GEOMETRIC QUALITY ASSURANCE

Gemai Chen and Jiahua Chen

IIQP Research Report
RR-92-06

June 1992

Geometric Quality Assurance

Gemai Chen and Jiahua Chen

Department of Statistics and Actuarial Science
University of Waterloo
Waterloo, Ontario, N2L 3G1

ABSTRACT

In modern industrial settings, a prototype product is often created in the form of a computer image via Computer Assisted Design (CAD). After the first group of prototype products are made, their geometric features are measured with a Coordinate Measuring Machine (CMM). The quality assurance problem here is to develop (statistical) methods to check the geometric integrity of the products against design specifications. An existing method (Chapman and Kim, 1992) uses spherical regression to analyze the directional features such as straight lines and planes. In this paper, we develop a new method in which the proposed measure of geometric quality is easy to calculate and has good statistical properties. In addition, our method can be used to check more general geometric qualities.

Key words and phrases: Geometric quality assurance, Computer assisted design, Coordinate measuring machine.

1 Introduction

In modern industrial settings, a prototype product is often created in the form of a computer image via Computer Assisted Design (CAD). The specifications of the prototype product are stored as a CAD file. However, when a product is actually made, it is always different from its design due to various manufacture errors. For instance, a manufactured door may fail to fit into the body of the car. Various efforts have been made in industry to solve this kind of problem. A direct approach is to control the geometric variations of the manufactured products. This gives rise to the problem of how to judge the geometric quality of manufactured products. We follow Chapman and Kim (1992) to call this the geometric quality assurance problem.

Traditionally, a high quality mould is made to examine whether a product is of satisfactory quality. In doing so, the product is placed against the mould to see whether or not it fits. This method is, as far as we are aware, still in wide use in practice.

Another technique depends on the availability of a Coordinate Measuring Machine (CMM). The product is held firmly and a number of points on the surface of the product are measured relative to the coordinate system of the CMM with high precision. We call these measurements the CMM data. The geometric features of the product are then calculated and compared with the CAD file.

In this paper, we are interested in developing methods to compare the CMM data with the CAD file. It is understood that the coordinate system for the CMM data and the coordinate system for the CAD file are different in general and are usually not linked in a known manner. Therefore, in order to make the CMM data and the CAD file comparable, some transformation has to be applied to the CMM data. Suppose that a transformation for this purpose can be found, the next question is how to judge the difference between the CMM data and the CAD file. In general, the CAD file can be used to locate the surface of the designed product. But, since no product is perfect, and there is always measurement error, points in the CMM data file may not all fall onto the surface in the CAD file. A measure of this departure is then necessary, and the problem of geometric quality assurance is essentially to define a good measure and use it to quantify the geometric quality of the product.

The first such effort is made by Chapman and Kim (1992). They propose to use the spherical regression method to analyze the directional features of a product such as the unit normal directions of planes. The quality of the product is judged by the closeness between the measured directions and the designed directions. This method can be applied to products whose directional features are most important. In this paper, we propose a new method that can be used to examine geometric features such as lines, planes, curves, or general surfaces. Computations associated with our method are easy to perform and our proposed measure of geometric quality has nice properties. Simulation results show that the new method is very sensitive to various simulated manufacture errors.

The organization of the paper is as below. The problem of geometric quality assurance will be described with more details in the next section, together with the proposed method and some properties of the proposed method. Simulation results will be presented in Section 3. Some remarks are given in Section 4 and a short proof in the appendix.

2 Problem Formulation

2.1 Description of the problem

Although the products in practice may have very complicated shapes, we use a simplex to explain our idea. Let us assume that the product has the shape as shown in Figure 1. The surface of the simplex consists of four planes, and we denote it as $S = S_1 \cup S_2 \cup S_3 \cup S_4$. In general, we assume that the surface of a product consists of m sub-surfaces and denote them by $S_1, \ldots, S_m$.

When a prototype product is made, its geometric features are measured with a CMM. To fix the idea, we denote the surface of the manufactured simplex as $\tilde{S} = \tilde{S}_1 \cup \tilde{S}_2 \cup \tilde{S}_3 \cup \tilde{S}_4$ and measure 6 points on each of $\tilde{S}_i$, $i = 1, \dots, 4$. In general, for m sub-surfaces, we denote the jth measurement on $\tilde{S}_i$ by y_{ij}, where $i = 1, \dots, m$ and $j = 1, \dots, n_i$. Note that usually, we cannot identify the corresponding points on S_i. However, it is reasonable to assume that

$$Ry_{ij} + T = x_{ij} + e'_{ij} + e''_{ij}, \tag{2.1}$$

where x_{ij} are the unidentifiable points on S_i that correspond to y_{ij}, e'_{ij} are the manufacture errors, e''_{ij} are the measurement errors, R is a 3×3 rotation matrix and T is a 3×1 translation vector. When there are no manufacture errors and measurement errors, R and T together will transform the coordinate system of the CMM data to that of the CAD file.

We have made an implicit assumption. That is, we can identify the sub-surfaces both in the CAD file and from the manufactured product. This is a reasonable assumption in practice. To simplify the problem, we further assume that e'_{ij} and e''_{ij} are independent of each other and are normally distributed. Note that in practice, it may not be possible to distinguish between e'_{ij} and e''_{ij}. But for our purpose, we write $e_{ij} = e'_{ij} + e''_{ij}$ and assume that e_{ij} are independent and identically distributed as trivariate normal $N_3(0, \sigma^2 I_3)$, where I_3 is the 3×3 identity matrix and σ^2 is a positive constant.

We do not discard a product simply because it has a tiny manufacture error. In fact, the real issue in practice is to decide how much manufacture error is acceptable in the presence of measurement error. We assume that the geometric quality of a product can be judged by the size of σ^2. That is, whether a product is of satisfactory quality depends on whether $\sigma^2 \leq \sigma_o^2$, where σ_o^2 is determined by the practical requirement on the manufacture precision and the precision of the CMM. In other words, we interpret the geometric quality assurance problem as a statistical hypothesis testing problem; the hypothesis to be tested is H_o: $\sigma^2 \leq \sigma_o^2$. It is noted that in Chapman and Kim (1992), a similar hypothesis has been stated using the concentration parameter of the Fisher-von Mises distribution on the sphere.

From model (2.1), one may try to estimate T, R and σ^2 by using the ordinary linear regression approach. This does not work directly because x_{ij} are unidentifiable points on S_i.

Instead, we suggest fitting the model by minimizing

$$Q(T,R) = \sum_{i=1}^{m}\sum_{j=1}^{n_i}\rho^2(Ry_{ij}+T,S_i) \tag{2.2}$$

over all possible T and R, where

$$\rho(y,S_i) = min_{x\in S_i} d(y,x)$$

and d is a distance which we choose as the Euclidean distance.

Let $\hat{T}$ and $\hat{R}$ be a translation and a rotation that produce the global minimum of $Q(T,R)$. We will study the properties of $Q(\hat{T},\hat{R})$ and its usage in the next subsection.

2.2 Some properties

Intuitively, $Q(\hat{T},\hat{R})$ should have a chi-square distribution when the errors are normal. However, this is not true in a strict mathematical sense. Nevertheless, we are able to claim the following properties:

Let $S = \bigcup_{i=1}^{m} S_i$ be the surface of a designed product in the CAD file. Assume that all S_i are planes. Then under model (2.1) and definition (2.2), we have

1. The distribution of $Q(\hat{T},\hat{R})/\sigma^2$ does not depend on σ^2.

2. Let R_o and T_o be one of the true transformation parameter sets between the coordinate systems of the CMM data and CAD file, then

$$Q(\hat{T},\hat{R})/\sigma^2 \leq Q(T_o,R_o)/\sigma^2 \sim \chi^2$$

where the degrees of freedom for χ^2 is $\sum_{i=1}^{m} n_i$.

3. If the normal directions of S_i do not fall into a single plane, then

$$Q(\hat{T},\hat{R})/\sigma^2 \leq \chi^2$$

where the degrees of freedom for χ^2 is $\sum_{i=1}^{m} n_i - 3$.

The proofs of these properties are straightforward except for that of property 3. We leave it in the appendix.

Property 1 ensures that $Q(\hat{T}, \hat{R})/\sigma^2$ is a pivotal. We can therefore test the hypothesis H_o as follows: Let F_α denote the upper α-percentile of the distribution of $Q(\hat{T}, \hat{R})/\sigma_o^2$. We reject H_o: $\sigma^2 \leq \sigma_o^2$ when $Q(\hat{T}, \hat{R})/\sigma_o^2 \geq F_\alpha$, where $0 \leq \alpha \leq 1$ is the predetermined significance level.

Although we do not know the exact size of F_α, Property 2 and 3 can give a good upper bound for it. Moreover, since T and R together have 6 free parameters, we expect $Q(T, R)/\sigma^2$ to lose 6 degrees of freedom in the chi-square distribution after minimization with respect to T and R. This conjecture is supported by our simulation results but not theoretically so far.

When some of S_i are not planes but curved surfaces, we expect the above properties to remain approximately true provided σ_o^2 is small. The rational is that smooth surfaces are flat in a small area. Also, we believe that the size of σ^2 compared to the size of a product must be small in practice; otherwise, we do not need very precise devices to detect the problem in the geometric quality of the product. Therefore, the above mentioned approximation is likely to be good even when some S_i are curved surfaces.

When an S_i is one dimensional, such as a curve, each measurement on S_i will contribute 2 degrees of freedom to the chi-square distribution instead of one degree of freedom. Thus, informally, the total degrees of freedom of a good approximate chi-square distribution for the most general case is

$$\nu = \sum_{i=1}^{m} n_i \times \{3 - dim(S_i)\} - 6, \tag{2.3}$$

where $dim(S_i) = 2$ if S_i is a surface, and $dim(S_i) = 1$ if S_i is a curve. However, when a product is invariant with respect to certain transformations, the loss of degrees of freedom will be less severe. For example, if the product is a sphere, then rotation is unnecessary, and the loss of degrees of freedom should be 3. Obviously, when ν is large, this difference is negligible.

2.3 Computation

Unlike linear regression, there exists no analytical solution to our minimization problem. We therefore consider solving the problem by numerical method. It is well known that almost all smooth functions are convex around its minimum points. The success of applying a nonlinear program to locate a minimum depends heavily on the choice of good initial values. Luckily, there is a very natural way to find a good initial value in our problem using its geometric characteristic.

Again, we use our simplex example to explain the idea. From the CAD file, we locate three points O, A and B. We choose O as a new origin, and set OA as a new x-axis. We then choose the upward normal direction to the plane containing O, A and B as a new z-axis, and choose our new y-axis by forming a right-hand coordinate system S_{CAD} with the new x and z axes. After this, we express our CAD file in terms of this new coordinate system S_{CAD}. For our CMM data, we can add three measurements $\tilde{O}$, $\tilde{A}$ and $\tilde{B}$ if such measurements do not exist yet. Then we can set up a new coordinate system S_{CMM} just as we do for our CAD file and express our CMM data in terms of S_{CMM}. (However, if there do not exist natural reference points O, A and B on the product, one may have to explore other features of the product. For example, if the product is a sphere, one can make use of the symmetry property.) Due to manufacture and measurement errors, system S_{CMM} is in general different from system S_{CAD}. But the physical closeness between (O, A, B) and $(\tilde{O}, \tilde{A}, \tilde{B})$ guarantees that S_{CMM} is very close to S_{CAD}. Therefore, in terms of the new coordinate systems, we can choose zero translation and identity rotation as our initial value for T and R, respectively.

More formally, since any rotation R can be written as the multiplication of three rotations R_1, R_2 and R_3, where

$$R_1 = \begin{pmatrix} \cos\theta_1 & -\sin\theta_1 & 0 \\ \sin\theta_1 & \cos\theta_1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, R_2 = \begin{pmatrix} \cos\theta_2 & 0 & -\sin\theta_2 \\ 0 & 1 & 0 \\ \sin\theta_2 & 0 & \cos\theta_2 \end{pmatrix}, R_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_3 & -\sin\theta_3 \\ 0 & \sin\theta_3 & \cos\theta_3 \end{pmatrix},$$

we can therefore minimize $Q(T, R)$ with respect to $\theta_1, \theta_2, \theta_3$ and T with initial values 0.

When all of S_i are planes, the expression for $Q(T, R)$ can be further simplified as

$$Q(T, R) = \sum_{i=1}^{m} \sum_{j=1}^{n_i} \langle Ry_{ij} + T - a_i, v_i \rangle^2, \tag{2.4}$$

where $\langle \cdot, \cdot \rangle$ denotes the usual inner product in three dimensional Euclidean space, a_i is an arbitrary point on S_i, and v_i is a unit normal vector to S_i.

3 Simulation

We conduct a limited simulation study to illustrate the method developed in section 2. Example 1 below deals with products that consist of planes only, while Example 2 deals with products that consist of various shapes.

Example 1. This is the example we have discussed. The assumed product consists of four planes. The unit normals of the four planes are $v_1 = (0, 0, -1)^t$ (the bottom plane), $v_2 = (-1, 0, 0)^t$ (the right plane), $v_3 = (0, -1, 0)^t$ (the left plane), and $v_4 = (1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3})^t$ (the top plane), respectively. Six points x_{ij} evenly distributed on each plane are measured. In our simulation, we use model (2.1) to generate y_{ij} for $i = 1, \ldots, 4$ and $j = 1, \ldots, 6$. The random number generator G05DDF from the NAG library is used to generate e_{ij} for a range of σ values.

In this example, the simplified formula (2.4) is used for minimization directly. To see whether our proposed method of choosing initial transformation works, we need to choose three points for setting up a new coordinate system for the CAD file, and another three points for setting up a new coordinate system for the CMM data. These points are chosen as $O = (0, 0, 0)^t$, $A = (1, 0, 0)^t$ and $B = (0, 1, 0)^t$ for the CAD file, and random errors are added to them to give $\tilde{O}$, $\tilde{A}$ and $\tilde{B}$ for the CMM data. The four other points on each of the four planes in (2.4) are chosen as $a_1 = a_2 = a_3 = (0, 0, 0)^t$ and $a_4 = (1, 0, 0)^t$.

For each value of σ, we repeat the simulation 10,000 times. The subroutine E04JAF from the NAG library is used to find the minimum of (2.4). It turns out that the subroutine can find the minimum for sure over 98% of times. In other cases, the subroutine is still able to find a minimum but it does not meet the criteria built in the subroutine. However, we find that these minima are good enough for our purpose. Also, Property 1 in section 1 is found to hold during the simulation. Therefore, we can concentrate on one σ value to study the distribution of $Q(\hat{T}, \hat{R})/\sigma^2$.

For $\sigma = 0.05$, we compare $Q(\hat{T}, \hat{R})/0.05^2$ with the chi-square distribution on 18 degrees of freedom, because $\nu = 4 \times 6 \times 1 - 6 = 18$. The following are some selected upper tail probabilities:

α	0.500	0.400	0.300	0.200	0.100	0.050	0.025	0.010
$P(Q/0.05^2 > \chi^2_{18,\alpha})$	0.5016	0.3981	0.3027	0.2006	0.0989	0.0516	0.0283	0.0117

where $P(\chi^2_{18} > \chi^2_{18,\alpha}) = \alpha$, and the second row is the proportion of $Q/0.05^2$ that exceed $\chi^2_{18,\alpha}$. Clearly, χ^2_{18} approximates the distribution of $Q/0.05^2$ very well. See also Figure 2 for a graphical comparison. Property 3 in section 2 concludes that χ^2_{21} is an upper bound for Q/σ^2 in this example. This is confirmed from the above simulation.

Instead of assuming that all manufacture errors are random, we now add some systematic errors as well. One way to do this is to assume that in (2.1), x_{ij} are points on new planes S_i' ($i = 1, \ldots, 3$) whose normal directions are $v_1 = (0, 0.04471018, -0.999)^t$, $v_2 = (-0.998, 0.06321392, 0)^t$, and $v_3 = (0.07740155, -0.997, 0)^t$. We repeat the above simulation with $\sigma = 0.05$ and find the following results:

α	0.500	0.400	0.300	0.200	0.100	0.050	0.025	0.010
$P(Q/0.05^2 > \chi^2_{18,\alpha})$	0.9541	0.9267	0.8889	0.8232	0.6973	0.5742	0.4634	0.3309.

See also Figure 3 for a graphical comparison. In terms of testing the hypothesis stated in section 2, H_o: $\sigma^2 \leq 0.05^2$ is more likely to be rejected than accepted. In other words, a product with systematic manufacture errors will more likely be judged to have unsatisfactory geometric quality by our method. Also, if a product has large undesirable manufacture variation ($\sigma > 0.05$), then according to property 1, it is likely that the product will be rejected using our method.

Example 2. In this example, the assumed product is like a screw; see Figure 4. The top S_1 is part of a sphere with a radius equal to 1.5 unit. The main body S_2 is a cylinder. Since it is possible that the locations and shapes of some curves are also important, we define S_3 as the bottom circle, S_4 as the circle on top of the cylinder and S_5 as the edge of S_1. The last three objects are one dimensional. Six evenly spaced points x_{ij} are taken from each of S_i, $i = 1, \ldots, 5$, and model (2.1) is again used to generate y_{ij} for $i = 1, \ldots, 5$ and $j = 1, \ldots, 6$. For a range of σ values, we repeat the simulation 10,000 times. For $\sigma = 0.05$, we have

α	0.500	0.400	0.300	0.200	0.100	0.050	0.025	0.010
$P(Q/0.05^2 > \chi^2_{43,\alpha})$	0.4964	0.3952	0.2968	0.1966	0.0964	0.0478	0.0238	0.0097
$P(Q/0.05^2 > \chi^2_{42,\alpha})$	0.5413	0.4374	0.3367	0.2269	0.1169	0.0591	0.0303	0.0119.

Because $\nu = 6+6+12+12+12-6 = 42$, one may expect χ^2_{42} to be as a reference distribution. However, χ^2_{43} approximates the distribution of $Q/0.05^2$ better than χ^2_{42} does. This is because the product is invariant to rotation about z-axis. This example illustrates that the distribution of $Q(\hat{T}, \hat{R})/\sigma^2$ also depends on the physical shape of the product.

Next, we introduce systematic manufacture errors into the product. Let $x_{1j}, j = 1, \ldots, 6$ be on a new sphere S'_1 whose radius is 1.6 unit. We repeat the simulation and find the following results:

α	0.500	0.400	0.300	0.200	0.100	0.050	0.025	0.010,
$P(Q/0.05^2 > \chi^2_{43,\alpha})$	0.9317	0.8958	0.8455	0.7669	0.6253	0.4896	0.3721	0.2515.

Also see Figure 6. In terms of testing the hypothesis stated in section 2, products with systematic manufacture errors are more likely to be judged to have bad geometric quality.

4 Some Remarks

From our simulation study, we can see that our method is rather successful in detecting undesirable large random or systematic manufacture errors. For the simple examples considered, it takes a fraction of a second to do the computation for one simulation.

Our simulation also reveals that χ^2_ν is indeed a good approximation to the distribution of $Q(\hat{T}, \hat{R})/\sigma^2$. However, our preliminary study (results are not shown here) shows that χ^2_ν is unlikely to be the true distribution. Some work is therefore needed to confirm that χ^2_ν is a good approximation in general.

Related to the geometric quality assurance problem, it is desirable to know which particular part of the product is not of satisfactory geometric quality. A possible answer is to break the total sum of squares into sums of squares due to each of $S_i, i = 1, \ldots, m$. Another more formal possibility is to drop the points on S_i in turn and fit the rest of data. This will reveal which part of the product is of the worst quality. Further research is necessary to develop diagnostic tools along these directions.

When the product has a sophisticated surface, the numerical calculation may not be as simple as it appears in our simulations. However, we believe that this is not a difficult problem for computer experts. We also believe that it should not be difficult to develop a computer program to implement our proposed method.

Acknowledgements

We want to thank Professors Robert Chapman and Peter Kim for their very helpful comments during the preparation of this paper.

Appendix

Proof of Property 3. Following the notation in section 2.3 and model (2.1), and noting that any vector lying on the plane S_i must be orthogonal to v_i, we can write

$$\begin{aligned} Q(T, R) &= \sum_{i=1}^{m} \sum_{j=1}^{n_i} \langle R y_{ij} + T - a_i, v_i \rangle^2 \\ &= \sum_{i=1}^{m} \sum_{j=1}^{n_i} \langle R(e_{ij} - e_{i\cdot}), v_i \rangle^2 + \sum_{i=1}^{m} \sum_{j=1}^{n_i} \langle R e_{i\cdot} - T, v_i \rangle^2 \end{aligned}$$

where $e_{i\cdot} = n_i^{-1} \sum_{j=1}^{n_i} e_{ij}$. Thus,

$$\min_{T,R} Q(T, R) \leq \sum_{i=1}^{m} \sum_{j=1}^{n_i} \langle R_o(e_{ij} - e_{i\cdot}), v_i \rangle^2 + \min_{T} \sum_{i=1}^{m} n_i \langle R_o e_{i\cdot} - T, v_i \rangle^2$$

where R_o is any fixed rotation matrix. Clearly, the last two terms in the above inequality are independent of each other. The first term has a chi-square distribution with degrees of freedom $\sum_{i=1}^{m}(n_i - 1)$.

Since R_o is fixed, without loss of generality, we can assume that $R_o = I_3$. The solution for T in the second term is then

$$\hat{T} = (\sum_{i=1}^{m} n_i v_i v_i^t)^{-1}(\sum_{i=1}^{m} n_i v_i v_i^t e_{i\cdot}).$$

It is not difficult to verify that $\sqrt{n_i} v_i^t e_{i\cdot}$ are independent and identically distributed standard normal random variables. Let $\mathbf{y}$ be the column vector with $\sqrt{n_i} v_i^t e_{i\cdot}$ as its ith component, and let C be the matrix with $n_i v_i^t (\sum_{i=1}^{m} n_i v_i v_i^t)^{-1} v_j$ as its (i,j)th element. Then the second term becomes $\mathbf{y}^t(I - C)\mathbf{y}$. Because the matrix $I - C$ is idempotent, it follows that the second term has a chi-square distribution with degrees of freedom equal to

$$\begin{aligned} tr(I - C) &= m - \sum_{k=1}^{m}\{n_k v_k^t (\sum_{i=1}^{m} n_i v_i v_i^t)^{-1} v_k\} \\ &= m - tr\{\sum_{k=1}^{m} n_k (\sum_{i=1}^{m} n_i v_i v_i^t)^{-1} v_k v_k^t\} = m - 3. \end{aligned}$$

This proves Property 3.

References

Chapman, G.R. and Kim, P.T. (1992), "Spherical regression in quality assurance", Research Report RR–92–01, The Institute for Improvement in Quality and Productivity, University of Waterloo, Waterloo, Ontario N2L 3G1.

The NAG Fortran Library, (1987), Numerical Algorithms Group Inc, 1101 31st Street, Suite 100, Downers Grove, IL 60515, U.S.A.

Figure 1: The simplex product, Example 1.

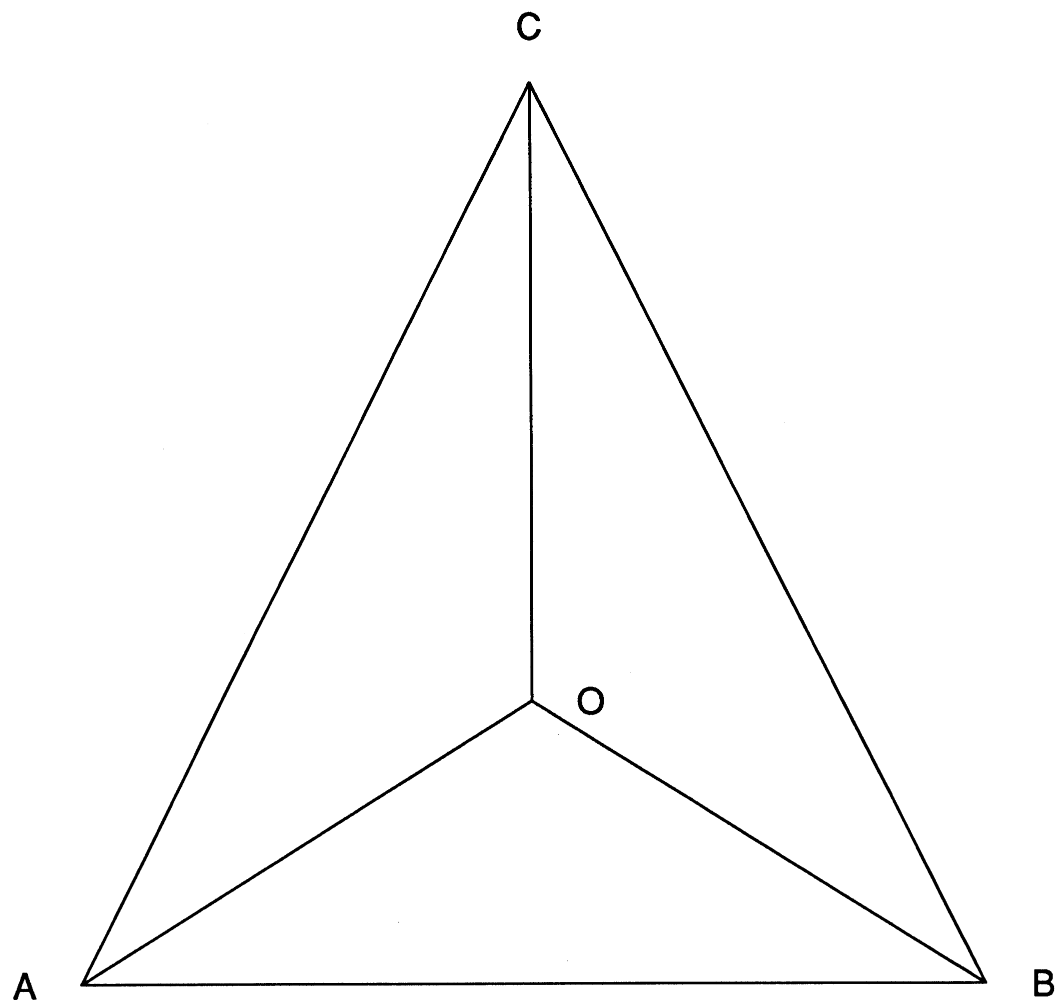


Figure 2: The simulated density (the histogram) when the manufacture errors are equal to the specified tolerance level 0.05, and the density for χ^2_{18} (the solid line), Example 1.

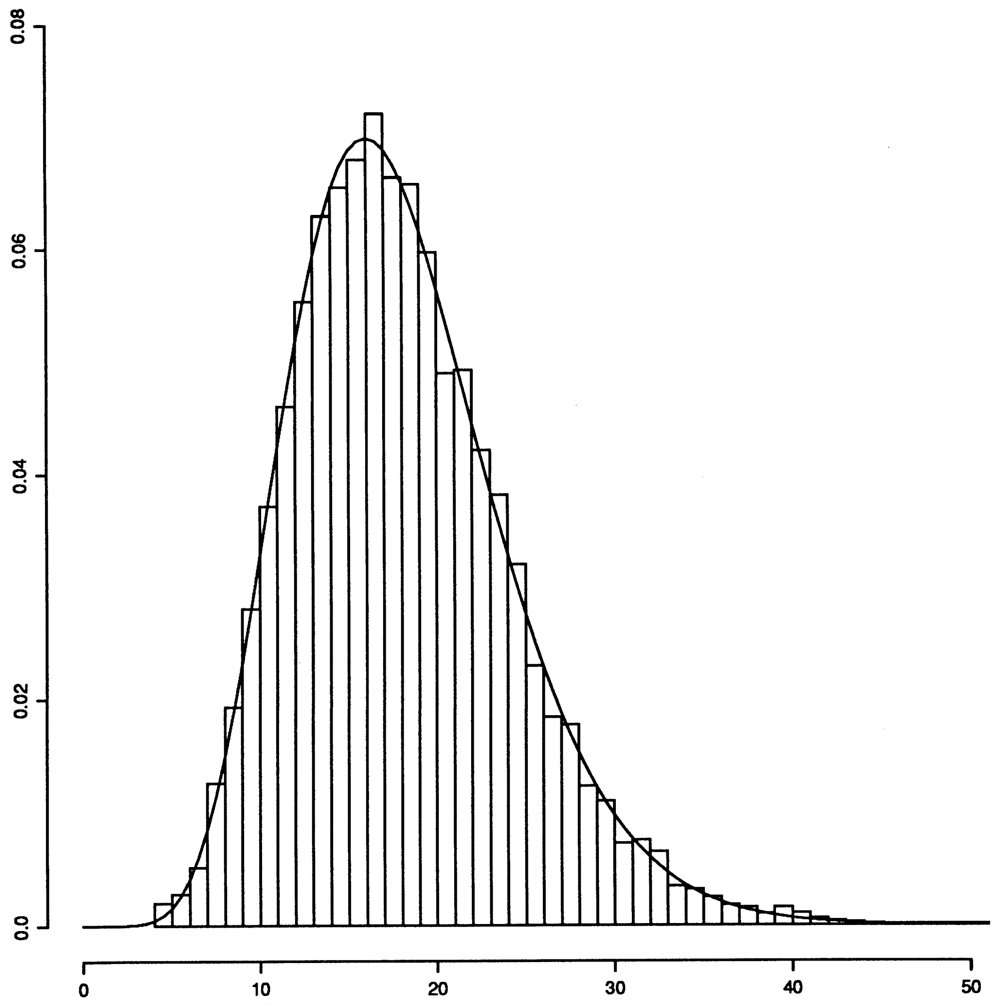


Figure 3: The simulated density (the histogram) when some systematic manufacture errors are added to the simplex product, and the density for χ^2_{18} (the solid line), Example 1.

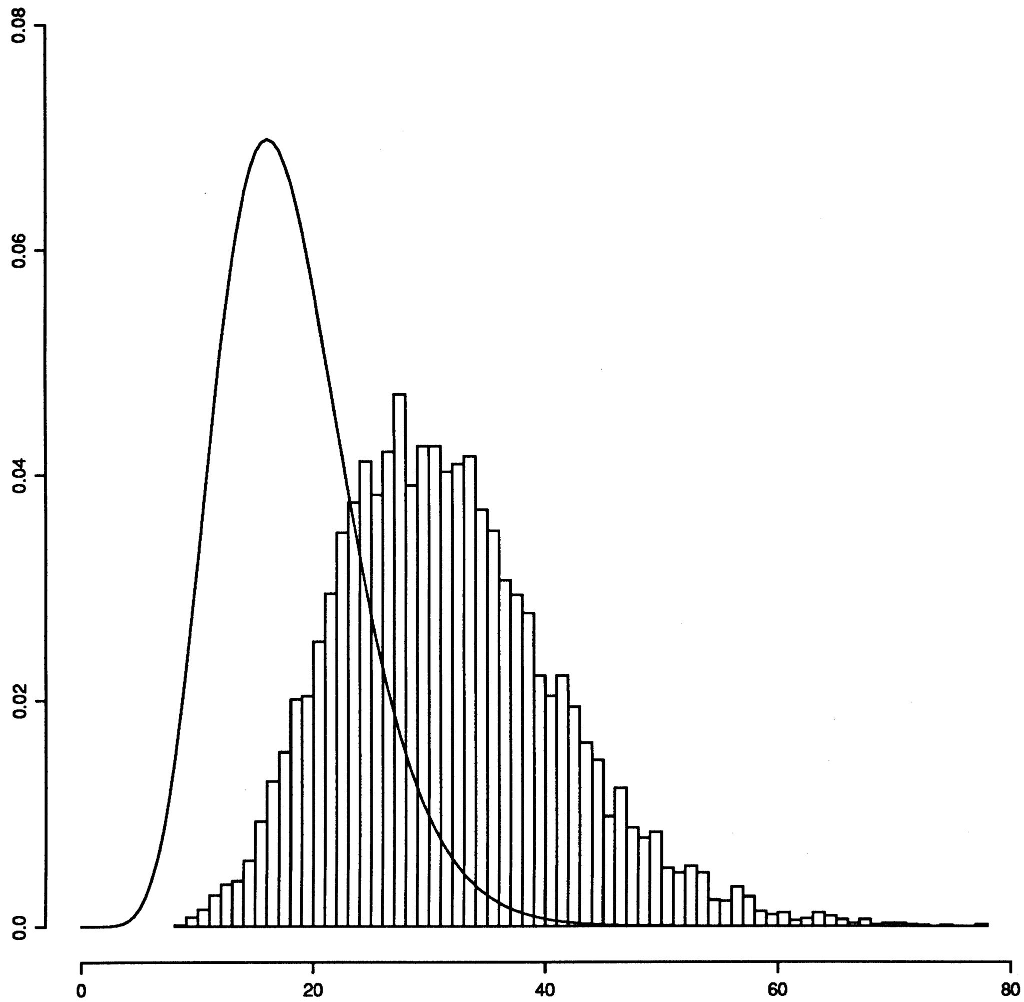


Figure 4: The screw-like product, Example 2.

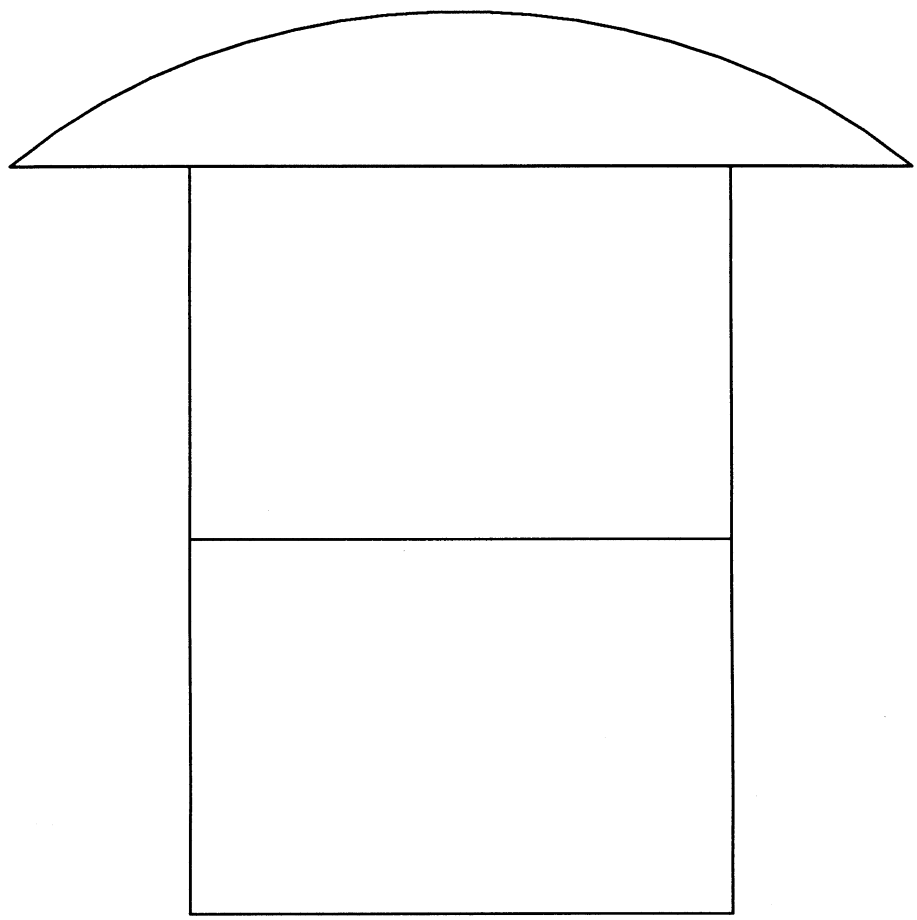
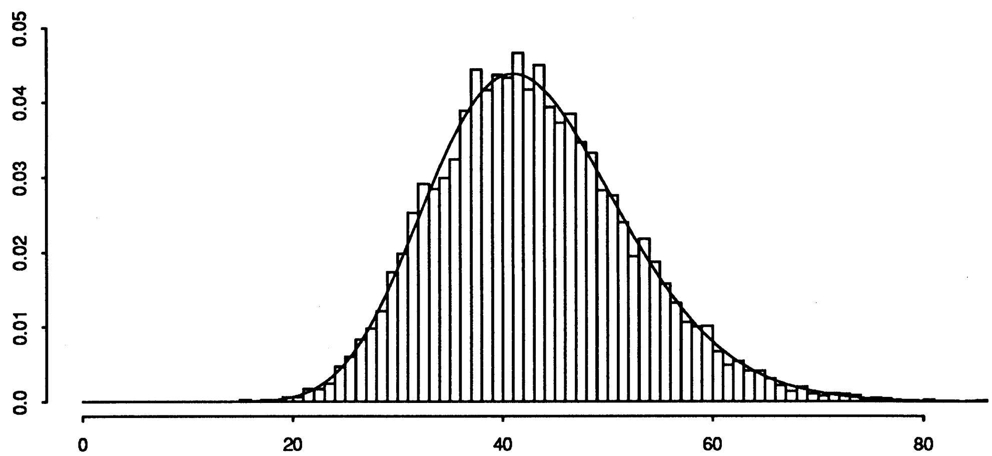


Figure 5: The simulated density (the histogram) when the manufacture errors are equal to the specified tolerance level 0.05, and the density for χ^2_{43} (the top plot), or the density for χ^2_{42} (the bottom plot), Example 2.



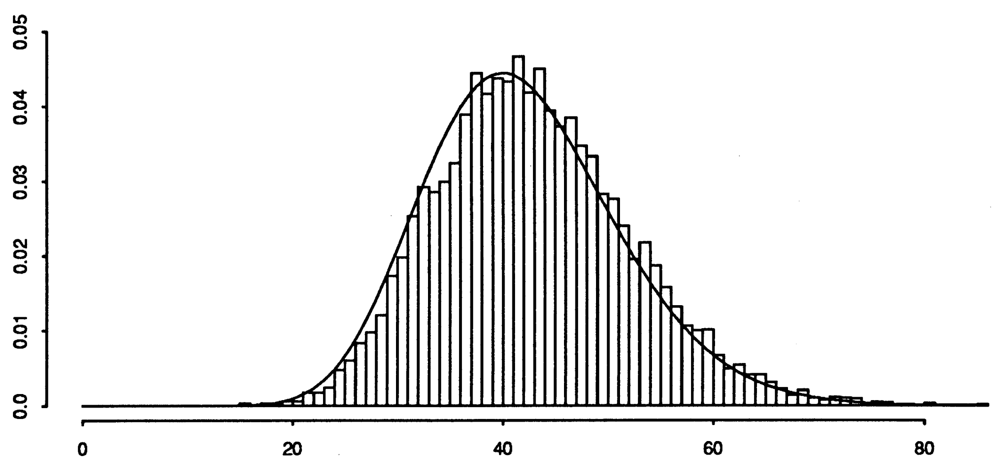


Figure 6: The simulated density (the histogram) when the radius of the sphere is made 0.1 unit larger than designed, and the density for χ^2_{43}, Example 2.