THE GAMMA DISTRIBUTION AS A MIXTURE OF EXPONENTIAL DISTRIBUTIONS

by

Leon Jay Gleser
Purdue University

Technical Report #87-28

Department of Statistics
Purdue University

June 1987

THE GAMMA DISTRIBUTION AS A MIXTURE OF EXPONENTIAL DISTRIBUTIONS

by

Leon Jay Gleser
Purdue University
West Lafayette, Indiana 47907

A gamma distribution with arbitrary scale parameter θ and shape parameter $r < 1$ can be represented as a scale mixture of exponential distributions. Arbitrary gamma distributions are thus mixtures of sums of independent exponential random variables.

1. INTRODUCTION AND MAIN RESULT

In a paper by Proschan (1963) on failure rate analysis some data are included concerning the time of successive failures of the air conditioning system of each member of a fleet of 13 Boeing 720 jet airplanes. Proschan tested the fit of the exponential distribution to this data using the Kolmogorov-Smirnov test of fit, and was unable to reject the hypothesis that the pooled data are exponentially distributed. However, Proschan remarks that the pooled data seem to exhibit a decreasing failure rate, and thus questions whether the exponential distribution really does provide an adequate model for the data.

In a later paper, Dahiya and Gurland (1972) use a test based on the sample moments to test the fit of the exponential distribution in Proschan's data against gamma alternatives. Their test rejects the null hypothesis of exponentiality at the .01 level of significance, confirming Proschan's doubts. They find that a gamma distribution with scale parameter $\theta = (122.56)^{-1}$ and shape parameter $r = 0.76$ provides a good fit to Proschan's data. They note that such a gamma distribution has a decreasing failure rate.

In Olkin, Gleser and Derman (1980), Proschan's data are used as an example of data which appear to follow an exponential distribution. In preparing a revision of this book, I came across Dahiya and Gurland's paper, and became interested in how I could explain their conclusions. Since Proschan had combined data from several airplanes, which might be subject to different uses and environments, it was natural to suspect (as Proschan had) that survival times might have different exponential distributions for different planes, and thus that Proschan's data would follow a mixture of exponential distributions. This led to the question of whether a gamma distribution can be represented as a scale mixture of exponential distributions. For values $r \leq 1$ of the shape parameter of the gamma distribution, the answer to this question is "Yes", as shown by the following theorem.

THEOREM. Let $f(x)$ be the density of a gamma distribution with scale parameter θ and shape parameter r, $0 < r \leq 1$. That is,

$$f(x) = \frac{\theta^r x^{r-1} e^{-\theta x}}{\Gamma(r)}, \quad x > 0. \tag{1}$$

Then

$$f(x) = \int_0^\infty p_\theta(\gamma)\gamma e^{-\gamma x} d\gamma \tag{2}$$

where

$$p_\theta(\gamma) = \begin{cases} \frac{(\gamma-\theta)^{-r}\theta^r}{\gamma\Gamma(1-r)\Gamma(r)}, & \text{if } \theta \geq \gamma, \\ 0, & \text{otherwise.} \end{cases} \tag{3}$$

The proof of this theorem is an elementary exercise in calculus (change variables from γ to $u = (\gamma - \theta)x$ in (2)). That $p_\theta(\gamma)$ is a legitimate probability density function can be shown by changing variables from γ to $v = (\gamma - \theta)/\theta$ in (3), and recognizing that the resulting density of v is that of $r^{-1}(1-r)$ times an F-distribution with $1-r$ and r degrees of freedom.

If X has a gamma distribution with scale parameter θ and arbitrary shape parameter $r > 1$, then

$$X \sim Z_0 + \sum_{i=1}^{[r]} Y_i, \quad [r] = \text{ integer part of } r,$$

where Z_0 has a gamma distribution with scale parameter θ and shape parameter $r - [r]$, and $Y_1, Y_2, \ldots, Y_{[r]}$ are i.i.d. exponential with scale parameter θ. Thus it follows from the theorem that the distribution of X (the gamma distribution) can be represented as a mixture of sums of independent exponential random variables. That is

$$X \sim (G/\theta)Y_0 + \sum_{i=1}^{[r]} Y_i, \tag{4}$$

where $Y_0, Y_1, \ldots, Y_{[r]}$ are i.i.d. exponential with scale parameter θ, and G is a random variable with density $p_\theta(\gamma)$ given by (3).

It might be asked whether the restriction $r \leq 1$ in the theorem is necessary. Again, the answer to this question is "Yes". That is, the density $f(x)$ of a gamma distribution with shape parameter $r > 1$ *cannot* be written as a scale mixture of exponentials. To see this, suppose that

$$f(x) = \frac{x^{r-1}\theta^r e^{-\theta x}}{\Gamma(r)} = \int_0^\infty q_\theta(\gamma)\gamma e^{-\gamma x} d\lambda(\gamma) \tag{5}$$

for some density $q_\theta(\gamma)$ relative to some (sigma-finite) measure λ on $[0, \infty)$. In this case

$$\frac{d}{dx} \log \int_0^\infty q_\theta(\gamma)\gamma e^{-\gamma x} d\lambda(\gamma) = \frac{-\int_0^\infty q_\theta(\gamma)\gamma^2 e^{-\gamma x} d\lambda(\gamma)}{\int_0^\infty q_\theta(\gamma)\gamma e^{-\gamma x} d\lambda(\gamma)} < 0$$

for all $x > 0$, since the existence of $f(x)$ for all $x > 0$ ensures that we can interchange the operations of differentiation and integration. [Note that the integral on the right side of (5) is a moment generating function.] On the other hand,

$$\frac{d}{dx} \log \left[\frac{x^{r-1}\theta^r e^{-\theta x}}{\Gamma(r)} \right] = \frac{r-1}{x} - \theta$$

which (when $r > 1$) is greater than 0 for $0 < x < \theta^{-1}(r-1)$. Consequently, the representation (5) is impossible.

2. DISCUSSION

We do not have to assume that the failure times X in Proschan's data have a gamma distribution with shape parameter $r < 1$ in order to explain the decreasing failure rate in Proschan's data. *Any scale mixture of exponentials has decreasing hazard rate.*

To see this, let

$$f(x) = \int_0^\infty p(\gamma)\gamma e^{-\gamma x} d\lambda(\gamma)$$

be any scale mixture of exponential densities. Note that here again we allow λ to be any (sigma-finite) measure on $[0, \infty)$, including counting measure, so that $f(x)$ can be a discrete mixture. It is easy to see that the survival function $1 - F(x)$ corresponding to $f(x)$ is

$$\begin{aligned} 1 - F(x) = \int_0^x f(t)dt &= \int_x^\infty \int_0^\infty p(\gamma)\gamma e^{-\gamma t} d\lambda(\gamma)dt \\ &= \int_0^\infty \int_x^\infty p(\gamma)\gamma e^{-\gamma t} dt d\lambda(\gamma) \\ &= \int_0^\infty p(\gamma) e^{-\gamma x} d\lambda(\gamma). \end{aligned}$$

The hazard function (conditional failure rate function) obtained from $f(x)$ is

$$h(x) = \frac{f(x)}{1 - F(x)}.$$

Note that for all $x \geq 0$, interchanging derivative and integral,

$$\begin{aligned} (6) \qquad \frac{d}{dx}h(x) &= \frac{\frac{d}{dx}f(x)(1 - F(x)) + (f(x))^2}{(1 - F(x))^2} \\ &= \frac{-\int_0^\infty p(\gamma)\gamma^2 e^{-\gamma x} d\lambda(\gamma) \int_0^\infty p(\gamma)e^{\gamma x} d\lambda(\gamma) + (\int_0^\infty p(\gamma)\gamma e^{-\gamma x} d\lambda(\gamma))^2}{(1 - F(x))^2} \end{aligned}$$

By the Cauchy-Schwartz Inequality,

$$\begin{aligned} (\int_0^\infty p(\gamma)\gamma e^{-\gamma x} d\lambda(\gamma))^2 &= (\int_0^\infty [p^{\frac{1}{2}}(\gamma)\gamma][p^{\frac{1}{2}}(\gamma)] e^{-\gamma x} d\lambda(\gamma))^2 \\ &\leq \int_0^\infty p(\gamma)\gamma^2 e^{-\gamma x} d\lambda(\gamma) \int_0^\infty p(\gamma) e^{-\gamma x} d\lambda(\gamma) \end{aligned}$$

so that it follows from (6) that $h(x)$ is a nonincreasing function of x.

Note: The above result was previously shown for finite mixtures of exponentials (or any other distributions with nonincreasing hazard function) by Proschan [1963; Theorem 2]. The slight generalization of his results to arbitrary mixtures proved here (by similar methods) could also be verified using Proschan's result, and the fact that any measurable function can be approximated by simple functions. The exponential distributions, of course, have constant hazard functions.

In Proschan's data, the mixing density (3) is apparently close enough to the actual mixing density (which is probably discrete, since data from 13 planes were pooled) so that any difference between the true mixed exponential distribution and a gamma distribution with $r = 0.76$ is undetectable by the test used by Dahiya and Gurland (1972).

My conclusion from this analysis [which echoes similar conclusions reached by Proschan (1963)] is that one can still entertain the hypothesis that times between failures of the air conditioning systems for individual planes in Proschan's data are exponentially distributed, even while accepting the fact that the combined data from all planes do not seem to follow an exponential distribution. Proschan (1963) suggests that scale mixtures of exponentials are likely candidates for modeling time to failure data exhibiting decreasing hazard rates. This note demonstrates that among such mixtures are the family of gamma distributions with shape parameter $r < 1$. As Proschan notes, mixtures are particularly likely to arise in pooled-data contexts. Because of the many useful properties of exponential distributions in reliability theory (e.g., lack of memory, constant failure rate), it is worth the effort to identify situations where lack of fit to the exponential distribution may be due to pooling (mixing) of data, rather than to the nature of the process underlying the data.

REFERENCES

Dahiya, Ram C., and Gurland, John (1972), "Goodness of Fit Tests for the Gamma and Exponential Distributions," *Technometrics*, **14**, 791–801.

Olkin, Ingram, Gleser, Leon J., and Derman, Cyrus (1980), *Probability Models and Applications,* New York: Macmillan.

Proschan, F. (1963), "Theoretical Explanation of Observed Decreasing Failure Rate," *Technometrics*, **5**, 375–383.