Monte Carlo Methods and Appl., Vol. 13, No. 2 (2007), pp. 219–225
DOI 10.1515 / MCMA.2007.007

Integral Evaluation Using the Δ^2-distribution. Simulation and Illustration

Trifon I. Missov

Abstract. The Δ^2-distribution is a multivariate distribution, which plays an important role in variance reduction of Monte Carlo integral evaluation. Selecting the nodes of random cubature formulae according to Δ^2 ensures an unbiased and efficient estimate of the studied integral regardless of the region it is solved over. The Δ^2 distribution is also relevant in problems such as separating errors in regression analysis and constructing D-optimal designs in multidimensional regions. Inefficient simulation of Δ^2 prevented the application of the underlying theory in real problems. Ermakov and Missov, [3], proposed an algorithm which combines all rejection, inversion, and mixture techniques. Its complexity allows simulating Δ^2 vectors of big lengths. Moreover, it works in the most general settings of the problem of integral evaluation. This article presents a modification of the simulation algorithm as well as its illustration for a popular integral in Reliability Theory.

Keywords. Variance reduction, Δ^2-distribution, Reliability Theory .

AMS classification. 65C05, 65C30 .

1. The Δ^2-distribution. A Simulation Insight

Consider a set of elements $X = \{x\}$, on which a σ-algebra A and a σ-finite measure μ are introduced. Note that μ can be discrete. We designate $\mu^n = \bigotimes_{k=1}^{n} \mu$ and assume that there exist n linearly independent on the support of the measure μ functions $\varphi_1(x), \varphi_2(x), \ldots, \varphi_n(x) \in L^2(\mu)$. The function

$$\Delta^2(Q) = \frac{1}{n!} \left(\det ||\varphi_i(x_j)|| _{i,j=1}^{n}\right)^2 , \; Q = (x_1, \ldots, x_n) \tag{1.1}$$

is called the density of the Δ^2-distribution with respect to the measure μ^n. The properties of $\Delta^2(Q)$ are discussed in depth by Ermakov, [2]. One of the important applications of $\Delta^2(Q)$ is associated with random cubature formula of the type:

$$\varkappa_n[f] = \sum_{i=1}^{n} A_i f(x_i) \approx \int\limits_X \varphi_1(x) f(x) \mu(dx) = I[f],$$

in which x_i, $A_i(x_1, \ldots, x_n)$ are random variables in X, and φ_1, f are given functions such that the product $\varphi_1 f$ is μ-integrable.
If we choose the nodes $x_1, \ldots, x_n$ according to the Δ^2-distribution and denote $\Delta(f; Q) = \det ||f(x_j), \varphi_2(x_j), \ldots, \varphi_n(x_j)|| _{j=1}^{n}$, $\Delta(Q) = \Delta(\varphi_1, Q)$, then

$$\varkappa_n[f] = \frac{\Delta(f; Q)}{\Delta(Q)} \tag{1.2}$$

will be an unbiased estimate of $I[f]$ with a variance

$$\text{Var}\, \varkappa_n[f] \leq \int \left[f(x) - \sum_{i=1}^{n} (f, \varphi_i) \varphi_i(x) \right]^2 \mu(dx).$$

The equality sign holds for regular systems $\{\varphi_i\}_{i=1}^{n}$. Note that this result is valid for μ-linearly independent functions $\varphi_1(x), \ldots, \varphi_n(x)$ (Ermakov, [2]). The orthogonality assumption is induced primarily to simplify calculations as a Gramm–Schmidt orthogonalization of $\{\varphi_i\}_{i=1}^{n}$ will not affect the determinant in (1.1).
The simulation method proposed in [3] and [8] works in the most general settings of the problem when X is a region in R^s, the functions $\varphi_i(x)\,, i = \overline{1, n}$ are bounded in X, and μ is the Lebesgue measure (or equivalent to it). It studies $\Delta^2(Q)$ as a product of n conditional densities. Each one of them represents a composition of other distribution densities. Moreover, it turns out compositions' coefficients can be derived iteratively one from another. In addition, the distributions themselves in each of the compositions can be simulated using either the rejection or the inversion method. The results are summarized in the following theorem (Ermakov and Missov, [3]):

Theorem 1.1. *The k-th conditional density $p\,(i_k | i_1 \ldots i_{k-1}), k = \overline{2, m}$ can be represented in the following way:*

$$\sum_{1 \leq i_1 < \ldots, i_k \leq n} \frac{\sum\limits_{j=1}^{k} \left(c^{(k)}_{\{i_1, i_2, \ldots, i_k\} \setminus \{i_j\}} \right)^2}{(n-k+1) \sum\limits_{1 \leq j_1 < \ldots, j_{k-1} \leq n} \left(c^{(k)}_{j_1 j_2 \ldots j_{k-1}} \right)^2}$$

$$\frac{\left(\sum\limits_{m=1}^{k}(-1)^{m+1}c^{(k)}_{\{i_1,i_2,\ldots,i_k\}\setminus\{i_m\}}\varphi_{i_m}(x_k)\right)^2}{\sum\limits_{j=1}^{k}\left(c^{(k)}_{\{i_1,i_2,\ldots,i_k\}\setminus\{i_j\}}\right)^2}, \tag{1.3}$$

where $c^{(1)}=1$ *and*

$$c^{(k)}_{\{i_1,i_2,\ldots,i_k\}\setminus\{i_j\}}=\sum_{s=1,s\neq j}^{k}(-1)^{s+1}c^{(k-1)}_{\{i_1,\ldots,i_k\}\setminus\{i_j,i_s\}}\varphi_{i_s}(x_{k-1}) \tag{1.4}$$

Note that (1.3) is a mixture of distributions. The first ratio under the sum assigns its coefficients, whereas the second one represents the distributions themselves. Coefficients (1.4) are calculated iteratively one from another.

The complexity of this simulation algorithm is (Ermakov and Missov, [3]):

$$C=O(n2^{n-1})\left(\max_{x_k\in X,i=\overline{1,n}}|\varphi_i(x_k)|\right)^2 \tag{1.5}$$

2. An Application of Δ^2 Multiple Integral Calculation

Let us implement the above simulation proposal and, using interpolation-cubature formulae, evaluate the following five-dimensional ($s=5$) integral:

$$\int\limits_{D_5}\frac{e^{-xyzst}}{1+x^2+y^2+z^2+s^2+t^2}\,dxdydzdsdt, \tag{2.1}$$

where D_5 is the 5-dimensional hypercube.

It plays an important role in Reliability Theory and its calculation is problematic as neither pure numerical methods nor Monte Carlo estimation work efficiently for $s\geq 2$. Let us first briefly focus our attention on the simulation algorithm itself.

Algorithm 2.1. Simulation of the Δ^2-Distribution

1. Consider the first (conditional) density

$$p_1(x_1) = \frac{1}{n}\sum_{i=1}^{n}\varphi_i^2(x_1) \tag{2.2}$$

Using a random number $\alpha_1 \in [0,1]$, we generate the discrete distribution

$$\begin{pmatrix} 1 & 2 & \dots & n \\ 1/n & 1/n & \dots & 1/n \end{pmatrix},$$

which determines composition's coefficients in (2.2). Having obtained a number j, we simulate by rejection $\varphi_j^2(x_1)$. Thus, we get x_1. Note that $c^{(1)} = 1$.

2. First, we calculate the coefficients $c_j^{(2)}$, $j = \overline{1,n}$ using (1.4):

$$c_j^{(2)} = 1.\varphi_j(x_1) = \varphi_j(x_1)$$

Then we study the second conditional density:

$$p_2(x_2|x_1) = \sum_{1\le i_1<i_2\le n} \frac{(c_{i_1}^{(2)})^2 + (c_{i_2}^{(2)})^2}{(n-1)\sum\limits_{l=1}^{n}(c_l^{(2)})^2} \frac{c_{i_1}^{(2)}\varphi_{i_2}(x_2) + c_{i_2}^{(2)}\varphi_{i_1}(x_2)}{(c_{i_1}^{(2)})^2 + (c_{i_2}^{(2)})^2}$$

Using $\alpha_2 \in [0,1]$ we simulate the respective discrete distribution

$$\left\{ \begin{array}{c} (i_1,i_2) \\ \frac{(c_{i_1}^{(2)})^2+(c_{i_2}^{(2)})^2}{(n-1)\sum\limits_{l=1}^{n}(c_l^{(2)})^2} \end{array} \right\}_{1\le i_1<i_2\le n}$$

and, having obtained a pair of indices (i_1, i_2), we simulate by rejection this distribution in the composition, which corresponds to these indices. The bounding constant equals

$$(|c_{i_1}^{(2)}| + |c_{i_2}^{(2)}|)^2 \left(\max_{x\in X,\, i=\overline{1,n}} |\varphi_i(x)| \right)^2$$

Thus, we get x_2.

3. For $k = \overline{3,n}$:

1. We calculate $c^{(k)}_{i_1,\ldots,i_{k-1}}$ according to (1.4).
2. We simulate the discrete distribution which determines the composition's coefficients.
3. Having uniquely determined $i_1,\ldots,i_k$, we simulate by rejection that distribution in the composition, which corresponds to these indices. Thus, we get x_k.

Taking advantage of Algorithm 1 we can simulate a Δ^2-distributed vector and construct an interpolation cubature formula (1.2) for the integral (2.1). The estimate for (2.1) is 12.851. The only question we have to answer is what value of n to choose in practice. Note that n is the length of the Δ^2-distributed vector we simulate, as well as the number of functions in the orthonormal set $\varphi_i(x)$, $i = \overline{1,n}$. If we want to attain accuracy up to the third decimal place, taking $n = 10$ is sufficient.
In such a case we choose an orthonormal polynomial system:

$$\varphi_1(x) = \frac{\sqrt{2}}{8};\ \ \varphi_2(x) = \frac{\sqrt{6}}{8}x;\ \ \varphi_3(x) = \frac{\sqrt{6}}{8}y;\ \ \varphi_4(x) = \frac{\sqrt{6}}{8}z;$$

$$\varphi_5(x) = \frac{\sqrt{6}}{8}s; \qquad \varphi_6(x) = \frac{\sqrt{6}}{8}t; \qquad \varphi_7(x) = \frac{\sqrt{10}}{16}(3x^2-1);$$

$$\varphi_8(x) = \frac{\sqrt{10}}{16}(3y^2-1); \qquad \varphi_9(x) = \frac{\sqrt{10}}{16}(3z^2-1);$$

$$\varphi_{10}(x) = \frac{\sqrt{10}}{16}(3s^2-1)$$

Table 1 compares the estimation results for pure rejection and Algorithm 1 for $n = 2,\ldots,10$, as well as the program execution times in each case. Note that for $n > 5$ rejection is practically inapplicable. As a result, the rejection execution times for $n > 5$ are estimated. Note also that by using Algorithm 1, we get an up-to-the-third-decimal precise result for less than a minute.

n	Rej Result	Rej Time	Cond Result	Cond Time	C/R Ratio (s)
2	12.549122	1.191	12.570709	0.476	0.400
3	12.698118	51.23	12.712320	0.913	0.017
4	12.726244	358.2	12.739003	1.225	0.003
5	12.787773	1873.2	12.796741	1.527	$8.1 \cdot 10^{-4}$
6		115340	12.830315	3.954	$3.4 \cdot 10^{-5}$
7		1312912	12.840299	12.87	$9.8 \cdot 10^{-6}$
8		$3.2 \cdot 10^{6}$	12.848023	29.59	$9.2 \cdot 10^{-6}$
9		$9.7 \cdot 10^{6}$	12.850376	41.07	$4.2 \cdot 10^{-6}$
10		$4.1 \cdot 10^{7}$	12.851290	55.09	$1.3 \cdot 10^{-6}$

Table 1. Pure Rejection and Algorithm 1 Simulation Results

Conclusion

Random interpolation-cubature formulas are a powerful instrument for multiple integral estimation. By selecting their nodes according to the Δ^2 distribution, we obtain an estimate with minimal variance. Moreover, having an efficient algorithm for the Δ^2 simulation, we can practically solve multiple integrals regardless of their order. Of course, the suggested method dwells upon simulation in the very general problem settings. Imposing different restrictions on X, μ, and $\varphi_i(x)$, $i = \overline{1,n}$, we can simplify the conditional densities. This can lead to Algorithm 1 modifications like, for instance, simulate the distribution densities in the compositions by inversion instead of rejection. Nevertheless, even in the general settings the method works efficiently, which is demonstrated on an example of a five-dimensional integral estimation.

References

1. Barlow,R.E., Proschan,F. *Statistical Theory of Reliability and Life Testing: Probability Models*, New York, Holt, Rinehart and Winston, Inc., 1975. 302p.
2. Ermakov S.M. *The Monte Carlo Method and Related Topics*, "Nauka", M., 1971. 327p.
3. Ermakov S.M., Missov T.I. *On Simulation of the Δ^2-distribution*. Vestnik St. Petersburg University, Vol.4, 2005. pp. 123-140.
4. Ermakov S.M., Missov T.I. *The Δ^2-distribution and its Application in Regression Analysis*. Journal of the Department of Statistical Simulation, 2006. pp.25-35 (in Russian).
5. Ermakov S.M., Schwabe R. *On Randomizing Estimators in Linear Regression Models*, Freie Universitaet Berlin Verlag, Serie A - Mathematik, 1999. 10p.
6. Ermakov S.M., Zolotoukhin V.G. *Polynomial approximations and the Monte Carlo Method*, Probability Theory and its Application (1960) 5, 4. pp.473-475 (in Russian).
7. Korn G.A., Korn T.M. *Mathematical handbook for scientists and engineers*, New York: McGraw-Hill, 2nd enl. and rev. edition, 1968. 1130p.
8. Missov T.I., Ermakov S.M. *Simulation of the Δ^2-distribution. The discrete case*. Proceedings of the 5th St. Petersburg Workshop on Simulation, 2005. pp. 499-502

Received

Author information

Trifon I. Missov, Dept. Mathematics and Mechanics, St. Petersburg State University, Bibliotechnaya Sq. 2 St. Petersburg, Petrodvorets 198904, and Max Planck Institute for Demographic Research, Konrad-Zuse-Str. 1, D-18057, Rostock, Russia, Germany.
Email: missov@demogr.mpg.de