Stochastic Inequality Probabilities for Adaptively Randomized Clinical Trials

John D. Cook[*,1] and **Saralees Nadarajah**[2]

[1] Department of Biostatistics & Applied Mathematics, M.D. Anderson Cancer Center, The University of Texas, Houston, Texas 77030, U.S.A.
[2] Department of Statistics, University of Nebraska Lincoln, Lincoln, Nebraska 68588, U.S.A.

Received 5 December 2005, revised 19 December 2005, accepted 23 December 2005

Summary

We examine stochastic inequality probabilities of the form $P(X > Y)$ and $P(X > \max(Y, Z))$ where X, Y, and Z are random variables with beta, gamma, or inverse gamma distributions. We discuss the applications of such inequality probabilities to adaptively randomized clinical trials as well as methods for calculating their values.

Key words: Adaptive randomization; Stochastic inequalities; Clinical trials; Hypergeometric functions.

1 Introduction

Let D denote a two-parameter distribution family. For integer $n \geq 2$ define

$$g_{n,D}(a_1, b_1, \ldots, a_n, b_n) = P(X_1 > \max_{i \geq 2} X_i)$$

where the X_i are independent random variables with distribution D and distribution parameters (a_i, b_i).

In this paper, we examine the functions $g_{n,D}$ for $n = 2, 3$ and distributions $D \in \{\mathrm{B}, \mathrm{G}, \mathrm{IG}\}$ where B stands for the beta distribution, G for the gamma, and IG for the inverse gamma distribution. We explain how these functions may be computed, and how they are used in adaptively randomized clinical trials.

In general, the functions $g_{n,D}$ must be evaluated numerically. Cook (2003) provides a general approach to computing such probabilities. However, for many special cases, these functions may be evaluated analytically. Nadarajah (2002, 2003b, 2004) and Nadarajah and Kotz (2003, 2004b) examine $g_{2,D}$ for many distribution families D. For some clinical trials, these analytical results are directly useful. Other trials require numerical evaluation of probabilities that cannot be computed in closed form. However, in that case the analytical results provide valuable test cases for the more general numerical software.

We will review the results for $g_{2,D}$ for the distributions of interest and present new results for $g_{3,D}$.

2 Motivating Applications

2.1 Adaptive randomization

The purpose of outcome-adaptive randomization is to treat patients more effectively by weighting randomization probabilities in favor of better performing arms. Berry and Eick (1995) and Berry (1994) discuss of the ethics and efficiency of adaptive randomization trials.

* Corresponding author: e-mail: cook@mdanderson.org, Phone: 713 794 4166, Fax: 713 563 4243

Fourteen adaptively randomized clinical trials are currently running at M. D. Anderson Cancer Center (MDACC). In these trials, the probability that a patient will be assigned to a given arm is proportional to r^{λ} where r is the probability that the arm is in some sense best and $\lambda > 0$ is a trial design parameter. (Often $\lambda = 1$, but one may choose a smaller value such as 1/2 to dampen the effect of the data on the randomization probabilities, or a larger value such as 2 to amplify the response to the data.) The value of r is determined by $g_{n,D}$ where n is the number of arms and D depends on the probability model being used. The most common value of n is 2, and the next most popular value is 3. Larger values are uncommon but possible. At MDACC, we have designed trials with n as large as 5. The most common distributions are beta and inverse gamma. More details will be given as we consider binary and time-to-event endpoints separately.

2.2 Binary outcomes

Suppose the probability of a binary response on a treatment is θ and that *a priori* θ has a beta(a, b) distribution. The beta distribution is very commonly used for the prior on binary outcomes because is it is very flexible – the probability density function can be U-shaped, J-shaped, flat, approximately normal, etc. – and because it is a conjugate distribution.

After observing s successes and f failures, the posterior probability distribution on θ, given this data, is beta($a + s$, $b + f$). This simple relationship is possible because the beta distribution is the conjugate prior to the binomial distribution.

If $X_i \sim$ beta(a_i, b_i) is the posterior probability of response on treatment i for $i = 1, 2, \ldots, n$, then the posterior probability that the first treatment has the best response is

$$r_1 = g_{n,\mathrm{B}}(a_1, b_1, a_2, b_2, \ldots, a_n, b_n)\,.$$

Given a fixed non-negative λ, we assign patients to the ith treatment with probability

$$\frac{r_i^{\lambda}}{\sum_{j=1}^{n} r_j^{\lambda}}$$

where r_i is the probability of the ith treatment having the best response.

The parameter λ controls the extent to which the randomization probabilities are influenced by the outcomes. The choice $\lambda = 0$ corresponds to equal randomization. As $\lambda \to \infty$, the trial approaches a deterministic "play the winner" strategy. In most adaptively randomized trials at MDACC, λ varies as a function of accrual. Specifically,

$$\lambda(n) = [n > N]\,\lambda_0 \tag{1}$$

where n is the patient accrual number and $[\cdot]$ is the indicator function of the expression inside the brackets. For example, a trial may set $N = 20$ and $\lambda_0 = 1/2$. Such a trial would start with a burn-in period of equal randomization for the first 20 patients before beginning the adaptive randomization. Wathen (2004) has proposed more general forms for $\lambda(n)$, gradually increasing λ rather than using a step function.

We also use the r_i values in stopping rules: if r_i falls below a specified threshold, we either suspend accrual to the i-th arm or permanently drop this arm from consideration. Wathen and Wooten (2005) provide software to simulate adaptively randomized trials according to this design. Thall et al. (1995) and Nguyen and Cook (2005) apply similar stopping rules by comparing outcome probabilities to a historical control.

2.3 Time-to-event outcomes

Let T be the time to a negative outcome event and assume that $T \mid \mu$ follows an exponential distribution with mean μ. Also, assume that *a priori* μ has an inverse gamma distribution IG(a, b). This

distribution has density function

$$f(x) = \left(\frac{b^a}{x^{a+1}\Gamma(a)} \right) \mathrm{e}^{-b/x}$$

for all positive x.

The inverse gamma distribution is used because it is a conjugate prior for the exponential distribution and leads to simple posterior calculations. After observing n events and total time on test t, the posterior distribution on μ is IG($a + n$, $b + t$). Let X_i be the the posterior mean time to event on treatment i. We assign patients to treatment i with probability proportional to r_i^{λ} where r_i is the probability that a sample from X_i is larger than independent samples from each of the other X's. If $X_i \sim \mathrm{IG}(a_i, b_i)$ for $i = 1, 2, \ldots n$, then

$$r_1 = g_{n,\mathrm{IG}}(a_1, b_1, a_2, b_2, \ldots, a_n, b_n)\,.$$

Refer to Thall et al. (2005) for a detailed description of a trial designed according to this probability model, and to Wathen and Wooten (2005) for software to simulate such trials.

One may also model exponential event times in terms of the median m rather than the mean μ. This model is equivalent to modeling the mean, but many physicians prefer to think in terms of medians rather than means. If the median is distributed *a priori* as IG(a, b) then after n events and time on test t, the posterior distribution is IG($a + n$, $b + \log(2)\, t$.).

Extensions to this model have been considered. Wathen (2005) has investigated time-to-event designs using different probability distributions. Thall et al. (2002), Thall and Wathen (2005), and Cheung et al. (2005) have used time-to-event models that add patient covariates.

3 Example Protocol

The following example is taken from a clinical trial design under development at MDACC for a protocol treating patients with recurrent glioblastoma multiforme (GBM), an aggressive form of brain cancer. The primary outcome is time to progression (TTP). Median time to progression for this disease has been around three months in previous trials. Patients are randomized between three treatment arms: Valproate + Accutane, Carboplatin + Accutane, and Valproate + Accutane + Carboplatin. A minimum of 30 and a maximum of 135 patients will be accrued.

We assume time to progression on the i-th arm is exponentially distributed with median m_i and that *a priori* each m_i is distributed as $\mathrm{IG}(a_0, b_0)$ where $a_0 = 2.009$ and $b_0 = 3.027$. These parameters were chosen so that each m_i has mean 3, matching the historical median TTP, and variance 1000, resulting in an uninformative prior.

The trial begins by randomizing the first 30 patients to each arm with fixed probability 1/3. After this initial burn-in period, randomization begins adapting. Patients are randomized to the i-th treatment with probability r_i, the probability that median time to progression on the ith treatment exceeds the median time to progression on the two other arms. In the notation of equation (1), $N = 30$ and $\lambda_0 = 1$. The result of this design is that each arm will be assigned patients with frequency proportional to the strength of evidence for that arm being superior.

We calculate r_i as follows. Let e_i be the number of patients who have had an "event" on arm i, that is, the number who have had disease progression after treatment. Also, let t_i be the total time on test for arm i, the sum of the times from treatment to either progression or latest observation for each patient on on i-th treatment. For a given treatment i, let j and k denote the other two treatments. Then

$$r_i = g_{3,\mathrm{IG}}(a_0 + e_i, b_0 + ct_i, a_0 + e_j, b_0 + ct_j, a_0 + e_k, b_0 + ct_k)$$

where $c = \log 2$.

After the first 30 patients, if at any point $r_i > 0.975$ the trial will be stopped for superiority of the i-th treatment. Similarly, if at any point $r_i < 0.025$ the i-th treatment arm will be dropped for inferiority

and the trial will continue as a two-arm trial. Also, arms will be dropped if they perform poorly relative to the historical TTP rate. Specifically, the ith arm is dropped if $P(m_i > 3 \mid \text{data}) < 0.01$.

If an arm is dropped, the trial will continue as a two-arm trial. In this case we redefine r_j and r_k as $g_{2,\text{IG}}(a_0 + e_j, b_0 + ct_j, a_0 + e_k, b_0 + ct_k)$ and $g_{2,\text{IG}}(a_0 + e_k, b_0 + ct_k, a_0 + e_j, b_0 + ct_j)$ respectively.

If the trial continues to maximum accrual, the ith treatment will be selected as superior if $r_i > 0.85$.

4 Beta Distributions

4.1 Two-variable case $g_{2,\text{B}}$

Cook (2003) shows that the function $g_{2,\text{B}}$ has numerous symmetries deriving from the basic relationships:

$$g_{2,\text{B}}(a_1, b_1, a_2, b_2) = g_{2,\text{B}}(b_1, a_2, b_1, a_1) = g_{2,\text{B}}(d_1, b_1, a_2, a_1) = 1 - g_{2,\text{B}}(a_2, d_1, a_1, b_1)\,.$$

Also that if we define

$$h(a_1, b_1, a_2, b_2) = \frac{\text{B}(a_1 + a_2, b_1 + b_2)}{\text{B}(a_1, b_1)\,\text{B}(a_2, b_2)}$$

then the following recurrence relations hold:

$$\begin{aligned}
g_{2,\text{B}}(a_1 + 1, b_1, a_2, b_2) &= g_{2,\text{B}}(a_1, b_1, a_2, b_2) + h(a_1, b_1, a_2, b_2)/a_1 \\
g_{2,\text{B}}(a_1, b_1 + 1, a_2, b_2) &= g_{2,\text{B}}(a_1, b_1, a_2, b_2) - h(a_1, b_1, a_2, b_2)/b_1 \\
g_{2,\text{B}}(a_1, b_1, a_2 + 1, b_2) &= g_{2,\text{B}}(a_1, b_1, a_2, b_2) - h(a_1, b_1, a_2, b_2)/a_2 \\
g_{2,\text{B}}(a_1, b_1, a_2, b_2 + 1) &= g_{2,\text{B}}(a_1, b_1, a_2, b_2) + h(a_1, b_1, a_2, b_2)/b_2
\end{aligned}$$

In another report, Cook (2005) shows that

$$g(a_1, b_1, a_2, 1) = \frac{\Gamma(a_1 + a_2)\,\Gamma(a_1 + b_1)}{\Gamma(a_1 + b_1 + a_2)}\,.$$

From this observation and the above symmetries and recurrence relations, we can compute $g_{2,\text{B}}$ in closed form if any one of its arguments is an integer. For example, if b_2 is an integer then

$$g_{2,\text{B}}(a_1, b_1, a_2, b_2) = \sum_{k=0}^{b_2 - 1} \frac{(a_2)_k}{k!}\,\frac{\text{B}(a_1 + a_2, b_1 + k)}{\text{B}(a_1, b_1)}\,. \tag{2}$$

Cook (2005) also shows how to express $g_{2,\text{B}}$ in terms of a the hypergeometric function ${}_3F_2$:

$$g_{2,\text{B}} = \frac{h(a_1, b_1, a_2, b_2)}{a_2}\;{}_3F_2\left(\begin{matrix} a_1 + a_2, a_2 + b_2, 1 \\ a_2 + 1, a_1 + b_1 + a_2 + b_2 \end{matrix}\;\middle|\; 1\right) \tag{3}$$

and from this derives two special cases. First, if $a_1 + b_1 + a_2 + b_2 = 1$ then

$$g_{2,\text{B}} = \frac{\sin(\pi a_1)\sin(\pi b_2)}{\sin(\pi(a_1 + b_1))\sin(\pi(b_1 + b_2))} \tag{4}$$

and secondly, if $a_1 + b_1 = a_2 + c_2 = 1$, $1 - a_1 > 0$, and $a_1 \neq b_1$, then

$$g_{2,\text{B}} = \frac{\Gamma(a_1 + a_2)\,\Gamma(b_1 + b_2)\sin(\pi a_1)\sin(\pi b_2)}{\pi^2}\,\frac{a_2(\psi(a_2) - \psi(1 - a_1))}{a_1 + a_2 - 1}\,.$$

The latter form is useful in cases where statisticians choose prior distribution parameters that sum to 1, a common choice in practice.

4.2 Three-variable case $g_{3,\mathrm{B}}$

If $X_i \sim \text{beta}(a_i, b_i)$ for $i = 1, 2, 3$ then

$$g_{3,\mathrm{B}} = \frac{1}{\mathrm{B}(a_1, b_1)} \int_0^1 x^{a_1-1}(1-x)^{b_1-1}\, I_x(a_2, b_2)\, I_x(a_3, b_3)\, \mathrm{d}x\,. \tag{5}$$

The only basic symmetries of $g_{3,\mathrm{B}}$ directly analogous to the symmetries of $g_{2,\mathrm{B}}$ are

$$g_{3,\mathrm{B}}(a_1, b_1, a_2, b_2, a_3, b_3) = g_{3,\mathrm{B}}(a_1, b_1, a_3, b_3, a_2, b_2)$$

and

$$g_{3,\mathrm{B}}(a_1, b_1, a_2, b_2, a_3, b_3) + g_{3,\mathrm{B}}(a_2, b_2, a_3, b_3, a_1, b_1) + g_{3,\mathrm{B}}(a_3, b_3, a_1, b_1, a_2, b_2) = 1\,.$$

These correspond to the trivial observations that $\max(X_2, X_3) = \max(X_3, X_2)$ and that if we sample from X_1, X_2, and X_3, then with probability 1, one of these samples is the largest.

The integral in (5) cannot be simplified in the general case. However, using known expansions of the incomplete beta function, we can express (5) as

$$g_{3,\mathrm{B}} = C \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \frac{(1-b_2)_k\,(a_2)_k\,(1-b_3)_l\,(a_3)_l}{(a_2+1)_k\,(a_3+1)_l\,k!\,l!}\, \mathrm{B}(a_1+a_2+a_3+k+l, b_1)\,,$$

where the constant C is given by

$$\frac{1}{C} = a_2 a_3 B(a_1, b_1)\, B(a_2, b_2)\, B(a_3, b_3)\,.$$

If the parameters take integer values then we can obtain several simpler forms of (5).

If a_2 and a_3 are integers then

$$g_{3,\mathrm{B}} = \frac{1}{\mathrm{B}(a_1, b_1)} \left[\mathrm{B}(a_1, b_1) - \sum_{k=0}^{a_2-1} \frac{(b_2)_k}{k!}\, \mathrm{B}(a_1+k, b_1+b_2) - \sum_{k=0}^{a_3-1} \frac{(b_3)_k}{k!}\, \mathrm{B}(a_1+k, b_1+b_3) + \sum_{k=0}^{a_2-1} \sum_{l=0}^{a_3-1} \frac{(b_2)_k\,(b_3)_l}{k!\,l!}\, \mathrm{B}(a_1+k+l, b_1+b_2+b_3) \right].$$

If a_2 and b_3 are integers then

$$g_{3,\mathrm{B}} = \frac{1}{\mathrm{B}(a_1, b_1)} \left[\sum_{k=0}^{b_3-1} \frac{(a_3)_k}{k!}\, \mathrm{B}(a_1+a_3, b_1+k) - \sum_{k=0}^{a_2-1} \sum_{l=0}^{b_3-1} \frac{(b_2)_k (a_3)_l}{k!\,l!}\, \mathrm{B}(a_1+a_3+k, b_1+b_2+l) \right].$$

If b_2 and b_3 are integers then

$$g_{3,\mathrm{B}} = \frac{1}{\mathrm{B}(a_1, b_1)} \sum_{k=0}^{b_2-1} \sum_{l=0}^{b_3-1} \frac{(a_2)_k\,(a_3)_l}{k!\,l!}\, \mathrm{B}(a_1+a_2+a_3, b_1+k+l)\,.$$

Define

$$H(a, b, c, d) = \frac{\mathrm{B}(a+c, b)}{c\, B(c, d)}\, {}_3F_2\left(\begin{matrix} 1-d, c, a+c \\ c+1, a+b+c \end{matrix} \,\middle|\, 1 \right).$$

If a_2 is an integer then

$$g_{3,\mathrm{B}} = \frac{H(a_1, b_1, a_3, b_3)}{B(a_1, b_1)} - \sum_{k=0}^{a_2-1} \frac{(b_2)_k}{k!}\, H(a_1+k, b_1+b_2, a_3, b_3)\,.$$

If b_2 is an integer then

$$g_{3,\mathrm{B}} = \sum_{k=0}^{b_2-1} \frac{(a_2)_k}{k!}\, H(a_1 + a_2, b_1 + k, a_3, b_3)\,.$$

Other forms can be obtained from the symmetries of $g_{3,\mathrm{B}}$.

5 Gamma and Inverse Gamma Distributions

We parameterize a gamma random variable G(a, b) by the PDF

$$\frac{a^b x^{b-1} \exp(-ax)}{\Gamma(b)}$$

and the inverse gamma IG(a, b) random variable as having PDF

$$\left(\frac{a^b}{x^{b+1}\Gamma(b)}\right) \exp(-b/x)\,.$$

With this parametrization $X \sim \mathrm{G}(a, b)$ if and only if $1/X \sim \mathrm{IG}(a, b)$.

5.1 Two-variable case $g_{2,G}$

If $X_i \sim \mathrm{G}(a_i, b_i)$ for $i = 1, 2$ then

$$f_{X_1}(x) = \frac{a_1^{b_1} x^{b_1-1} \exp(-a_1 x)}{\Gamma(b_1)}$$

and

$$F_{X_2}(x) = \frac{\gamma(a_2 x, b_2)}{\Gamma(b_2)}$$

and so we can write

$$\begin{aligned} P(X_1 > X_2) &= g_{2,\mathrm{G}}(a_1, b_1, a_2, b_2) \\ &= \frac{a_1^{b_1}}{\Gamma(b_1)\Gamma(b_2)} \int_0^\infty x^{b_1-1} \exp(-a_1 x)\, \gamma(b_2, a_2 x)\, \mathrm{d}x\,. \end{aligned} \tag{6}$$

Cook (2003) shows that

$$g_{2,\mathrm{G}}(a_1, b_1, a_2, b_2) = I_{b_2/(b_1+b_2)}(a_2, a_1)\,. \tag{7}$$

Using the hypergeometric representation for the incomplete beta function, we have

$$g_{2,\mathrm{G}}(a_1, b_1, a_2, b_2) = \left(\frac{b_2}{b_1 + b_2}\right)^{a_2-1} {}_2F_1\left(\begin{matrix} a_2, 1 - a_1 \\ a_2 + 1 \end{matrix} \,\middle|\, \frac{b_2}{b_1 + b_2}\right).$$

If b_2 is an integer, then using the identity that

$$\gamma(n, z) = \Gamma(n)\left\{1 - \exp(-z) \sum_{k=0}^{n-1} \frac{z^k}{k!}\right\}, \tag{8}$$

we can rewrite (6) as

$$g_{2,\mathrm{G}} = \frac{a_1^{b_1}}{\Gamma(b_1)} \int_0^\infty x^{b_1-1} \exp(-a_1 x) \left\{ 1 - \exp(-a_2 x) \sum_{k=0}^{b_2-1} \frac{(a_2 x)^k}{k!} \right\} \mathrm{d}x$$

$$= 1 - \frac{a_1^{b_1}}{\Gamma(b_1)} \sum_{k=0}^{b_2-1} \frac{a_2^k}{k!} \int_0^\infty x^{k+b_1-1} \exp(-a_1 x - a_2 x)\, \mathrm{d}x$$

$$= 1 - \frac{a_1^{b_1}}{\Gamma(b_1)} \sum_{k=0}^{b_2-1} \frac{a_2^k \Gamma(b_1+k)}{k!} (a_1+a_2)^{-b_1-k}.$$

5.2 Two-variable case $g_{2,\mathrm{IG}}$

If X_i are distributed as IG(a_i, b_i) for $i = 1, 2$ then $Y_i = 1/X_i$ are distributed as G(a_i, b_i). One then has

$$g_{2,\mathrm{IG}}(a_1, b_1, a_2, b_2) = P(X_1 > X_2) = P(Y_2 > Y_1) = g_{2,\mathrm{G}}(a_2, b_2, a_1, b_1).$$

5.3 Three-variable case $g_{3,\mathrm{G}}$

If $X_i \sim \mathrm{G}(a_i, b_i)$ for $i = 1, 2, 3$ then $g_{3,\mathrm{G}}(a_1, b_1, a_2, b_2, a_3, b_3)$ is given by

$$\frac{a_1^{b_1}}{\Gamma(b_1)\Gamma(b_2)\Gamma(b_3)} \int_0^\infty x^{b_1-1} \exp(-a_1 x)\, \gamma(b_2, a_2 x)\, \gamma(b_3, a_3 x)\, \mathrm{d}x. \tag{9}$$

Applying Eq. (2.10.6.3) from the work of Prudnikov et al. [1986] to evaluate the above integral, we obtain

$$g_{3,\mathrm{G}} = \frac{\Gamma(b_1+b_2+b_3)\, a_2^{b_2} a_3^{b_3}}{b_2 b_3 \Gamma(b_1)\, \Gamma(b_2)\, \Gamma(b_3)\, a_1^{b_2+b_3}} F_2\left(\begin{matrix} b_1+b_2+b_3, b_2, b_3 \\ b_2+1, b_3+1 \end{matrix} \middle| -\frac{a_2}{a_1}, -\frac{a_3}{a_1} \right) \tag{10}$$

where F_2 denotes the Appell function of the second kind. Using results from section 206D of Itô (1996), we can express F_2 as a double integral obtaining

$$g_{3,\mathrm{G}} = \frac{\Gamma(b_1+b_2+b_3)\, a_2^{b_2} a_3^{b_3}}{\Gamma(b_1)\, \Gamma(b_2)\, \Gamma(b_3)} \int_0^1 \int_0^1 \frac{u^{b_2-1} v^{b_3-1}}{(1 + (a_2/a_1)\, u + (a_3/a_1) v)^{b_1+b_2+b_3}}\, \mathrm{d}u\, \mathrm{d}v \tag{11}$$

which is amenable to efficient numerical evaluation.

We define

$$E(a, b, c, d) = \frac{c^d \Gamma(a+d)}{\mathrm{d}b^{a+d}}\, {}_2F_1\left(\begin{matrix} d, a+d \\ d+1 \end{matrix} \middle| -\frac{c}{b} \right).$$

Then by applying identity (8) and Eq. (2.10.3.2) from Prudnikov et al. (1986) we can reduce (10) to simpler forms for certain integer arguments.

If b_2 is an integer then

$$g_{3,\mathrm{G}} = \frac{a_1^{b_1}}{\Gamma(b_1)\, \Gamma(b_3)} \left[E(b_1, a_1, a_3, b_3) - \sum_{k=0}^{b_2-1} \frac{a_2^k E(b_1+k, a_1+a_2, a_3, b_3)}{k!} \right] \tag{12}$$

and, if both b_2 and b_3 are integers then

$$g_{3,\mathrm{G}} = 1 - \frac{a_1^{b_1}}{\Gamma(b_1)}\left[\sum_{k=0}^{b_2-1}\frac{a_2^k\Gamma(b_1+k)}{k!}(a_1+a_2)^{-(b_1+k)} + \sum_{k=0}^{b_3-1}\frac{a_3^k\Gamma(b_1+k)}{k!}(a_1+a_3)^{-(b_1+k)} - \sum_{k=0}^{b_2-1}\sum_{l=0}^{b_3-1}\frac{a_2^k a_3^l\Gamma(b_1+k+l)}{k!\,l!}(a_1+a_2+a_3)^{-(b_1+k+l)}\right]. \quad (13)$$

5.4 Three-variable case $g_{3,\mathrm{IG}}$

Suppose $X_i \sim \mathrm{IG}(a_i, b_i)$ for $i = 1, 2, 3$ and define $Y_i = 1/X_i$. Then $Y_i \sim \mathrm{G}(a_i, b_i)$ and

$$P(X_1 > \max(X_2, X_3)) = P(Y_1 < \min(Y_2, Y_3)).$$

Now we have

$$P(Y_1 < \min(Y_2, Y_3)) = 1 - P(Y_1 > Y_2) - P(Y_1 > Y_3) + P(Y_1 > \max(Y_2, Y_3))$$

and so $g_{3,\mathrm{IG}}(a_1, b_1, a_2, b_2, a_3, b_3)$ is equal to

$$1 - g_{2,\mathrm{G}}(a_1, b_1, a_2, b_2) - g_{2,\mathrm{G}}(a_1, b_1, a_3, b_3) + g_{3,G}(a_1, b_1, a_2, b_2, a_3, b_3).$$

6 Applications

In this section we discuss how the results in this paper have been applied to develop and to test numerical algorithms. We also discuss how these algorithms have been applied in clinical trials.

6.1 Implementing numerical algorithms

It is quite simple to estimate $g_{n,D}$ by simulation, drawing samples from each of the X_i and seeing how often the draw from C_1 is the largest. However, even with efficient random number generation routines, it can take an unacceptable amount of time to evaluate $g_{n,D}$ by simulation when this function is itself part of a larger simulation, as when simulating an adaptively randomized clinical trial.

Equation (7) for $g_{2,\mathrm{G}}$ and its counterpart for $g_{2,\mathrm{IG}}$ are used in the numerical computation of these functions. In Cook (2003) we give experimental results that demonstrate that computing $g_{2,\mathrm{G}}$ by this method is orders of magnitude faster than using simulation if one wants to obtain two decimal places of accuracy.

When the software Adaptive Randomization (Wathen and Wooten, 2005) was first released, $g_{3,\mathrm{IG}}$ was being evaluated by numerically integrating Eq. (9), with simulation being used as a back-up if the integration error was estimated to be too large. Simulating three-arm time-to-event trials strained the patience of our users. We switched to an implementation using (11) and greatly increased the speed of the simulations. (We do not use (11) directly but a closely related form that first integrates by parts twice to guarantee that the exponents of u and v are positive.)

The time savings from the new method depended on how often the original numerical integration was insufficiently accurate, which in turn depended on the scenarios being simulated. To estimate the possible effect on run time, we evaluated $g_{3,\mathrm{IG}}$ 100 000 times with uniform random arguments drawn from the interval (0.1, 90) using three methods: simulation, the original integration method, and integration based on (11). The times for the three methods were 2088 seconds, 12 seconds, and 4 seconds respectively. This suggests the new algorithm cuts the time required to simulate a 3-arm time-to-event trial by at least a factor of three and possibly by much more.

6.2 Testing numerical algorithms

Cook (2003) shows how to calculate $g_{n,D}$ using numerical integration. However, no numerical integration routine is perfectly accurate and one must thoroughly test applications that rely on numerical integration. Although we cannot often assume in practice that parameters to $g_{n,D}$ have special properties, it is nevertheless valuable to have exact results for special parameters to use as test cases. For example, we validated our implementation of $g_{2,\mathrm{B}}$ for large arguments using Eq. (2) and for small arguments using equation (8).

To test the algorithm used in the software of Wathen et al. (2005) for beta distributions, we generated one million 4-tuples of the form $(a_1, b_1, a_2, \lfloor b_2 \rfloor)$ where a_i and b_i are uniform samples from (1, 100) and $\lfloor x \rfloor$ denotes the largest integer no greater than x. We then computed $g_{2,\mathrm{B}}$ using Eq. (2) and using numerical integration. The average difference between the two values was 1.19373×10^{-8} and the maximum difference was 0.000966536.

We also generated one million 4-tuples $(u_1, u_2, u_3, u_4)/(u_1 + u_2 + u_3 + u_4)$ where each u_i a uniform (0,1) random variable. We calculated $g_{2,\mathrm{B}}$ both using Eq. (4) and numerical integration. The average difference between the values was 5.24256×10^{-14} and the maximum difference was 1.82965×10^{-12}.

6.3 Statistical applications of numerical algorithms

The software of Wathen et al. (2005) implements $g_{2,D}$ for D representing beta, gamma, inverse gamma, normal, Weibull, and log normal distributions. While not a clinical trial method *per se*, this software is frequently used while preparing interim analyses of clinical trials.

The software of Nguyen and Cook (2005) requires evaluating $g_{2,\mathrm{B}}$ to determine stopping boundaries for the safety monitoring rule of Thall, Simon, and Estey Thall et al. (1995) and is used regularly for designing single-arm phase II clinical trials.

The software of Wathen and Wooten (2005) and previous versions have been used to design many clinical trials. A companion web-based trial conduct application, using the same underlying numerical software, is currently being used to run 14 clinical trials at MDACC. These trials require evaluation of $g_{\mathrm{n,B}}$ and $g_{n,\mathrm{IG}}$ for their randomization probabilities and stopping rules.

7 Conclusions

In this paper we have defined the functions $g_{n,D}$ and explained how they are used in adaptively randomized trials MDACC. For the cases of $n = 2, 3$ and D representing beta, gamma, and inverse gamma distributions, we show how to express $g_{n,D}$ in terms of hypergeometric functions and report simplifications for special arguments.

References

Berry, D. A. (2004). Bayesian statistics and the ethics of clinical trials. *Statistical Science* **19**, 175–187.

Berry, D. A. and Eick, S. G. (1995). Adaptive assignment versus balanced randomization in clinical trials: a decision analysis. *Statistics in Medicine* **14**, 231–246.

Cheung, Y. K., Inoue, L. Y. T., Wathen, J. K., and Thall, P. F. (2005) Continuous Bayesian adaptive randomization based on event times with covariates. *Statistics in Medicine*, in press.

Cook, J. D. (2003). Numerical Computation of Stochastic Inequality Probabilities. MDACC technical report UTMDABTR-008-03.

Cook, J. D. (2005). Exact calculation of beta inequalities. MDACC technical report UTMDABTR-005-05.

Mathematical Society of Japan, Kiyosi Itô editor (1996). *Encyclopedic Dictionary of Mathematics*, second edition. MIT Press, Cambridge.

Nguyen, H. Q. and Cook, J. D. (2005). Multc Lean [Computer software].

Nadarajah, S. (2002). Reliability for Beta Models. *Serdica Mathematical Journal* **28**, 276–282.

Nadarajah, S. (2003). Reliability for extreme value distributions. *Mathematical and Computer Modelling* **37**, 915–922.
Nadarajah, S. (2003). Reliability for lifetime distributions. *Mathematical and Computer Modelling* **37**, 683–688.
Nadarajah, S. and Kotz, S. (2003). Reliability for Pareto models. *Metron* **LXI**, 191–204.
Nadarajah, S. (2004). Reliability for Laplace distributions. *Mathematical Problems in Engineering* **10**, 169–183.
Nadarajah, S. and Kotz, S. (2004) Reliability for logistic distributions. *Engineering Simulation* **26**, 81–98.
Prudnikov, A. P., Brychkov, Y. A. and Marichev, O. I. (1986). *Integrals and Series* (volume 2). Gordon and Breach Science Publishers, Amsterdam.
Thall, P. F., Inoue, L. Y. T., and Martin, T. G. (2002). Adaptive decision making in a lymphocyte infusion trial. *Biometrics* **58**, 560–568.
Thall, P. F., Simon, R. M., and Estey, E. H. (1995). Bayesian sequential monitoring designs for single-arm clinical trials with multiple outcomes, *Statistics in Medicine* **14**, 357–379.
Thall, P. F. and Wathen, J. K. (2005). Covariate-adjusted adaptive randomization in a sarcoma trial with multi-stage treatments. *Statistics in Medicine* **27**, 1947–1964.
Thall, P. F., Wooten, L. H., and Tannir, N. M. (2005). Monitoring Event Times in Early Phase Clinical Trials: Some Practical Issues, *Clinical Trials* **2**, 467–478.
Wathen, J. K. (2005) "Bayesian Doubly Optimal Group Sequential Designs for Randomized Clinical Trials." Dissertation, University of Texas Health Science Center at Houston, 2005.
Wathen, J. K. (2004) Personal correspondence.
Wathen, J. K., Nguyen, H. Q., and Cook, J. D. (2005). Inequality Calculator [Computer software].
Wathen, J. K. and Wooten, O. C. (2005) Adaptive Randomization 3.1.1 [Computer software].
Technical reports listed above are available at http://tinyurl.com/743xv. Software applications are available at http://biostatistics.mdanderson.org/SoftwareDownload.

