Continious Random Variables

Dr Ahmad Khonsari
ECE Dept.
The University of Tehran

Laplace transform of a sum

Let X and Y be independent random variables with L-transforms $F_X(s)$, $F_Y(s)$(from now we use $f_X^*(s)$ and $f_Y^*(s)$).

$$f_{X+Y}^*(s) = E[e^{-s(X+Y)}]$$

$$= E[e^{-sX}e^{-sY}]$$

$$= E[e^{-sX}]E[e^{-sY}] \qquad \text{(independence)}$$

$$= f_X^*(s)f_Y^*(s)$$

thus

$$f_{X+Y}^*(s)=f_X^*(s)f_Y^*(s)$$

Calculating moments with the aid of Laplace transform

By derivation one sees

$$f^{*'}(s) = \frac{d}{ds}E[e^{-sX}] = E[-Xe^{-sX}]$$

Similarly, the *nth* derivative is

$$f^{*(n)}(s) = \frac{d^n}{ds^n}E[e^{-sX}] = E[(-X)^n e^{-sX}]$$

Evaluating these at *s* = 0 one gets

$$E[X] = -f^{*'}(0)$$

$$E[X^2] = +f^{*''}(0)$$

$$E[X^n] = (-1)^n f^{*(n)}(0)$$

Laplace transform of a random sum

Consider the random sum

$Y = X_1 + \cdot \cdot \cdot + X_N$

where the X_i are *i.i.d.* with the common L-transform $f_X^*(s)$ and

$N \geq 0$ is a integer-valued r.v. with the generating function $\mathcal{G}_N(z)$.

$f_Y^*(s) = E[e^{-sY}]$

$= E[E[e^{-sY}|N]]$ (outer expectation with respect to variations of N)

$= E[E[e^{-s(X_1 + \cdot \cdot \cdot + X_N)}|N]]$ (in the inner expectation N is fixed)

$= E[E[e^{-s(X_1)}] \cdot \cdot \cdot E[e^{-s(X_N)}]|N]$ (independence)

$= E[(f_X^*(s))^N]$

$= \mathcal{G}_N(f_X^*(s))$ (by the definition $E[z^N] = \mathcal{G}_N(z)$)

Uniform distribution X ~ U(a, b) models the fact that any interval of length δ between *a* and *b* is equally likely.

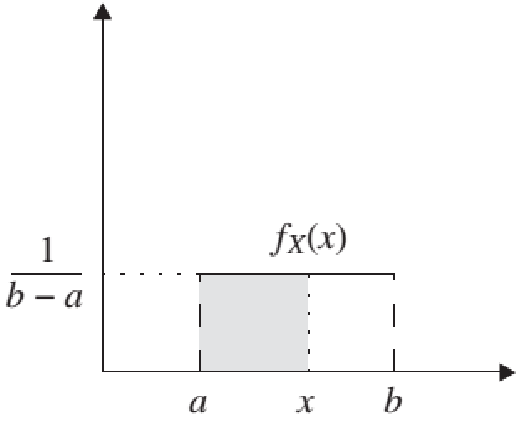
Definition: The pdf of *X* is constant in the interval (*a, b*):

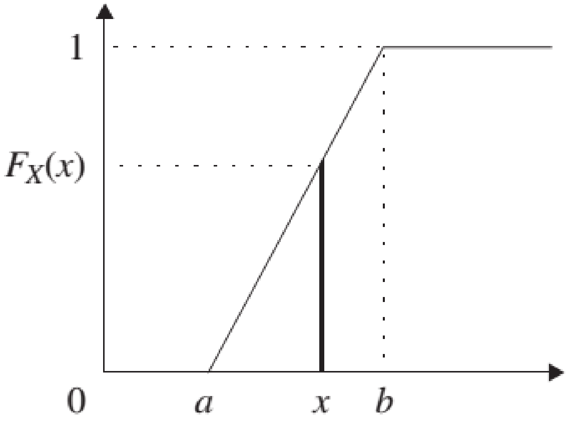
$$f_X(x) = \frac{1}{b-a} \quad a \le x \le b \qquad F_X(x) = \frac{x-a}{b-a} \quad a \le x < b \;, \; F_X(x) = 1 \quad b \le x$$

i.e. the value *X* is drawn randomly in the interval (*a, b*) (115)

Mean and variance are as follows:

$$E[X] = \int_{-\infty}^{\infty} x f(x) dx = \frac{a+b}{2} \quad V[X] = \int_{-\infty}^{\infty} \left(x - \frac{a+b}{2}\right)^2 f(x) dx = \frac{1}{12}(b-a)^2 \qquad (116)$$





Uniform distribution cntd.

Uniform distribution (continued)

Let $U_1, \ldots, U_n$ be independent uniformly distributed random variables, $U_i \sim \mathrm{U}(0, 1)$.

- The number of variables which are $\le x$ ($0 \le x \le 1$)) is $\sim \mathrm{Bin}(n, x)$
  - the event $\{U_i \le x\}$ defines a Bernoulli trial where the probability of success is x

Uniform distribution cntd.

Uniform distribution (continued)

- Let $U_{(1)}, \ldots, U_{(n)}$ be the ordered sequence of the values.

Define further $U_{(0)} = 0$ and $U_{(n+1)} = 1$.
It can be shown that all the intervals are identically distributed and

$$P\{U_{(i+1)} - U_{(i)} > x\} = (1 - x)^n \qquad i = 1, \ldots, n$$

- for the first interval $U_{(1)} - U_{(0)} = U_{(1)}$ the result is obvious because $U_{(1)} = \min(U_1, \ldots, U_n)$

7.5. Exponential(λ) distribution

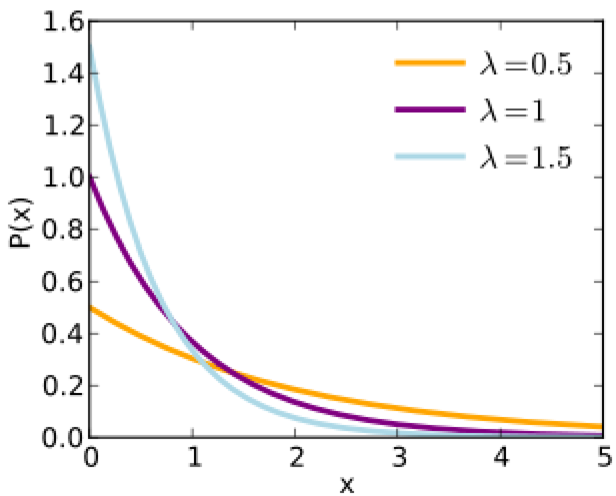
(Note that sometimes the shown parameter is $\frac{1}{\lambda}$, i.e. the mean of the distribution)

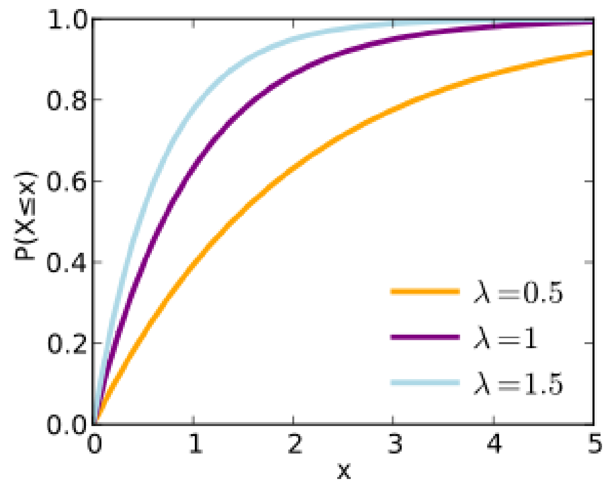
Definition: $X \sim \text{Exp}(\lambda)$, if:

$$f_X(x) = \lambda e^{-\lambda x} \qquad x \geq 0 \qquad\qquad F_X(x) = 1 - e^{-\lambda x}, \qquad x \geq 0. \tag{111}$$

Mean and variance are as follows:

$$E[X] = \frac{1}{\lambda} \qquad\qquad V[X] = \frac{1}{\lambda^2} \tag{112}$$





7.5. Exponential distribution cntd.

Laplace transform and moments of exponential distribution

The Laplace transform of a random variable with the distribution Exp(λ) is

$$f^*(s) = \int_{-\infty}^{\infty} e^{-st} \lambda e^{-\lambda t} dt = \frac{\lambda}{\lambda + s}$$

With the aid of this one can calculate the moments:

$$E[X] = -f^{*\prime}(0) = \left.\frac{\lambda}{(\lambda + s)^2}\right|_{s=0} = \frac{1}{\lambda}$$

$$E[X^2] = +f^{*\prime\prime}(0) = \left.\frac{2\lambda}{(\lambda+s)^3}\right|_{s=0} = \frac{2}{\lambda^2}$$

$$V[X] = E[X^2] - E[X]^2 = \frac{1}{\lambda^2}$$

$$E[X] = \frac{1}{\lambda} \qquad V[X] = \frac{1}{\lambda^2}$$

7.5. Exponential distribution cntd.

The memoryless property of exponential distribution

Assume that $X \sim \text{Exp}(\lambda)$ represents e.g. the duration of a call.

What is the probability that the call will last at least time *x* more given that it has already lasted the time *t*:

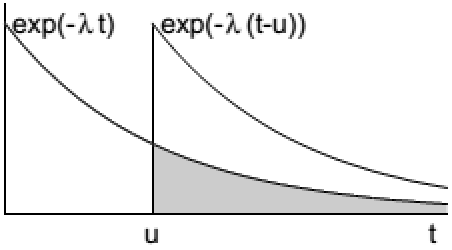
$$\text{P}\{X > t + x \,|\, X > t\} = \frac{\text{P}\{X > t + x, X > t\}}{\text{P}\{X > t\}}$$

$$= \frac{\text{P}\{X > t + x\}}{\text{P}\{X > t\}}$$

$$= \frac{e^{-\lambda(t+x)}}{e^{-\lambda t}} = e^{-\lambda x} = \text{P}\{X > x\}$$

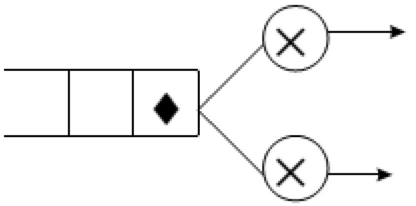
$$\text{P}\{X > t + x, X > t\} = \text{P}\{X > x\}$$

- The distribution of the remaining duration of the call does not at all depend on the time the call has already lasted has the same $\text{Exp}(\lambda)$ distribution as the total duration of the call.



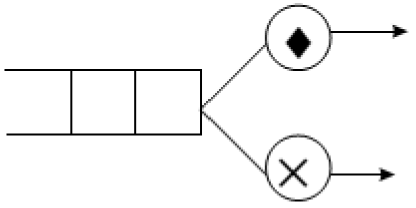
Perf Eval of Comp Systems

Example of the use of the memoryless property



A queueing system has two servers. The service times are assumed to be exponentially distributed (with the same parameter). Upon arrival of a customer (◊) both servers are occupied (×) but there are no other waiting customers.

The question: what is the probability that the customer (◊) will be the last to depart from the system?



The next event in the system is that either of the customers (×) being served departs and the customer enters (◊) the freed server.

By the memoryless property, from that point on the (remaining) service times of both customers (◊) and (×) are identically (exponentially) distributed.
The situation is completely symmetric and consequently the probability that the customer (◊) is the last one to depart is 1/2.

The ending probability of an exponentially distributed interval

Assume that a call with Exp(λ) distributed duration has lasted the time t.

What is the probability that it will end in an infinitesimal interval of length h?

$P\{X \le t + h \mid X > t\} = P\{X \le h\}$ (memoryless)

$= 1 - e^{-\lambda h}$

$= 1 - (1 - \lambda h + \frac{1}{2}(\lambda h)^2 - \cdots)$

$= \lambda h + o(h)$

The ending probability per time unit = λ (constant!)

The minimum and maximum of exponentially distributed random variables

Let $X_1 \sim \cdots \sim X_n \sim \mathrm{Exp}(\lambda)$ (*i.i.d.*)

The tail distribution of the minimum is

$P\{\min(X_1, \ldots, X_n) > x\} = P\{X_1 > x\} \cdots P\{X_n > x\}$ (independence)

$$\left(e^{-\lambda x}\right)^n = e^{-n\lambda x}$$

The minimum obeys the distribution $\mathrm{Exp}(n\lambda)$.

The cdf of the maximum is

$P\{\max(X_1, \ldots, X_n) \le x\} = \left(1 - e^{-\lambda x}\right)^n$

Perf Eval of Comp Systems

Calculating E[max(X_1, . . . , X_n)] Method 1:

$$E[N]=\sum_{k=0}^{\infty}P(N>k) \qquad E[X]=\int_0^{\infty}(1-F(x))dx=\int_0^{\infty}R_X(t)dt$$

E[max(X_1, . . . , X_n)]= $\displaystyle\int_0^{\infty}R_X(t)dt=\int_0^{\infty}\left(1-\left(1-e^{-\lambda t}\right)^n\right)dt=$

$$\frac{1}{\lambda}\int_0^1\frac{1-u^n}{1-u}du=\quad\frac{1}{\lambda}\int_0^1\left(\sum_{i=1}^{n}u^{i-1}\right)du=$$

$$\left(\frac{1}{\lambda}\sum_{i=1}^{n}\int_0^1 u^{i-1}du\right)=\frac{1}{\lambda}\sum_{i=1}^{n}\frac{u^i}{i}\Bigg|_0^1$$

$$=\frac{1}{\lambda}\sum_{i=1}^{n}\frac{1}{i}=\frac{1}{n\lambda}+\frac{1}{(n-1)\lambda}+\cdots+\frac{1}{\lambda}$$

$u=1-e^{-\lambda t}$ $\quad \frac{(1-u)^n}{1-u}=\sum_{i=1}^{n}u^{i-1}$

$1-e^{-\lambda t}\Big|_{t=0}^{t=\infty}=1$

$$1-e^{-\lambda t}\Big|_{t=0}^{\infty}\longrightarrow u\Big|_0^1$$

$1-e^{-\lambda t}\Big|_{t=0}=0$

$du=\lambda e^{-\lambda t}dt=\lambda(1-u)dt$

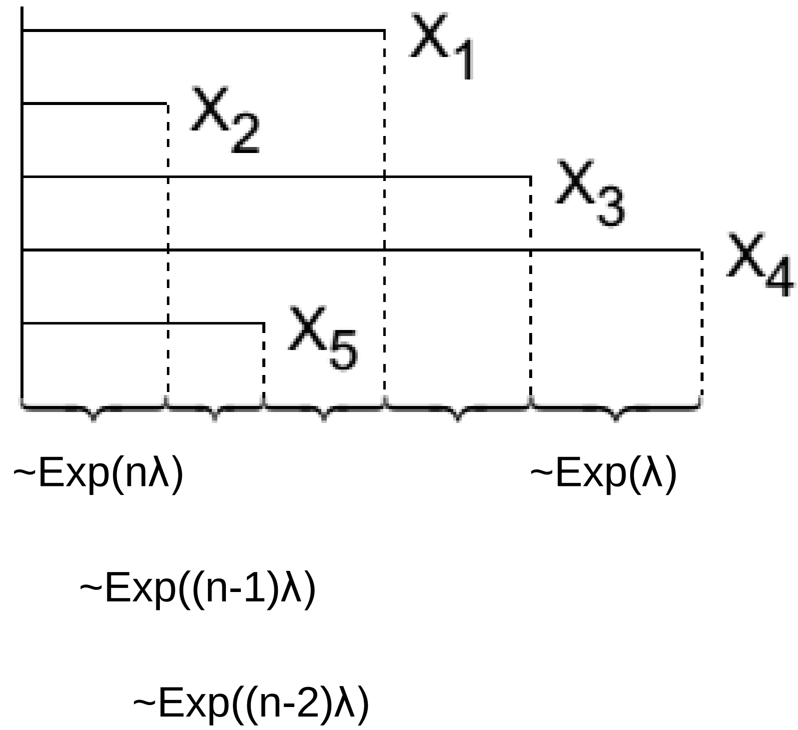
$$dt\Big|_0^{\infty}\longrightarrow\frac{du}{\lambda(1-u)}\Bigg|_0^1$$

Calculating E[max(X_1, . . . , X_n)] Method 2:

The ending intensity of the minimum = $n\lambda$

n parallel processes each of which ends with intensity λ independent of the others

The expectation can be deduced by inspecting the figure



$$E[\max(X_1, \ldots, X_n)]$$

$$= \frac{1}{n\lambda} + \frac{1}{(n-1)\lambda} + \cdots + \frac{1}{\lambda}$$

Performability : background

Q1: find distribution of Y=aX+b

A: $f_Y(y) = \begin{cases} \frac{1}{|a|} f_X(\frac{y-b)}{a}) & \text{If } x \in \text{Y=aX+b} \\ 0 & \text{otherwise} \end{cases}$

Q2: if $X \sim$ Exp(λ), what is distribution of Y=rX

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

Using Q1

$$f_Y(y) = \frac{1}{r}\lambda e^{-\lambda\frac{y}{r}}$$

06-Q2

Thus Y is an exponential with parameter λ/r

Performability : Example

Consider a system with *n* processors. In the beginning we let all *n* processors be active, performing different computations,

Thus the total computing capacity is *n* (where a unit of computing capacity corresponds to that of one active processor).

Performability : Example

Let $X_1, X_2, \ldots, X_n$ be the times to failure of the n processors.

After a period of time $Y_1 = \min\{X_1, X_2, \ldots, X_n\}$ only n - 1 processors will be active and the computing capacity of the system will have dropped to n - 1.

The cumulative computing capacity that the system supplies until all processors have failed is then given by the random variable:

$C_n = nY_1 + (n-1)(Y_2 - Y_1) + \cdots + (n-J)(Y_{j+1} - Y_j) + \cdots + (Y_n - Y_{n-1})$.

Performability : Example

we note that C_n is the area under the curve.

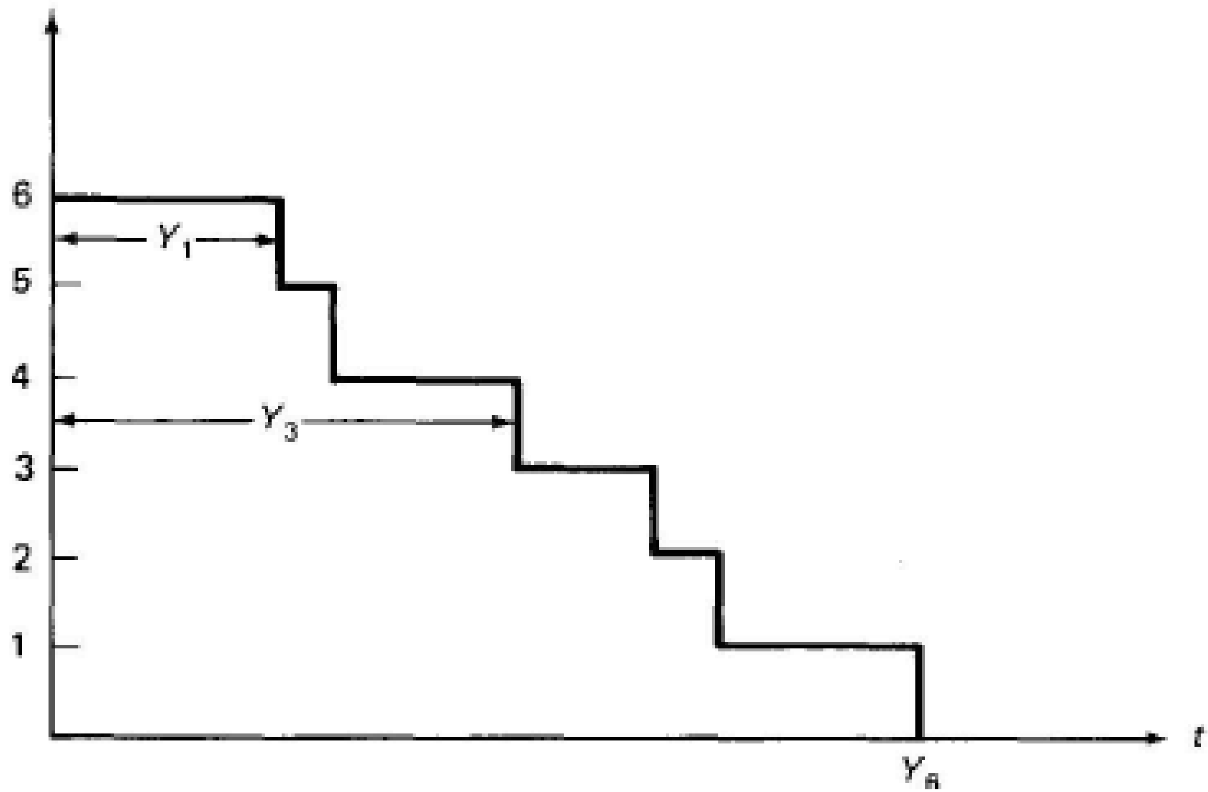


Fig: Computing capacity as a function of time

Performability : Example

C_n "computation before failure" Beaudry [BEAU 1978]

"performability" Meyer [MEYE 1980]

Distribution of C_n

first we obtain the distribution of $Y_{j+1} - Y_j$ ·

assume processor lifetimes are mutually independent EXP (λ) RVs. we claim that the distribution of

$Y_{j+1} - Y_j$ is $\mathrm{Exp}((\mathrm{n}-\mathrm{j})\lambda)$

Performability : Example

Define $Y_0 = 0$.

we know that $Y_1 = \min\{X_1, X_2, \ldots, X_n\} \sim \text{Exp}(n\lambda)$,

Thus our claim holds for $j = 0$.

After j processors have failed,(n-j) processors remain.

the residual lifetimes of the remaining $(n - j)$ processors, denoted by $W_1, W_2, \ldots, W_{n-j}$ are each exponentially distributed with parameter λ due to the memoryless property of the exponential distribution.

Performability : Example (cntd.)

Note that Y_{j+1} - Y_j is simply the time between the (j + 1)th and the jth failure; that is,

$Y_{j+1} - Y_j = \min \{W_1, W_2, . . . , W_{n-j}\}$

It follows that Y_{j+1} - Y_j is $\sim \mathrm{Exp}((\mathrm{n}-\mathrm{j})\lambda)$.

Performability : Example (cntd.)

Hence, using 06-Q2 (i.e. if $X \sim \mathrm{Exp}(\lambda)$, dist. of Y=rX is $f_Y(y) = \frac{1}{r}\lambda e^{-\lambda\frac{y}{r}}$)we get:

$X \sim \mathrm{Exp}((\text{n-j})\lambda)$, dist. of Y=(n-j)X is (r=n-j)

$$f_Y(y) = \frac{1}{r}(n-j)\lambda e^{-(n-j)\lambda\frac{1}{r}} \Big|_{r=n-j}$$

Thus $(n\text{-}j)$ $(Y_{j+1} - Y_j)$ is $\sim \mathrm{Exp}(\lambda)$.

Therefore, C_n is the sum of *n* independent identically distributed exponential RVs,

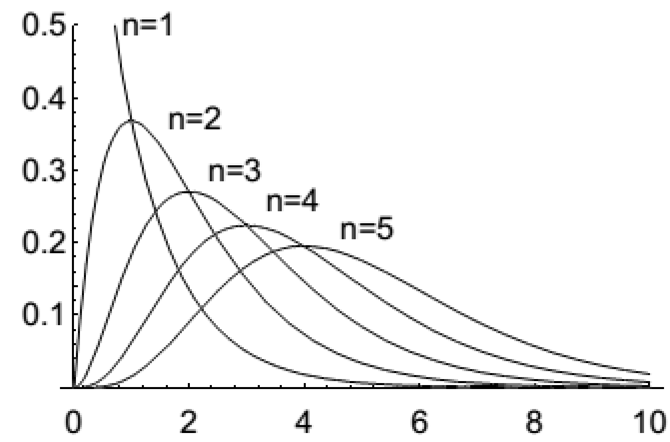
That is C_n is n-stage Erlang distributed with parameter λ.

Erlang distribution *X* ∼ Erlang(*n, λ*) Also denoted Erlang-*n*(*λ*).
X is the sum of *n* independent random variables with the distribution Exp(*λ*)

$X = X_1 + \cdots + X_n$ $\qquad X_i \sim \text{Exp}(\lambda)$ (*i.i.d.*)
The Laplace transform is
$$f^*(s) = \left(\frac{\lambda}{\lambda+s}\right)^n$$

By inverse transform (or by recursively convoluting the density function) one obtains the pdf of the sum *X*
$$f(x) = \frac{(\lambda x)^{n-1}}{(n-1)!}\lambda e^{-\lambda x} \qquad x \geq 0$$
The expectation and variance of Erlang
are *n* times those of the Exp(*λ*) distribution:

$$E[X] = \frac{n}{\lambda} \qquad V[X] = \frac{n}{\lambda^2}$$



Note:
Erlang(*1, λ*) **(**Erlang-*1*(*λ*)= $E_{1,\lambda}$) is exp(*λ*).

Erlang distribution (continued): gamma distribution

The formula for the pdf of the Erlang distribution can be generalized, from the integer parameter *n*, to arbitrary real numbers by replacing the factorial $(n-1)!$ by the gamma function $\Gamma(n)$:

$$f(x) = \frac{(\lambda x)^{n-1}}{\Gamma(n)} \lambda e^{-\lambda x}$$

Gamma(n, λ) distribution

Gamma function $\Gamma(p)$ is defined by

$$\Gamma(p) = \int_0^\infty e^{-u}\, u^{p-1}\, du$$

By partial integration it is easy to see that when *p* is an integer (i.e. p=n) then, indeed, $\Gamma(n) = (n-1)!$

Erlang distribution (continued)

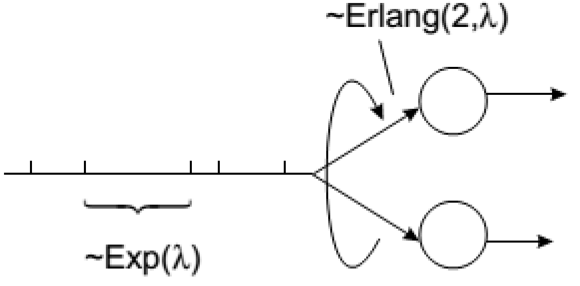
Example. The system consists of two servers. Customers

arrive with Exp(λ) distributed interarrival times.

Customers are alternately sent to servers 1 and 2.

The interarrival time distribution of customers arriving at

a given server is Erlang(2, λ).

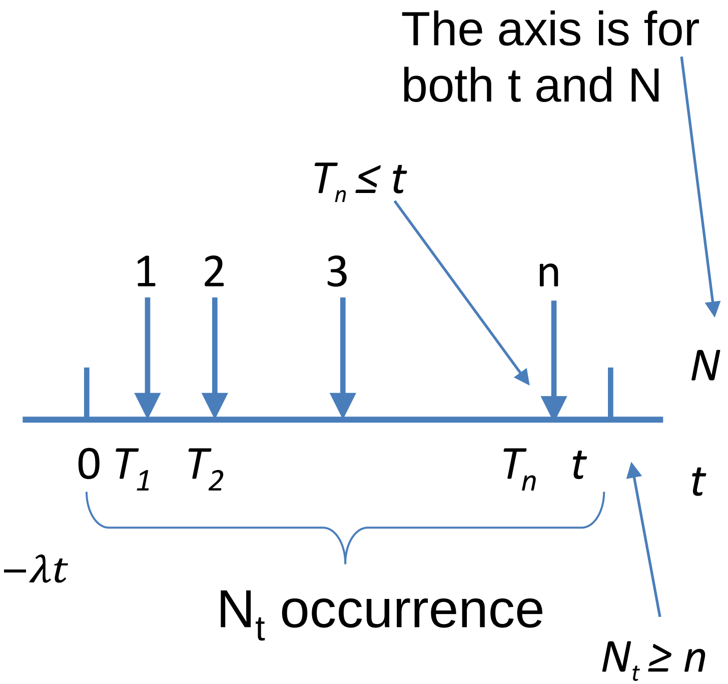


Erlang distribution (continued)

Proposition. Let N_t, the number of events in an interval of length t, obey the Poisson distribution:

$N_t \sim \text{Poisson}(\lambda t)$

Then the time T_n from an arbitrary event to the n^{th} event thereafter obeys the distribution Erlang(n, λ).



<u>Proof.</u>

$F_{T_n}(t) = \text{P}\{T_n \leq t\} = \text{P}\{N_t \geq n\}$

$\sum_{i=n}^{\infty} \text{P}\{N_t = i\} = \sum_{i=n}^{\infty} \frac{(\lambda t)^i}{i!} e^{-\lambda t}$

$$f_{T_n} = \frac{d}{dt} F_{T_n}(t) = \sum_{i=n}^{\infty} \frac{i\lambda(\lambda t)^{i-1}}{i!} e^{-\lambda t} - \sum_{i=n}^{\infty} \frac{(\lambda t)^i}{i!} \lambda e^{-\lambda t}$$

$$= \sum_{i=n}^{\infty} \frac{(\lambda t)^{i-1}}{(i-1)!} \lambda e^{-\lambda t} - \sum_{i=n}^{\infty} \frac{(\lambda t)^i}{i!} \lambda e^{-\lambda t}$$

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

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

Erlang distribution (continued) method II

Proof.

let W denote the **waiting time until the n^{th} event** occurs and find the distribution of W

$F_{T_n}(t)$ = P{$T_n \le t$} = 1-P{$T_n > t$}

the waiting time T_n is greater than some value t only if there are fewer than n events in the interval [0,t]. i.e.:

P{$T_n > t$} = P(fewer than n events in [0,w])

A more specific way of writing that is:

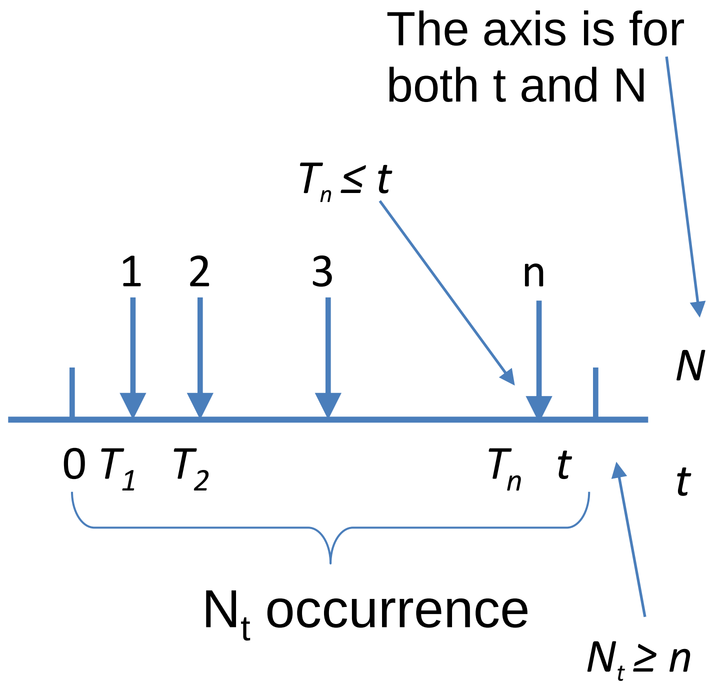
= P(0 events **or** 1 event **or** … **or** (n−1) events in [0,t])

mutually exclusive "ors" mean that we need to add up the probabilities of having 0 events occurring in the interval [0,t], 1 event occurring in the interval [0,t], ..., up to (n−1) events in [0,t].

$$\sum_{i=0}^{n-1} P\{N_t = i\} = \sum_{i=0}^{n-1} \frac{(\lambda t)^i}{i!} e^{-\lambda t}$$

Thus:

$$F_{T_n}(t) = 1-\sum_{i=0}^{n-1} \frac{(\lambda t)^i}{i!} e^{-\lambda t} = \sum_{i=n}^{\infty} \frac{(\lambda t)^i}{i!} e^{-\lambda t}$$



Normal distribution $X \sim$ N(μ, σ^2)

The pdf of a normally distributed random variable *X* with parameters μ and σ^2 is

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Parameters μ and σ^2 are expectation and variance are of the distribution
E[*X*] = μ , V[*X*] = σ^2

Normal Standard $Z \sim$ N(0, *1*)

$$f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}$$

0 and 1 are expectation and variance are of the distribution E[Z] = 0 , V[Z] =1

Proposition: If $X \sim N(\mu, \sigma^2)$, then $Y = \alpha X + \beta \sim N(\alpha\mu + \beta, \alpha^2\sigma^2)$.

Proof:

$$F_Y(y) = P\{Y \le y\} = P\{X \le \frac{y-\beta}{\alpha}\} = F_X\left(\frac{y-\beta}{\alpha}\right)$$

$$= \int_{-\infty}^{\frac{y-\beta}{\alpha}} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$

$z = \alpha x + \beta$

Specifically if we set ($\alpha = 1/\sigma$, $\beta = -\mu/\sigma$)

Then we find $z = \frac{x-\mu}{\sigma} \sim N(0, 1)$

$$F_Y(y) = \int_{-\infty}^{\frac{y-\beta}{\alpha}} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$

$$= \int_{-\infty}^{y} \frac{1}{\sqrt{2\pi}(\alpha\sigma)} e^{-\frac{1}{2}(z-(\alpha\mu+\beta))^2/(\alpha\sigma)^2} dz$$

$$= \int_{-\infty}^{y} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz$$

$$x = \frac{z-\beta}{\alpha} = \frac{z+\mu/\sigma}{1/\sigma} = \sigma Z + \mu$$

$$x\Big|_{-\infty}^{\frac{y-\beta}{\alpha}} \to z\Big|_{-\infty}^{\alpha\frac{y-\beta}{\alpha}+\beta} \to z\Big|_{-\infty}^{y}$$

$$dx \to \frac{dz}{\alpha}$$

$$\frac{(x-\mu)^2}{2\sigma^2} = \frac{(\sigma Z+\mu-\mu)^2}{2\sigma^2} = \frac{z^2}{2}$$

$$\alpha\sigma = \frac{1}{\sigma}\sigma = 1$$

Note: Denote the PDF of a N(0,1) random variable by $\Phi(x)$. Then

$$F_X(x) = P\{X \le x\} = P\{\sigma Z + \mu \le x\} = P\{Z \le \frac{x-\mu}{\sigma}\} = \Phi\left(\frac{x-\mu}{\sigma}\right)$$

Multivariate Gaussian (normal) distribution

Let $X_1, \ldots, X_n$ be a set of Gaussian (i.e. normally distributed) random variables with expectations $\mu_1, \ldots, \mu_n$ and covariance matrix

$$\mathbf{\Gamma} = \begin{bmatrix} \sigma_{11}^2 \cdots \sigma_{1n}^2 \\ \ldots \ldots \ldots \\ \sigma_{n1}^2 \cdots \sigma_{nn}^2 \end{bmatrix} \qquad \sigma_{ij}^2 = \mathrm{Cov}[X_i, X_j] \qquad \sigma_{ii}^2 = \mathrm{V}[X_i]$$

Denote $\mathbf{X} = (X_1, \ldots, X_n)^{\mathrm{T}}$.
The probability density function of the random vector **X** is

$$f(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^n |\mathbf{\Gamma}|}} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^{\mathrm{T}}\mathbf{\Gamma}^{-1}(\mathbf{x}-\boldsymbol{\mu})}$$

where $|\mathbf{\Gamma}|$ is the determinant of the covariance matrix.

By a change of variables one sees easily that the pdf of the random vector $\mathbf{Z} = \mathbf{\Gamma}^{-1/2}(\mathbf{X} - \boldsymbol{\mu})$ is

$$(2\pi)^{-n/2}\exp\left(-\frac{1}{2}\mathbf{z}^{\mathrm{T}}\mathbf{z}\right) = \sqrt{2\pi}\, e^{-z_1^2/2} \cdots \sqrt{2\pi}\, e^{-z_n^2/2}$$

Thus the components of the vector **Z** are independent N(0,1) distributed random variables.
Conversely, $\mathbf{X} = \boldsymbol{\mu} + \mathbf{\Gamma}^{1/2}\mathbf{Z}$ by means of which one can generate values for **X** in simulations.

Appendix

Power Laws

- Two quantities y and x are related by a power law if

$$y \approx x^{-\alpha}$$

- Compare it to $y \approx e^{-x}$ that decays fast.
- A (continuous) random variable X follows a power-law distribution if it has density function

$$f(x) = Cx^{-\alpha}$$

- Integrating above we find : Cumulative function

$$P[X \geq x] = \frac{C}{\alpha - 1} x^{-(\alpha-1)}$$

Normalization constant

- Assuming a minimum value x_{min}

$$C = (\alpha - 1) x_{min}^{\alpha - 1}$$

- The density function becomes

$$f(x) = \frac{(\alpha - 1)}{x_{min}} \left(\frac{x}{x_{min}} \right)^{-\alpha}$$

- Ref:http://tuvalu.santafe.edu/~aaronc/courses/7000/csci7000-001_2011_L2.pdf

Pareto distribution

- Pareto distribution is pretty much the same but we have

$$P[X \geq x] = C' x^{-\beta}$$

- and we usually we require

$$x \geq x_{min}$$

Zipf's Law

- A random variable X follows Zipf's law if the r-th largest value x_r satisfies

$$x_r \approx r^{-\gamma}$$

- Same as requiring a Pareto distribution

$$P[X \geq x] \approx x^{-1/\gamma}$$

Zipf &Pareto:
what they have to do with power-laws

- Zipf
  - George Kingsley Zipf, a Harvard linguistics professor, sought to determine the 'size' of the 3rd or 8th or 100th most common word.
  - Size here denotes the frequency of use of the word in English text, and not the length of the word itself.
  - Zipf's law states that the size of the r'th largest occurrence of the event is inversely proportional to its rank:

    $y \sim r^{-\beta}$, with β close to unity.

So how do we go from Zipf to Pareto?

- The phrase "The ***r*** th largest city has ***n*** inhabitants" is equivalent to saying "***r*** cities have ***n*** or more inhabitants".
- This is exactly the definition of the Pareto distribution, except the x and y axes are flipped. Whereas for Zipf, ***r*** is on the x-axis and ***n*** is on the y-axis, for Pareto, ***r*** is on the y-axis and ***n*** is on the x-axis.
- Simply inverting the axes, we get that if the rank exponent is β, i.e.

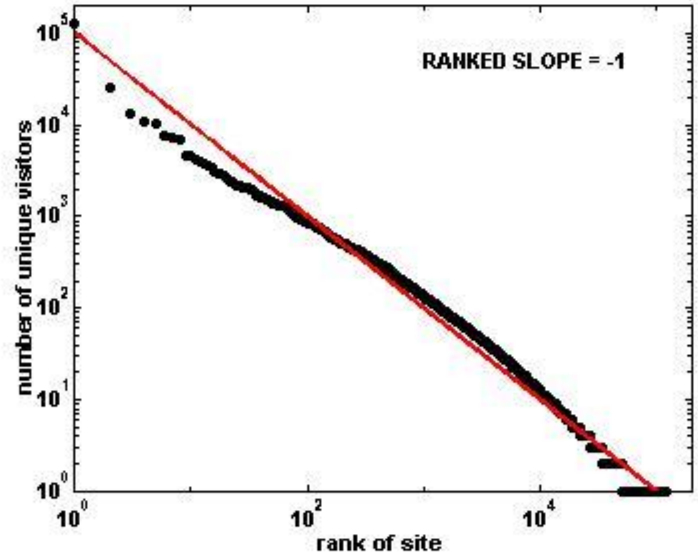
  $n \sim r^{-\beta}$ for Zipf, (n = income, r = rank of person with income n)

  then the Pareto exponent is $1/\beta$ so that

  $r \sim n^{-1/\beta}$ (n = income, r = number of people whose income is n or higher)

Zipf’s law & AOL site visits

- Deviation from Zipf’s law
  - slightly too few websites with large numbers of visitors:



Exponents and averages

- In general, power law distributions do not have an average value if $\alpha < 2$ (but the sample will!)
- This is because the average is given by (for integer values of k)

for a finite sample this will only go to the largest observed value

$$\sum_{k=k_{\min}}^{\infty} kp(k) = \sum_{k=k_{\min}}^{\infty} kk^{-\alpha} = \sum_{k=k_{\min}}^{\infty} \frac{1}{k^{\alpha-1}}$$

- The harmonic series diverges… $1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\ldots$

- Same holds for continuous values of k

80/20 rule

- The fraction W of the wealth in the hands of the richest P of the the population is given by

$$W = P^{(\alpha-2)/(\alpha-1)}$$

- Example: US wealth: $\alpha = 2.1$
  - richest 20% of the population holds 86% of the wealth ($0.2^{\left(\frac{0.1}{1.1}\right)} = 0.2^{0.0909} = 0.86$)

Generative processes for power-laws

- Many different processes can lead to power laws
- There is no one unique mechanism that explains it all

- Next class: Yule’s process and preferential attachment

What does it mean to be scale free?

- A power law looks the same no mater what scale we look at it on (2 to 50 or 200 to 5000)
- Only true of a power-law distribution!
- p(bx) = g(b) p(x) – shape of the distribution is unchanged except for a multiplicative constant
- $p(bx) = (bx)^{-\alpha} = b^{-\alpha} x^{-\alpha}$