Generating Normally Distributed Random Numbers by Inverting the Normal Distribution Function

By Friedrich Gebhardt

1. Introduction and Summary. Some calculations on digital computers require a multitude of normally distributed random numbers, for instance in connection with Monte Carlo methods. Because of the great quantity, a fast way of generating them is desired, which would naturally be done at some expense of precision. Several methods are described and compared by M. E. Muller in [3]. In order to transform uniformly distributed random numbers into normally distributed ones, he proposes in [2] to approximate the inverse function of the cumulative normal distribution function by polynomials. The interval $0 \leqq t \leqq 1$ is divided into 128 parts of equal lengths, and for all of them except the first and the last one, where the inverse function becomes infinite, polynomials of first, second, and fourth degree respectively, are given, approximating the inverse distribution with a maximum error of 0.0004. The division into 128 parts is appropriate for binary electronic computers. We give in Section 2 similar approximations for decimal computers dividing the whole interval into 100 parts for maximum errors of .0004 and .0001. Near its singularities, the inverse function can be approximated by rational functions and by an iterative method based on the semiconvergent series of the normal distribution function. This will be done in Section 3.

2. Approximation by Polynomials. Let $\Phi(x)$ be the cumulative normal distribution function and $\psi(t)$ its inverse, $0 \leqq t \leqq 1$. Because of the relation

$$\psi(1 - t) = -\psi(t),$$

we restrict our attention to $t \leqq .5$. We want to approximate $\psi(t)$ in appropriate intervals by polynomials. In order to facilitate address modification on decimal computers, the whole range of t is divided into subintervals of length $\frac{1}{100}$, the subinterval I_n being

$$I_n: \qquad \frac{n}{100} \leqq t < \frac{n+1}{100}, \qquad n = 0, 1, \cdots 49.$$

To get a maximum error ϵ_M of less than 0.0004, approximation of $\psi(t)$ by linear functions $a_n + b_n t$ is sufficient for $8 \leqq n \leqq 49$. For $n \geqq 16$, the maximum error is even less than 0.0001. The coefficients a_n and b_n are shown in Table 1. Quadratic functions are sufficient for $\epsilon_M = .0001$, $n \geqq 3$, and $\epsilon_M = .0004$, $n \geqq 2$, respectively, polynomials of third order for $\epsilon_M = .0001$, $n = 2$, and for $\epsilon_M = .0004$, $n = 1$, and a fourth order polynomial for $\epsilon_M = .0001$, $n = 1$. The coefficients are listed in Table 2. The first polynomial in Table 2 ($n = 0$) is correct to 0.001 for $t \geqq .0020093$ only, corresponding to $\psi(t) = 2.8767$; for smaller values of t, the error becomes fairly large, and the function value at $t = 0$ is $-3.283\ 258$ instead of $-\infty$. This approxi-

Received July 23, 1963. Work was done in part while the author was at the Statistische Arbeitsgruppe Prof. Dr. H. Kneser, University Tübingen.

TABLE 1

Linear approximation of $\psi(t)$. *Maximum error* $\epsilon_M < .0004$ *for all values of* n, $\epsilon_M < .0001$ *for* $n \geqq 16$

n	$-a_n$	b_n	n	$-a_n$	b_n
8	1.919 253	6.4317	29	1.393 892	2.8984
9	1.873 295	5.9203	30	1.380 905	2.8551
10	1.831 555	5.5024	31	1.368 507	2.8151
11	1.793 280	5.1541	32	1.356 829	2.7786
12	1.757 969	4.8596	33	1.345 743	2.7450
13	1.725 181	4.6072	34	1.335 272	2.7142
14	1.694 582	4.3885	35	1.325 474	2.6862
15	1.665 962	4.1976	36	1.316 259	2.6606
16	1.639 047	4.0293	37	1.307 603	2.6372
17	1.613 677	3.8800	38	1.299 624	2.6162
18	1.589 728	3.7469	39	1.292 216	2.5972
19	1.567 050	3.6275	40	1.285 417	2.5802
20	1.545 557	3.5200	41	1.279 268	2.5652
21	1.525 151	3.4228	42	1.273 684	2.5519
22	1.505 752	3.3346	43	1.268 782	2.5405
23	1.487 311	3.2544	44	1.264 515	2.5308
24	1.469 771	3.1813	45	1.260 872	2.5227
25	1.453 075	3.1145	46	1.257 974	2.5164
26	1.437 140	3.0532	47	1.255 719	2.5116
27	1.422 024	2.9972	48	1.254 233	2.5085
28	1.407 578	2.9456	49	1.253 449	2.5069

mation might be used, if not too many normally distributed random numbers are needed, or if a lower accuracy is sufficient near $t = 0$ and $t = 1$.

The linear functions are computed such, that the error is the same at both ends of the proper interval and has the same absolute value and opposite sign in its middle. The maximum absolute error is then slightly greater. The polynomials of second and higher degree, with the exception of the first one ($n = 0$), are Chebyshev approximations, which again almost minimize the maximum absolute error. The coefficients were computed first for a linear transform, y, of t, such that $y = \pm 1$ at the end points of the corresponding interval. The rounding error of each term then did not exceed 10^{-6}. Transforming to t as independent variable, the rounding error remains of the same order of magnitude, although the high order coefficients seem to provide an accuracy of four decimals only. I.e., any rounding error of the high order term is adjusted by corresponding alterations of the other coefficients to yield a total rounding error not much greater than 10^{-6}. Much accuracy is lost if all coefficients are rounded to four decimal places. The first polynomial ($n = 0$) yields an error of absolute value 0.001 and opposite signs at six points of the interval $0.002\ 0093 \leqq t \leqq 0.01$. The first end point is chosen so as to get as large an interval as possible with an absolute error less than 0.001.

3. Special Methods for the End Intervals. If the accuracy of the first polynomial in Table 2 is not sufficient, then other methods must be used to invert the normal distribution function in this interval. An approximation closer than by a poly-

TABLE 2

Approximation of $\psi(t)$ by polynomials of second, third and fourth degree

<table>
<tr><th>n</th><th>$-a_n$</th><th>b_n</th><th>$-c_n$</th><th>d_n</th><th>$-e_n$</th></tr>
<tr><td colspan="6">Maximum error $<.001$ for $t \geqq .002\ 0093$</td></tr>
<tr><td>0</td><td>3.283 258</td><td>268.6351</td><td>38 726.92</td><td>3 186 023</td><td>104 377·10³</td></tr>
<tr><td colspan="6">Maximum error $<.0004$</td></tr>
<tr><td>1</td><td>2.862 543</td><td>73.768 93</td><td>2 360.586</td><td>34 721.3</td><td></td></tr>
<tr><td>2</td><td>2.573 835</td><td>31.862 55</td><td>292.280</td><td></td><td></td></tr>
<tr><td colspan="6">Maximum error $<.0001$</td></tr>
<tr><td>1</td><td>2.940 578</td><td>95.840 37</td><td>4 654.285</td><td>138 586.9</td><td>1 731 093</td></tr>
<tr><td>2</td><td>2.686 913</td><td>45.711 66</td><td>851.840</td><td>7 460.8</td><td></td></tr>
<tr><td>3</td><td>2.454 298</td><td>23.714 938</td><td>153.1653</td><td></td><td></td></tr>
<tr><td>4</td><td>2.363 419</td><td>19.115 067</td><td>94.8907</td><td></td><td></td></tr>
<tr><td>5</td><td>2.289 468</td><td>16.134 053</td><td>64.8247</td><td></td><td></td></tr>
<tr><td>6</td><td>2.226 823</td><td>14.034 602</td><td>47.2253</td><td></td><td></td></tr>
<tr><td>7</td><td>2.172 275</td><td>12.469 763</td><td>35.9980</td><td></td><td></td></tr>
<tr><td>8</td><td>2.123 837</td><td>11.255 154</td><td>28.3813</td><td></td><td></td></tr>
<tr><td>9</td><td>2.080 175</td><td>10.282 668</td><td>22.9647</td><td></td><td></td></tr>
<tr><td>10</td><td>2.040 482</td><td>9.487 217</td><td>18.9787</td><td></td><td></td></tr>
<tr><td>11</td><td>2.003 891</td><td>8.820 893</td><td>15.9447</td><td></td><td></td></tr>
<tr><td>12</td><td>1.970 004</td><td>8.255 343</td><td>13.5847</td><td></td><td></td></tr>
<tr><td>13</td><td>1.938 422</td><td>7.768 893</td><td>11.7113</td><td></td><td></td></tr>
<tr><td>14</td><td>1.908 714</td><td>7.344 010</td><td>10.1920</td><td></td><td></td></tr>
<tr><td>15</td><td>1.880 811</td><td>6.971 602</td><td>8.9493</td><td></td><td></td></tr>
</table>

TABLE 3

Approximation of $\psi(t)$ by rational functions $R_i(t) = A_i t + B_i + C_i/t + D_i/t^2$ in the intervals $\alpha_i \leqq t \leqq \beta_i$

<table>
<tr><th>α_i</th><th>β_i</th><th>A_i</th><th>$-B_i$</th><th>-10^3C_i</th><th>10^6D_i</th></tr>
<tr><td colspan="6">Maximum error $<.0004$</td></tr>
<tr><td>.080 000</td><td>.030 000</td><td>4.382 71</td><td>1.535 054</td><td>19.621 90</td><td>159.3904</td></tr>
<tr><td>.030 000</td><td>.010 000</td><td>9.462 73</td><td>1.978 144</td><td>6.165 05</td><td>17.4057</td></tr>
<tr><td>.010 000</td><td>.003 153</td><td>23.914 4</td><td>2.402 951</td><td>1.782 268</td><td>1.608 30</td></tr>
<tr><td>.003 153</td><td>.000 962</td><td>66.124 0</td><td>2.794 870</td><td>.500 034</td><td>.138 558</td></tr>
<tr><td colspan="6">Maximum error $<.0001$</td></tr>
<tr><td>.010 000</td><td>.004 4649</td><td>20.958 56</td><td>2.342 2273</td><td>2.179 6742</td><td>2.436 095</td></tr>
<tr><td>.004 4649</td><td>.001 9609</td><td>42.183 87</td><td>2.624 0660</td><td>.897 8071</td><td>.441 8733</td></tr>
<tr><td>.001 9609</td><td>.000 8480</td><td>87.840 36</td><td>2.890 1189</td><td>.366 167 73</td><td>.078 138 077</td></tr>
<tr><td>.000 8480</td><td>.000 3624</td><td>187.579 65</td><td>3.141 6999</td><td>.148 422 09</td><td>.013 573 432</td></tr>
</table>

nomial of fourth degree is given by the functions

$$R_i(t) = A_i t + B_i + \frac{C_i}{t} + \frac{D_i}{t^2}, \qquad \alpha_i \leqq t \leqq \beta_i$$

which need about the same computation time on digital computers. Table 3 shows the coefficients of four such functions with a maximum error of 0.0004 in the interval $.000\ 962 \leqq t \leqq 0.08$ and of four functions with a maximum error of 0.0001 in the interval $.000\ 362\ 4 \leqq t \leqq 0.01$. In the first case, the use of all four rational functions eliminates the need of quadratic and cubic polynomials as described in Section 2. This simplifies the program; however, the rational functions need about one and a half times the computation time of a quadratic polynomial.

In the remaining interval, $t < .000\ 962$, $t < .000\ 362\ 4$ respectively, i.e. $x < -3.1016$, $x < -3.3800$ respectively, the approximation

$$\Phi(x) \approx S(x) = \frac{\phi(x)}{T(x)},$$

$$T(x) = -x - \frac{1}{x} + \frac{2}{x^3} - \frac{6}{x^5}$$

may be used. It is derived from the semi-convergent series for $\Phi(x)/\phi(x)$; the coefficient of x^{-5}, however, is altered in order to yield a smaller absolute error. Let t be given, and x, x^*, Δx be the solutions of

$$t = \Phi(x),$$

$$t = S(x^*),$$

$$x^* = x + \Delta x.$$

Then the first terms of the power expansion yield

$$S(x^*) = \Phi(x^* - \Delta x) \approx \Phi(x^*) - \Delta x \phi(x^*),$$

$$\Delta x \approx \frac{\Phi(x^*)}{\phi(x^*)} - \frac{1}{T(x^*)}.$$

Numerical evaluations of this expression show that $|\Delta x| < .000\ 067$ for $t < .001$ $(x < -3.09)$.

The solution x^* is obtained by an iterative procedure. Let x_k be the kth approximation and

$$Q = \frac{S(x_k) - t}{S(x_k)},$$

$$\Delta x_k = \frac{Q + Q^2/2 + Q^3/3}{x_k + \dfrac{1}{x_k}}$$

$$x_{k+1} = x_k + \Delta x_k.$$

This procedure converges considerably faster than Newton's method for $S(x)$, and was derived from Newton's method applied to log $S(x)$. Numerical calculations showed: If $t < .001$ and $| \Delta x_k | < .02$, then $| x_{k+1} - x^* | < .000\ 35$ and, as $| x_{k+1} - x^* |$ and $| x^* - x |$ depend on t and assume their maxima at different points, $| x_{k+1} - x | < .0004$ (even $<.000\ 37$). Thus, the iteration may be ended as soon as $| \Delta x_k | < .02$, if an accuracy of .0004 is sufficient. In the same way, if $t < .001$ and $| \Delta x_k | < .0035$, then $| x_{k+1} - x^* | < .000\ 52$, $| x_{k+1} - x | < .0001$. As initial values, x_0, of the iteration process are recommended in the first case ($\epsilon_M < .0004$) $\psi(.000\ 962) - .020 = -3.122$ and in the second case ($\epsilon_M < .0001$) $\psi(.000\ 362\ 4) - .0035 = -3.3835$ in order to cover the greatest possible t-interval with a single iteration. In the first case, 3 iterations are needed for $t = 0.0001$ and 4 for $t = 0.000\ 01$.

A program to compute $\psi(t)$ was written for the Siemens 2002 computer, using linear approximations, rational functions $R_i(t)$ and the iteration process, $\epsilon_M = 0.0004$. One iteration needed about the time of 40 multiplications. However, in most values of t, the iteration process is not involved, and the average computation time was approximately that of four and a half multiplications. This program was part of a multidimensional integration problem [1], where more than 900 000 normal deviates were computed.

University of Connecticut
Storrs, Connecticut

1. Friedrich Gebhardt, "On the risk of some strategies for outlying observations," submitted to *Ann. Math. Statist.*

2. Mervin E. Muller, "An inverse method for the generation of random normal deviates on large-scale computers," *MTAC*, v. 12, 1958, p. 167–174.

3. Mervin E. Muller, "A comparison of methods for generating normal deviates on digital computers," *J. Assoc. Comput. Mach.*, v. 6, 1959, p. 376–383.

Optimum Runge-Kutta Methods

By T. E. Hull and R. L. Johnston

Abstract. The optimum Runge-Kutta method of a particular order is the one whose truncation error is a minimum. Various measures of the size of the truncation error are considered. The optimum method is practically independent of the measure being used. Moreover, among methods of the same order which one might consider using the difference in size of the estimated error is not more than a factor of 2 or 3. These results are confirmed in practice insofar as the choice of optimum method is concerned, but they underestimate the variation in error between different methods.

1. Introduction. For the solution of

$$y' = f(x, y), \qquad y(x_0) = y_0 \tag{1}$$

the general Runge-Kutta method of order m uses the formula

$$y_{n+1} = y_n + \sum_{i=1}^{m} w_i k_i \tag{2}$$

Received September 24, 1963, revised November 19, 1963. This research was supported in part by the Defence Research Board of Canada.