Standard Gamma model: introduction

We consider the following Gamma model, with the two parameters α and β unknown

$$\left\{ Ga(\alpha, \beta); \alpha > 0,\ \beta > 0 \right\}$$

The Probability Density Function (PDF) of the Gamma distribution is given by

$$f_{\alpha,\beta}(x) = \frac{\beta^{\alpha}}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}\ \mathbb{1}_{\mathbb{R}_+^*}(x)$$

The gamma function $\Gamma(\cdot)$, also called the Euler integral of the second kind is defined, for a nonnegative integer α, as $\Gamma(\alpha) = \int\limits_0^\infty x^{\alpha-1} e^{-x} dx$.

Method of Moments (MoM) estimation

Method of Moments (MoM): If $E[\varphi(X)] = h(\theta)$, then we have that $\hat{\theta}_{MOM} = h^{-1}\left(\frac{1}{n}\sum_{i=1}^{n}\varphi(x_i)\right)$ is a method of moments estimator for θ.

To obtain a **Method of Moments estimator** for the parameter vector $\boldsymbol{\theta} = (\alpha\ ,\ \beta)$ from a sample $x_1, ..., x_n$ of n $i.i.d.$ realizations of a Gamma distributed r.v X, we need the moments of X, given by

$$E[X] = \frac{\alpha}{\beta} \qquad\qquad E[X^2] = \frac{\alpha(\alpha+1)}{\beta^2}$$

Derivation of the first moment

$$
\begin{aligned}
E[X] &= \int_{-\infty}^{+\infty} x \ f_{\alpha,\beta}(x) \ dx = \int_0^{\infty} x \ \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} \ dx \\
&= \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^{\infty} x \ x^{\alpha-1} e^{-\beta x} \ dx \quad (u = \beta x \Leftrightarrow x = u/\beta, \quad du = \beta dx \Leftrightarrow dx = du/\beta) \\
&= \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^{\infty} \left(\frac{u}{\beta}\right)^\alpha e^{-u} \ \frac{du}{\beta} \\
&= \frac{\beta^\alpha}{\Gamma(\alpha)} \frac{1}{\beta^\alpha} \frac{1}{\beta} \int_0^{\infty} u^\alpha e^{-u} \ du \qquad \int_0^{\infty} x^\alpha e^{-x} \ dx = \Gamma(\alpha+1) = \alpha\Gamma(\alpha) \\
&= \frac{\beta^\alpha}{\Gamma(\alpha)} \frac{1}{\beta^{\alpha+1}} \alpha\Gamma(\alpha) \\
&= \frac{\alpha}{\beta}
\end{aligned}
$$

Derivation of the second moment

$$E[X^2] = \int_{-\infty}^{+\infty} x^2 \ f_{\alpha,\beta}(x) \ dx = \int_0^{\infty} x^2 \ \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} \ dx$$

$$= \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^{\infty} x^2 \ x^{\alpha-1} e^{-\beta x} \ dx \qquad (u = \beta x \Leftrightarrow x = u/\beta, \quad du = \beta dx \Leftrightarrow dx = du$$

$$= \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^{\infty} \left(\frac{u}{\beta}\right)^{\alpha+1} e^{-u} \ \frac{du}{\beta}$$

$$= \frac{\beta^\alpha}{\Gamma(\alpha)} \frac{1}{\beta^{\alpha+1}} \frac{1}{\beta} \int_0^{\infty} u^\alpha e^{-u} \ du \quad \int_0^{\infty} x^{\alpha+1} e^{-x} \ dx = \Gamma(\alpha+2) = (\alpha+1)\alpha\Gamma(\alpha)$$

$$= \frac{\beta^\alpha}{\Gamma(\alpha)} \frac{1}{\beta^{\alpha+2}} (\alpha+1)\alpha\Gamma(\alpha)$$

$$= \frac{(\alpha+1)\alpha}{\beta^2}$$

Gamma MoM estimators

So we set $\overline{x} = u$ and $\overline{x^2} = v$ and solve for α and β in the following equation:

$$h\binom{\alpha}{\beta} = \begin{pmatrix} \alpha/\beta \\ \alpha(\alpha+1)/\beta^2 \end{pmatrix} = \begin{pmatrix} u \\ v \end{pmatrix}$$

$$\begin{cases} u = \alpha/\beta \\ v = \alpha(\alpha+1)/\beta^2 \end{cases} \quad \Leftrightarrow \quad v = u^2 + \tfrac{u}{\beta} \quad \Leftrightarrow \quad v - u^2 = \tfrac{u}{\beta}$$

Then it is easy to show that $\alpha = u\beta \quad \Leftrightarrow \quad \alpha = \frac{u^2}{v-u^2}$
So we get as our Method of Moments estimators:

$$\hat{\boldsymbol{\theta}} = h^{-1}\left(\frac{1}{n}\sum_{i=1}^{n}\varphi(x_i)\right) = h^{-1}\begin{pmatrix} \overline{x} \\ \overline{x^2} \end{pmatrix} = \begin{pmatrix} \frac{\overline{x}^2}{\overline{x^2}-\overline{x}^2} \\ \frac{\overline{x}}{\overline{x^2}-\overline{x}^2} \end{pmatrix} = \begin{pmatrix} \hat{\alpha}_{MOM} \\ \hat{\beta}_{MOM} \end{pmatrix}$$

Example 1 - Lifetime of a product

The Gamma distribution is often used in reliability analysis to model time until failure kind of events, in particular Poisson distributed events. Suppose that we test the lifetime of a specific product on a sample of size 100 and obtain the following data, in weeks (20 first entries of the dataset)

$$\begin{aligned} x_i = {} & 17.49, 14.93, 16.01, 17.85, 7.37, 5.96, 12.27, 16.44, 28.34, 30.28, \\ & 10.66, 25.77, 17.21, 34.99, 28.53, 12.98, 9.32, 22.67, 12.55, 16.81 \end{aligned}$$

(i) From which Gamma distribution have the data been generated if we use the Method of Moment estimators for the parameters?

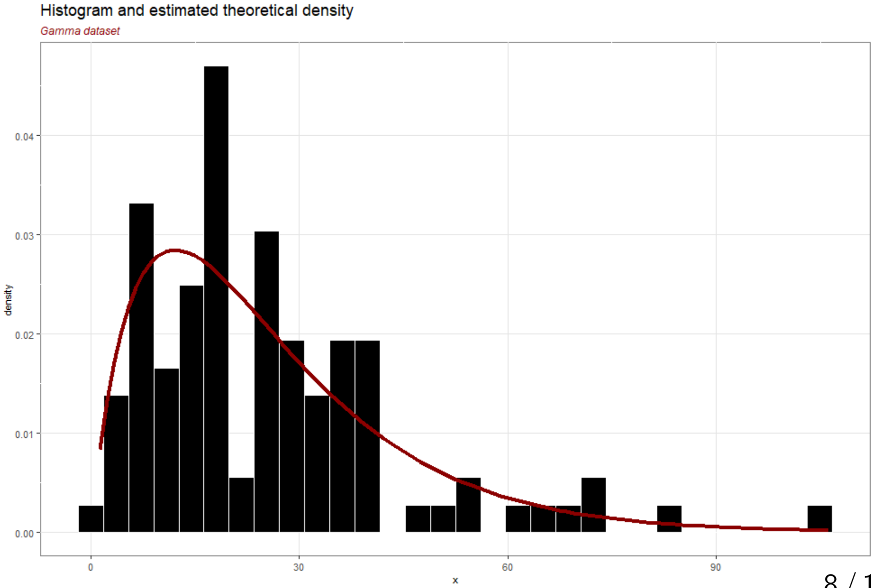
(ii) What are the cumulative hazard rates after 12, 24 and 36 weeks (or the total number of times we expect a failure for one product)?

Example 1 - Estimation

The full dataset is available here:
https://github.com/JRigh/Gamma-distribution-estimation/blob/main/dataset.csv

The data have been generated from a $Ga(\alpha = 2, \beta = 0.08)$ and computing the MoM estimators in R and Python, we can obtain a $Ga(\alpha = 1.900579, \beta = 0.07399452)$ with R and a $Ga(\alpha = 1.900579, \beta = 0.073995)$ with Python, so equivalent results up to rounding.

Histogram and theoretical density



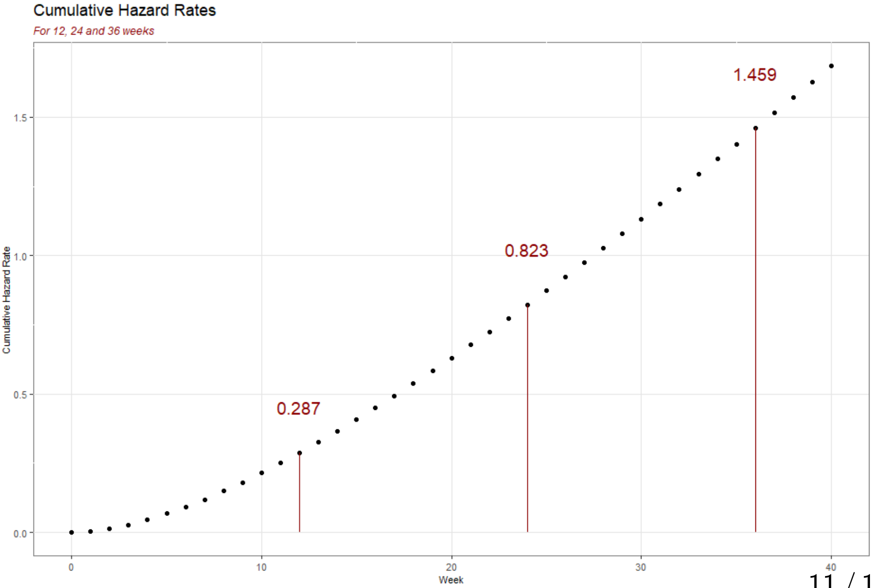
R code

```r
# function that implements the estimators above
MoMgamma <- function(x) {

  n <- length(x)
  sample_moment_1 <- sum(x) / n
  sample_moment_2 <- sum(x^2) / n

  alpha_mom <- sample_moment_1^2 / (sample_moment_2 - sample_moment_1^2)
  beta_mom <- sample_moment_1 / (sample_moment_2 - sample_moment_1^2)

  output <- NULL
  output$alpha_mom <- alpha_mom
  output$beta_mom <- beta_mom

  return(output)
}

# generate artificial data
set.seed(2023)
x <- round(rgamma(n = 100, shape = 2, rate = 0.08), 2)

# perform MoM estimation
MoMgamma(x = x)
# $alpha_mom
# [1] 1.900579
# $beta_mom
# [1] 0.07399452
```

Python code

```
import numpy as np
import pandas as pd
from scipy.stats import gamma

# function that implements the estimators above
def MoMgamma(x):
    n = len(x)
    sample_moment_1 = np.sum(x) / n
    sample_moment_2 = np.sum(x**2) / n

    alpha_mom = sample_moment_1**2 / (sample_moment_2 - sample_moment_1**2)
    beta_mom = sample_moment_1 / (sample_moment_2 - sample_moment_1**2)

    output = {
        'alpha_mom': alpha_mom,
        'beta_mom': beta_mom
    }

    return output

# read the dataset
x = pd.read_csv("C:/Users/julia/OneDrive/Desktop/github/37. momgamma/dataset.csv")

# perform MoM estimation
result = MoMgamma(x)
print(result)
{'alpha_mom': x    1.900579
dtype: float64, 'beta_mom': x    0.073995
dtype: float64}
```

Cumulative Hazard Rates



References

R.V.Hogg and E.A.Tanis: Probability and Statistical Inference, Sixth Edition, Prentice Hall, Upper Saddle River, N.J., 2001.

The R Project for Statistical Computing:
https://www.r-project.org/

Python:
https://www.python.org/

course notes