Rank Constrained Distribution and Moment Computations

Seksan Kiatsupaibul[a,*], Anthony J. Hayter[b], Wei Liu[c]

[a]*Department of Statistics, Chulalongkorn University, Bangkok 10330, Thailand*
[b]*Department of Business Information and Analytics, University of Denver, Denver, USA*
[c]*School of Mathematics and S3RI, University of Southampton, Southampton, UK*

Abstract

Consider a set of independent random variables with specified distributions or a set of multivariate normal random variables with a product correlation structure. This paper shows how the distributions and moments of these random variables can be calculated conditional on a specified ranking of their values. This can be useful when the ordering of the variables can be determined without observing the actual values of the variables, as in ranked set sampling, for example. Thus, prior information on the distributions and moments from their individual specified distributions can be updated to provide improved posterior information using the known ranking. While these calculations ostensibly involve high dimensional integral expressions, it is shown how the previously developed general recursive integration methodology can be applied to this problem so that they can be evaluated in a straightforward manner as a series of one-dimensional or two-dimensional integral calculations. Furthermore, the proposed methodology possesses a self-correction mechanism in the computation that prevents any serious growth of the errors. Examples illustrate how different kinds of ranking information affect the distributions, expectations, variances, and covariances of the variables, and how they can be employed to solve a decision making problem.

Keywords:
Conditional distribution, Moments, Normal distribution, Order restriction,

*Corresponding author, seksan@cbs.chula.ac.th

Ranked set sampling, Recursive integration.

1. Introduction

The computations of conditional probabilities and conditional moments are often the basis for a statistical analysis in the sciences, social sciences, and engineering. This can be particularly true with a Bayesian approach. When updated information is in a qualitative form, the distribution of interest will be conditioned on a polytope. In general, such a problem is intractable numerically (see, for example, Khachiyan (1989)) and both theorists and practitioners usually resort to Monte Carlo methods, e.g., Smith (1984), Lovász (1999), Lovász and Vempala (2006) and Kiatsupaibul et al. (2011). However, an efficient numerical method, when it is available, can have many advantages. In this work, the envelope of numerical methods is expanded to the computation of probability distribution and moments, that are conditioned on the important class of polytopes that are formed by rankings.

Consider a set of independent continuous random variables X_i with specified probability density functions $f_i(x_i)$, $1 \leq i \leq n$. The moments of these random variables are thus determined by their specified distributions. Suppose that information becomes available which indicates the ordering

$$X_1 \leq X_2 \leq \ldots \leq X_n. \tag{1}$$

The objective of this paper is to show how the information provided by this ranking can be used to provide updated distributions and moments for the random variables X_i.

This problem has applications to many areas where the ranking of the variables can be determined without observing the actual values of the variables. The literature on ranked set sampling provides discussions of many situations where this is the case (see McIntyre (1952), Patil (2002), Chen et al. (2004) and Wolfe (2004) for example). However, in ranked set sampling the information on the ordering of a set of variables is used to determine which variables to include in the sample and to subsequently observe. In contrast, this paper considers the problem where the ordering of a set of variables is used to augment prior information on their distributions, and the variables may never actually be observed. While these applications of the ranking information are different, both cases are similar in that they utilize

information on the ranking of the variables without the full realizations of the variables being available.

As an example, suppose that prior distributions for the levels X_i of a particular medical condition may be available for a set of n patients based upon covariate values of the patients. While the actual levels of this condition may be very difficult or impossible to measure, there may be an ancillary variable that can be measured for the patients and which is sufficiently correlated with the condition of interest so that it can be used to infer the ranking of the X_i. It is then useful to obtain updated expectations and variances, say, of the levels X_i of the medical condition of interest based upon the information provided by the rankings.

Alternatively, Patil (2002) discusses a problem where a hazardous waste site inspector may be able to reliably rank areas of soil with respect to concentrations of a toxic contaminant, based on features like surface staining, discoloration, or the appearance of stressed vegetation. Thus, the actually contaminant levels X_i may have specified prior distributions, but their moments can be updated based upon the ranking provided by the soil features.

In general, additional information on ranking may be derived from different sources. Besides being derived from the observation of a covariate, rankings may be derived from conditions of economic systems, as in Topkis (1998), Milgrom and Roberts (1994), and Milgrom and Shannon (1994), for example. Furthermore, rankings can also be subjectively derived from a systematic preference aggregation process, as in Kemeny and Snell (1962), Young (1995), and Ali and Meilă (2012), for example. The methodology described in this paper is an important tool to directly incorporate such ranking information into the statistical inference process, as discussed by Chiarawongse et al. (2012), which can be applied in these various areas of study.

Calculations of the conditional distributions and moments of the X_i ostensibly involve the evaluation of an n-dimensional integral expression. However, it will be seen that by employing the technique of recursive integration (discussed in Hayter (2006)) the calculations can be performed easily as a series of 1-dimensional integral calculations. In fact, it will be seen that the recursive integration technique can also be employed for more general problems when the distributions and moments are conditioned on information more complex than just a simple ranking, and when the variables X_i have a multivariate normal distribution with a product correlation structure.

It is important that the recursive integration methodology does not suffer

from a growth of errors that are compiled in high dimensions. In this application of recursive integration to the particular problem of the computations of conditional probabilities and conditional expectations, it is demonstrated that there exists a self-correction mechanism in the computation that prevents any serious growth of the errors. This condition has never been discussed before in the literature of recursive integration, and it confirms that the recursive integration technique is useful for high dimensional computations such as these.

Some examples are provided to show how the information provided by the ranking can affect the distributions, expectations, and variances of the random variables X_i. In particular, a reinforcing ranking can be considered to be one which is consistent with the rankings of the expectations of the distributions $f_i(x_i)$ (the prior expectations of the X_i), while various degrees of opposing rankings have some discrepancies with the rankings of these prior expectations. As illustrated in the examples, these different kinds of rankings will have different kinds of effects on the expectations and variances of the variables. It is also illustrated how ranking information can be useful for an important problem in portfolio selection, and applications of the proposed methodology to a real data set of asset returns are provided.

The layout of this paper is as follows. The theoretical discussion of how recursive integration can be used to calculate the conditional distributions and moments is provided in section 2 for independent random variables. An extension to random variables X_i with a multivariate normal distribution with a product correlation structure is also provided in section 2. Section 3 contains algorithms and details of the implementation of the procedure. A self-correction mechanism is discussed together with error rates and computational times. Some illustrative examples are provided in section 4, and finally a conclusion is provided in section 5.

2. General theory

The general theory concerning how to use recursive integration to calculate quantities such as moments conditional on information such as rankings is presented in this section, first for independent random variables and then for random variables with a multivariate normal distribution with a product correlation structure. Finally, some extensions are also discussed.

2.1. Independent random variables

Consider the set $S \subseteq \Re^n$ of values $\boldsymbol{X} = (X_1, \ldots, X_n)$ defined by

$$S = S_{1,2} \cap S_{2,3} \cap \ldots \cap S_{n-1,n}$$

where the set $S_{i,i+1}$ places restrictions on only X_i and X_{i+1}. Thus, the set S corresponds to the simple ordering in equation (1), for example, with

$$S_{i,i+1} = \{\boldsymbol{X} : X_i \leq X_{i+1}\}$$

for $1 \leq i \leq n-1$.

For any intervals (l_i, u_i), $1 \leq i \leq n$, it follows that

$$P(l_i \leq X_i \leq u_i; 1 \leq i \leq n \mid \boldsymbol{X} \in S) = \frac{A_1}{B} \tag{2}$$

where

$$A_1 = \int_{\boldsymbol{X} \in S^*} \cdots \int \prod_{i=1}^{n} f_i(x_i)\ dx_1 \ldots dx_n$$

and

$$B = P(\boldsymbol{X} \in S) = \int_{\boldsymbol{X} \in S} \cdots \int \prod_{i=1}^{n} f_i(x_i)\ dx_1 \ldots dx_n$$

with

$$S^* = S^*_{1,2} \cap S^*_{2,3} \cap \ldots \cap S^*_{n-1,n}$$

for

$$S^*_{1,2} = S_{1,2} \cap \{\boldsymbol{X} : l_1 \leq X_1 \leq u_1, l_2 \leq X_2 \leq u_2\}$$

and

$$S^*_{i,i+1} = S_{i,i+1} \cap \{\boldsymbol{X} : l_{i+1} \leq X_{i+1} \leq u_{i+1}\}$$

for $2 \leq i \leq n-1$.

Similarly, for any functions $g_i(x_i)$, $1 \leq i \leq n$, it follows that

$$E[g_1(X_1)g_2(X_2)\ldots g_n(X_n) \mid \boldsymbol{X} \in S] = \frac{A_2}{B} \tag{3}$$

where

$$A_2 = \int_{\boldsymbol{X} \in S} \cdots \int \prod_{i=1}^{n} (g_i(x_i)f_i(x_i))\ dx_1 \ldots dx_n.$$

While A_1, A_2, and B are each ostensibly n-dimensional integrals, they are each of the form of the integral in section 1 of Hayter (2006) with $d = 1$, and so they can each be evaluated in a straightforward manner with a series of 1-dimensional integral computations using recursive integration, regardless of the value of n. Thus, the probability in equation (2) and the expectation in equation (3) can both be evaluated in a straightforward manner with a series of 1-dimensional integral computations.

Notice that the conditional joint cumulative distribution function of the X_i can be obtained from equation (2) with $l_i = -\infty$, $1 \leq i \leq n$, and the conditional marginal distribution of a particular variable can be obtained by taking $l_i = -\infty$ and $u_i = \infty$ for all of the other variables. Also, the conditional moments of X_i can be calculated with $g_i(x_i) = x_i^k$ and with all the other functions $g_j(x_j)$ equal to one, while the conditional covariance of X_{i_1} and X_{i_2}, say, can be calculated with $g_{i_1}(x_{i_1}) = x_{i_1}$ and $g_{i_2}(x_{i_2}) = x_{i_2}$ and again with all the other functions $g_j(x_j)$ equal to one.

2.2. Multivariate normal distribution with a product correlation structure

If the random variables X_i have a multivariate normal distribution with means μ_i, variances σ_i^2, and covariances $\rho_i\rho_j$, then it is possible to write

$$X_i = \mu_i + \rho_i M + \sqrt{\sigma_i^2 - \rho_i^2} Z_i, \quad 1 \leq i \leq n, \tag{4}$$

where M and the Z_i are independent standard normal random variables. Conditional on the value of M, the random variables X_i are thus independent normal random variables.

For the evaluation of equations (2) and (3), conditioning on the value of M requires a 1-dimensional integral computation over the values m of M, with the integrand being the equation evaluated at each given value m. Since the integrand can be evaluated each time as a series of 1-dimensional integral computations, the overall computational intensity will consequently be equivalent to a series of 2-dimensional integral computations, regardless of the value of n.

It can also be noted that if the covariances are all equal and positive, so that the ρ_i are all equal to ρ, say, then for the simple ordering given in equation (1) the set S depends only on the Z_i and not on M. In this case, for moment calculations, the overall computational intensity may only be that of a 1-dimensional integral computation, depending on the functions $g_i(x_i)$. This reduction in computational intensity is possible when evaluating

the conditional expectations of the X_i, for example, since the conditional expectations of the X_i will be equal to the conditional expectations of $\mu_i + \sqrt{\sigma_i^2 - \rho^2} Z_i$.

2.3. Extensions

In addition to the simple ordering in equation (1), the set S upon which the expressions are conditioned can encompass other types of information, such as the "umbrella" ordering

$$X_1 + c_1 \leq X_2 + c_2 \leq \ldots \leq X_u + c_u \geq \ldots \geq X_{n-1} + c_{n-1} \geq X_n + c_n \qquad (5)$$

for example, for any constants c_i, which has received considerable attention in the statistical literature (see Hans and Dunson (2005), Singh and Liu (2006), Nakas and Alonzo (2007), and Gaur et al. (2012), for example). Extensions can also be made to orderings of the random variables which form a tree structure, as discussed in section 4 of Hayter (2006).

The simplicity of the evaluations of equations (2) and (3) as a series of 1-dimensional integral computations using recursive integration is because of two conditions. These are firstly that the set S only places restrictions on "adjacent" variables X_i (although it should be remembered that any labeling of the n variables is permissable), and secondly that the integrand factors into the product of separate terms for each of the variables. These conditions are seen to be met for the simple ordering in equation (1) and the umbrella ordering in equation (5), and for independent variables X_i where the expectation is required of the product of the functions $g_i(x_i)$. In fact, conditioning on equation (1), the term B is just the probability of this simple ordering, and for independent random variables its evaluation by a series of 1-dimensional integral computations using recursive integration was first shown in Hayter and Liu (1996).

If the random variables X_i are not independent, or if the conditioning information imposes restrictions on non-adjacent X_i, then A_1, A_2, and B cannot necessarily be evaluated as a series of 1-dimensional integral computations. However, recursive integration of a higher order, in which the evaluation can be performed as a series of r-dimensional integral computations (with $r \geq 2$), say, may be possible depending upon the form of the expressions for A_1, A_2, and B.

It is also worth noting that for any set $T \subseteq \Re^n$ of values $\boldsymbol{X} = (X_1, \ldots, X_n)$ defined by

$$T = T_{1,2} \cap T_{2,3} \cap \ldots \cap T_{n-1,n}$$

where the set $T_{i,i+1}$ places restrictions on only X_i and X_{i+1}, then the conditional probability $P(\boldsymbol{X} \in T \mid \boldsymbol{X} \in S)$ is also equal to A_1/B with $S^*_{i,i+1} = S_{i,i+1} \cap T_{i,i+1}$. Thus, this conditional probability can also be evaluated as a series of 1-dimensional integral computations using recursive integration.

The utility of the methodology that is presented here is paramount when the random variables X_i have different distributions. This is because if the X_i are independent and identically distributed, then for the purpose of obtaining the conditional distribution and moments of a specific X_i, say, the information provided by the simple ordering in equation (1) is just equivalent to the information that X_i is the ith order statistic (the actual ordering of the $i-1$ variables less than X_i and the $n-i$ variables larger than X_i is irrelevant). In this case, the standard literature on order statistics (such as Arnold et al. (1992), Harter and Balakrishnan (1996), and David and Nagaraja (2003), for example) can be used to obtain the conditional information on X_i. However, when the random variables X_i are not identically distributed, then the simple ordering in equation (1) provides much more information than that X_i is simply the ith order statistic, and the methodology presented here allows all of that information to be utilized.

It may be the case that the ranking provided to the experimenter is incorrect, due to errors in its construction or simply perceived uncertainties. In fact, in section 6 of Chiarawongse et al. (2012) it is pointed out with respect to financial applications that "When an analyst offers a qualitative view but is uncertain about its validity, it is useful for the decision maker to be provided with a measure of confidence. This could take the form of a probability that the view is valid."

In this case of an "imperfect ranking", Chiarawongse et al. (2012) proposed the shrinkage model

$$P(\boldsymbol{X} \in C \mid \boldsymbol{\mathcal{S}} = S) = \kappa P(\boldsymbol{X} \in C \mid \boldsymbol{X} \in S) + (1-\kappa) P(\boldsymbol{X} \in C) \qquad (6)$$

where C is any event, $\boldsymbol{\mathcal{S}}$ is the "imperfect" ranking and S is the observed ranking. Equation (6) simply states that the probability of an event given an imperfect ranking is a convex combination of the probability with the observed ranking and the prior probability. The additional parameter κ represents the probability that the observed ranking is valid.

Notice that in this case the expectation of a variable can be expressed as a convex combination of the expectation with the observed ranking and that

without ranking information

$$E[g(\boldsymbol{X}) \mid \mathcal{S} = S] = \kappa E[g(\boldsymbol{X}) \mid \boldsymbol{X} \in S] + (1 - \kappa)E[g(\boldsymbol{X})]. \tag{7}$$

Also, it can be seen that once the expressions under the observed ranking are obtained, equations (6) and (7) follow readily without additional computational efforts, and can be used to estimate the expressions under the imperfect ranking. In what follows, we consider distribution and moment computations only for the observed ranking. The computations for the case of an imperfect ranking under this shrinkage model then follow naturally from these computations.

3. Implementation with recursive integration

In this section first some formulas are provided for the implementation of the methodology with recursive integration, and the algorithms are explicitly provided. Finally, a discussion is provided of a self-correction mechanism and computational times.

3.1. Formulas for the recursive integration

To evaluate equation (2) for the independent random variables case in section 2.1, B and A_1 can be evaluated according to the following recursive integration methodology. To evaluate B, the intermediate functions $b_1, \ldots, b_{n-1}$ can be sequentially evaluated, each with a one-dimensional integration. Let $b_0(z) = 1$ and for $i = 1, \ldots, n-1$ evaluate for each $z \in \Re$

$$b_i(z) = \int_{-\infty}^{z} b_{i-1}(x) f_i(x)\, dx \tag{8}$$

where f_i is the density of X_i. Then

$$B = \int_{-\infty}^{\infty} b_{n-1}(z) f_n(z)\, dz. \tag{9}$$

To evaluate A_1 in a similar manner, the intermediate functions $a_1, \ldots, a_{n-1}$ are sequentially evaluated. Here $a_0(z) = 1$ and for $i = 1, \ldots, n-1$, evaluate for each $z \in \Re$

$$a_i(z) = \int_{l_i}^{\max\{\min\{z, u_i\}, l_i\}} a_{i-1}(x) f_i(x)\, dx. \tag{10}$$

so that

$$A_1 = \int_{l_n}^{u_n} a_{n-1}(z) f_n(z)\, dz. \tag{11}$$

To evaluate equation (3) for the independent random variables case in section 2.1, B can be evaluated as above. To evaluate A_2, the intermediate functions $h_1, \ldots, h_{n-1}$ can be sequentially evaluated, each with a one-dimensional integration. Let $h_0(z) = 1$ and for $i = 1, \ldots, n-1$ evaluate for each $z \in \Re$

$$h_i(z) = \int_{-\infty}^{z} h_{i-1}(x) g_i(x) f_i(x)\, dx. \tag{12}$$

Then

$$A_2 = \int_{-\infty}^{\infty} h_{n-1}(z) g_n(z) f_n(z)\, dz. \tag{13}$$

To evaluate equation (2) for the multivariate normal case in section 2.2, B and A_1 can be evaluated according to the following recursive integration methodology. To evaluate B, the intermediate functions $b_1, \ldots, b_{n-1}$ can be sequentially evaluated, each with a two-dimensional integration complexity. Let $b_0(m, z) = 1$ and for $i = 1, \ldots, n-1$, evaluate for each $m, z \in \Re$

$$b_i(m, z) = \int_{-\infty}^{z} b_{i-1}(m, x) \phi_i(m, x)\, dx, \tag{14}$$

where ϕ_i is the density of a $N(\mu_i + \rho_i m, \sigma_i^2 - \rho_i^2)$ random variable. Then with ϕ as the standard normal density

$$B = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \phi(m) b_{n-1}(m, z) \phi_n(m, z)\, dz\, dm. \tag{15}$$

To evaluate A_1 in a similar manner with the intermediate functions $a_1, \ldots, a_{n-1}$, let $a_0(m, z) = 1$ and for $i = 1, \ldots, n-1$, evaluate for each $m, z \in \Re$

$$a_i(m, z) = \int_{l_i}^{\max\{\min\{z, u_i\}, l_i\}} a_{i-1}(m, x) \phi_i(m, x)\, dx. \tag{16}$$

Then

$$A_1 = \int_{m=-\infty}^{\infty} \int_{z=l_n}^{u_n} \phi(m) a_{n-1}(m, z) \phi_n(m, z)\, dz\, dm. \tag{17}$$

To evaluate equation (3), the formulas for evaluating A_2 with the intermediate functions $h_1, \ldots, h_{n-1}$ are $h_0(m, z) = 1$, and for $i = 1, \ldots, n-1$

$$h_i(m, z) = \int_{-\infty}^{z} h_{i-1}(m, x) g_i(x) \phi_i(m, x)\, dx \tag{18}$$

with

$$A_2 = \int_{m=-\infty}^{\infty} \int_{z=-\infty}^{\infty} \phi(m) h_{n-1}(m, z) g_n(z) \phi_n(m, z)\, dz\, dm. \tag{19}$$

3.2. Algorithms

The evaluations of equations (9), (11) and (13) are accomplished with a sequence of n 1-dimensional numerical integrations, with each integration being evaluated by a sum on a truncated-discretized real line. Algorithm 1 can be used to compute equation (13) with the integrals being performed with a first order Newton-Cotes formula (the trapezoidal rule). To compute equations (9) and (11), simply replace $g_i, i = 1, \ldots, n$ in the algorithm with appropriate indicator functions or with a constant function equal to one, respectively. R code is available from the authors to implement this algorithm.

The evaluations of equations (15), (17) and (19), require a sequence of 2-dimensional numerical integrations. Algorithm 2 can be used to compute (19) with the integrals being performed with the first order Newton-Cotes formula. To compute equations (15) and (17), simply replace $g_i, i = 1, \ldots, n$ in Algorithm 2 with appropriate indicator functions or with a constant function equal to one, respectively. Again, R code is available from the authors to implement this algorithm.

3.3. Self-correction mechanism

It is useful to point out that the conditional probability and the conditional expectation computations based on the recursive integration methodology possess a self-correction mechanisim. To see this, observe that the target value is of the form

$$f(A, B) = \frac{A}{B}$$

where A is either A_1 in equation (2) or A_2 in equation (3). Let $\hat{A}$ and $\hat{B}$ denote the computed values of A and B, respectively. The computed value of the target quantity is then $f(\hat{A}, \hat{B})$, and with a first order approximation

$$\begin{aligned} f(\hat{A}, \hat{B}) &\approx f(A, B) + f_1'(A, B)(\hat{A} - A) + f_2'(A, B)(\hat{B} - \hat{B}) \\ &= f(A, B) + \frac{1}{B}(\hat{A} - A) - \frac{A}{B^2}(\hat{B} - \hat{B}) \end{aligned} \tag{20}$$

Algorithm 1 Computation of A_2 in equation (13)

1: Assume n variables with ranking $X_1 \leq X_2 \cdots \leq X_n$.
2: Discretization grid size Δ with lower bound x_0, forming $N+1$ grid points

$$\{x_0, x_1, \ldots, x_N\},$$

where $x_j = x_{j-1} + \Delta$ for $j = 1, \ldots, N$.
3: Let $h_0(x_j) = 1$ for $j = 0, 1, \ldots, N$.
4: **for** $i = 1$ to n **do**
5: Let, for $j = 1, \ldots, N$,

$$\overline{h}_j = \frac{h_{i-1}(x_{j-1})g_i(x_{j-1})f_i(x_{j-1}) + h_{i-1}(x_j)g_i(x_j)f_i(x_j)}{2}.$$

6: Let, for $j = 1, \ldots, N$,

$$h_i(x_j) = \sum_{k=1}^{j} \overline{h}_k \Delta.$$

and let $h_i(x_0) = h_i(x_1)$.
7: **end for**
8: **return** $A_2 = h_n(x_N)$.

where f_1' and f_2' are the partial derivatives of f with respect to its first and second arguments. If we let $\varepsilon, \varepsilon_A$ and ε_B be the computational error of the target value, that of A and that of B, respectively, so that

$$\begin{aligned} f(\hat{A}, \hat{B}) &= f(A, B) + \varepsilon \\ \hat{A} &= A + \varepsilon_A \\ \hat{B} &= B + \varepsilon_B, \end{aligned}$$

then equation (20) implies that

$$\varepsilon \approx \frac{1}{B}\varepsilon_A - \frac{A}{B^2}\varepsilon_B. \tag{21}$$

Thus, when A and both B are positive, and when ε_A and ε_B have the same sign, the computational errors of A and B tend to cancel each other in producing the total error of the target value. Notice that for the applications in this paper, B is positive since it is a probability, and A is positive when it is a probability and may be positive when it is an expectation.

Algorithm 2 Computation of A_2 in equation (19)

1: Assume n variables with ranking $X_1 \leq X_2 \cdots \leq X_n$.
2: Discretization grid size δ with lower bound m_0, forming $M+1$ grid points

$$\{m_0, m_1, \ldots, m_M\},$$

where $m_l = m_{l-1} + \delta$ for $l = 1, \ldots, M$.
3: **for** $l = 0$ to M **do**
4: Discretization grid size Δ with lower bound x_0, forming $N+1$ grid points

$$\{x_0, x_1, \ldots, x_N\},$$

where $x_j = x_{j-1} + \Delta$ for $j = 1, \ldots, N$.
5: Let $h_0(m_l, x_j) = 1$ for $j = 0, 1, \ldots, N$
6: **for** $i = 1$ to n **do**
7: Let, for $j = 1, \ldots, N$,

$$\overline{h}_j = \frac{h_{i-1}(m_l, x_{j-1}) g_i(x_{j-1}) \phi_i(m_l, x_{j-1}) + h_{i-1}(m_l, x_j) g_i(x_j) \phi_i(m_l, x_j)}{2}.$$

8: Let, for $j = 1, \ldots, N$,

$$h_i(m_l, x_j) = \sum_{k=1}^{j} \overline{h}_k \Delta.$$

and let $h_i(m_l, x_0) = h_i(m_l, x_1)$.
9: **end for**
10: $\tilde{h}(m_l) = h_n(m_l, x_N)$.
11: **end for**
12: Let, for $l = 1, \ldots, M$,

$$\overline{h}_l = \frac{\phi(m_l)\tilde{h}(m_l) + \phi(m_{l-1})\tilde{h}(m_{l-1})}{2}.$$

13: **return** $A_2 = \sum_{l=1}^{M} \overline{h}_l \delta$.

This error cancellation effect or self-correction mechanism is strongest when the values of A and B are comparable, and when the values of ε_A and

ε_B are comparable. For the problems considered in this paper, the processes of computing $\hat{A}$ and $\hat{B}$ share some common numerical integration sequences, and hence ε_A and ε_B will tend to have the same sign. However, the level of the error cancellation depends upon the relative values of A and B, and the relative values of ε_A and ε_B. This is illustrated and further discussed in the following numerical examples.

This self-correction mechanism caused by the cancellation of the errors from the numerator and the denominator is a useful property of the implementation of the recursive integration methodology for the problems discussed in this paper. In fact, even in the case where the exact value of the denominator B might be known, according to (21) it may be better that $\hat{B}$ is computed and employed in evaluating the ratio since the self-correction mechanism applies. The estimate obtained by employing $\hat{B}$ can be interpreted as an estimate that has been formed by a legitimate discrete distribution induced by the discretization procedure. The resulting estimate approaches the true value when the discretization becomes finer. In the following examples it is shown that not taking advantage of this self-correction mechanism causes a significant increase in the error of the estimate if the discretization is not fine enough.

Some calculations are now presented to demonstrate the errors for problems with independent identically distributed random variables where the solutions are known. Specifically, consider the cases of $n = 101$ independent uniform $[0, 1]$ random variables or independent standard normal random variables. In both cases the cumulative distribution at three points of X_{70}, under the condition $X_1 \leq \ldots \leq X_{101}$, was evaluated by the recursive integration methodology for various grid sizes, and the computed values of the probability, the computational errors, and the computational times are shown in Tables 1 and 2 (note that the true values can be obtained from the cumulative distribution of a binomial distribution). The computations were implemented in R on a 64-bit Windows machine with an Intel Core i5-2500 3.30GHz CPU.

Next, the expectations of X_{17} and X_{51}, under the condition $X_1 \leq \ldots \leq X_{101}$, were evaluated by the recursive integration methodology for various grid sizes, and the computed values of the expectations, the computational errors, and the computational times are shown in Table 3 for the case of independent uniform $[0, 1]$ variables and in Table 4 for the case of independent standard normal variables. In the first case the true values of $E[X_{17}]$ and $E[X_{51}]$ are known to be $1/6$ and 0.5, while in the second case the true value

Table 1: The numerical errors and the computational times for the cumulative distribution function evaluated at three points for the 70^{th} order statistic of $n = 101$ independent $U[0, 1]$ random variables.

Grid size	True value	Comp. value	Error	Comp. time (seconds)	Comp. value With exact denominator	Error With exact denominator
0.01	0.03382186	0.09552326	6.170e-02	0.00	231.1455	$\gg 1$
	0.60791267	0.65772572	4.981e-02	0.00	1591.5334	$\gg 1$
	0.99628437	0.99005003	6.234e-03	0.00	2395.7061	$\gg 1$
0.001	0.03382186	0.03422536	4.035e-04	0.00	0.03739016	3.568e-03
	0.60791267	0.60800312	9.045e-05	0.02	0.66422472	5.631e-02
	0.99628437	0.99620053	8.385e-05	0.00	1.08831847	9.203e-02
0.0001	0.03382186	0.03382585	3.987e-06	0.06	0.03385576	3.390e-05
	0.60791267	0.60791339	7.206e-07	0.07	0.60845101	5.383e-04
	0.99628437	0.99628353	8.411e-07	0.08	0.99716462	8.802e-04
0.00001	0.03382186	0.03382190	3.982e-08	1.36	0.03382220	3.387e-07
	0.60791267	0.60791268	7.037e-09	1.20	0.60791805	5.380e-06
	0.99628437	0.99628437	8.414e-09	1.42	0.99629317	8.797e-06

Table 2: The numerical errors and the computational times for the cumulative distribution function evaluated at three points for the 70^{th} order statistic of $n = 101$ independent standard normal random variables.

Grid size	True value	Comp. value	Error	Comp. time (seconds)	Comp. value With exact denominator	Error With exact denominator
0.01	0.03382186	0.03766693	3.845e-03	0.05	0.08004251	4.622e-02
	0.60791267	0.61814333	1.023e-02	0.05	1.31355904	7.056e-01
	0.99628437	0.99663459	3.502e-04	0.03	2.11785572	1.122e+00
0.001	0.03382186	0.03367669	1.452e-04	0.39	0.03393352	1.117e-04
	0.60791267	0.60697199	9.407e-04	0.36	0.61160101	3.688e-03
	0.99628437	0.99632079	3.642e-05	0.39	1.00391914	7.635e-03
0.0001	0.03382186	0.03379485	2.701e-05	4.38	0.03379548	2.638e-05
	0.60791267	0.60791453	1.860e-06	4.36	0.60792588	1.320e-05
	0.99628437	0.99628272	1.657e-06	4.39	0.99630131	1.694e-05
0.00001	0.03382186	0.03382366	1.796e-06	38.64	0.03382172	1.365e-07
	0.60791267	0.60791215	5.155e-07	36.95	0.60787742	3.525e-05
	0.99628437	0.99628432	5.078e-08	38.14	0.99622740	5.697e-05

of $E[X_{17}]$ is unknown but is about $\Phi^{-1}(1/6) \approx 0.97$, and the true value of $E[X_{51}]$ is known to be zero.

In these tables the second column shows the known true values, while the third column shows the estimates from the proposed methodology. The fourth and fifth columns show the errors (from the true values) and the computational time of the proposed methodology. Also, both for the independent uniform $[0, 1]$ variables and the independent standard normal variables the

Table 3: The numerical errors and the computational times for the expectations of $X_{(17)}$ and $X_{(51)}$ of $n = 101$ independent $U[0, 1]$ random variables.

Grid size	True value	Comp. value	Error	Comp. time (seconds)	Comp. value With exact denominator	Error With exact denominator
0.01	1/6	0.14369985	2.297e-02	0.00	347.7224	$\gg 1$
	0.5	0.48606226	1.394e-02	0.00	1176.1651	$\gg 1$
0.001	1/6	0.16661772	4.894e-05	0.00	0.18202474	1.536e-02
	0.5	0.49997063	2.937e-05	0.00	0.54620256	4.620e-02
0.0001	1/6	0.16666624	4.272e-07	0.05	0.16681363	1.470e-04
	0.5	0.49999974	2.563e-07	0.06	0.50044193	4.419e-04
0.00001	1/6	0.16666666	4.215e-09	0.92	0.16666814	1.469e-06
	0.5	0.50000000	2.529e-09	0.89	0.50000442	4.416e-06

Table 4: The numerical errors and the computational times for the expectations of $X_{(17)}$ and $X_{(51)}$ of $n = 101$ independent standard normal random variables.

Grid size	True value	Comp. value	Error	Comp. time (seconds)	Comp. value With exact denominator	Error With exact denominator
0.01	NA	-0.9672	NA	0.04	-2.0552	NA
	0.0000	0.0000	1.397e-14	0.05	0.0000	2.970e-14
0.001	NA	-0.9779	NA	0.38	-0.9854	NA
	0.0000	0.0000	1.734e-16	0.39	0.0000	1.747e-16
0.0001	NA	-0.9780	NA	4.33	-0.9781	NA
	0.0000	0.0000	9.826e-18	4.25	0.0000	9.826e-18
0.00001	NA	-0.9780	NA	38.55	-0.9780	NA
	0.0000	0.0000	1.281e-17	37.22	0.0000	1.281e-17

true values of the denominators B in equations (2) and (3) are known to be $1/n!$. The sixth and seventh columns then show the estimates and errors when the true values of the denominators are employed and the recursive integration is used only for computing the numerators A_1 and A_2.

First of all, it can be seen from these tables that these calculations which involve 101 successive one-dimensional numerical integrations attain a small error with a very reasonable computation time. In fact, with the potential optimization of the coding on a low level computer programming language such as C++ or Java, the computation can be expected to be accelerated even more.

Furthermore, special attention should be given to the results in the sixth and seventh columns. In the sixth column the estimates computed by employing the true values for the denominators do not take advantage of the self-correction mechanism feature of the proposed methodology given in equa-

tion (21). It can be seen that when the discretization grid sizes are not very small, the errors can become so large that the estimates are unreasonable. Specifically, some estimates for the conditional probabilities in Table 1 and in Table 2 when the grid sizes are 0.01 and 0.001 are much greater than one. Note that such unreasonable values of the estimates do not occur when the methodology is applied appropriately to both the numerators and the denominators and the self-correction mechanism applies (as shown in the third column of each table).

In almost all cases in Table 1 and in Table 2 the errors from the proposed methodology with self-correction mechanism are much smaller than the corresponding values without the feature. The exceptions are row 4, row 7 and row 10 in Table 2. The errors with the self-correction mechanism are not smaller than those without the feature in these cases, although they are close. One explanation for this is that the distribution function is evaluated at a low quantile, so that A is much smaller than B in equation (21). Furthermore, at this low quantile $\hat{A}$ and $\hat{B}$ do not share a lot of common integration sequences, so that ε_A and ε_B may be quite different. Consequently, the error cancellation in equation (21) does not apply substantially, although in these exceptional cases the differences between the errors are very small.

In Table 3 the estimates with the self-correction mechanism are much more accurate than those without the self-correction mechanism. According to equation (21), this is because the value of A is comparable to that of B, causing a strong error cancellation. In Table 4 when the true value of $E[X_{51}]$ is known to be zero, A is zero. According to equation (21) there is therefore no error cancellation, and consequently the errors of the estimates with or without self-correction mechanism are quite similar. In the case where the true value of $E[X_{17}]$ is unknown it can be observed that the estimate converges to a certain value as the grid size becomes smaller. The estimate with the self-correction mechanism seems to converge more quickly than the estimate without this feature.

4. Examples

Three examples are presented in this section. The first is a heathcare example where the random variables are taken to have independent gamma distributions with equal shape parameters but different scale parameters. The second is a soil contamination example where the random variables are taken to have a multivariate normal distribution with different means but

equal variances and covariances. It is shown how information on both reinforcing rankings and opposing rankings affects the distributions, expectations, variances, and covariances of the variables. The third example concerns portfolio selection in finance.

4.1. Healthcare

Suppose that the levels X_i of a medical condition of $n = 5$ patients are of interest, but that they cannot be directly measured. However, the levels can be modelled with a gamma distribution

$$f(x; 3, \theta) = \frac{1}{2\theta^3} x^2 e^{-x/\theta}$$

with a shape parameter $k = 3$ and with a scale parameter θ that depends upon some covariate values of the patients. Specifically, suppose that the five patients have scale parameters $\theta_i = i$, $1 \leq i \leq 5$, so that based upon these distributions the prior expectations and variances are $E(X_i) = 3i$ and $Var(X_i) = 3i^2$, $1 \leq i \leq 5$. It should also be noted that the variables X_i are modelled to be independent.

Now suppose that an ancillary measurement becomes available for the five patients that provides the information that

$$X_1 \leq X_2 \leq X_3 \leq X_4 \leq X_5$$

(or equivalently, this same ranking with strict inequalities). This is a reinforcing ranking since it matches the ranking of the prior expectations of the X_i. It is interesting to note that under the prior distributions this reinforcing ranking has a probability of 0.107 (this is B in equations (2) and (3)). Using the recursive integration techniques discussed in section 2, the conditional expectations, standard deviations, and correlations of the X_i (conditionally the X_i are no longer independent) are shown in Table 3.

It can first be noted that with this reinforcing ranking the conditional expectations of the X_i have maintained their ordering but are now more spread out than the prior expectations. Furthermore, the conditional standard deviations are each smaller than the prior standard deviations. Also, it can be seen that the correlations are largest for adjacent variables.

In addition, suppose that it has been decided that urgent corrective action needs to be taken whenever the level of this deterioration condition is less than 5. Table 3 also shows how these probabilites change under the

Table 5: Example 1 - Medical conditions of five patients.

Reactor	Expectation	Standard Deviation	Correlation Matrix					$P(X_i \leq 5)$
Prior								
1	3	1.73	1	0	0	0	0	0.876
2	6	3.46		1	0	0	0	0.456
3	9	5.20			1	0	0	0.234
4	12	6.93				1	0	0.132
5	15	8.66					1	0.080
Reinforcing ranking $X_1 \leq X_2 \leq X_3 \leq X_4 \leq X_5$								
1	2.32	1.22	1	0.36	0.17	0.08	0.03	0.968
2	4.89	2.03		1	0.49	0.24	0.10	0.578
3	8.19	3.05			1	0.52	0.23	0.132
4	12.86	4.63				1	0.45	0.010
5	21.41	8.34					1	0.000
Opposing ranking $X_5 \leq X_4 \leq X_3 \leq X_2 \leq X_1$								
1	7.99	2.18	1	0.82	0.69	0.57	0.40	0.064
2	6.68	1.86		1	0.85	0.70	0.49	0.184
3	5.64	1.69			1	0.82	0.58	0.382
4	4.58	1.56				1	0.71	0.641
5	3.24	1.45					1	0.881

knowledge provided by the reinforcing ranking. It can be seen that the probabilities that urgent corrective action needs to be taken become larger for patients 1 and 2, and become smaller for patients 3, 4 and 5.

Now consider the opposing ranking

$$X_5 \leq X_4 \leq X_3 \leq X_2 \leq X_1$$

which is completely opposite to the ranking of the prior expectations of the X_i. In fact, under the prior distributions this opposing ranking has a very small probability of 0.00003. Under this opposing ranking the conditional expectations, standard deviations, and correlations of the X_i are also shown in Table 3.

It can be seen that with this opposing ranking the order of the conditional expectations has switched to match this ranking, and that the conditional expectations are less spread out than the prior expectations. The conditional standard deviations are also much smaller than the prior standard deviations, and their order also matches the opposing ranking. Again, the correlations

are largest for the adjacent variables, and they are all much larger than the correlations for the reinforcing ranking. Also, the probabilities that urgent corrective action needs to be taken are now ordered to match the opposing ranking.

In summary, this example illustrates how knowledge of the ranking can result in important changes in the distributions and moments of the variables, which will be important information for practitioners.

4.2. Hazardous waste sites

Suppose that based upon knowledge of polluting activities, scientists originally model the unknown toxic contamination levels X_i at $n = 6$ locations with a multivariate normal distribution with means $\boldsymbol{\mu} = (10, 10, 12, 15, 18, 20)$, standard deviations all equal to 3, and correlations all equal to 0.4. Then suppose that subsequently surface features indicate the reinforcing ranking

$$X_1 \leq X_2 \leq X_3 \leq X_4 \leq X_5 \leq X_6.$$

Under the prior distribution this reinforcing ranking has a probability of 0.332. In this case the conditional expectations, standard deviations, and correlations of the toxic contamination levels are shown in Table 4.

It can be seen that with this reinforcing ranking the conditional expectations are quite similar to the prior expectations, although for location 1 the expectation has decreased from 10 to 8.21, while for location 6 the expectation has increased from 20 to 20.94. The standard deviations have all decreased and are all fairly similar, while the correlations have increased and are largest for the adjacent variables.

Also, suppose that it has been decided that decontamination needs to be taken whenever the toxic contamination level is larger than 18. It can be seen from Table 2 that the reinforcing ranking has increased the probability that decontamination needs to be taken at location 6 from 0.748 to 0.865, while these probabilities have fallen at the other 5 locations.

Now suppose that subsequently surface features indicate the partially opposing ranking

$$X_1 \leq X_2 \leq X_5 \leq X_4 \leq X_3 \leq X_6$$

where the ordering of the toxic contamination levels at locations 3, 4, and 5 is opposite to their prior expectations. This partially opposing ranking has a probability of 0.007 under the prior distribution.

Table 6: Example 2 - Toxic contamination levels at six locations.

Location	Expectation	Standard Deviation	Correlation Matrix						$P(X_i \geq 18)$
Prior									
1	10	3	1	0.40	0.40	0.40	0.40	0.40	0.004
2	10	3		1	0.40	0.40	0.40	0.40	0.004
3	12	3			1	0.40	0.40	0.40	0.023
4	15	3				1	0.40	0.40	0.159
5	18	3					1	0.40	0.500
6	20	3						1	0.748
Reinforcing ranking $X_1 \leq X_2 \leq X_3 \leq X_4 \leq X_5 \leq X_6$									
1	8.21	2.59	1	0.79	0.67	0.59	0.56	0.52	0.000
2	10.28	2.45		1	0.78	0.65	0.60	0.55	0.001
3	12.47	2.45			1	0.74	0.62	0.56	0.012
4	15.12	2.50				1	0.71	0.58	0.125
5	17.97	2.54					1	0.69	0.494
6	20.95	2.71						1	0.863
Partially opposing ranking $X_1 \leq X_2 \leq X_5 \leq X_4 \leq X_3 \leq X_6$									
1	8.52	2.64	1	0.76	0.60	0.61	0.62	0.48	0.000
2	10.88	2.52		1	0.65	0.68	0.69	0.51	0.002
3	16.13	2.36			1	0.92	0.86	0.60	0.214
4	15.08	2.31				1	0.92	0.59	0.104
5	14.05	2.34					1	0.58	0.046
6	20.34	2.83						1	0.794

Table 4 shows that the conditional expectations are now ordered in the same way as this ranking. In addition, the standard deviations have decreased, but now there are very high correlations of 0.92 between locations 3 and 4, and between locations 4 and 5, which are the locations where the ranking contradicts the prior expectations. The probabilities that decontamination needs to be taken are now ordered in the same way as the partially opposing ranking, and specifically the probability at location 3 has risen from 0.023 to 0.214.

As with Example 1, this example illustrates how the different rankings can result in important changes in the distributions and moments of the variables, and the methodology presented in this paper allows practitioners to calculate those changes.

4.3. Portfolio Selection

The celebrated mean-variance portfolio selection model requires two sets of moments (the means and the variance-covariance matrix) of the returns on N assets in order to recommend the proportions of capital, or the portfolio weights, to be invested in these N assets. The objective is to optimize the risk-return trade-off of the entire portfolio.

The problem can be expressed as

$$\max_{\boldsymbol{w}_t} \quad \boldsymbol{\mu}_t^\top \boldsymbol{w}_t - \tfrac{\gamma}{2} \boldsymbol{w}_t^\top \boldsymbol{\Sigma}_t \boldsymbol{w}_t \tag{22}$$

where $\boldsymbol{\mu}_t$ and $\boldsymbol{\Sigma}_t$ are the mean vector and the variance-covariance matrix of the return (in excess of risk free) at time t, γ is a parameter representing risk aversion of the investor which we set equal to 1 in this analysis, and $\boldsymbol{w}_t$ is the vector of decision variables that represent the portfolio weights at time t. The objective function in (22) is the certainty-equivalent return which is a utility function interpreted as the return that is penalized by risk. The solution to the optimization problem is

$$\boldsymbol{w}_t^* = \frac{\boldsymbol{\Sigma}_t^{-1} \boldsymbol{\mu}_t}{\mathbf{1}^\top \boldsymbol{\Sigma}_t^{-1} \boldsymbol{\mu}_t} \tag{23}$$

(see, for example, DeMiguel et al. (2009) for more details). The main difference between one portfolio selection strategy and another is how the parameter estimates $\hat{\boldsymbol{\mu}}_t$ and $\hat{\boldsymbol{\Sigma}}_t$ are obtained.

In this example the estimate $\hat{\boldsymbol{\mu}}_t$ is based on the rank constrained statistical estimates proposed by Chiarawongse et al. (2012), where the estimates are obtained by a Markov chain Monte Carlo. Here the calculations are replaced by the recursive integration methodology presented in this paper in order to obtain more accurate estimates. Furthermore, in Chiarawongse et al. (2012) the authors perform their experiments based on simulated data sets, whereas in this example a real data set is considered instead of ten industry monthly asset returns during the period of 01/1999 - 06/2014 obtained from Kenneth French's web site.

Our objective is the same as that in Chiarawongse et al. (2012), which is to demonstrate the potential benefit of adopting a rank constrained statistical estimate. Two experiments are performed, one using *prior* parameter estimates $\tilde{\boldsymbol{\mu}}$ and the other using rank constrained statistical estimates $\hat{\boldsymbol{\mu}}$. The performances of the two models are then compared.

In the analysis that follows the rolling-sample approach appearing in DeMiguel et al. (2009) is employed. An estimation window of length 66 months is fixed. Then, starting from $t = 66$ and applying the capital asset pricing model, the data from month $t - 65$ to month t are used to estimate the *prior* parameters $(\tilde{\boldsymbol{\mu}}_t, \tilde{\boldsymbol{\Sigma}}_t)$. These are then applied to (23) to obtain the weight vector $\tilde{\boldsymbol{w}}_t$ for $t = 66, 67, \ldots, 185$. The weight at time t is applied to the out-of-sample returns at $t + 1$ and summed across assets to provide 120 out-of-sample portfolio returns $\tilde{r}_t, t = 67, 68, \ldots, 186$. Subsequently, the 120 portfolio returns are then divided into 10 twelve-month periods. In each period a certainty equivalent return is estimated as

$$\tilde{u}_i = \bar{\tilde{r}}_i - \tilde{s}_i^2/2, \quad i = 1, \ldots, 10,$$

where $\bar{\tilde{r}}_i$ and $\tilde{s}_i^2$ are the averages and the sample variances of the $\tilde{r}_t$ in period i. These certainty equivalent returns are the performance measurements of the prior model which were compared with those obtained from the following rank constrained statistical counterpart.

The experiment was repeated with $(\hat{\boldsymbol{\mu}}_t, \tilde{\boldsymbol{\Sigma}}_t)$ in (23) to obtain different portfolio weights, where $\hat{\boldsymbol{\mu}}_t$ is the rank constrained estimates based on Chiarawongse et al. (2012) and $\tilde{\boldsymbol{\Sigma}}_t$ is the same as in the prior model. To estimate $\hat{\boldsymbol{\mu}}_t$, $(\tilde{\boldsymbol{\mu}}_t, \tilde{\boldsymbol{\Sigma}}_t)$ are first estimated as in the prior model, and then for each $t = 66, 67, \ldots, 185$, $\hat{\boldsymbol{\mu}}_t$ is estimated as the expectation of a multivariate normal distribution parameter $(\tilde{\boldsymbol{\mu}}_t, 0.1\tilde{\boldsymbol{\Sigma}}_t)$ conditioned on the ranking obtained from that of the ten industry returns at time $t+1$. This can be interpreted as a process to improve the quality of prior mean estimates by a one-step ahead ranking. Note that in practice the ranking would usually be obtained from another database of investor views. However, this approach affords an understanding of the potiential benefit of the one-step ahead ranking methodology. As with the prior model, the portfolio returns $\hat{r}_t, t = 67, 68, \ldots, 186$ are computed for the rank constrained model, and finally the certainty equivalent returns $\hat{u}_i$ are computed for the 10 twelve-month periods.

The certainty equivalence returns for the ten periods obtained from the prior model and the rank constrained model are plotted against each other in Figure 1. The certainty equivalence returns from the rank constrained model outperform those from the prior model consistently in every period. This analysis indicates the potential benefit of the rank constrained model and demonstrates the advantages in this area of study available from employing the recursive integration methodology discussed in this paper.

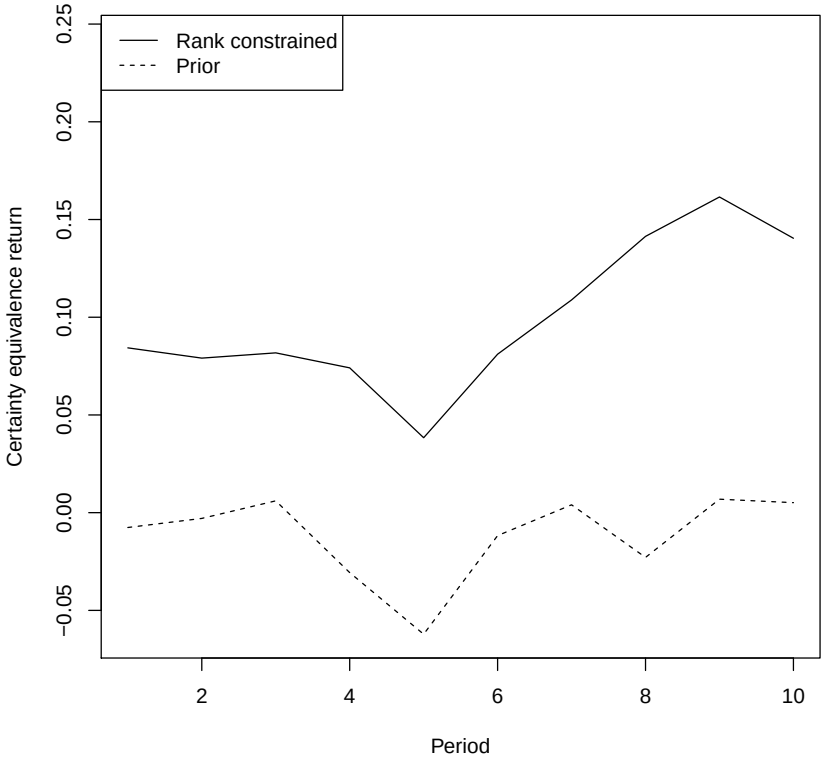

Figure 1: Certainty equivalence returns of the 10 twelve-month period from the rank constrained model (solid line) and the prior model (dashed line). The returns from the rank constrained model consistently outperform those from the prior model.

5. Conclusion

This paper has addressed the situation where a set of variables have independent specified prior distributions, and where some information becomes available on the ordering of the variables. This is a common phenomenon which has been discussed and utilized in other statistical methodologies such as ranked set sampling. In this paper it is shown how updated distributions and moments of the variables can be calculated conditional on the knowledge provided by the ranking. It has been shown how the technique of recursive integration can be used to perform these calculations in a straightforward

manner as a series of one-dimensional integral computations regardless of the number of variables.

For these particular problems of conditional probability and conditional expectation computations, it has been demonstrated that the errors in the numerators and the denominators can partially cancel each other, providing a self-correction mechanism that improves the accuracy of the recursive integration methodology.

The methodology presented in this paper has been implemented for a simple ordering of the variables. The methodology has also been generalized to variables with a multivariate normal distribution with a product correlation structure. In principle, the methodology can be extended to more complicated orderings such as an umbrella ordering or tree orderings, which are topics planned for future research.

Examples have been presented which illustrate how different kinds of rankings, such as reinforcing rankings and opposing rankings, can have different and substantial effects on the distributions, expectations, standard deviations, and correlations of the variables. This can be valuable information for practitioners, and the methodology presented in this paper allows this information to be obtained. Finally, the methodology has been applied to a decision problem in portfolio selection with ranking information, where it has been shown to provide a potential benefit. R code is available from the authors to replicate the tables and examples in this paper, and to implement the algorithms discussed

Acknowledgements

We would like to thank the reviewers whose comments have substantially improved this work. This work was supported by the Thailand Research Fund and Chulalongkorn University [RSA5780005].

References

Ali, A., Meilă, M., 2012. Experiments with Kemeny ranking: What works when? Mathematical Social Sciences 64, 28–40.

Arnold, B. C., Balakrishnan, N., Nagaraja, H. N., 1992. A First Course in Order Statistics. Wiley, New Jersey.

Chen, Z., Bai, Z., Sinha, B. K., 2004. Ranked Set Sampling. Springer, New York.

Chiarawongse, A., Kiatsupaibul, S., Tirapat, S., Roy, B. V., 2012. Portfolio selection with qualitative input. Journal of Banking and Finance 36, 489–496.

David, H. A., Nagaraja, H. N., 2003. Order Statistics, 3rd Edition. Wiley, New Jersey.

DeMiguel, V., Garlappi, L., Uppal, R., 2009. Optimal versus naive diversification: how inefficient is the $1/n$ portfolio strategy? The Review of Financial Studies 22, 1915–1953.

Gaur, A., Mahajan, K. K., Arora, S., 2012. New nonparametric tests for testing homogeneity of scale parameters against umbrella alternative. Statistics and Probability Letters 82, 1681–1689.

Hans, C., Dunson, D. B., 2005. Bayesian inferences on umbrella orderings. Biometrics 64 (4), 1018–1026.

Harter, H. L., Balakrishnan, N., 1996. C.R.C. Handbook of Tables for the Use of Order Statistics in Estimation. CRC Press Inc., Boca Raton.

Hayter, A. J., 2006. Recursive integration methodologies with statistical applications. Journal of Statistical Planning and Inference 136, 2284–2296.

Hayter, A. J., Liu, W., 1996. A note on the calculation of $\mathrm{pr}(x_1 < \cdots < x_k)$. The American Statistician 50 (4), 365.

Kemeny, J. L., Snell, J. G., 1962. Mathematical Models in the Social Sciences. Blaisdell, New York.

Khachiyan, L. G., 1989. The problem of computing the volume of polytopes is NP-hard. Uspekhi Mat. Nauk 44 (3), 199–200.

Kiatsupaibul, S., Smith, R. L., Zabinsky, Z. B., 2011. An analysis of a variation of hit-and-run for uniform sampling from general regions. ACM Transactions on Modeling and Computer Simulation 21 (3), Article number 16.

Lovász, L., 1999. Hit-and-run mixes fast. Mathematical Programming 86, 443–461.

Lovász, L., Vempala, S., 2006. Fast algorithms for logconcave functions: Sampling, rounding, integration and optimization. In: Proc. of the 47th IEEE Symposium on Foundations of Computer Science (FOCS 06). pp. 57–68.

McIntyre, G. A., 1952. A method for unbiased selective sampling, using ranked sets. Australian Journal of Agricultural Research 3, 385–390.

Milgrom, P., Roberts, J., 1994. Comparing equilibrium. American Economic Review 84 (3), 441–459.

Milgrom, P., Shannon, C., 1994. Monotone comparative statics. Econometrica 62 (1), 157–180.

Nakas, C. T., Alonzo, T. A., 2007. ROC graphs for assessing the ability of a diagnostic marker to detect three disease classes with an umbrella ordering. Biometrics 63, 603–609.

Patil, G. P., 2002. Ranked set sampling. In: Encyclopedia of Environmetrics. Vol. 3. John Wiley & Sons, Ltd., Chichester, pp. 1684–1690.

Singh, P., Liu, W., 2006. A test against an umbrella ororder alternative. Computational Statistics & Data Analysis 51, 1957–1964.

Smith, R. L., 1984. Efficient Monte Carlo procedures for generating points uniformly distributed over bounded regions. Operations Research 32, 1296–1308.

Topkis, D. M., 1998. Supermodularity and Complementarity. Princeton University Press.

Wolfe, D. A., 2004. Ranked set sampling: an approach to more efficient data collection. Statistical Science 19, 636–643.

Young, P., 1995. Optimal voting rules. Journal of Economic Perspectives 9, 51–64.