The Gamma Distribution
Properties, Proofs and Applications

Etienne Collin, Marie Ouellet & Rania Yahyaoui
Probability and Statistics - 201-BNM-LW

Section 20108

Professor Vincent Carrier

St. Lawrence Cégep Champlain

Due on May 25, 2022 at 23:59

St.Lawrence

CEGEP CHAMPLAIN

1958

Table of Contents

1   General Applications and Family

The gamma distribution is used in many fields, since it is directly related to the Erlang, normal and exponential distributions whose contributions extend to several disciplines. Here, we will present some of those applications.

1.1   Insurance Companies

First, the gamma distribution is of great use in the field of insurance services, given its direct relation to the exponential distribution. For example, an analyst could use this distribution to specify the amount of time a product lasts if one uses it at a constant average rate, thus modeling how reliable it is (and how much insurance should be charged for it). Then, the size of loan defaults and the cost of insurance claims are also often modeled according to a gamma distribution.

1.2   Natural Events Prediction

Then, the gamma distribution can also be used to model the amount of rainfall accumulated in a given reservoir. Indeed, this distribution fits positive data, represents rainfall distribution well and its two parameters -shape and scale- give it sufficient flexibility to fit various climates.[1]

1.3   Customer Satisfaction

Service time can also be modeled using the gamma distribution. For example, if one is waiting in line for a meal, the waiting time until one receives the long-awaited food can be modelled using an exponential distribution. Using the same principle, the Erlang distribution can allow one to determine the total length of a process, that is, the duration of a sequence of independent events. For instance, if a large number of people are waiting in line to be served,

the distribution of each of their individual waiting times (the sum of several independent exponentially distributed variables) will correspond to the time it takes for the employee to serve everyone in it. Therefore, the gamma distribution lies at the heart of what is called queuing theory: the mathematical study of the congestion of waiting lines. Since waiting lines are found in countless places such as banks, restaurants and hospitals, as well as on web servers or multistep manufacturing and distribution processes, the gamma distribution provides very useful applications in everyday life.

1.4   Call Centers

One of the most famous of those applications concerns phone queuing, on which A. K. Erlang famously worked. The Erlang distribution has indeed been developed in the goal to model the time in between incoming calls at a call center, along with the expected number of calls, thus allowing call centers to know what their staffing capacity should be depending on the time of day.

Beyond waiting lines, the Erlang distribution is often used by retailers to model the frequency of interpurchase times by consumers, which gives them an idea of how often a given consumer is expected to purchase a product from them and helps them control inventory and staffing.

1.5   Oncology

Finally, in the field of oncology, the age distribution of cancer incidence also follows a gamma distribution. Although the factors underlying cancer development are not yet fully understood, it has been hypothesized that cancers arise after several successive "driver events", that is, after some number of mutations occurs in a cell. Analyses of cancer statistics suggest that the incidence of the most prevalent cancer types with respect to the patients' age closely follows the gamma probability distribution or, more specifically, the Erlang distribution. This may be due to the fact that, more broadly, the Erlang distribution can be used

to model cell cycle time distribution [2] .

$$Y = A \cdot x^{k-1} \frac{e^{\frac{-x}{b}}}{b^k} \cdot \Gamma(k) \tag{1}$$

In this case, the shape parameter α predicts the number of carcinogenic driver events, whereas the shape parameter λ predicts the average time between those events for each cancer type. Using an additional amplitude parameter A, the maximal population susceptibility to a given type of cancer can even be predicted. [2] Given that experimental research on cancer development is crucial for the lives of many people, numerical references such as that provided by the gamma distribution are of paramount importance in our society. The gamma distribution can save lives, if it is used wisely!
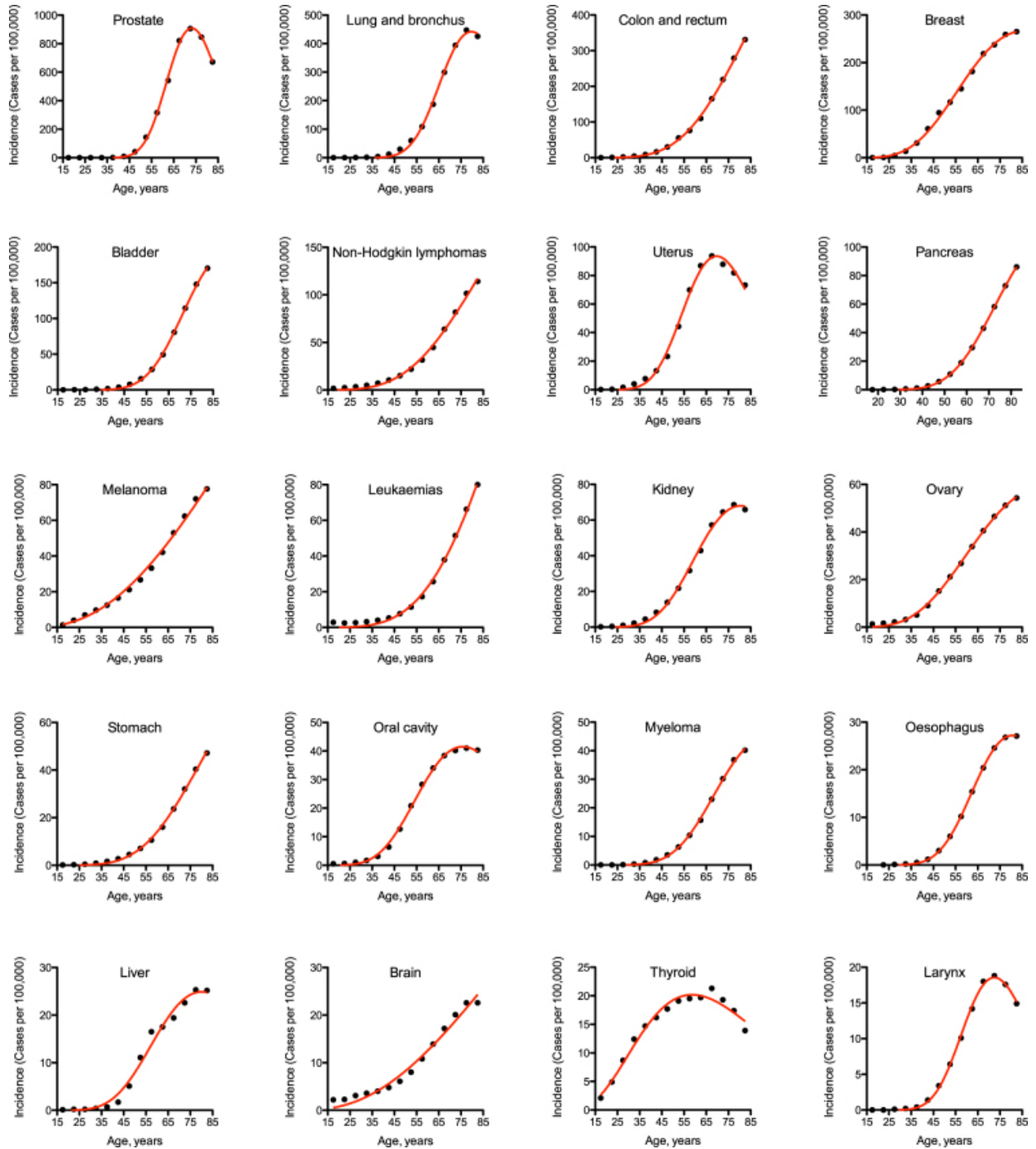
FIGURE 1: "The Erlang distribution approximates cancer incidence by age for 20 most prevalent cancer types. Dots indicate actual data for 5-year age intervals, curves indicate the PDF of the Erlang distribution fitted to the data [...]. The middle age of each age group is plotted. Cancer types are arranged in the order of decreasing incidence" [2] .

2   Special Cases of the Gamma Distribution

The gamma distribution has many common parameterizations. Hence, in order to make working with them easier, they were given a specific name such as the chi-square distribution, the exponential distribution and more. In this section, we will present some of these special cases and prove their relationship.

2.1   The Erlang Distribution

In real world applications, it is often useful to limit the number of occurrences to a positive integer (it may not be pertinent to determine, for example, the waiting time until 2.665 events take place). In this context, the special name "Erlang distribution" has been attributed to the specific gamma distribution in which the shape parameter α only takes positive integer values and where the rate parameter λ is a positive real number.

$$f_{X_\alpha}(x) = \begin{cases} \dfrac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x} & \alpha \in \mathbb{Z}^+; x \geq 0 \\ 0 & x < 0 \end{cases}$$

This distribution takes its name from A. K. Erlang, the man who initially popularized its use in order to examine the number of telephone calls which can be made at the same time to the operators of the switching stations, in the field of phone traffic engineering. Indeed, the Erlang distribution's original purpose was to measure the time between incoming calls, a statistic which can be used along with the expected duration of incoming calls to analyze phone traffic loads. His work has since been expanded to consider queuing systems in general, as well as the load on web servers, interpurchase times and cancer incidence (see section 1). [3]

2.2   The Exponential Distribution

The gamma distribution can model the elapsed time between random and independent events. However, what should one do when they want to strictly model the time to wait before a new event?

If a unique event is modeled, it follows that the shape parameter is limited to $\alpha = 1$. In such a case, the gamma distribution is said to follow an exponential distribution: more specifically, if $X \sim \mathcal{G}\text{amma}(1, \lambda)$, then it can also be said that $X \sim \mathcal{E}\text{xp}(\lambda)$. X thus follows an exponential distribution with rate parameter λ, where the rate parameter represents how quickly events occur. More concretely:

$$X \sim \mathcal{G}\text{amma}(1, \lambda) \equiv \mathcal{E}\text{xp}(\lambda)$$

MAYBE HERE INCLUDE GRAPH OF GAMMA DISTRIBUTION WITH SHAPE PARAMETER 1 COMPARED TO OTHER SHAPE PARAMETERS ? Indeed, using a simple substitution of the variables, one gets that:

$$\begin{aligned}
f(x) &= \frac{\lambda^\alpha x^{\alpha-1}}{\Gamma(\alpha)} e^{-\lambda x} \\
&= \frac{\lambda^1 x^0}{\Gamma(1)} e^{-\lambda x} \\
&= \lambda e^{-\lambda x}
\end{aligned} \tag{2}$$

$$\therefore X \sim \mathcal{G}\text{amma}(1, \lambda) \equiv X \sim \mathcal{E}\text{xp}(\lambda)$$

Accordingly, it can be demonstrated that the sum of exponential random variables is a gamma random variable. That is:

If X_1, X_2, ..., X_n are i.i.d random variables following $\mathcal{E}\text{xp}(\lambda)$, then $Y = \sum_{i=1}^{k} [X_i] \sim \mathcal{G}\text{amma}(k, \lambda)$

Proof. **ADD HERE PROOF FOR THE DENSITY FUNCTION OF THE SUM**

OF EXPONENTIAL RANDOM VARIABLES (ETIENNE'S PART)

$$XYZ$$

QED

2.3   The Chi-Squared Distribution

Then, the gamma distribution with parameters $\alpha = {}^{n}\!/_{2}$ and $\lambda = {}^{1}\!/_{2}$ is called the chi-squared distribution with n degrees of freedom such that

$$X \sim \mathcal{G}\text{ammma}({}^{n}\!/_{2}, {}^{1}\!/_{2}) \equiv \chi^2_n$$

Proof. First, to show that $\mathcal{G}\text{ammma}({}^{n}\!/_{2}, {}^{1}\!/_{2}) \equiv \chi^2_n$, let's prove that $\mathcal{G}\text{ammma}({}^{1}\!/_{2}, {}^{1}\!/_{2}) \equiv \chi^2_1$, a chi-squared distribution with one (1) degree of freedom, using a change of variables with $Y = Z^2$ for $Y \sim \chi^2_1$.

$$
\begin{aligned}
G(y) &= P(Y \leq y) \\
&= P(Z^2 \leq y) \\
&= P(-\sqrt{y} \leq z \leq \sqrt{y}) \\
&= \Phi(\sqrt{y}) - \Phi(-\sqrt{y}) \\[6pt]
g(y) &= \frac{f(\sqrt{y})}{2\sqrt{y}} + \frac{f(-\sqrt{y})}{2\sqrt{y}} \\
&= \frac{e^{-y/2}}{2\sqrt{2\pi y}} + \frac{e^{-y/2}}{2\sqrt{2\pi y}} \\
&= \frac{\cancel{2}e^{-y/2}}{\cancel{2}\sqrt{2\pi y}} \\
&= \frac{1}{\sqrt{2}\Gamma(\frac{1}{2})} y^{-1/2} e^{-y^{1}/2} \\
&= \frac{(\frac{1}{2})^{1/2}}{\Gamma(\frac{1}{2})} y^{1/2-1} e^{-(1/2)y}
\end{aligned}
\tag{3}
$$

$$\therefore Y \sim \mathcal{G}\text{ammma}({}^{1}\!/_{2}, {}^{1}\!/_{2}) \equiv \chi^2_1$$

One may generalize this result by using that $Z_1, Z_2, \ldots, Z_n$ are i.i.d. random variables

following $\mathcal{N}(0,1)$ to find the distribution of

$$Y = \sum_{i=1}^{n} Z_i^2$$

Hence, using the moment generating function,

$$
\begin{aligned}
M_{\sum_{i=1}^{n} Z_i^2}(t) &= \prod_{i=1}^{n} M_{Z_i^2}(t) \\
&= M(t)^n \\
&= \left(\frac{1/2}{1/2 - t}\right)^{n/2}
\end{aligned}
\tag{4}
$$

This last line is equal to the moment generating function of a random variable following the distribution $\mathcal{G}\mathrm{amma}(n/2, 1/2)$. Hence,

$$\therefore Y \sim \mathcal{G}\mathrm{amma}(n/2, 1/2) \equiv \chi_n^2$$

QED

Using the result proved previously, the density function of the chi-squared distribution may be determined by substituting the parameters $\alpha = n/2$ and $\lambda = 1/2$ into the density function of the gamma distribution:

$$
\begin{aligned}
f(x) &= \frac{\lambda^{\alpha}}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x} \\
&= \frac{\left(\frac{1}{2}\right)^{n/2}}{\Gamma\left(\frac{n}{2}\right)} x^{\frac{n}{2}-1} e^{-\frac{1}{2}x} \\
&= \frac{1}{2^{n/2}\Gamma\left(\frac{n}{2}\right)} x^{\frac{n}{2}-1} e^{-x/2}
\end{aligned}
\tag{5}
$$

HERE INCLUDE GRAPH OF CHI SQUARE DISTRIBUTIONS

The chi square distribution is often used to describe the distribution of a sum of squared random variables, to test the fit of a distribution of data or to determine whether data series are independent or dependent.

2.4   The Normal Distribution

Text on the normal distribution

2.5   The Beta Distribution

Text on the Beta distribution

2.6   The Wishart Distribution

Text on the Wishart distribution

3   Introduction

The gamma distribution is part of the two-parameters family of continuous probability distributions. Indeed, it may be parameterized with two different parameterizations[4] :

Parameterization 1:

$$\text{Shape: } \alpha > 0 \qquad\qquad \text{Rate: } \lambda > 0 \tag{6}$$

Parameterization 2:

$$\text{Shape: } k > 0 \qquad\qquad \text{Scale: } \theta > 0 \tag{7}$$

In the second parameterization, the scale parameter θ corresponds to the inverse of the rate parameter λ, allowing flexibility in the modelization of the distribution. Indeed, the application of the scale parameter is pertinent when one wants to use the value of the mean time between events in the probability density function, as opposed to the total time until a given number of events occurs.

That being said, in this document, as the two parameterizations only exist for the sake of convenience and are identical in their results, only the parameterization 1 presented in Equation 6 will be considered and used for proofs.

3.1   Properties to prove

Support: $x \in (0, \infty)$

Probability density function:

$$f(x) = \frac{\lambda^{\alpha}}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x} \tag{8}$$

Proof. The probability density function of the gamma distribution will be proved using induction. First, take

$$W = \sum_{i=1}^{n} X_i$$

where W is a sum of n X i.i.d. random variables following the exponential distribution

$\mathcal{E}\mathrm{xp}(\lambda)$.

For $n = 1$:

$$W = X_1$$

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

$$M(t) = \frac{\lambda}{\lambda - t}$$

Then, with the induction that $n = k$ works such that:

$$W = \sum_{i=1}^{k} X_i$$

$$M_{\sum_{i=1}^{n} X_i}(t) = E(e^{t \sum_{i=1}^{k} X_i})$$

$$= E(\prod_{i=1}^{k} e^{tX_i})$$

$$= \prod_{i=1}^{k} E(e^{tX_i})$$

$$= \prod_{i=1}^{k} M_{X_i}(t)$$

$$= \left(\frac{\lambda}{\lambda - t}\right)^k$$

Here is the proof that $n = k + 1$ works such that:

$$W = \sum_{i=1}^{k+1} X_i$$

$$= \sum_{i=1}^{k} [X_i] + X_{k+1}$$

$$M_{\sum_{i=1}^{k} [X_i] + X_{k+1}}(t) = E\left(e^{t\left(\sum_{i=1}^{k}[X_i]+X_{k+1}\right)}\right)$$

$$= E\left(e^{t\left(\sum_{i=1}^{k} X_i\right)} e^{tX_{k+1}}\right)$$

$$= E\left(e^{t\left(\sum_{i=1}^{k} X_i\right)}\right) E\left(e^{tX_{k+1}}\right)$$

$$= M_{\sum_{i=1}^{n} X_i}(t) M(t)$$

$$= \left(\frac{\lambda}{\lambda - t}\right)^{k} \left(\frac{\lambda}{\lambda - t}\right)$$

$$= \left(\frac{\lambda}{\lambda - t}\right)^{k+1}$$

QED

Cumulative distribution function:

$$F(x) = \frac{1}{\Gamma(\alpha)} \gamma(\alpha, \, \lambda x) \tag{9}$$

Expected value, also known as the theoretical mean:

$$\mu = E(x) = \frac{\alpha}{\lambda} \tag{10}$$

There is no simple closed form equation for the median of a gamma distribution.

Mode:

$$\text{Mode} = \frac{(\alpha - 1)}{\lambda} \text{ for } \alpha \geq 1 \tag{11}$$

Variance:

$$\text{Var}(x) = \frac{\alpha}{\lambda^2} \tag{12}$$

Skewness:

$$\text{Skewness} = \frac{2}{\sqrt{\alpha}} \tag{13}$$

Excess kurtosis:

$$\text{Kurtosis} = \frac{6}{\alpha} \tag{14}$$

Entropy:

$$\text{Entropy} = \alpha + \ln\lambda + \ln\Gamma(\alpha) + (1-\alpha)\psi(\alpha) \tag{15}$$

Moment generating function:

$$M(t) = \left(1 - \frac{t}{\lambda}\right)^{-\alpha} \quad \text{for } t < \lambda \tag{16}$$

Characteristic function:

$$\text{CF} = \left(1 - \frac{it}{\lambda}\right)^{-\alpha} \tag{17}$$

Methods of moments:

$$\alpha = \frac{E(X)^2}{\text{Var}(X)}$$
$$\lambda = \frac{E(X)}{\text{Var}(X)} \tag{18}$$

List of Figures

List of Tables

References

[1]   Gregory J. Husak, Joel Michaelsen, and Chris Funk. "Use of the Gamma Distribution to Represent Monthly Rainfall in Africa for Drought Monitoring Applications". In: *International Journal of Climatology* 27.7 (2007), pp. 935–944. ISSN: 1097-0088. DOI: 10.1002/joc.1441. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/joc.1441 (visited on 03/22/2022).

[2]   Aleksey V. Belikov. "The Number of Key Carcinogenic Events Can Be Predicted from Cancer Incidence". In: *Sci Rep* 7 (2017-09-22), p. 12170. ISSN: 2045-2322. DOI: 10.1038/s41598-017-12448-7. pmid: 28939880. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5610194/ (visited on 03/22/2022).

[3]   ZACH. *What Is the Erlang Distribution?* Statology. 2020-12-20. URL: https://www.statology.org/erlang-distribution/ (visited on 03/22/2022).

[4]   Wikipedia. *Gamma Distribution.* In: *Wikipedia.* 2022-02-23. URL: https://en.wikipedia.org/w/index.php?title=Gamma_distribution&oldid=1073512326 (visited on 03/02/2022).