The Gamma Distributions

1. Characterisation

   - Probability Density Function (*pdf*):

$$f(x;\alpha,\theta) = \frac{1}{\theta^{\alpha}\Gamma(\alpha)} x^{\alpha-1} e^{-x/\theta}$$

or

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

   where

$$\Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1} e^{-x} dx.$$

   - Moments:

$$E_{\alpha,\theta}(X) = \alpha\theta \quad E_{\alpha,\theta}(X^2) = \alpha(\alpha+1)\theta^2 \quad Var_{\alpha,\theta}(X) = \alpha\theta^2$$

or

$$E_{\alpha,\lambda}(X) = \frac{\alpha}{\lambda} \quad E_{\alpha,\lambda}(X^2) = \frac{\alpha(\alpha+1)}{\lambda^2} \quad Var_{\alpha,\lambda}(X) = \frac{\alpha}{\lambda^2}$$

1. Characterisation *(Continued)*

   - Moment Generating Function:

$$
\begin{aligned}
m_X(t) = E_{\alpha,\theta}(e^{tX}) &= \int_0^\infty e^{tx} \frac{1}{\theta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-x/\theta} dx \\
&= \frac{1}{(1-t\theta)^\alpha} \int_0^\infty \frac{(1-t\theta)^\alpha}{\theta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-\frac{x(1-t\theta)}{\theta}} dx \\
&= \frac{1}{(1-t\theta)^\alpha},
\end{aligned}
$$

   as long as $t < \theta^{-1}$.

1. Characterisation *(Continued)*

   - Moment Generating Function *(Continued)*:

     RECALL:

     · $m_X^{(k)}(t) = \frac{d^k}{dt^k} m_X(t) = \frac{d^k}{dt^k} E(e^{tX}) = E\big(\frac{d^k}{dt^k} e^{tX}\big) = E(X^k e^{tX})$

     $\implies m_X^{(k)}(0) = E(X^k)$

     For example, for the Gamma distribution

     $$\frac{d}{dt} m_X(t) = \alpha\theta(1-t\theta)^{-\alpha-1}, \quad \frac{d^2}{dt^2} m_X(t) = \alpha(\alpha+1)\theta^2(1-t\theta)^{-\alpha-2}.$$

     So, evaluating at $t = 0$,

     $$m_X^{(1)}(0) = \alpha\theta = E_{\alpha,\theta}(X), \qquad m_X^{(2)}(0) = \alpha(\alpha+1)\theta^2.$$

1. Characterisation *(Continued)*
   - Moment Generating Function *(Continued)*:

     RECALL *(Continued)*:

     · the *mgf* uniquely determines a distribution. For example, if $Y_1, \ldots, Y_r \overset{iid}{\sim} \text{Exp}(\theta)$, then $X = \sum_{i=1}^r Y_i$ has *mgf*:

$$\begin{aligned} m_X(t) &= E_\theta\left\{ \exp\left(t \sum_{i=1}^r Y_i \right) \right\} = E_\theta\left(\prod_{i=1}^r e^{tY_i} \right) = \prod_{i=1}^r E_\theta(e^{tY_i}) \\ &= \{(1 - t\theta)^{-1}\}^r = (1 - t\theta)^{-r}, \end{aligned}$$

     since $m_{Y_i}(t) = E_\theta(e^{tY_i}) = (1 - t\theta)^{-1}$, because an exponential is just a Gamma with $\alpha = 1$.

     So, the sum of an *iid* collection of r exponential random quantities has a Gamma distribution with parameters r and θ.

1. Characterisation *(Continued)*

   - Quantiles (Percentiles):

     $$x_p \text{ solves : } \int_0^{x_p} \frac{1}{\theta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-x/\theta} dx = p$$

     No closed-form solution

     Needs iterative (computer-based) solution methods.

2. Estimation of Parameters based on $x_1, \ldots, x_n$:

   - (Standard) Method of Moments (*MOM*):

     Solve system:

$$\alpha\theta = \overline{x} \qquad \text{and} \qquad \alpha(\alpha+1)\theta^2 = \overline{x^2} = \frac{1}{n}\sum_{i=1}^{n} x_i^2,$$

     The MOM estimates are:

$$\hat{\alpha}_{MOM} = \frac{\overline{x}^2}{\overline{x^2} - \overline{x}^2} = \frac{n\overline{x}^2}{(n-1)s^2}, \quad \hat{\theta}_{MOM} = \frac{\overline{x^2} - \overline{x}^2}{\overline{x}} = \frac{(n-1)s^2}{n\overline{x}},$$

     since $s^2 = \frac{n}{n-1}(\overline{x^2} - \overline{x}^2)$.

     For our example, $\hat{\alpha}_{MOM} = 0.1922$, $\hat{\theta}_{MOM} = 15558.26$.

2. Estimation of Parameters based on $x_1, \ldots, x_n$ *(Continued)*:

   - Method of Percentiles (*MOP*):

     Needs iterative (computer-based) solution methods.

     For our data, *MOP* estimates based on upper and lower quartiles ($p_1 = 0.25$, $p_2 = 0.75$) are:

$$\hat{\alpha}_{MOP} = 0.7236 \qquad \hat{\theta}_{MOP} = 2848.38$$

2. Estimation of Parameters based on $x_1, \ldots, x_n$ *(Continued)*:

   - Maximum Likelihood Estimate (*MLE*)

     Likelihood Function:

$$
\begin{aligned}
L(\alpha, \theta; x_1, \ldots, x_n) &= \prod_{i=1}^{n} f_X(x_i; \alpha, \theta) \\
&= \frac{1}{\theta^{n\alpha}\{\Gamma(\alpha)\}^n} \left(\prod_{i=1}^{n} x_i \right)^{\alpha-1} \exp\left(-\frac{1}{\theta} \sum_{i=1}^{n} x_i \right)
\end{aligned}
$$

     Log-Likelihood Function:

$$
\begin{aligned}
l(\alpha, \theta) &= \sum_{i=1}^{n} \ln \left\{ \frac{1}{\theta^{\alpha}\Gamma(\alpha)} x_i^{\alpha-1} \exp\left(-\frac{x_i}{\theta} \right) \right\} \\
&= \sum_{i=1}^{n} \big[-\alpha \ln \theta - \ln\{\Gamma(\alpha)\} + (\alpha - 1) \ln x_i - x_i \theta^{-1} \big] \\
&= -n\alpha \ln \theta - n \ln\{\Gamma(\alpha)\} + (\alpha - 1) \sum_{i=1}^{n} \ln x_i - \theta^{-1} \sum_{i=1}^{n} x_i .
\end{aligned}
$$

     Thus, the score equations are:

$$
\begin{aligned}
\frac{\partial l(\alpha, \theta)}{\partial \alpha} &= -n \ln \theta - n\psi(\alpha) + \sum_{i=1}^{n} \ln x_i = 0, \\
\frac{\partial l(\alpha, \theta)}{\partial \theta} &= -\frac{n\alpha}{\theta} + \theta^{-2} \sum_{i=1}^{n} x_i = 0,
\end{aligned}
$$

     where $\psi(\alpha) = d \ln\{\Gamma(\alpha)\}/d\alpha = \Gamma'(\alpha)/\Gamma(\alpha)$, which is often called the *digamma function*.

2. Estimation of Parameters based on $x_1, \ldots, x_n$ *(Continued)*:

   - Maximum Likelihood Estimates *(Continued)*

     No closed-form solution to *MLE* equations.

     Need computer-based iterative solution methods.

     However,

     $$-\frac{n\alpha}{\theta} + \theta^{-2} \sum_{i=1}^{n} x_i = 0 \quad \Longrightarrow \quad \hat{\alpha}_{MLE} \hat{\theta}_{MLE} = \overline{x}$$

     In other words, *MLE* of $\mu = E_{\alpha,\theta}(X) = \alpha\theta$ is $\hat{\mu}_{MLE} = \overline{x}$. This property is called *functional equivariance*.

2. Estimation of Parameters based on $x_1, \ldots, x_n$ *(Continued)*:

   - Maximum Likelihood Estimates *(Continued)*

     Reparameterisation from (α, θ) to (α, μ) where

$$\mu = \alpha\theta \quad \Longrightarrow \quad \theta = \frac{\mu}{\alpha}$$

Log-likelihood is now

$$l(\alpha, \mu) = n\alpha \ln \alpha - n\alpha \ln \mu - n \ln\{\Gamma(\alpha)\} + (\alpha - 1) \sum_{i=1}^{n} \ln x_i - \alpha\mu^{-1} \sum_{i=1}^{n} x_i$$

New score equations are:

$$\frac{\partial l(\alpha, \mu)}{\partial \alpha} = n \ln \alpha + n - n \ln \mu - n\psi(\alpha) + \sum_{i=1}^{n} \ln x_i - \mu^{-1} \sum_{i=1}^{n} x_i = 0,$$

$$\frac{\partial l(\alpha, \mu)}{\partial \mu} = -n\alpha\mu^{-1} + \alpha\mu^{-2} \sum_{i=1}^{n} x_i = 0,$$

Second equation yields $\hat{\mu}_{MLE} = \overline{x}$.

First equation still needs iterative solution. ($\hat{\alpha}_{MLE} = 0.6258$)

2. Estimation of Parameters based on $x_1, \ldots, x_n$ *(Continued)*:

   - Maximum Likelihood Estimates *(Continued)*

     Information matrix:

$$I(\alpha, \mu) = -\left\{ \begin{array}{cc} E_{\alpha,\mu}\left(\frac{\partial^2 l(\alpha,\mu)}{\partial \alpha^2}\right) & E_{\alpha,\mu}\left(\frac{\partial^2 l(\alpha,\mu)}{\partial \alpha \partial \mu}\right) \\ E_{\alpha,\mu}\left(\frac{\partial^2 l(\alpha,\mu)}{\partial \alpha \partial \mu}\right) & E_{\alpha,\mu}\left(\frac{\partial^2 l(\alpha,\mu)}{\partial \mu^2}\right) \end{array} \right\}$$

$$= \begin{bmatrix} n\alpha^{-1}\left\{\alpha \frac{d\psi(\alpha)}{d\alpha} - 1\right\} & 0 \\ 0 & n\alpha\mu^{-2} \end{bmatrix},$$

     Variance of *MLE*'s:

$$I^{-1}(\alpha, \mu) = \begin{bmatrix} \frac{\alpha}{n}\left\{\alpha \frac{d\psi(\alpha)}{d\alpha} - 1\right\}^{-1} & 0 \\ 0 & \frac{\mu^2}{n\alpha} \end{bmatrix}$$

     99% confidence interval for μ:

$$\hat{\mu}_{MLE} \pm 2.575\left(\frac{\hat{\mu}_{MLE}}{\sqrt{n\hat{\alpha}_{MLE}}}\right) = 2989.83 \pm 2.575\left\{\frac{2989.83}{\sqrt{96(0.6258)}}\right\}$$
$$= (1996.55,\ 3983.11).$$

2. Estimation of Parameters based on $x_1, \ldots, x_n$ *(Continued)*:

- Maximum Likelihood Estimates *(Continued)*

  99% confidence interval for α:

$$
\begin{aligned}
\hat{\alpha}_{MLE} &\pm 2.575\sqrt{\frac{\hat{\alpha}_{MLE}}{n}\{\hat{\alpha}_{MLE}\psi'(\hat{\alpha}_{MLE}) - 1\}^{-1}} \\
&= 2989.83 \pm 2.575\sqrt{\frac{0.6258}{96}\{0.6258(3.3941) - 1\}^{-1}} \\
&= (0.4297,\ 0.8219).
\end{aligned}
$$

  $\psi'(\alpha)$ is the *trigamma function*. Calculate $\psi'(0.6258) = 3.3941$ using computer.

  Note that $\alpha = 1$ is not in the interval. Additional evidence that exponential was not correct distribution.

3. Goodness-of-Fit Testing:

- Pearson Chi-Squared Test:

  · Use equal-count bins from previous exponential calculations. (Exact-equal count bins requires computer for calculations, since *CDF* of Gamma not available in closed form.)

Table 2.3: Observed and Expected Claim Amounts (in £'s) using the Gamma Distribution

Bin Range	O_i	$E_{i,MLE}$	$E_{i,MOM}$	Bin Range	O_i	$E_{i,MLE}$	$E_{i,MOM}$
0-260	12	17.0	47.4	2072-2618	5	6.1	2.8
260-545	18	9.4	7.1	2618-3285	6	6.1	2.7
545-860	10	7.9	4.8	3285-4145	6	6.1	2.8
860-1212	8	7.1	3.8	4145-5357	3	6.3	3.0
1212-1612	7	6.6	3.3	5357-7429	4	6.9	3.7
1612-2072	10	6.3	3.0	7429+	7	10.3	11.6

$X^2_{MLE} = 16.52$, $df = 12 - 1 - 2 = 9$, p-value $= 0.057$

$X^2_{MOM} = 85.55$, $df = 12 - 1 - 2 = 9$, p-value $= 1.266 \times 10^{-14}$.

3. Goodness-of-Fit Testing *(Continued)*:

   - Pearson Chi-Squared Test *(Continued)*:

     · Calculating E_i's:

$$E_{(a,b)} = nPr_{\alpha,\theta}(a < X \leq b) = n\left\{ \int_a^b \frac{1}{\theta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-x/\theta} dx \right\}$$

     Using MLEs, computer calculation yields

$$\begin{aligned} E_{(0,260.15),MLE} &= nPr\{G(0.6258, 4777.61) \leq 260.15\} \\ &= 96(0.1767) = 16.97 \end{aligned}$$

3. Goodness-of-Fit Testing *(Continued)*:

   - Pearson Chi-Squared Test *(Continued)*:

     · Calculating E_i's *(Continued)*:

     ASIDE: Change of variable formula shows

$$X \sim G(\alpha, \theta) \qquad \Longrightarrow \qquad Y = 2\theta^{-1}X \sim G(\alpha, 2) = \chi^2_{2\alpha},$$

since, letting $x(y) = \theta y/2$:

$$\begin{aligned} f_Y(y) &= f_X\{x(y)\}\left|\frac{dx(y)}{dy}\right| \\ &= \frac{1}{\theta^\alpha \Gamma(\alpha)}\{x(y)\}^{\alpha-1}\exp\left\{-\frac{x(y)}{\theta}\right\}\left|\frac{dx(y)}{dy}\right| \\ &= \frac{1}{\theta^\alpha \Gamma(\alpha)}\left\{\frac{\theta y}{2}\right\}^{\alpha-1}\exp\left\{-\frac{\theta y}{2\theta}\right\}\frac{\theta}{2} \\ &= \frac{1}{2^\alpha \Gamma(\alpha)}y^{\alpha-1}e^{-y/2}, \end{aligned}$$

which is the chi-squared density with 2α degrees of freedom.

So,

$$\begin{aligned} E_{(0,260.15),MLE} &= nPr\{G(0.6258, 4777.61) \le 260.15\} \\ &= nPr\left\{\frac{2G(0.6258, 4777.61)}{4777.61} \le \frac{2(260.15)}{4777.61}\right\} \\ &= nPr\{\chi^2_{(1.2516)} \le 0.1089\} \\ &\approx n[0.7484Pr\{\chi^2_{(1)} \le 0.1089\} \\ &\qquad\qquad + 0.2516Pr\{\chi^2_{(2)} \le 0.1089\}] \\ &= 96\{0.7484(0.2586) + 0.2516(0.0530)\} \\ &= 19.86 \end{aligned}$$

3. Goodness-of-Fit Testing *(Continued)*:

- Pearson Chi-Squared Test *(Continued)*:

  · Equal-count bin construction, 12 bins each with $E_i = 8$:

  · For first bin, $(0, b)$, solve:

$$n\left\{\int_0^b \frac{1}{\theta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-x/\theta} dx\right\} = 8.$$

Using $\hat{\alpha}_{MLE}$ and $\hat{\theta}_{MLE}$ and a computer yields $b = 76.44$

  · For next bin, $(76.44, b)$, solve:

$$n\left\{\int_{76.44}^b \frac{1}{\theta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-x/\theta} dx\right\} = 8.$$

Using *MLE*s and computer yields $b = 236.16$. Continue.

  · $X^2 = 20$, $df = 12 - 1 - 2 = 9$, p-value $= 0.0179$