1 × 40, λC = 0.9 × 40, PC, result},
  PC = PoissonProcess[λC];
  result = Expectation[x[1], x ≈ PC]
```

36.

Part c) We have to consider the number of ways the cars and trucks could be arranged.

```
Module[{nVehicles = 50, nCars = 45, nTrucks = 5, pCar = 0.9, pTruck = 0.1},
  Binomial[50, 5] pTrucknTrucks (1 - pTruck)nVehicles - nTrucks]
```

0.184925

Part d)

```
Module[{λT = 4, λC = 36, T, ENumberCars, ETimeOneTruck},
  T = Expectation[T, T ≈ ExponentialDistribution[λT]];
  Print["ETimeOneTruck = ", T];
  ENumberCars = Expectation[NC[T], NC ≈ PoissonProcess[λC]];
  Print["ENumberCars = ", ENumberCars]
```

$$\mathbb{E}\text{TimeOneTruck} = \frac{1}{4}$$

$$\mathbb{E}\text{NumberCars} = 9$$

14 Light bulb failures

Raic, Toman : Marking, Thinning, Superposition - 2

2. The life time of a light bulb follows the exponential distribution with expectation 200 days. When the bulb blows out, it is replaced immediately by a maintainer. Meanwhile, another maintainer replaces the bulbs regardless of their condition, according to a homogeneous Poisson process with intensity 0.01 replacement per day. Of course, we assume that the bulbs, actually their life times, are independent of each other.

- How frequently is a bulb replaced?
- For a longer period, compute the percentage of the bulbs replaced because of blowing out and the bulbs replaced because of 'precaution.'

a) What is the minimum of the two expected times?

```
Module[{λBurnout = 1/200, λReplacement = 1/100},
  Expectation[Min[x, y], {x ≈ ExponentialDistribution[λBurnout],
    y ≈ ExponentialDistribution[λReplacement]}] // N
```

66.6667

b) The expected time to burn out is 200 days. The expected time to be replaced is 100 days.

```
Module[{λBurnout = 1/200, λReplacement = 1/100},
  Probability[x < y, {x ≈ ExponentialDistribution[λBurnout],
    y ≈ ExponentialDistribution[λReplacement]}]]
```

$\frac{1}{3}$

15 Three actors

Raic, Toman : Marking, Thinning, Superposition - 3

3. A director is searching for three actors, one man and two women. Men apply according to a homogeneous Poisson process with intensity 2 per day, while women apply according to a homogeneous Poisson process with intensity 1 per day, independently of the men. Compute the expected time needed for the director to get the man as well as the two women. We assume that all candidates are acceptable.

The independent expected time required to hire 1 man is

`Expectation[M, M ≈ ExponentialDistribution[λM]]`

$$\frac{1}{\lambda M}$$

while the independent expected time to hire two women is

`Expectation[W1 + W2,
{W1 ≈ ExponentialDistribution[λW], W2 ≈ ExponentialDistribution[λW]}]`

$$\frac{2}{\lambda W}$$

In this case the events are dependent. The quantity of interest is the maximum of M and W1 + W2

`w15[1] = Expectation[Max[M, W1 + W2], {M ≈ ExponentialDistribution[λM],
W1 ≈ ExponentialDistribution[λW], W2 ≈ ExponentialDistribution[λW]}]`

$$\frac{2 \lambda M^3 + 4 \lambda M^2 \lambda W + 2 \lambda M \lambda W^2 + \lambda W^3}{\lambda M \lambda W (\lambda M + \lambda W)^2}$$

For the supplied intensities

`w15[2] = w15[1] /. {λM → 2, λW → 1} // N`

2.05556

I could also use the fact that the sum of exponentially distributed times is distributed like an Erlang distribution

`w15[3] = Expectation[Max[M, W],
{M ≈ ExponentialDistribution[λM], W ≈ ErlangDistribution[2, λW]}]`

$$\frac{2 \lambda M^3 + 4 \lambda M^2 \lambda W + 2 \lambda M \lambda W^2 + \lambda W^3}{\lambda M \lambda W (\lambda M + \lambda W)^2}$$

16 Used car

Raic, Toman : Marking, Thinning, Superposition - 4

4. A married couple is searching for a used car. Each one of them is looking up offers for their favourite brand. The offers for wife's brand are coming according to a homogeneous Poisson process with intensity λ , while the offers for husband's brand are coming according to a homogeneous Poisson process with intensity μ . The wife is ready to buy a car when she encounters the third offer for her brand, while the husband is ready when he encounters the second offer for his brand. The couple buys a car when either of them is ready to buy. We assume that the offer processes for both brands are independent.
- Compute the probability that the couple buys a car according to wife's choice.
 - Compute the expected time needed to buy a car.

Part a) The expected time for the wife to buy is (by two different methods)

$$\{\text{Expectation}[TW, TW \approx \text{ErlangDistribution}[3, \lambda W]],$$

$$\text{Expectation}[T1 + T2 + T3, \{T1 \approx \text{ExponentialDistribution}[\lambda W],$$

$$T2 \approx \text{ExponentialDistribution}[\lambda W], T3 \approx \text{ExponentialDistribution}[\lambda W]\}]\}$$

$$\left\{ \frac{3}{\lambda W}, \frac{3}{\lambda W} \right\}$$

The expected time for the husband to buy is

$$\text{Expectation}[TH, TH \approx \text{ErlangDistribution}[2, \lambda H]]$$

$$\frac{2}{\lambda H}$$

The probability that the wife buys is the probability of the event $TW > TH$

$$\text{Probability}[TW > TH,$$

$$\{TW \approx \text{ErlangDistribution}[3, \lambda W], TH \approx \text{ErlangDistribution}[2, \lambda H]\}]$$

$$\frac{\lambda H^2 (\lambda H^2 + 4 \lambda H \lambda W + 6 \lambda W^2)}{(\lambda H + \lambda W)^4}$$

Part b)

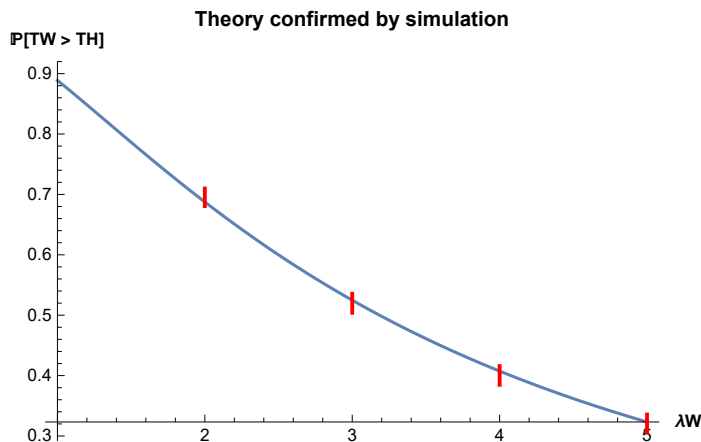
$$\text{Expectation}[\text{Min}[TW, TH],$$

$$\{TW \approx \text{ErlangDistribution}[3, \lambda W], TH \approx \text{ErlangDistribution}[2, \lambda H]\}]$$

$$\frac{2 \lambda H^3 + 8 \lambda H^2 \lambda W + 12 \lambda H \lambda W^2 + 3 \lambda W^3}{(\lambda H + \lambda W)^4}$$

I check the probability calculation against simulation

```
Module[{kW = 3, kH = 2, λW = 3, λH = 2, nSamples = 4000, simResults, lab, Bar},
  Bar[{\λ_, mean_, error_}] := Line[{\λ, mean - 2 error}, {\λ, mean + 2 error}];
  simResults =
    Table[Flatten@{\λW, Simulation[kW, kH, λW, λH, nSamples]}, {\λW, 2, 5}] // N;
  lab = Stl["Theory confirmed by simulation"];
  Plot[
$$\frac{\lambda H^2 (\lambda H^2 + 4 \lambda H \lambda W + 6 \lambda W^2)}{(\lambda H + \lambda W)^4}$$
,
    {\λW, 1, 5}, AxesLabel → {Stl["λW"], Stl["P[TW > TH]"]},
    Epilog → {RED, Bar /@ simResults}, PlotLabel → lab]]
```



```
Clear[Simulation];
Simulation[kW_, kH_, λW_, λH_, nSamples_] :=
  Module[{\εW, εH, timePairs, simResults, simMean, simError},
    εW = ExponentialDistribution[λW];
    εH = ExponentialDistribution[λH];
    timePairs =
      Table[Total /@ {RandomVariate[εW, kW], RandomVariate[εH, kH]}, {nSamples}];
    simResults = Boole[#[[1]] > #[[2]]] & /@ timePairs;
    {simMean, simError} =
      {Mean[simResults], StandardDeviation[simResults] / Sqrt[nSamples]} // N]
```

17 Comparing two Poisson processes

Raic, Toman : Marking, Thinning, Superposition - 5

5. Consider two parallel Poisson processes with intensities λ and μ . Find:
- the probability that in the second process, exactly one arrival in occurs before the first arrival in the first process;
 - the expected number of the arrivals in the second process before the first arrival in the first process.

Part a)

For two merged processed, the probability that T1 occurs before T2 is

$$w[1] = p1 == \text{Probability}[T1 < T2, \\ \{T1 \approx \text{ExponentialDistribution}[\lambda1], T2 \approx \text{ExponentialDistribution}[\lambda2]\}]$$

$$p1 = \frac{\lambda1}{\lambda1 + \lambda2}$$

and similarly,

$$w[2] = p2 == \text{Probability}[T2 < T1, \\ \{T1 \approx \text{ExponentialDistribution}[\lambda1], T2 \approx \text{ExponentialDistribution}[\lambda2]\}]$$

$$p2 = \frac{\lambda2}{\lambda1 + \lambda2}$$

Thus the probability that an event in process2 occurs before an event in process1 is just $p2$ $p1$, in analogy to Bernoulli trials.

$$p1 p2 /. (w[1] // ER) /. (w[2] // ER)$$

$$\frac{\lambda1 \lambda2}{(\lambda1 + \lambda2)^2}$$

Part b)

Consider that arrival events occurring in the second process are “failures” that occur before a process 1 arrival event that constitutes a “success”. The probability of success is $p = \left(\frac{\lambda1}{\lambda1 + \lambda2}\right)$. The GeometricDistribution describes the number of such failures.

$$\text{Expectation}[x, x \approx \text{GeometricDistribution}[p]] /. p \rightarrow \lambda1 / (\lambda1 + \lambda2) // \text{Simplify}$$

$$\frac{\lambda2}{\lambda1}$$

It is also constructive to work through Raic’s second method of solving this problem. Consider the event

$$X2 = \{\text{number of arrivals in process 2 before first arrival in process 1}\}$$

Then denote by $T1$ the first arrival time in process 1. We are interested in

$$P[X2 = 1 | T1]$$

The probability of exactly 1 arrival is

$$wb[1] = \text{Probability}[X2[T1] == 1, X2 \approx \text{PoissonProcess}[\lambda2]]$$

$$e^{-T1 \lambda2} T1 \lambda2$$

But $T1$ is exponentially distributed.

$$\text{wb}[2] = \text{Expectation}[\text{wb}[1], T1 \approx \text{ExponentialDistribution}[\lambda1]]$$

$$\frac{\lambda1 \lambda2}{(\lambda1 + \lambda2)^2}$$

Similar logic is used for the expected number of arrivals

$$\text{wb}[3] = \text{Expectation}[X2[T1], X2 \approx \text{PoissonProcess}[\lambda2]]$$

$$T1 \lambda2$$

$$\text{wb}[4] = \text{Expectation}[\text{wb}[3], T1 \approx \text{ExponentialDistribution}[\lambda1]]$$

$$\frac{\lambda2}{\lambda1}$$

18 Arriving students

Raic, Toman : Marking, Thinning, Superposition - 5

6. From a given time on, students of financial and general mathematics have an opportunity to lookup their corrected test papers. The students of financial mathematics arrive according to a homogeneous Poisson process with intensity 4 students an hour, while the students of general mathematics arrive according to a homogeneous Poisson process with intensity 2 students an hour. Assume that both groups of students arrive independently of each other.

Suppose that during the first half an hour, exactly one student came to lookup his/her test paper. Compute the conditional arrival time of the first student of *financial* mathematics. The time is measured from the beginning, assuming that the students are arriving infinitely long in the future.

This is an exercise in the logic and manipulation of conditional expectations. I follow Raic's argument. There are two possibilities — a financial student arrives first with probability p_F or a general student arrives first with probability p_G . Define the event A that one student arrived in the first half hour, A_F the event that that student was a financial student, etc. Then

$$\mathbb{P}[A_F | A] = p_F = \frac{\lambda_F}{\lambda_F + \lambda_G}, \quad \mathbb{P}[A_G | A] = p_G = \frac{\lambda_G}{\lambda_F + \lambda_G}$$

Denote T_F the time that the first financial student arrived. If we are given that A_F is true, then the arrival could have occurred anytime before $T_S = 1/2$. So

$$T_F | A_F \approx \text{UniformDistribution}[0, T_S]$$

On the other hand, if A_G is true then T_F is exponentially distributed.

$$(T_F - T_S) | A_G \approx \text{ExponentialDistribution}[\lambda_F]$$

From these observations, the specific time T_S enters the the calculations.

The specific calculation proceeds

$$\begin{aligned} \mathbb{E}[T_F | A] &= \frac{\mathbb{E}[T_F \cap A]}{\mathbb{P}[A]} = \frac{\mathbb{E}[T_F \cap A_F] + \mathbb{E}[T_F \cap A_G]}{\mathbb{P}[A]} \\ &= \frac{\mathbb{P}[A_F]}{\mathbb{P}[A]} \mathbb{E}[T_F | A_F] + \frac{\mathbb{P}[A_G]}{\mathbb{P}[A]} \mathbb{E}[T_F | A_G] \end{aligned}$$

Next note

$$\mathbb{P}[A_F | A] = \frac{\mathbb{P}[A_F \cap A]}{\mathbb{P}[A]} = \frac{\mathbb{P}[A_F]}{\mathbb{P}[A]} \quad (\text{because } A_F \text{ is part of } A)$$

So

$$\begin{aligned} \mathbb{E}[T_F | A] &= \frac{\mathbb{P}[A_F]}{\mathbb{P}[A]} \mathbb{E}[T_F | A_F] + \frac{\mathbb{P}[A_G]}{\mathbb{P}[A]} \mathbb{E}[T_F | A_G] \\ &= \mathbb{P}[A_F | A] \mathbb{E}[T_F | A_F] + \mathbb{P}[A_G | A] \mathbb{E}[T_F | A_G] \\ &= p_F \mathbb{E}[T_F | A_F] + p_G \mathbb{E}[T_F | A_G] \end{aligned}$$

The first expectation is

$$\begin{aligned} \mathbb{E}["T_F | A_F"] &= \text{Expectation}[TF, TF \approx \text{UniformDistribution}[\{0, TS\}]] \\ \mathbb{E}[T_F | A_F] &= \frac{TS}{2} \end{aligned}$$

Note that

$$\mathbb{E}[T_F | A_G] = \mathbb{E}[T_S + (T_F - T_S) | A_G] = T_S + \mathbb{E}[(T_F - T_S) | A_G]$$

This second expectation is

$$\begin{aligned} \mathbb{E}["T_F - T_S | A_G"] &= \text{Expectation}[TFS, TFS \approx \text{ExponentialDistribution}[\lambda_F]] \\ \mathbb{E}[T_F - T_S | A_G] &= \frac{1}{\lambda_F} \end{aligned}$$

Thus

$$\begin{aligned}
\mathbb{E}[T_F | A] &= p_F \mathbb{E}[T_F | A_F] + p_G \mathbb{E}[T_F | A_G] \\
&= p_F \left(\frac{T_S}{2} \right) + p_G \left(T_S + \frac{1}{\lambda_F} \right) \\
&= \frac{\lambda_F}{\lambda_F + \lambda_G} \left(\frac{T_S}{2} \right) + \frac{\lambda_G}{\lambda_F + \lambda_G} \left(T_S + \frac{1}{\lambda_F} \right)
\end{aligned}$$

19 Distribution of relative arrivals of two Poisson processes

Raic, Toman : Marking, Thinning, Superposition - 7

7. Take two independent Poisson processes with intensities λ and μ . Let $n \in \mathbb{N}$. Determine the distribution of the number of arrivals in the first process before the n -th arrival in the second process.

Since the arrivals in the two processes can be considered to be Bernoulli trials with probabilities

$$p_1 = \frac{\lambda_1}{\lambda_1 + \lambda_2} \quad p_2 = \frac{\lambda_2}{\lambda_1 + \lambda_2}$$

this is just an exercise in constructing a binomial distribution. For there to be n process 1 events before m process 2 events.

$$\mathbb{P}[n \text{ process 1 events before } m \text{ process 2 events}] = \sum_{k=n}^{n+m-1} \text{Binomial}[n+m-1, k] p_1^k p_2^{n+m-1-k}$$

20 Probability of reaching a point in space

8. Consider two homogeneous Poisson processes with intensities λ and μ . Denote by $N_t^{(1)}$ the number of arrivals up to time t in the first, and with $N_t^{(2)}$ the number of arrivals up to time t in the second process. Compute the probability that the two-dimensional walk $(N_t^{(1)}, N_t^{(2)})$ reaches the point (i, j) .

This is easy. It takes $n_{\text{Total}} = i + j$ steps to get to $\{i, j\}$. There are $\text{Binomial}[i + j, i]$ possible paths. The probability of an X step is $p_X = \lambda X / (\lambda X + \lambda Y)$ and the analogous for Y. The probability of reaching $\{i, j\}$ is

$$\mathbb{P}[X(t) = i, Y(t) = j] = \text{Binomial}[i + j, i] \left(\frac{\lambda X}{\lambda X + \lambda Y} \right)^i \left(\frac{\lambda Y}{\lambda X + \lambda Y} \right)^j$$

I check against simulation.

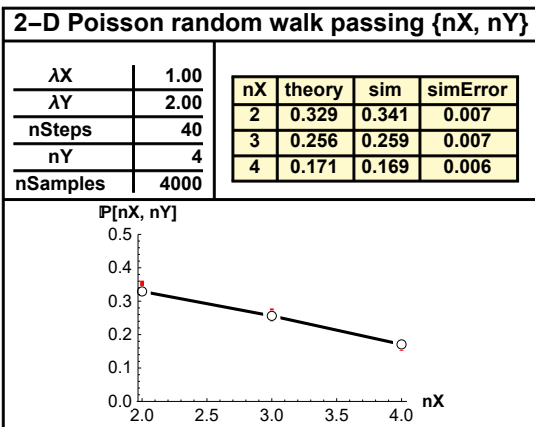
```
Clear [PTheoretical];
PTheoretical[nX_, nY_, pX_, pY_] := Binomial[nX + nY, nX] pX^nX pY^nY
```

```

Module[{λX = 1, λY = 2, nSteps = 40, nY = 4, nSamples = 4000, pX, pY, results},
  {pX, pY} = {λX, λY} / (λX + λY);
  results = Table[Flatten[{nX, N@PTheoretical[nX, nY, pX, pY],
    SimModel[nX, nY, λX, λY, nSteps, nSamples]}],
    {nX, 2, 4}];

Module[{info, params, bars, Bar, G},
  Bar[{λ_, th_, mean_, error_}] := Line[{{λ, mean - 2 error}, {λ, mean + 2 error}}];
  info = ({#[[1]], NF3@#[[2]], NF3@#[[3]], NF3@#[[4]]}) & /@ results;
  PrependTo[info, {"nX", "theory", "sim", "simError"}];
  params = {{{"λX", λX, {5, 2}}, {"λY", λY, {5, 2}}, {"nSteps", nSteps, {5, 0}},
    {"nY", nY, {5, 0}}, {"nSamples", nSamples, {5, 0}}};
  G[0] = Stl[StringForm["2-D Poisson random walk passing {nX, nY}"];
  G[1] = ParameterTable[params];
  G[2] = LGrid[info, ""];
  bars = {RED, Bar /@ results};
  G[3] = ListLinePlot[results[[All, {1, 2}]], PlotMarkers →
    {ChartElementData["SimpleMarkers"][[6]]}, PlotStyle → Black, Prolog → bars,
    AxesLabel → {Stl["nX"], Stl["P[nX, nY]"]}, PlotRange → {Automatic, {0, 0.5}}];
  Grid[{{G[0], SpanFromLeft}, {G[1], G[2]}, {G[3], SpanFromLeft}}, Frame → All]]

```



```

Clear[SimModel];
SimModel[nX_, nY_, λX_, λY_, nSteps_, nSamples_] :=
Module[{pX, pY, P, paths, results, mean, errorMC},
  {pX, pY} = {λX, λY} / (λX + λY);
  {P, P} = {PoissonProcess[pX], PoissonProcess[pY]};
  paths = Table[GenerateMergedPath[λX, λY, pX, pY, P, P, nSteps], {nSamples}];
  results = TargetPathQ[#, nX, nY] & /@ paths;
  {mean, errorMC} = {Mean[results], StandardDeviation[results] / Sqrt[nSamples]} // N;
  {mean, errorMC}]

```

TargetPathQ is true if the path touches {nX, nY}

```
Clear[TargetPathQ];
TargetPathQ[path_, nX_, nY_] :=
  Boole[Length@Select[path[[1 ;; nX + nY]], (#[[2]] == "X") & ] == nX]
```

```
Clear[GenerateMergedPath];
GenerateMergedPath[λX_, λY_, pX_, pY_, ρX_, ρY_, nSteps_] :=
  Module[{pathX, pathY, pathMerged},
    pathX = RandomFunction[ρX, {0, nSteps}]["Paths"][[1]][[2 ;; -2]];
    pathX = ({#[[1]], "X"}) & /@ pathX;
    pathY = RandomFunction[ρY, {0, nSteps}]["Paths"][[1]][[2 ;; -2]];
    pathY = ({#[[1]], "Y"}) & /@ pathY;
    (* Join the paths and sort by arrival time *)
    pathMerged = Sort[Join[pathX, pathY], (#1[[1]] < #2[[1]] & )]
```

21 London Bombs

From *Introduction to Probability*, Grinstead and Snell (available on web)

Example 5.4 In his book,¹ Feller discusses the statistics of flying bomb hits in the south of London during the Second World War.

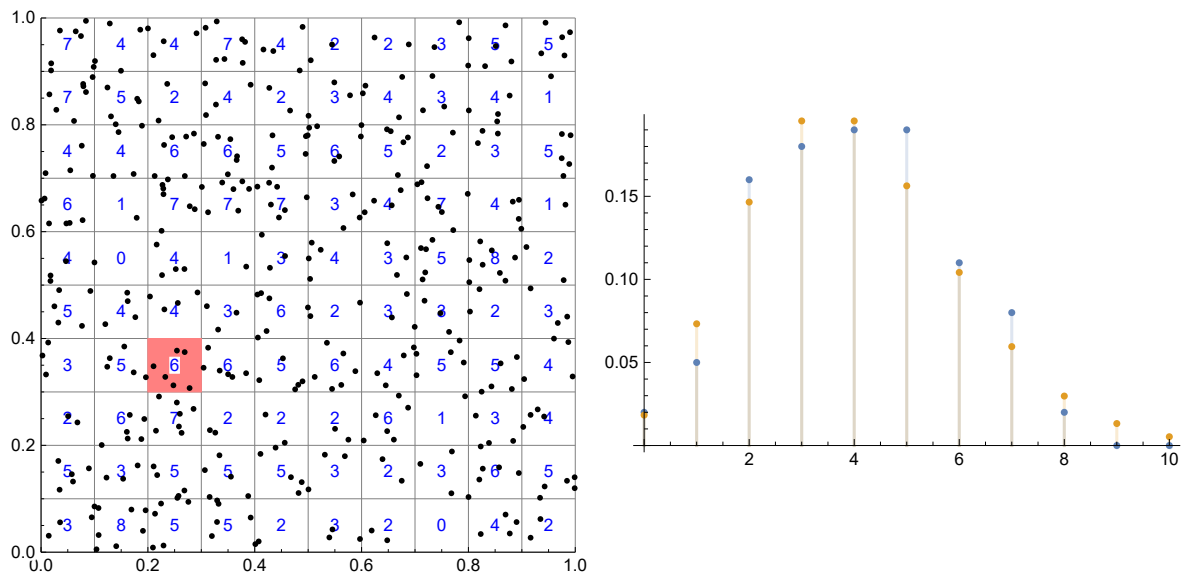
Assume that you live in a district of size 10 blocks by 10 blocks so that the total district is divided into 100 small squares. How likely is it that the square in which you live will receive no hits if the total area is hit by 400 bombs?

We assume that a particular bomb will hit your square with probability $1/100$. Since there are 400 bombs, we can regard the number of hits that your square receives as the number of *successes* in a Bernoulli trials process with $n = 400$ and $p = 1/100$. Thus we can use the **Poisson** distribution with $\lambda = 400 \cdot 1/100 = 4$ to approximate the probability that your square will receive j hits. This probability is $p(j) = e^{-4}4^j/j!$. The expected number of squares that receive exactly j hits is then $100 \cdot p(j)$. It is easy to write a program **LondonBombs** to simulate this situation and compare the expected number of squares with j hits with the observed number. In Exercise 26 you are asked to compare the actual observed data with that predicted by the **Poisson** distribution.

```

Module[{p = 1/100, nSamples = 400, λ, u, H, values, counts, rect, markers, G},
  λ = p nSamples;
  u = UniformDistribution[{{0, 1}, {0, 1}}];
  values = RandomVariate[u, nSamples];
  counts = Table[Length@
    Select[values, ((0.1 i < #[[1]] ≤ 0.1 (i + 1)) && (0.1 j < #[[2]] ≤ 0.1 (j + 1))) &],
    {i, 0, 9}, {j, 0, 9}];
  markers = {Blue, Table[Text[ToString[counts[[i, j]]], {0.1 i - 0.05, 0.1 j - 0.05}],
    {i, 1, 10}, {j, 1, 10}]}];
  rect = {Pink, Rectangle[{0.2, 0.3}, {0.3, 0.4}]}];
  G[1] =
  Graphics[{rect, markers, {Point /@ values}}, Axes → Automatic, AspectRatio → 1,
    GridLines → {Table[0.1 i, {i, 0, 10}], Table[0.1 i, {i, 0, 10}]},
    PlotRange → {{0, 1}, {0, 1}}, ImageSize → 300];
  H = HistogramDistribution[Flatten@counts];
  G[2] = DiscretePlot[{PDF[H, k],  $\frac{e^{-\lambda} \lambda^k}{k!}$ },
    {k, 0, 10}, PlotLegends → {"sim", "Poisson"}, ImageSize → 300];
  Grid[{{G[1], G[2]}}]]

```



22 Store Credit (compound Poisson process)

From Introductory Probability - Gravner UCD p 137 (available on web)

5. Assume that a Poisson number with expectation 10 of customers enters the store. For promotion, each of them receives an in-store credit uniformly distributed between 0 and 100 dollars. Compute the expectation and variance of the amount of credit the store will give.

This is purely an application of formulae derived for expectation and variance of a composite Poisson process. See *Poisson Process Basics 04-06-17*

Expectation and variance

$$\mathbb{E}[S] = \sum_{n=0}^{\infty} n \mu_X \mathbb{P}[N = n] = \mu_X \sum_{n=0}^{\infty} n \mathbb{P}[N = n] = \mu_X \mathbb{E}[N] = \mathbb{E}[X_1] \mathbb{E}[N]$$

$$\begin{aligned} \mathbb{V}[S] &= \sigma_X^2 \mathbb{E}[N] + \mu_X^2 \mathbb{V}[N] \\ &= \mathbb{V}[X] \mathbb{E}[N] + \mathbb{E}[X]^2 \mathbb{V}[N] \end{aligned}$$

`w[1] = {E[N] → Expectation[x, x ≈ PoissonDistribution[λ]],
V[N] → Variance[PoissonDistribution[λ]]}`

`{E[N] → λ, V[N] → λ}`

`w[2] = {E[X] → Expectation[x, x ≈ UniformDistribution[{a, b}]],
V[X] → Variance[UniformDistribution[{a, b}]]}`

`{E[X] → $\frac{a+b}{2}$, V[X] → $\frac{1}{12} (-a+b)^2$ }`

`w[3] = {E[N] E[X], V[X] E[N] + E[X]^2 V[N]}`

`{E[X] E[N], E[N] V[X] + E[X]^2 V[N]}`

`w[4] = w[3] /. w[1] /. w[2]`

`{ $\frac{1}{2} (a+b) \lambda$, $\frac{1}{12} (-a+b)^2 \lambda + \frac{1}{4} (a+b)^2 \lambda$ }`

`w[5] = w[4] /. {a → 0, b → 100, λ → 10} // N`

`{500., 33333.3}`

which agrees with Gravner

`(1002/12.) 10 + 502 × 10`

`33333.3`

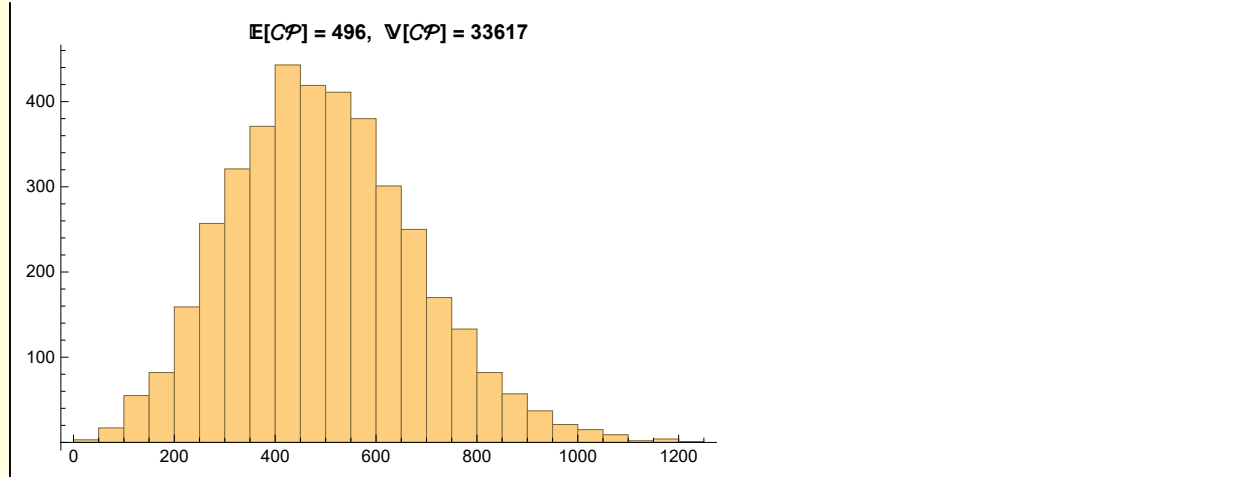
Mathematica has built-in functionality for compound Poisson processes

```
Module[{D},
  D = CompoundPoissonDistribution[λ, UniformDistribution[{a, b}]];
  {{Mean[D], Variance[D]},
   {Mean[D], Variance[D]} /. {a → 0, b → 100, λ → 10} // Simplify
```

```
{{ $\frac{1}{2} (a+b) \lambda$ ,  $\frac{1}{3} (a^2 + a b + b^2) \lambda$ }, {500., 33333.3}}
```


Since this class of problems is important, I'll perform a simulation.

```
Module[{a = 0, b = 100, λ = 10, nSamples = 4000, CP, pathList, lab},
  CP = CompoundPoissonDistribution[λ, UniformDistribution[{a, b}]];
  pathList = RandomVariate[CP, nSamples];
  lab = Stl@StringForm["E[CP] = ``, V[CP] = ``",
    Round@Mean[pathList], Round@Variance[pathList]];
  Histogram[pathList, {50}, PlotLabel → lab]
```



23 Poisson and the Law

Introductory Probability - Gravner UCD p 52 Famous People vs Collins case in LA 1968. (available on web)

Example 5.11. *Poisson distribution and law.* Assume a crime has been committed. It is known that the perpetrator has certain characteristics, which occur with a small frequency p (say, 10^{-8}) in a population of size n (say, 10^8). A person who matches these characteristics has been found at random (e.g., at a routine traffic stop or by airport security) and, since p is

so small, charged with the crime. *There is no other evidence.* We will also assume that the authorities stop looking for another suspect once the arrest has been made. What should the defense be?

I just follow Gravner's treatment. From a large number of people, n , suppose some modest number of people, \mathcal{N} , have with probability p the characteristics identified by the witness. Those people form the pool of suspects. One of them, C , is the actual criminal. Choose \mathcal{A} to be a person randomly arrested from the pool of suspects. What is the probability that $C = \mathcal{A}$?

$$\mathbb{P}[C = \mathcal{A} \mid \mathcal{N} \geq 1] = \frac{\mathbb{P}[C = \mathcal{A} \cap \mathcal{N} \geq 1]}{\mathbb{P}[\mathcal{N} \geq 1]}$$

Now

$$\mathbb{P}[C = \mathcal{A} \cap N \geq 1] = \sum_{k=1}^{\infty} \mathbb{P}[C = \mathcal{A} | N = k] \mathbb{P}[N = k]$$

Since there is no other evidence, each of suspects is equally guilty,

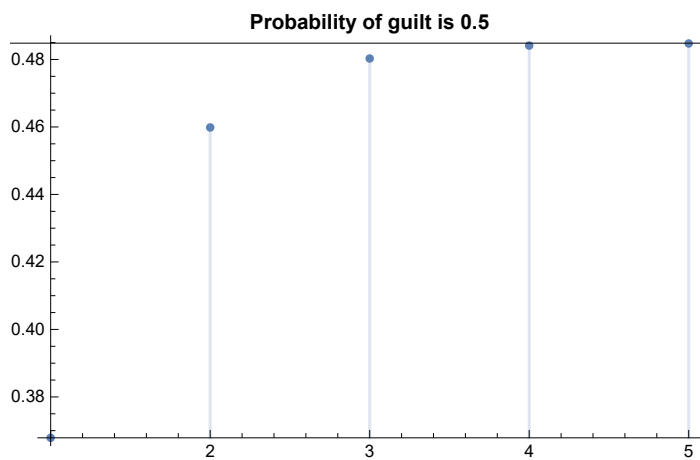
$$\mathbb{P}[C = \mathcal{A} | N = k] = \frac{1}{k}$$

The probability $\mathbb{P}[N = k]$ is binomial, but for small p can be approximated by Poisson with $\lambda = np$

$$\mathbb{P}[C = \mathcal{A} \cap N \geq 1] = e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^k}{k \cdot k!}$$

As a representative parameter value for λ given small p and large n , choose $\lambda = 1$

```
Module[{λ = 1, Plim, lab},
  Plim = Exp[-λ] Sum[ $\frac{\lambda^k}{k \cdot k!}$ , {k, 1, 10}];
  lab = Stl@StringForm["Probability of guilt is ``", NF1@N[Plim]];
  DiscretePlot[Exp[-λ] Sum[ $\frac{\lambda^k}{k \cdot k!}$ , {k, 1, kMax}], {kMax, 1, 5},
    Epilog -> Line[{{0, Plim}, {20, Plim}}], PlotLabel -> lab ]]
```



This result disagrees with Gravner — but there appears to be an error in the formula

so

$$P(C = A, N \geq 1) = \sum_{k=1}^{\infty} \frac{1}{k} \cdot \frac{\lambda^k}{k!} \cdot e^{-\lambda}.$$

The probability that the arrested person is guilty then is

$$P(C = A | N \geq 1) = \frac{e^{-\lambda}}{1 - e^{-\lambda}} \cdot \sum_{k=1}^{\infty} \frac{\lambda^k}{k \cdot k!}.$$

Where does that $1 - \text{Exp}[-\lambda]$ is the denominator come from ???

This problem is called the “prosecutor’s fallacy”

<https://www.youtube.com/watch?v=s2G5MQIT6Jk>