University of Texas, MD Anderson Cancer Center

UT MD Anderson Cancer Center Department of Biostatistics
Working Paper Series

Year 2012 *Paper* 77

Fast Approximation of Inverse Gamma Inequalities

John D. Cook*

*The University of Texas MD Anderson Cancer Center, jdcook@mdanderson.org

http://biostats.bepress.com/mdandersonbiostat/paper77

Fast Approximation of Inverse Gamma Inequalities

John D. Cook

Abstract

Bayesian clinical trial methods sometimes use a conjugate exponential-inverse gamma model for event times. Random inequalities between posterior inverse gamma distributions are used to determine stopping conditions, for example in [1]. Computing these inequalitiy probabilities accounts for nearly all of the computation time used in simulating such trials. This report presents an approximation that could reduce this time by two orders of magnitude.

Fast Approximation of Inverse Gamma Inequalities

John D. Cook
Department of Biostatistics
P. O. Box 301402, Unit 1409
The University of Texas, M. D. Anderson Cancer Center
Houston, Texas 77230-1402, USA
cook@mdanderson.org

October 17, 2012

Abstract

Bayesian clinical trial methods sometimes use a conjugate exponential-inverse gamma model for event times. Random inequalities between posterior inverse gamma distributions are used to determine stopping conditions, for example in [1]. Computing these inequalitiy probabilities accounts for nearly all of the computation time used in simulating such trials. This report presents an approximation that could reduce this time by two orders of magnitude.

1 Approximation

When X and Y are independent inverse gamma random variables, the inequality

$$P(X > Y)$$

can be computed in closed form [2]. However, the probability

$$P(X > Y + \delta)$$

requires numerical integration when $\delta > 0$.

The idea presented here is simply to approximate the distribution on $Y + \delta$ by the distribution on an inverse gamma random variable with the

same and variance. That is,

$$P(X > Y + \delta) \approx P(X > Y_\delta),$$

evaluating the later exactly using the closed-form solution mentioned above.

If Y has mean μ and variance σ^2, the Y_δ has mean $\mu + \delta$ and variance σ^2. This implies the shape parameter of Y_δ is

$$\alpha = \frac{(\mu + \delta)^2}{\sigma^2} + 2$$

and the scale parameter is

$$\beta = (\alpha - 1)(\mu + \delta).$$

Matching moments to define Y_δ assumes that the first two moments of Y exist. In practice, Y often has a large shape parameter and so this is not a concern. (In the safety monitoring application developed in [1], Y represents what is known regarding a historical control. The shape parameter is the effective sample size and so is typically large, say on the order of 100 or larger.)

2 Error estimation

Let f_X denote the PDF of X and F_Y the CDF of Y. Let F_δ be the CDF of Y_δ. Then

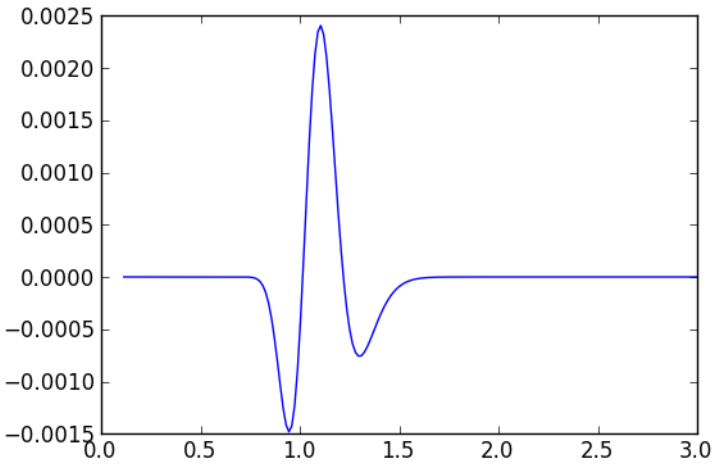
$$\begin{aligned} |P(X > Y + \delta) - P(X > Y_\delta)| &= \left| \int_\delta^\infty f_X(x) \left(F_Y(x - \delta) - F_\delta(x)\right) \, dx \right| \\ &\leq \int_\delta^\infty f_X(x) \left|F_Y(x - \delta) - F_\delta(x)\right| \, dx \\ &\leq \max_x \left|F_Y(x - \delta) - F_\delta(x)\right| \end{aligned}$$

This gives an upper bound on the approximation error independent of X. However, as we will see in the next section, it is a pessimistic error bound.

3 Illustration

To illustrate the accuracy of the proposed approximation, let Y inverse gamma with shape 100 and scale 99. This makes $E(Y) = 1$. Rescaling does not effect the accuracy, so we can always rescale to make the mean 1. The variance of Y is $1/98$. We pick $\delta = 0.1$, approximately the standard deviation of Y.

The following graph plots $F_Y(x - \delta) - F_\delta(x)$.



The maximum absolute difference between the two functions is about 0.0025. However, to achieve this error bound, f_X would have to be a point-mass concentrated near 1.1. In practice, the distribution on Y would be compared to distributions on X that are fairly dispersed, no more concentrated than Y. In this case the positive and negative differences between $F_Y(x - \delta)$ and $F_\delta(x)$ would largely cancel.

We computed the error in approximating $P(X > Y + \delta)$ by $P(X > Y_\delta)$, varying the shape and scale of X. We let the shape vary from 1 to 100 and the scale from 1 to 200. The maximum error occurs when the shape is 100 and the scale is 88.489. At that point the true inequality value is 0.06194 and the approximate value is 0.06240, a difference of 0.00046, about 5 times smaller than the upper bound on error given in the previous section. The average error over the same region is 0.0000453 which is about 10 times

smaller than the maximum error.

4 References

References

[1] Peter F. Thall, Leiko H. Wooten, and Nizar M. Tannir. Monitoring Event Times in Early Phase Clinical Trials: Some Practical Issues, Clinical Trials 2, 467-478 (2005).

[2] John D. Cook. Numerical Computation of Stochastic Inequality Probabilities. MDACC technical report UTMDABTR-008-03, (2003), revised 2005.

Available at http://biostats.bepress.com/mdandersonbiostat/paper46/