

On the Inverse Gamma as a Survival Distribution

Andrew G. Glen

Abstract This paper presents properties of the inverse gamma distribution and how it can be used as a survival distribution. A result is included that shows that the inverse gamma distribution always has an upside-down bathtub (UBT) shaped hazard function, thus, adding to the limited number of available distributions with this property. A review of the utility of UBT distributions is provided as well. Probabilistic properties are presented first, followed by statistical properties to demonstrate its usefulness as a survival distribution. As the inverse gamma distribution is discussed in a limited and sporadic fashion in the literature, a summary of its properties is provided in an appendix.

Keywords Classification of distributions • Hazard functions • Transformations of random variables

This paper, originally published in *The Journal of Quality Technology*, Volume 43 in 2011, is another paper that relied very heavily on APPL procedures. The procedure **Transform**, which derives the PDF of transformations of random variables was key to exploring various uses for the inverse of the gamma distributions. The procedure **Product** also helped derive the newly found distribution, which the author calls the Gamma Ratio distribution, a new one-parameter lifetime distribution. Creating distributions and determining their many probabilistic properties is exactly what APPL excels at.

A.G. Glen (✉)

Department of Mathematics and Computer Science, The Colorado College,
Colorado Springs, CO, USA

e-mail: andrew.glen.1984@gmail.com

2.1 Introduction

The inverse gamma distribution with parameters α and β , $IG(\alpha, \beta)$, is mentioned infrequently in statistical literature, and usually for a specific purpose. Also called the inverted gamma distribution or the reciprocal gamma distribution, it is most often used as a conjugate prior distribution in Bayesian statistics. This article has three primary contributions regarding the IG distribution. First it proves that the IG distribution falls into the class of upside-down bathtub shaped (UBT) distributions, an unreported result. Secondly this paper demonstrates how the IG distribution can be used as a survival distribution, a use that appears unreported. Thirdly this paper collects the scattered properties of the IG distribution into one source, similar to the type of summary found in popular works like Evans et al. [50].

The number of established UBT distributions is relatively small compared to the more common increasing failure rate distributions. However, UBT distributions have been shown to provide excellent models for certain situations. For example, Aalen and Gjessing [1] show that the UBT class of distributions are good models of absorption times for stochastic models called Wiener processes. Also, Bae et al. [5] show that degradation paths often are modeled by a restricted set of distributions, with certain paths requiring UBT distributions. Crowder et al. [38] have a conjecture that UBT distributions best model a particular data set of ball bearing failure times. Lai and Xie [85] point out that UBT distributions model certain mixtures as well as failure time models in which failure is primarily due to fatigue or corrosion. To date, the well known UBT distributions are the inverse Gaussian, log-logistic, and the log-normal. Thus, a need for more UBT distributions is certainly recognized. We will show that the IG distribution fits in a moment ratio diagram in between the log-logistic and the log-normal, thus, filling a void in the UBT area (see Figure 2.3).

One primary use of the IG distribution is for Bayesian estimation of the mean of the one parameter exponential distribution (see for example Johnson et al. [72, p. 524] or Phillips and Sweeting [131, p. 777]), as well as estimating variance in a normal regression (see for example Gelman et al. [57]). It is one of the Pearson Type V distributions, as is the Wald distribution (a.k.a. the inverse Gaussian distribution, see Johnson et al. [72, p. 21]). A number of brief descriptions of the properties of the distribution are available, mostly in text books on Bayesian methods, often in the econometrics literature, e.g., Poirier [133] and Koch [81]. Kleiber and Kotz [78] list some basic properties of the IG distribution and also model incomes with the distribution. Milevsky and Posner [115] discuss the inverse gamma distribution and point out that estimation by the method of moments is tractable algebraically. There is a different distribution with, coincidentally, the same name in Zellner [175] that is derived with the square root of the inverse of a gamma random variable. Witkovsky [169, 170] derived the characteristic function of the inverse gamma. The most complete listing of some of the

properties of the inverse gamma distribution is found on the Wikipedia website at http://en.wikipedia.org/wiki/Inverse_gamma_distribution, an anonymously written summary that does not list any sources or references. That summary is limited to the basic properties, the PDF, CDF, MGF, characteristic function, median, entropy, and the first four moments. There appears to be no comprehensive effort in the literature to employ the *IG* distribution as a survival distribution. This article will do that, specifically exploring the probabilistic and statistical properties of the *IG* distribution. Further, it is shown that the *IG* distribution fills a void in the available UBT distribution list, as seen in a moment ratio diagram to follow.

2.2 Probabilistic Properties

This section presents a number of probabilistic properties that are useful when considering the *IG* distribution as a survival distribution. An inverse gamma random variable X can be derived by transforming a random variable $Y \sim \text{gamma}(\alpha, \beta)$ with the multiplicative inverse, i.e., $X = 1/Y$. Thus, for the gamma PDF

$$f_Y(y) = \frac{\beta^\alpha}{\Gamma(\alpha)} y^{\alpha-1} e^{-\beta y} \quad y, \alpha, \beta > 0,$$

the resulting distribution, the *IG*(α, β), has PDF, CDF, survivor function, and hazard function (HF)

$$\begin{aligned} f(x) &= \frac{\beta^\alpha}{\Gamma(\alpha)} \left(\frac{1}{x}\right)^{\alpha+1} e^{-\beta/x}, \\ F(x) &= \frac{\Gamma(\alpha, \beta/x)}{\Gamma(\alpha)}, \\ S(x) &= 1 - \frac{\Gamma(\alpha, \beta/x)}{\Gamma(\alpha)}, \\ h(x) &= \frac{f(x)}{1 - F(x)} = \frac{\beta^\alpha}{\Gamma(\alpha, \beta/x)} \left(\frac{1}{x}\right)^{\alpha+1} e^{-\beta/x}, \end{aligned}$$

all with x , shape α , and scale $\beta > 0$. Recall that $\Gamma(\cdot)$ is the Euler gamma function, $\Gamma(\cdot, \cdot)$ is the incomplete gamma function, and $\Gamma(\cdot, \cdot, \cdot)$ is the generalized incomplete gamma function (see Wolfram [172] for example). Figures 2.1 and 2.2 show various PDFs and HFs for certain parameter values to indicate the shape of these functions. One of the more important aspects of a survival distribution is the shape of its hazard function. The four main classes of hazard functions for survival distributions are increasing failure rate (IFR), decreasing failure rate (DFR), bathtub-shaped (BT) and upside-down bathtub shaped (UBT). Appendix 1 shows that the *IG* distribution will always have a UBT hazard function.

In addition to the *IG* distribution always being in the UBT class, there are a number of important properties that the *IG* distribution has as a survival distribution.

- Moments are calculable, and the r th moment about the origin is as follows

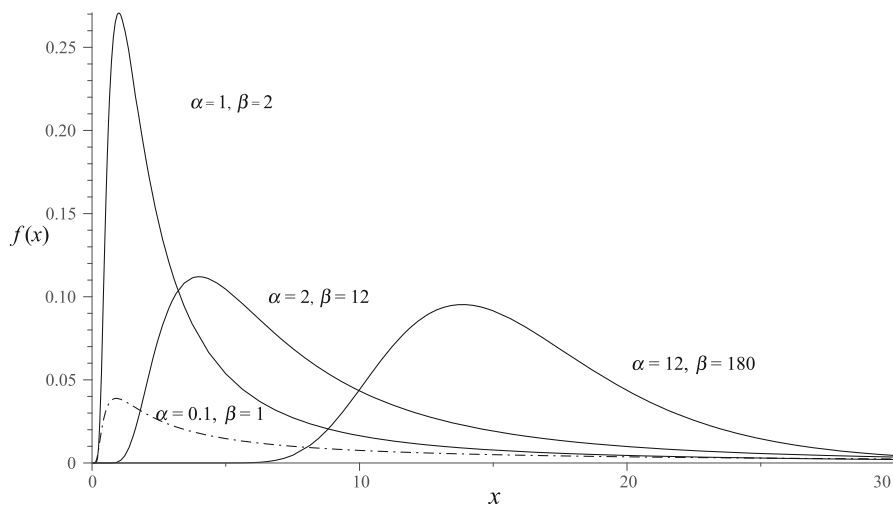


Fig. 2.1. Examples of PDFs of the inverse gamma distribution. Note the dashed PDF has a heavy right-hand tail, resulting in an undefined mean because $\alpha < 1$

$$E(X^r) = \int_0^\infty x^r f(x) dx = \frac{\beta^r \Gamma(\alpha - r)}{\Gamma(\alpha)}, \quad \alpha > r.$$

This function is sometimes referred to as the moment function, and while it is typical that $r = 1, 2, \dots$, the function holds true for non negative real values of r . Some asymptotic results also are calculable:

$$\lim_{\alpha \rightarrow \infty} E(X^r) = 0 \quad \text{and} \quad \lim_{\alpha \rightarrow 0} E(X^r) = \infty.$$

- Method of moments estimation techniques are straightforward, because the mean and variance are expressed in closed-form. The first two moments about the mean are

$$\mu = \frac{\beta}{\alpha - 1}, \quad \alpha > 1 \quad \text{and} \quad \sigma^2 = \frac{\beta^2}{(\alpha - 1)(\alpha - 2)^2}, \quad \alpha > 2,$$

so the method of moments estimators can be found by algebraically inverting the set of equations in which the sample moments are set equal to the distribution moments, i.e.,

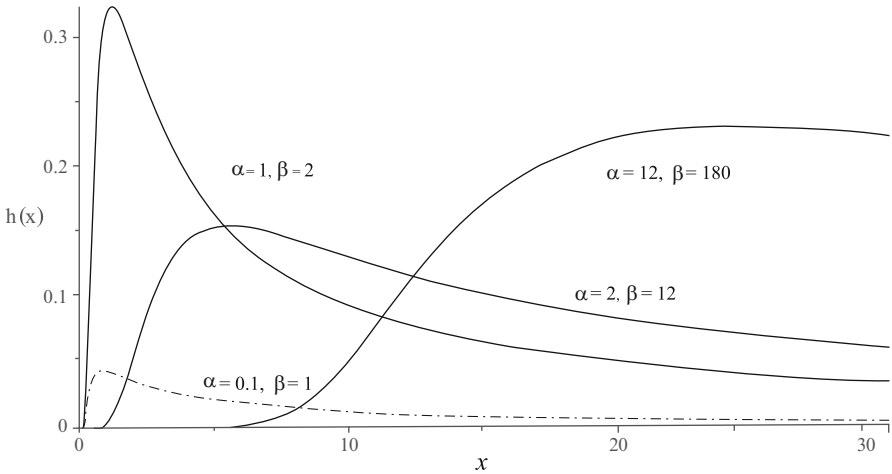


Fig. 2.2. Examples of the inverse gamma distribution hazard functions for various parameters. Note the dashed HF has a heavy tail, resulting in an undefined mean because $\alpha < 1$

$$\hat{\alpha} = \frac{\bar{x}^2 + 2s}{s} \quad \text{and} \quad \hat{\beta} = \frac{\bar{x}(\bar{x}^2 + s)}{s},$$

in which \bar{x} is the sample mean and s is the sample standard deviation. This relationship is useful for finding initial estimate for numeric solutions of the MLEs.

- The limiting distribution as $\alpha \rightarrow \infty$ is degenerate at $x = 0$.
- One comparison technique when choosing among survival distributions is to plot the coefficient of variation $\gamma_2 = \sigma/\mu$ versus skewness $\gamma_3 = \frac{E((X - \mu)^3)}{\sigma^3}$, see Cox and Oakes [35, p. 27]. A modeler would plot $(\hat{\gamma}_2, \hat{\gamma}_3)$ on such a graph to see which curve appears to be closest to the data as a start to model selection. Lan and Leemis [86] expand that graph to include the logistic-exponential distribution. Figure 2.3 takes their figure and adds the *IG* curve to the set of distributions that now include the Weibull, gamma, log-logistic, log-normal, log-exponential, inverse gamma, and the exponential distributions. The curve for the *IG* distribution falls in between the log-logistic and the log-normal distributions, in effect, helping to fill the gap between those two well-known UBT survival distributions. A more complete listing of moment ratio diagrams can be found in Vargo et al. [163], which includes the inverse gamma distribution in its figures.
- Closed-form inverse distribution functions do not exist for the *IG* distribution, so calculating quantiles, to include the median, must be done with numerical methods.
- Variate generation for the *IG* distribution can be done by inverting a variate from a gamma distribution. However, gamma variate generation

is not straightforward (as the gamma inverse distribution function (IDF) also does not exist in closed-form). Leemis and Park [96], Chap. 7, provide an explanation of various algorithms for generating gamma variates.

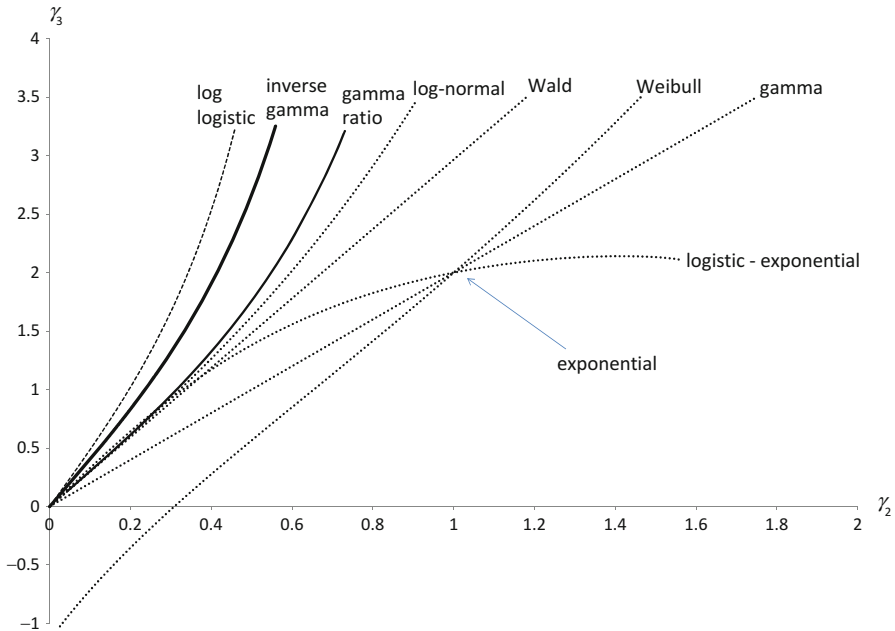


Fig. 2.3. Various two-parameter survival models with corresponding coefficient of variation versus skewness plotted. The *heavier solid line* is the inverse gamma distribution and the *lighter solid line* is the gamma ratio distribution

- Special cases and transformations of the *IG* distribution are equivalent to other known distributions (beside the obvious inverse transformation back to a gamma distribution). The $IG(1, \lambda)$ distribution is the inverse exponential distribution with PDF

$$f(x) = \frac{\lambda e^{-\lambda/x}}{x^2} \quad \lambda, x > 0.$$

The $IG(\nu/2, 1/2)$ distribution is the inverse χ^2_ν distribution with PDF

$$f(x) = \frac{(\nu/2)^{\nu/2} x^{-1-\nu/2} e^{-\nu/(2x)}}{\Gamma(\nu/2)} \quad x, \nu > 0.$$

The $IG(1/2, c/2)$ distribution is the Levy distribution with PDF

$$f(x) = \frac{(c/(2\pi))^{1/2} e^{-c/(2x)}}{x^{3/2}} \quad x, c > 0.$$

The $IG(\nu/2, s^2/2)$ distribution is also called the scaled inverse χ^2 distribution, which is the form of the IG distribution that is typically used for Bayesian estimation of σ^2 in normal regression (see for example Robert [137]).

- The negative log-gamma distribution is obtained by letting $Y \sim IG(\alpha, \beta)$ and deriving the PDF of $X = \ln Y$ to be

$$f(x) = \frac{\beta^\alpha e^{-\alpha x - \beta e^{-x}}}{\Gamma(\alpha)} \quad -\infty < x < \infty.$$

The log-gamma distribution is a well-behaved distribution with calculable moments and derivable inference, e.g., see Lawless [90]. Clearly, the negative log-gamma distribution is similarly behaved.

- An interesting new one-parameter survival distribution, which will be called the gamma ratio distribution, is derived as follows. Let $Y \sim \text{gamma}(\alpha, \beta)$ and $X \sim IG(\alpha, \beta)$ be independent random variables. The distribution of $V = XY$ has PDF

$$f(v) = \frac{v^{\alpha-1} \Gamma(\alpha + \frac{1}{2}) (\frac{1}{4} + \frac{1}{2}v + \frac{1}{4}v^2)^{-\alpha}}{2\sqrt{\pi}\Gamma(\alpha)} \quad \alpha, v > 0.$$

Note, this distribution is alternately formed by the ratio of two iid IG distributed random variables (which is the same as the ratio of two iid gamma distributed random variables). The r th moments about the origin for V are calculable,

$$E(V^r) = \int_0^\infty v^r f(v) dv = \frac{2^{2\alpha-1} \sqrt{2} \Gamma(\alpha - r) \Gamma(\alpha + r)}{2^{2\alpha-\frac{1}{2}} (\Gamma(\alpha))^2} \quad \alpha > r$$

and the mean is close to one, as is expected by its construction,

$$\mu_V = \frac{\alpha}{\alpha - 1} \quad \alpha > 1$$

with variance

$$\sigma_V^2 = \frac{(2\alpha - 1)\alpha}{(\alpha - 2)(\alpha - 1)^2} \quad \alpha > 2.$$

The CDF and hazard function are calculable, but require special functions in Maple. The gamma ratio distribution is of interest because it joins the exponential and the Rayleigh distributions as a one parameter survival distribution. For parameter values of $\alpha > 1$ it can be shown to have a UBT failure rate, however for $0 < \alpha \leq 1$ it appears to have a decreasing failure rate. This conjecture still needs to be proven, but can be shown anecdotally. When $\alpha \leq 1$ the distribution has a very heavy right tail, further indicating that no first moment exists. Furthermore, the PDF is hump-shaped for $\alpha > 1$, as is the Rayleigh, a shape the exponential can not attain. The gamma ratio distribution fits nicely in the moment ratio

diagram, see Figure 2.3, as it further fills the gap between the log-logistic and the log-normal distributions. A disadvantage to this distribution being used as a survival distribution is that the parameter α is a shape parameter, but not a scale parameter, thus, limiting its flexibility as units of measure change. This distribution warrants further research as another survival distribution, as it is in a small class of one-parameter distributions as well as in the UBT class.

- The distribution of a sum of iid IG random variables does not produce a closed-form PDF. However, products are better behaved. Let $X_i \sim IG(\alpha, \beta)$, for $i = 1, 2$ be iid, then $Z = X_1 X_2$ has PDF

$$f(z) = \frac{2z^{-1-\alpha}\beta^{2\alpha}\text{BesselK}\left(0, \frac{2\beta}{\sqrt{z}}\right)}{(\Gamma(\alpha))^2} \quad z > 0.$$

2.3 Statistical Inference

In order for a survival distribution to be viable for empirical modeling, statistical methods must be reasonably tractable. The inverse gamma distribution can produce statistical inference for both complete and right-censored data sets. Some likelihood theory and examples of fitting each type of data set are presented.

2.3.1 Complete Data Sets

For the uncensored case, let t_1, t_2, \dots, t_n be the failure times from an experiment. The likelihood function is

$$L(\alpha, \beta) = \prod_{i=1}^n f(t_i, \alpha, \beta) = \prod_{i=1}^n \frac{\beta^\alpha}{\Gamma(\alpha)} \left(\frac{1}{t_i}\right)^{\alpha+1} e^{-\beta/t_i}.$$

Taking the natural logarithm and simplifying produces

$$\ln L(\alpha, \beta) = n\alpha \ln \beta - n \ln(\Gamma(\alpha)) + (\alpha + 1) \sum_{i=1}^n \ln\left(\frac{1}{t_i}\right) - \beta \sum_{i=1}^n \frac{1}{t_i}.$$

The first partial derivatives of $\ln L(\alpha, \beta)$ with respect to the two parameters are

$$\frac{\partial \ln L(\alpha, \beta)}{\partial \alpha} = n \ln \beta - n\Psi(\alpha) + \sum_{i=1}^n \ln\left(\frac{1}{t_i}\right)$$

and

$$\frac{\partial \ln L(\alpha, \beta)}{\partial \beta} = \frac{n\alpha}{\beta} - \sum_{i=1}^n \frac{1}{t_i},$$

for $\Psi(\alpha) = \frac{d}{d\alpha} \ln \Gamma(\alpha)$ is the digamma function. Equating these two partial derivatives to zero and solving for the parameters does not yield closed-form solutions for the maximum likelihood estimators $\hat{\alpha}$ and $\hat{\beta}$ but the system of equations is well behaved in numerical methods. If initial estimates are needed, the method of moments can be used. When the equations are set equal to zero, one finds $\beta = n\alpha(\sum t_i^{-1})^{-1}$, which reduces the problem to a single parameter equation

$$n \ln n + n \ln \alpha - \ln \left(\sum_{i=1}^n t_i^{-1} \right) - n\Psi(\alpha) - \sum_{i=1}^n \ln t_i = 0,$$

which must be solved by iterative methods.

Confidence intervals for the MLEs can be obtained with the observed information matrix, $O(\hat{\alpha}, \hat{\beta})$. Cox and Oakes [35] show that it is a consistent estimator of the Fisher information matrix. Taking the observed information matrix

$$O(\hat{\alpha}, \hat{\beta}) = \begin{pmatrix} \frac{-\partial^2 \ln L(\alpha, \beta)}{\partial \alpha^2} & \frac{-\partial^2 \ln L(\alpha, \beta)}{\partial \alpha \partial \beta} \\ \frac{-\partial^2 \ln L(\alpha, \beta)}{\partial \beta \partial \alpha} & \frac{-\partial^2 \ln L(\alpha, \beta)}{\partial \beta^2} \end{pmatrix}_{\alpha=\hat{\alpha}, \beta=\hat{\beta}},$$

one then inverts the matrix and uses the square root of the diagonal elements as estimates of the standard deviations of the MLEs to form confidence intervals.

To illustrate the use of the inverse gamma distribution as a survival distribution, consider Lieblein and Zelen’s [101] data set of $n = 23$ ball bearing failure times (each measurement in 10^6 revolutions):

17.88	28.92	33.00	41.52	42.12	45.60	48.48
51.84	51.96	54.12	55.56	67.80	68.64	68.64
68.88	84.12	93.12	98.64	105.12	105.84	127.92
128.04	173.40					

This is an appropriate example because Crowder et al. [38, p. 63] conjectured that UBT shaped distributions might fit the ball bearing data better than IFR distributions based on the values of the log likelihood function at the maximum likelihood estimators. Using Maple’s numeric solver `fsolve()`, the MLEs are $\hat{\alpha} = 3.6785, \hat{\beta} = 202.5369$. Figure 2.4 gives a graphical comparison of the survivor functions for the Weibull and inverse gamma distributions fit to the empirical data.

The observed information matrix can be calculated from the MLEs of the ball bearing set:

$$O(\hat{\alpha}, \hat{\beta}) = \begin{pmatrix} 7.1783 & -0.1136 \\ -0.1136 & 0.0021 \end{pmatrix}.$$

The inverse of this matrix gives the estimate of the variance–covariance matrix for the MLEs:

$$O^{-1}(\hat{\alpha}, \hat{\beta}) = \begin{pmatrix} 1.080 & 59.472 \\ 59.472 & 3759.361 \end{pmatrix}.$$

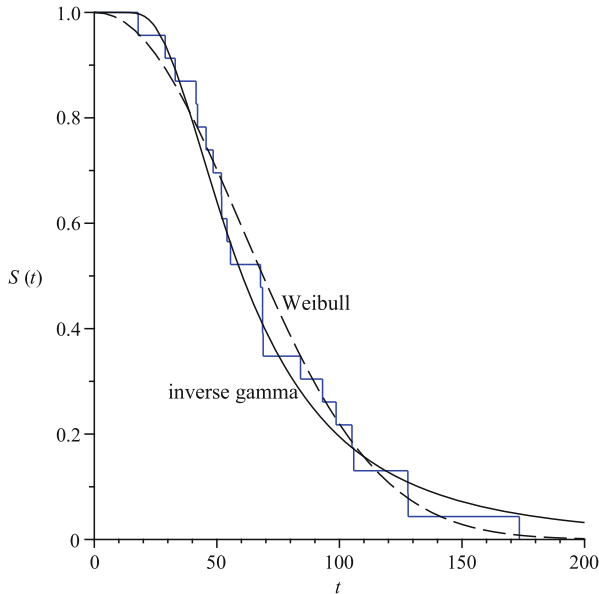


Fig. 2.4. The empirical, fitted inverse gamma, and fitted Weibull distributions for the ball bearing data set

The square roots of the diagonal elements give estimates of standard deviations for the MLEs $SE_{\hat{\alpha}} = \sqrt{1.080} \cong 1.039$ and $SE_{\hat{\beta}} = \sqrt{3759} \cong 61.31$. Thus, the approximate 95% confidence intervals for the estimates are

$$1.550 < \alpha < 5.807 \quad \text{and} \quad 85.511 < \beta < 319.563.$$

The off-diagonal elements give us the covariance estimates. Note the positive covariance between the two estimates, a fact that is made evident in the following method for joint confidence regions.

There are a number of different methods to get joint confidence regions for these two estimates. Chapter 8 of Meeker and Escobar [113] gives a good summary of many of these techniques. One such method, finding the joint confidence region for α and β , relies on the fact that the likelihood ratio statistic, $2(\ln L(\hat{\alpha}, \hat{\beta}) - \ln L(\alpha, \beta))$, is asymptotically distributed as a χ_2^2 random variable. Because $\chi_{2,0.05}^2 = 5.99$, the 95% confidence region is the set of all (α, β) pairs satisfying

$$2(\ln L(\hat{\alpha}, \hat{\beta}) - \ln L(\alpha, \beta)) = 2(-114.154 - \ln L(\alpha, \beta)) < 5.99.$$

The boundary of the region, found by solving this inequality numerically as an equality, is displayed in Figure 2.5 for the ball bearing failure times. Note in this figure the positive correlation between the two parameter estimates, displayed by the positive incline to the shape of the confidence region.

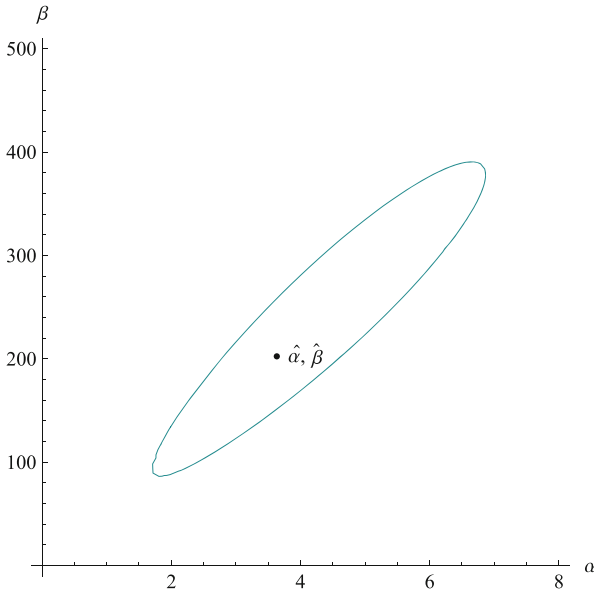


Fig. 2.5. The 95% joint confidence region for α and β for the ball bearing data

Comparisons can be made with other survival distributions. Lan and Leemis [86], as well as Glen and Leemis [61], compare the Kolmogorov–Smirnov (K–S) goodness of fit statistic D_{23} at their MLE values for a number of typical survival distributions. Table 2.1 inserts the inverse gamma into that comparison. Note that the inverse gamma has a better fit than any of the IFR class of distributions (to include its ‘parent,’ the gamma distribution). The inverse gamma also fits similarly to the other UBT distributions, giving more credence to Crowder’s conjecture that ball bearing data is better fit by UBT models.

2.3.2 Censored Data Sets

It is important that survival distributions are capable of producing inference for censored data as well. The statistical methods are similar to uncensored data, but the derivatives of the likelihood equation are not in closed-form, thus, the numerical methods require some more assistance in the form of initial

estimates. Mirroring the process of Glen and Leemis [61] and Lan and Leemis [86], initial estimates will be derived from a “method of fractiles” estimation. The data set to be used comes from Gehan’s [56] test data of remission times for leukemia patients given the drug 6-MP of which $r = 9$ observed remissions were combined with 12 randomly right censored patients. Denoting the right censored patients with an asterisk, the remission times in weeks are:

6 6 6 6* 7 9* 10
 10* 11* 13 16 17* 19* 20*
 22 23 25* 32* 32* 34* 35*

Table 2.1. Kolmogorov–Smirnov goodness-of-fit Statistics for the ball bearing data at the respective MLEs

Distribution	Class	D_{23}
Exponential	IFR	0.301
Weibull	IFR	0.152
Gamma	IFR	0.123
Logistic-exponential	UBT	0.109
Inverse gamma	UBT	0.104
Inverse Gaussian	UBT	0.099
Log-logistic	UBT	0.094
Arctangent	UBT	0.093
Log-normal	UBT	0.090

To fit the inverse gamma distribution to this data, let t_1, t_2, \dots, t_n be the n remission times and let c_1, c_2, \dots, c_m be the m associated censoring times. Our maximum likelihood function is derived as follows:

$$L(\alpha, \beta) = \prod_{i \in U} f(t_i, \alpha, \beta) \prod_{i \in C} S(c_i, \alpha, \beta)$$

in which U and C are the sets of indices of uncensored and censored observations, respectively. The log likelihood function does not simplify nicely as natural logarithms of $S(x)$ are functions that cannot be expressed in closed-form. Programming environments like Maple and Mathematica will produce the log likelihood function, but it is not compact, thus, it is not shown here. It is, however, available from the author. A method of fractiles initial estimate sets the empirical fractiles equal to the $S(x)$ evaluated at the appropriate observation in a 2×2 set of equations. In the case of the 6-MP data set, two reasonable choices for the equations are

$$S(10) \cong 0.80 = 1 - \frac{\Gamma(\alpha, \frac{\beta}{10})}{\Gamma(\alpha)}$$

and

$$S(21) \cong 0.55 = 1 - \frac{\Gamma(\alpha, \frac{\beta}{21})}{\Gamma(\alpha)},$$

which produce initial estimates $\hat{\alpha}_0 = 0.9033$ and $\hat{\beta}_0 = 14.6507$. Initial values of S (i.e., in this case $S(10)$ and $S(21)$) should be chosen to adequately represent a reasonable spread of the data, not being too close together, and not being too close to an edge. The goal is to find two values that will act as initial estimates for the numeric method for finding the true MLEs that will cause the numerical method to converge. One may have to iterate on finding two productive S values. Then one must use these initial estimates required by numerical methods, take the two partial derivatives of the log likelihood with respect to α and β , set them equal to zero, and solve numerically for the MLEs, which are $\hat{\alpha} = 0.9314$ and $\hat{\beta} = 15.4766$. A plot of the inverse gamma and the arctangent distributions (see Glen and Leemis [61]) along with the Kaplan–Meier non parametric estimator are presented in Figure 2.6. The inverse gamma has a superior fit in the earlier part of the distribution.

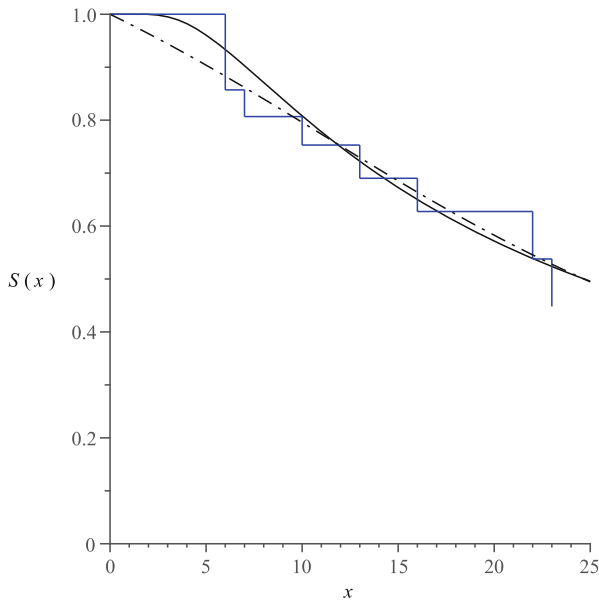


Fig. 2.6. The MLE fits of the inverse gamma distribution (*solid curve*) and the arctangent distribution (*dashed curve*) against the censored 6-MP data set

2.4 Conclusions

This paper presents the inverse gamma distribution as a survival distribution. A result is proved that shows that the inverse gamma distribution is another survival distribution in the UBT class. The well-known UBT distributions are relatively few, so it is helpful to have an alternative when dealing with UBT models. Since sporadic mention is made in the literature of the distribution, as well as very little probability or statistical information, this paper helps fill this gap in the literature. Probabilistic properties and statistical methods are provided to assist the practitioner with using the distribution. Finally, many properties, previously published in a scattered manner, are combined into this one article.

Appendix 1

In this appendix it is shown that the hazard function of the inverse gamma distribution can only have the UBT shape and the mode, x^* , of the hazard function is bounded by $0 < x^* < \frac{2\beta}{\alpha+1}$.

In order to show that the hazard function is always UBT in shape, the following is sufficient: (1) $\lim_{x \rightarrow 0} h(x) = 0$, (2) $\lim_{x \rightarrow \infty} h(x) = 0$, and (3) $h(x)$ is unimodal. Because all hazard functions are positive and have infinite area underneath the hazard curve, Rolle's theorem, taken with (1) and (2) guarantees at least one maximum point (mode). So it must be shown that there is only one maximum value of $h(x)$ on the interval $0 < x < \infty$ to prove that h is UBT in shape.

First consider (1) that $\lim_{x \rightarrow 0} h(x) = 0$. If we make a change of variable $z = 1/x$ then the limit to be evaluated becomes

$$\lim_{z \rightarrow \infty} \frac{\beta^\alpha e^{-\beta z} z^{\alpha+1}}{\Gamma(\alpha, 0, \beta z)}.$$

The denominator is an incomplete gamma function which in the limit becomes the complete gamma function, $\Gamma(\alpha) = \Gamma(\alpha, 0, \infty)$, a constant. The numerator can be rewritten as

$$\frac{\beta^\alpha z^{\alpha+1}}{e^{z\beta}}$$

so that it has the form ∞/∞ in the limit and can therefore be evaluated with $\lceil \alpha \rceil$ (the next highest integer if α is not an integer) successive applications of L'Hospital's rule. Thus, the limit of the numerator of h is $\lim_{z \rightarrow \infty} \beta^\alpha e^{-z\beta} z^{\alpha+1} = 0$. Consequently $\lim_{x \rightarrow 0} h(x) = 0$ and (1) is satisfied.

Next, consider (2) that $\lim_{x \rightarrow \infty} h(x) = 0$. Again applying L'Hospital's rule, this time only once, the first derivative of the numerator with respect to x is

$$-\beta^\alpha e^{-\beta/x} \left(\frac{1}{x}\right)^{3+\alpha} (-\beta + x + \alpha x).$$

The first derivative of the denominator with respect to x is

$$-\frac{\beta^\alpha e^{-\beta/x} \left(\frac{1}{x}\right)^{-1+\alpha}}{x^2}.$$

Dividing the numerator's derivative by the denominator's derivative and simplifying we get

$$\lim_{x \rightarrow \infty} \frac{-\beta + x + \alpha x}{x^2} = 0$$

and (2) is satisfied.

Rolle's Theorem allows us to conclude that because the left and right limits of $h(x)$ are both zero, and because h is a positive function, there must be some value ξ such that $h'(\xi) = 0$ on the interval $0 < \xi < \infty$. Because h has a left limit at the origin, it cannot have a DFR or a BT shape. Because h has a right limit at zero, it cannot have an IFR shape. Also, because h is a positive function, there is at least one critical point at $x = \xi$ that must be a maximum.

Finally, considering element (3) of the proof, one must rely on the Lemma E.1 and Theorem E.2 of Marshall and Olkin [107, pp. 134–135]. That theorem and lemma establish that

$$\rho(x) = -\frac{f'(x)}{f(x)}$$

and $h(x)$ have the same number of sign changes and in the same order. Further the critical point of ρ is an upper bound for the critical point of h . For this distribution,

$$\rho(x) = \frac{-\beta + x + \alpha x}{x^2},$$

which has first derivative

$$\rho'(x) = \frac{2\beta - (1 + \alpha)x}{x^3},$$

and can be shown to be positive from $0 < x < \frac{2\beta}{\alpha+1}$ and negative from $\frac{2\beta}{\alpha+1} < x < \infty$. Thus, h has only one change in the sign of its slope, is therefore UBT in shape, and has a single mode $x^* < \frac{2\beta}{\alpha+1}$.

Appendix 2

This appendix summarizes the key properties of the inverse gamma distribution that have not already been reported in the rest of the paper. The purpose of this appendix is ensure that all the major properties of the IG distribution

are in one document, as to date no manuscript appears to be this comprehensive. Most of the properties reported in the appendix are found in the works previously cited or are derived with straightforward methods.

- Mode: $\frac{\beta}{\alpha + 1}$
- Coefficient of variation: $\sigma/\mu = (\alpha - 1)^{-1/2}$
- Skewness: $E \left[\left(\frac{X - \mu}{\sigma} \right)^3 \right] = \frac{4\sqrt{\alpha - 1}}{\alpha - 3}$ for all $\alpha > 3$
- Excess kurtosis: $E \left[\left(\frac{X - \mu}{\sigma} \right)^4 \right] = \frac{30\alpha - 66}{(\alpha - 3)(\alpha - 4)}$ for all $\alpha > 4$
- Entropy: $\alpha - (1 + \alpha)\Psi(\alpha) + \ln(\beta\Gamma(\alpha))$
- Moment generating function: $\frac{-2\beta t^{\alpha/2} \text{BesselK}(\alpha, 2\sqrt{-\beta t})}{\Gamma(\alpha)}$
- Characteristic function: $\frac{-2i\beta t^{\alpha/2} \text{BesselK}(\alpha, 2\sqrt{-i\beta t})}{\Gamma(\alpha)}$
- Laplace transform: $\frac{2\beta s^{\alpha/2} \text{BesselK}(\alpha, 2\sqrt{-\beta s})}{\Gamma(\alpha)}$
- Mellin transform: $\frac{\beta s^{-1} \Gamma(1 - s + \alpha)}{\Gamma(\alpha)}$