

Numerical Integration of Bivariate Gaussian Distribution

S. H. Derakhshan and C. V. Deutsch

The bivariate normal distribution arises in many geostatistical applications as most geostatistical techniques rely on two-point statistics. This paper addresses an algorithm to calculate the bivariate normal probabilities (i.e. cumulative distribution function for different cutoffs for both variables). Various algorithms and approximations have been proposed to calculate the bivariate normal probabilities. In this paper, Simpson's 1/3 rule is used to calculate the bivariate normal probabilities. The efficiency and accuracy of the method are validated. A FORTRAN subroutine called `getbnp` is written for this purpose.

1. Introduction

The goal in this paper is to numerically calculate the probability that a normal random variable Y_1 with mean zero and variance one is greater than y_p , and another normal random variable Y_2 with mean zero and variance one is greater than y_q , while the correlation between Y_1 and Y_2 is ρ . This probability is related to standard bivariate normal cumulative distribution function. The two cutoffs of $y_1 = y_p$ and $y_2 = y_q$ divide the whole 2D Cartesian space into four regions p_{00} , p_{01} , p_{10} and p_{11} :

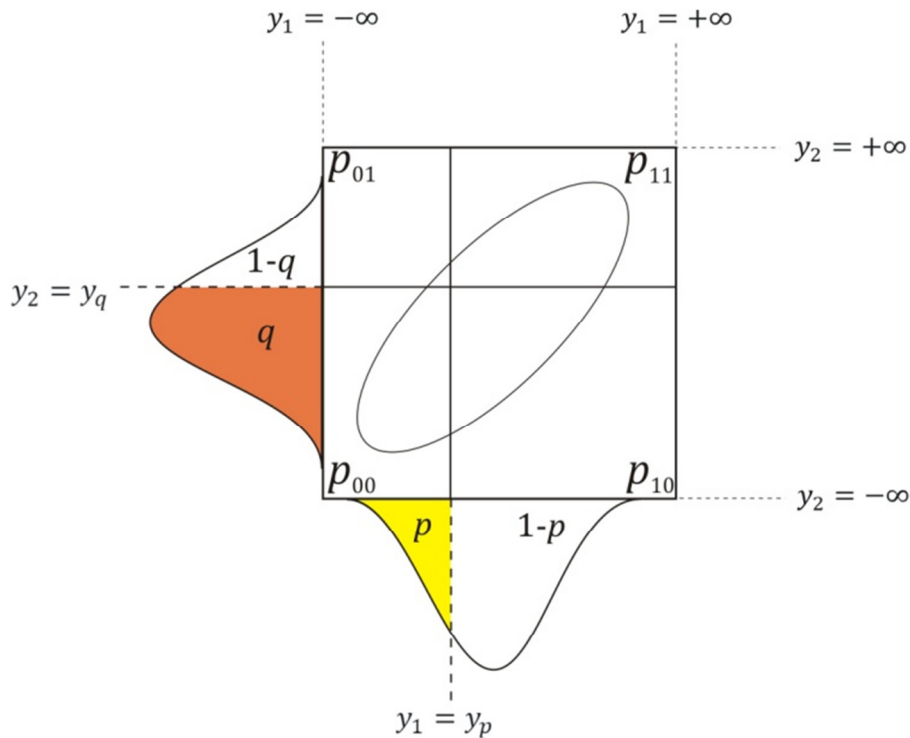


Figure 1 standard bivariate normal distribution between Y_1 and Y_2 ; the two cutoffs of $y_1 = y_p$ and $y_2 = y_q$ divide the Cartesian space into four regions

The probability definitions of p_{00} , p_{01} , p_{10} and p_{11} are as below:

$$\begin{aligned} p_{00} &= \text{Prob}\{Y_1 \leq y_p, Y_2 \leq y_q\} \\ p_{01} &= \text{Prob}\{Y_1 \leq y_p, Y_2 > y_q\} \\ p_{10} &= \text{Prob}\{Y_1 > y_p, Y_2 \leq y_q\} \\ p_{11} &= \text{Prob}\{Y_1 > y_p, Y_2 > y_q\} \end{aligned}$$

p_{00} and p_{11} have specific meaning, p_{00} is the bivariate normal cumulative distribution function and p_{11} is the survival function.

Another ways of showing p_{00} and p_{11} as functions of y_p, y_q and ρ are:

$$p_{00} = G(y_p, y_q; \rho)$$

$$p_{11} = L(y_p, y_q; \rho)$$

In literature the focus is more on p_{11} . There are different formulas, tables and algorithms in literature to calculate p_{11} (e.g. Abramowitz et al. 1965 and Balakrishnan et al. 2009). Generally there is no closed form analytical formula for p_{11} but there are some approximations and numerical tables to calculate $L(y_p, y_q; \rho)$ in the literatures. For analytical and approximation purposes and also tabulating standard values for $L(y_p, y_q; \rho)$ in the literature, one can use below identity (Abramowitz et al. 1965, Balakrishnan et al. 2009 and Lin 1995):

$$p_{11} = L(y_p, y_q; \rho) = L(y_p, 0; \rho_{y_p, y_q}) + L(y_q, 0; \rho_{y_q, y_p}) - \frac{1}{2} \delta_{y_p, y_q}$$

The right hand side still contains the L function but the second input parameter for that is equal to zero. The new parameters in above equation are defined as below:

$$\rho_{y_p, y_q} = \text{sign}(y_p) \cdot \frac{(\rho y_p - y_q)}{\sqrt{y_p^2 - 2\rho y_p y_q + y_q^2}}$$

$$\rho_{y_q, y_p} = \text{sign}(y_q) \cdot \frac{(\rho y_q - y_p)}{\sqrt{y_p^2 - 2\rho y_p y_q + y_q^2}}$$

$$\delta_{y_p, y_q} = \begin{cases} 0 & ; \text{sign}(y_p) \cdot \text{sign}(y_q) = 1 \\ 1 & ; \text{otherwise} \end{cases}$$

$$\text{sign}(y_p) = \begin{cases} +1 & ; y_p \geq 0 \\ -1 & ; y_p < 0 \end{cases} ; \quad \text{sign}(y_q) = \begin{cases} +1 & ; y_q \geq 0 \\ -1 & ; y_q < 0 \end{cases}$$

ρ_{y_p, y_q} and ρ_{y_q, y_p} are newly calculated correlation coefficients and are functions of y_p, y_q and ρ . By using the relations in methodology part of this paper, other bivariate normal probabilities (p_{01}, p_{10} and p_{00}) can also be calculated. The above identity helps to convert the trivariate function of $L(y_p, y_q; \rho)$ to three bivariate functions of $L(y_p, 0; \rho_{y_p, y_q})$, $L(y_q, 0; \rho_{y_q, y_p})$ and δ_{y_p, y_q} . Usually the values for $L(y_p, 0; \rho_{y_p, y_q})$ and $L(y_q, 0; \rho_{y_q, y_p})$ are tabulated in handbooks and literature and by using above identity $L(y_p, y_q; \rho)$ is calculated.

A subroutine called `getbnp` is written to get p, q and ρ and calculate p_{00}, p_{01}, p_{10} and p_{11} . It uses Simpson's 1/3 rule to calculate the bivariate normal integrals.

2. Methodology

The four bivariate normal probabilities are related to each other. The marginal normal distributions can help to assess this relationship. The following relations are obtained (see Figure 1):

- The summation of p_{00} and p_{01} must be p , that is $p_{00} + p_{01} = p$
- The summation of p_{00} and p_{10} must be q , that is $p_{00} + p_{10} = q$
- All of the four bivariate normal probabilities must sum to 1, that is $p_{00} + p_{01} + p_{10} + p_{11} = 1$

Therefore, by knowing one of the probabilities (e.g. p_{11}), the three others can be calculated:

			p_{00}	p_{01}	p_{10}	p_{11}
1	p_{00} is known	$p_{00} = \int_{-\infty}^{y_q} \int_{-\infty}^{y_p} g(y_1, y_2; \rho) dy_1 dy_2$	p_{00}	$p - p_{00}$	$q - p_{00}$	$1 - p - q + p_{00}$
2	p_{01} is known	$p_{01} = \int_{y_q}^{+\infty} \int_{-\infty}^{y_p} g(y_1, y_2; \rho) dy_1 dy_2$	$p - p_{01}$	p_{01}	$q - p + p_{01}$	$1 - q - p_{01}$
3	p_{10} is known	$p_{10} = \int_{-\infty}^{y_q} \int_{y_p}^{+\infty} g(y_1, y_2; \rho) dy_1 dy_2$	$q - p_{10}$	$p - q + p_{10}$	p_{10}	$1 - p - p_{10}$
4	p_{11} is known	$p_{11} = \int_{y_q}^{+\infty} \int_{y_p}^{+\infty} g(y_1, y_2; \rho) dy_1 dy_2$	$p + q - 1 + p_{11}$	$1 - q - p_{11}$	$1 - p - p_{11}$	p_{11}

Where $g(y_1, y_2; \rho)$ is the standard bivariate normal probability density function and formulated as:

$$g(y_1, y_2; \rho) = \frac{1}{2\pi\sqrt{1-\rho^2}} \cdot e^{-\frac{(y_1^2 - 2\rho y_1 y_2 + y_2^2)}{2(1-\rho^2)}}$$

p and q are the univariate normal probabilities and are related to y_p and y_q by:

$$p = G(y_p) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{y_p} e^{-\frac{y_1^2}{2}} dy_1 \Leftrightarrow y_p = G^{-1}(p)$$

$$q = G(y_q) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{y_q} e^{-\frac{y_2^2}{2}} dy_2 \Leftrightarrow y_q = G^{-1}(q)$$

$G(\cdot)$ is the standard univariate normal cumulative distribution function and $G^{-1}(\cdot)$ is its inverse.

Special cases

Depending on the values of p , q and ρ the above equations can be simplified and easy to calculate:

1. If there is no correlation between Y_1 and Y_2 then $\rho = 0$. In this case the integrals are separable. Since the two variables are independent, the bivariate normal probabilities are obtained using the multiplication rule for probability:

$$p_{00} = p \cdot q ; p_{01} = p \cdot (1 - q); p_{10} = (1 - p) \cdot q \text{ and } p_{11} = (1 - p) \cdot (1 - q)$$

2. If the correlation between Y_1 and Y_2 is $\rho = -1$, then two different cases might happen:
 - a. The summation of p and q is greater than or equal to 1, that is $p + q \geq 1$. In this case the bivariate normal probabilities are:

$$p_{00} = p + q - 1 ; p_{01} = 1 - q; p_{10} = 1 - p \text{ and } p_{11} = 0$$

- b. The summation of p and q is less than 1, that is $p + q < 1$. In this case the bivariate normal probabilities are:

$$p_{00} = 0 ; p_{01} = p; p_{10} = q \text{ and } p_{11} = 1 - p - q$$

3. If the correlation between Y_1 and Y_2 is $\rho = +1$, then two different cases might happen:
 - a. p is greater than or equal to q , that is $p \geq q$. In this case the bivariate normal probabilities are:

$$p_{00} = q ; p_{01} = p - q ; p_{10} = 0 \text{ and } p_{11} = 1 - p$$

- b. p is less than q , that is $p < q$. In this case the bivariate normal probabilities are:

$$p_{00} = p ; p_{01} = 0; p_{10} = q - p \text{ and } p_{11} = 1 - q$$

4. If p and q are the same and equal to 0.5, that is $p = q = 0.5$. The double integral for the bivariate normal probabilities has closed form analytical solution:

$$p_{00} = p_{11} = \frac{1}{4} + \frac{\text{ArcSin}(\rho)}{2\pi} \text{ and } p_{01} = p_{10} = \frac{\text{ArcCos}(\rho)}{2\pi}$$

Simpson's 1/3 rule

Simpson's 1/3 rule is used to calculate the bivariate normal integrals. The interval of integration in Simpson's 1/3 rule is divided into an even number of equal intervals (or an odd number of nodes). Each node has a weight. The integral by Simpson's rule is equal to the weighted summation (Simpson's 1/3 weights) of the values of function at each node multiplied by one third of the subinterval length. For one dimensional integral, the integration starts at the first two equal intervals; the two equal intervals are constructed by three points of start, middle and end of the two intervals together. A polynomial of degree two can be fitted such that all of these three points are honored. In other words the interested function (integrand) is approximated by polynomial of degree two in these two equal intervals. If this approximation is done for the rest of subsequent intervals then below formulas are obtained using Simpson's 1/3 rule (for two dimensional integrals) for p_{11} .

For calculating p_{11} first infinity should be approximated by a value. In Gaussian units usually -5.0 and +5.0 are used for minus and plus infinity. For better approximation of p_{11} , minus and plus infinity in Gaussian units are set to -10.0 and +10.0 in this paper. It will be shown that with this assumption for infinity, the numerical integration is still fast. Simpson's 1/3 rule is applied for both variables, y_1, y_2 therefore the two dimensional Simpson's 1/3 rule can be condensed in a double summation by introducing weights for each node:

$$p_{11} = \int_{y_q}^{+\infty} \int_{y_p}^{+\infty} g(y_1, y_2; \rho) dy_1 dy_2 \cong \int_{b_1}^{b_{2m+1}} \int_{a_1}^{a_{2n+1}} g(y_1, y_2) dy_1 dy_2 \cong \frac{h^2}{9} \cdot \sum_{j=1}^{2m+1} \sum_{i=1}^{2n+1} 2^{w_{ij}} \cdot g(a_i, b_j; \rho)$$

$$g(y_1, y_2; \rho) = \frac{1}{2\pi\sqrt{1-\rho^2}} \cdot e^{-\frac{(y_1^2 - 2\rho y_1 y_2 + y_2^2)}{2(1-\rho^2)}}$$

$$n = 1 + \text{int}\left(\frac{A - y_p}{h}\right) - \text{int}\left[\frac{1}{2} + \frac{1}{2}\text{int}\left(\frac{A - y_p}{h}\right)\right] \Rightarrow a_i = y_p + (i - 1) \cdot h \quad \text{for } i = 1, \dots, 2n + 1$$

$$m = 1 + \text{int}\left(\frac{A - y_q}{h}\right) - \text{int}\left[\frac{1}{2} + \frac{1}{2}\text{int}\left(\frac{A - y_q}{h}\right)\right] \Rightarrow b_j = y_q + (j - 1) \cdot h \quad \text{for } j = 1, \dots, 2m + 1$$

$$w_{ij} = 4 - \text{mod}(i, 2) - \text{mod}(j, 2) - \delta_{i,1} - \delta_{i,2n+1} - \delta_{j,1} - \delta_{j,2m+1}$$

$$\delta_{i,j} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$$

Note that h is the length of the subinterval, A is a big number in Gaussian space that represents plus infinity (in this paper A is set to 10.0) and a_i and b_j are nodes for integration. m and n are chosen in such a way that a_{2n+1} and b_{2m+1} are greater than A and there exists an odd number of nodes for both variables. A $(2n + 1) \times (2m + 1)$ matrix can be used to show the weights, w_{ij} , for each node of the integration (each array in the matrix is considered as one node):

$$W = [w_{ij}; i = 1, \dots, 2n + 1, j = 1, \dots, 2m + 1] = \begin{bmatrix} 1 & 4 & 2 & 4 & \dots & 2 & 4 & 1 \\ 4 & 16 & 8 & 16 & \dots & 8 & 16 & 4 \\ 2 & 8 & 4 & 8 & \dots & 4 & 8 & 2 \\ 4 & 16 & 8 & 16 & \dots & 8 & 16 & 4 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 2 & 8 & 4 & 8 & \dots & 4 & 8 & 2 \\ 4 & 16 & 8 & 16 & \dots & 8 & 16 & 4 \\ 1 & 4 & 2 & 4 & \dots & 2 & 4 & 1 \end{bmatrix}_{(2n+1) \times (2m+1)}$$

Once p_{11} is calculated, the other four bivariate normal probabilities are obtained using the relations previously presented.

3. Implementation

To calculate the integral numerically using Simpson's rule and approximate the four bivariate normal probabilities, five parameters should be specified in the subroutine `getbnp(rho, p, q, h, lgauss, p00, p01, p10, p11, ierr)`. They are as follows:

- `rho` is the correlation coefficient between the two normal variables, ρ . It is a real number between -1 and 1.
- `p` is the univariate normal probability for the first variable. In other words it is the area under the univariate normal probability density function from minus infinity to y_p .
- `q` is the univariate normal probability for the second variable. In other words it is the area under the univariate normal probability density function from minus infinity to y_q .
- `h` is the interval length for Simpson's 1/3 rule numerical integration. Small h results in more precise calculation for the integral but increases CPU time. Choosing 0.01 for h is reasonable as it gives a precise answer for integral calculation within a small amount of CPU time.
- `lgauss` is a big number in Gaussian units that represents plus infinity. In the calculation for this section it is chosen to be equal to 10.0.
- `p00, p01, p10` and `p11` are the output of the subroutine and are equal to bivariate normal probabilities for the four regions under the bivariate normal curve (see Figure 1).

- `ierr` is the indicator for error situation. If `ierr=1` then it means that `p` and `q` are out of range. If `ierr=0` then it means that `p` and `q` are in range.

The results of the algorithm are validated for different sets of p, q and ρ . The interval length for integration is 0.01. The tables below show a few approximated values compared to exact values up to 4 decimal places (Based on Lin, J. 1995, the exact values are provided up to 4 decimal places).

p	y_p	ρ	Exact $L(y_p, 0; \rho)$ up to 4 decimal places	Calculated $L(y_p, 0; \rho)$ up to 4 decimal places
0.5793	0.2	-0.5	0.1289	0.1289
0.6915	0.5	-0.6	0.0657	0.0657
0.7881	0.8	-0.7	0.0233	0.0233
0.8643	1.1	-0.8	0.0037	0.0037

The computational efficiency of the algorithm is also checked. Four new parameters are used to show the computational efficiency. They are as below:

1. Number of bivariate bins (nodes) for Simpson's 1/3 rule integration, N_{bins} :

$$N_{bins} = (2n + 1) \times (2m + 1)$$

m and n are calculated using the formulas in previous section.

2. Number of similar digits (from left) of the calculated integration result with respect to exact value for integration (converged value), N_{s_dig} . For example if the calculated integration result is equal to 0.065145685 and the exact value is 0.065135585 then N_{s_dig} is equal to 4.
3. Error of the integration result with respect to exact value for integration (converged value);

$$Error (\%) = \left| \frac{\text{calculated result} - \text{converged result}}{\text{converged result}} \right| \times 100$$

Simpson's 1/3 rule finally converges to the exact value provided the interval length for integration is sufficiently small.

4. CPU time (in seconds, t_{CPU}) that elapsed for executing the subroutine, `getbnp`, for a specific p, q, ρ and h .

The length of the interval, h , is varied from 0.0001 to 1.0 with increment of 0.0001. Infinity is represented by 10.0 in Gaussian units. p, q and ρ are equal to 0.70, 0.55 and -0.5 respectively. The converged value up to 27 decimal places with interval length of 0.0001 for integration is calculated as 0.065355531871318817138671875. The integral result for other interval length is compared to this converged value. A 3.19 GHz processor with 3.0 GB of RAM was used to calculate the CPU time. Figure 2 shows the computational efficiency of the algorithm. As the number of bivariate bins increases, the number of similar digits and CPU time increases but error decreases. Figure 2 does not show the computational efficiency for bivariate bins with orders of magnitudes between 6 and 10 ($10^6 < N_{bins} < 10^{10}$). Figure 3 is the same as Figure 2 but for large values of N_{bins} . The relationship between N_{bins} and CPU time for large N_{bins} is:

$$t_{CPU} \cong N_{bins} \times 10^{-6.36} \text{ for } 10^6 < N_{bins} < 10^{10}$$

The first value that results in zero error is at $h = 0.0332$, $N_{bins} = 85813$. The CPU time for this point is 0.0313 seconds. The table below shows the results for some specific N_{bins} :

p	y_p	q	y_q	ρ	A	h	N_{bins}	N_{s_dig}	Error (%)	t_{CPU} (seconds)
0.7	0.5244	0.55	0.1257	-0.5	10.0	1.0000	121	2	5.0027	0.0000
0.7	0.5244	0.55	0.1257	-0.5	10.0	0.1000	9797	6	0.0005	0.0156
0.7	0.5244	0.55	0.1257	-0.5	10.0	0.0100	938561	27	0.0000	0.0463
0.7	0.5244	0.55	0.1257	-0.5	10.0	0.0010	93604329	27	0.0000	40.0938
0.7	0.5244	0.55	0.1257	-0.5	10.0	0.0001	9356779965	27	0.0000	3999.8906

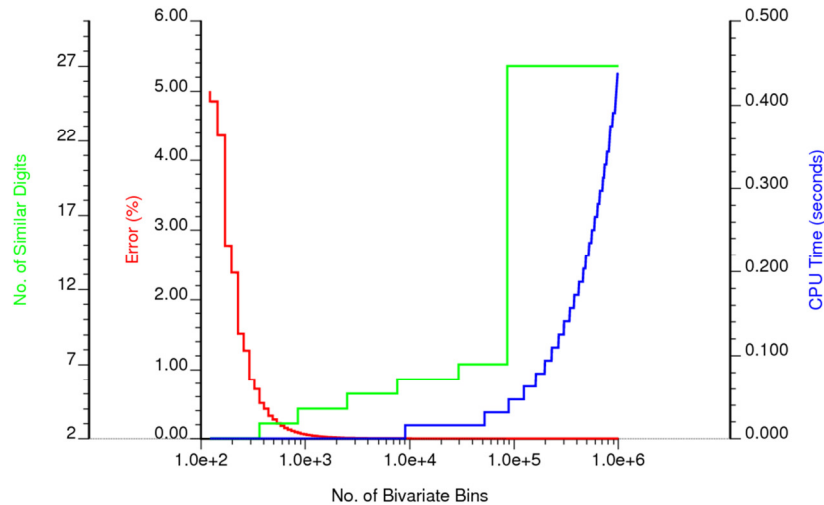


Figure 2 the horizontal axis is the number of bivariate bins (nodes) for integration, N_{bins} ; there are three vertical axes, (1) number of similar digits from left, (2) error and (3) CPU time, they are distinguished by colors of green, red and blue respectively. CPU time axis is in arithmetic scale (not log scale)

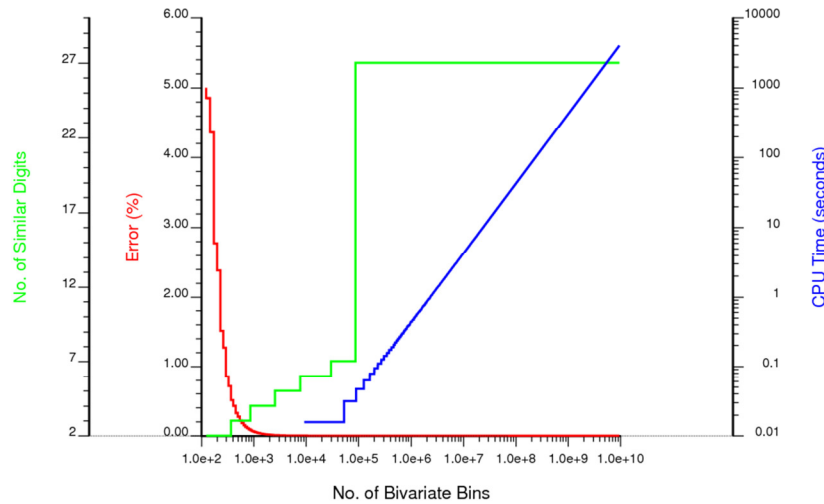


Figure 3 the computational efficiency of the Simpson's 1/3 rule for calculating the bivariate normal integrals; the focus in this figure is more on large N_{bins} and also the linear relationship between N_{bins} and CPU time in log-log.

4. Conclusion

Simpson's 1/3 rule is used to calculate the bivariate normal probabilities. With knowing the two univariate normal probabilities and the correlation coefficient between the two variables, the four bivariate normal probabilities are obtained. Subroutine `getbnp` is written in FORTRAN to get p , q and ρ and calculate p_{00} , p_{01} , p_{10} and p_{11} , the four bivariate normal probabilities. As future work; a fast algorithm to calculate the numerical integration of multivariate Gaussian distribution at different cut-offs can be considered.

References

- Abramowitz, M. and Stegun, I. A., 1965, *Handbook of Mathematical Functions*, Dover, New York
- Balakrishnan, N. and Lai C. D., 2009, *Continuous Bivariate Distributions*, Springer, 2nd edition
- Cox, D. R. and Wermuth, N., 1991, *A simple approximation for bivariate and trivariate normal integrals*, International Statistical Review, Vol. 59, Issue 2, p. 263-269
- Lin, J., 1995, *A simple approximation for bivariate normal integrals*, Journal of Probability in the Engineering and Information Sciences, Vol. 9, p. 317-321
- Owen, D.B., 1956, *Tables for computing bivariate normal probabilities*, Annals of Mathematical Statistics, Vol. 27, p. 1075-1090
- Zelen, M. and Severo, N. C., 1960, *Graphs for bivariate normal probabilities*, Annals of Mathematical Statistics, Vol. 31, p. 619-624