

Robust Estimation of Optimal Sample Size for CMM Measurements with Statistical Tolerance Limits

Petr Chelishchev¹, Aleksandr Popov² and Knut Sørby¹

¹Department of Mechanical and Industrial Engineering, NTNU, NO-7491 Trondheim, Norway

²Department of Mathematics, Baltic State Technical University, St. Petersburg, Russia

Abstract. The paper proposes the kernel probability density function approach to estimate the distribution of measurements on a part which is measured in a coordinate measuring machine (CMM). The study is based on the experimental data derived from internal cylinder measurements. The distribution free model suggested by Wilks was used as a reference for the selection of the sample size. Three cross sections of a cylinder were measured regarding to this reference. The work defines the minimum required sample size for obtaining at least 0.95 proportion of radius variation for particular studied cylindrical part with 95% confidence level.

1 Introduction

The main goal of Geometric Dimensioning and Tolerancing (GD&T) inspection of a part is to assess if the geometry and dimensions of the part are inside of the specified tolerance limits to verify that an assembly fits, or that the intended functionality of the part is guaranteed. This paper deals with verification of radius size variation of internal cylinder.

Coordinate measuring machines (CMMs) are universal and widely employed automated measuring systems in industry [1]. One of the most critical parameter of measuring strategy with CMM is the number of measuring points that is used to extract data from the part features. Obviously, a greater number of measuring points provides a better accuracy, however it leads to higher time consumption and costs. Since the accuracy requirement in a design specification is defined by the tolerance interval, within which the part dimension or geometry may vary, then evidently, it should exist a certain number of points sufficient enough to confirm with some given probability if the size is inside of the tolerance limits or not.

The influence of sample size on the measurement result has been widely discussed, and several different approaches has been used to estimate the contribution to the measurement uncertainty. Approaches such as statistical methods [2,3] (for normal distribution), fuzzy logic [4], genetic algorithm [5], extended zone model optimization [6], adaptive sample strategy with use of Kriging models [7] and analytical methods with implementation uncertainty simulations [8, 9] have been suggested. However, a standard guide or criterion for sample strategy with CMM GD&T inspection has not been established yet.

According to [10] “a statistical tolerance interval is an estimated interval, based on a sample, which can be asserted with confidence level $1-\alpha$ to contain at least a specified proportion p of the items in the population. The limits of a statistical tolerance interval are called statistical tolerance limits.” Theoretically, the statistical tolerance interval for the case of normal distributed data set is the most well developed method [11, 12]. There are tabulated data in international standard ISO16269-6 for calculating both one-side and two-sided statistical tolerance intervals for sample size $n \geq 2$ and the at least population proportion p for such confidence levels $100(1-\alpha)\%$ as 90%, 95%, 99%, 99.9%. However, if other tolerance intervals (in form of $\pm k\sigma$) need to be found, which are not provided by the standard, for example other p and/or not typical $1-\alpha$ value, or other sample size, then the *K.factor* function from the “tolerance” package in R programming language may be employed to calculate the factor k .

This paper provides an approach for estimation of an optimal number of the measuring points for a two-sided statistical tolerance interval based on a distribution-free model. A continuous probability density function (pdf) from measurements from a workpiece was approximated by kernel density estimator (KDE). The estimated continuous pdf was further used to simulate different sample strategies and evaluation of the confidence level for detecting at least 0.95 content of total radius variation range.

2 Data and experimental study

An internal cylindrical hole of an aluminum workpiece produced by a turning operation with nominal diameter 60 mm and length 130 mm was inspected in a CMM (Leitz PMM-C-600) with an analogue probe to detect a largest possible deviation of radius variable. The cylinder axis was aligned with the vertical axis (z axis) of the CMM. Three cross sections of the cylinder are measured: the first close to the top (Section A), the second in the middle (Section B), and the third on the bottom (Section C).

Table 1. Shapiro-Wilk normality test for 473 points sample

Sections	Section A	Section B	Section C
P-value	5.392e-07	0.007842	0.01435

There are uniformly distributed 473 points coordinates (x_i, y_i) measured around the circle in each section, and a least squares circle (LSC) method was utilized to calculate the circle centre. The radius variable r_i , for each measured point was calculated by:

$$r_i = \sqrt{(x_i - x_c)^2 + (y_i - y_c)^2} \tag{1}$$

The uncertainty of the CMM itself is about 10 times less than inspected radius variance range, and thus it is not considered in the analysis of sample size.

Intuitively it is clear that less number of points may provide lower measurement accuracy due to the probability that extreme points on the feature are missing in the extracted data set. We have used MATLAB source for our simulation approach to investigate the degree of influence of sample size on the inspection confidence level and the detected radius variation range.

It is always advisable to evaluate the normality of a distribution of the original data set in the very beginning. The Shapiro-Wilk normality test [13] was applied to the measured data sets by use of the *shapiro.test* function in the R programming language. The results are shown in Table 1. The extremely lower p-value (especially Section A and Section B) yields us a reason to reject the assumption about normal distribution of the measurements.

2.1 Distribution-free model

The distribution of the radius variable of the part is not known before we start the measurements. We will therefore suggest to use the Wilks criterion [14, 15] to define the minimum sample size. The criterion is based on the order statistic. It postulates the following: *if an investigated random characteristic belongs to a population of any unknown continuous distribution function, then at least a content p of the population included between the smallest observation r_{min} and the largest observation r_{max} of the data sample with confidence level $(1-\alpha)$, and a required minimum sample size n_{min} .* For the two-sided tolerance interval with the conditions determined above, can be expressed by following [10]:

$$n_{min} \cdot p^{n_{min}-1} - (n_{min} - 1) \cdot p^{n_{min}} \leq \alpha . \tag{2}$$

Results computed by (2) of minimal sample size for the two-sided statistical tolerance limits (between the first and n -th order of sample order statistic) with unknown continuous distribution, and predefined $(1-\alpha)$ and p , are shown in Table 2. As long as the number of measuring points supposed to be the natural numbers ($n_{min} \in \mathbb{N}$), negative solutions were not considered, and all results of Table 2 were rounded to the nearest upper integer. This particular fact together with the distribution independency of (2) provides a robust property of the method, which is further going to be confirmed by the experiment data and a simulation model.

Table 2. Minimal sample size n_{min} for proportion p and confidence level $1-\alpha$

Confidence level, $100(1-\alpha)\%$	Proportion of population, p				
	0.500	0.750	0.900	0.950	0.990
50	3	7	17	34	168
75	5	10	27	53	269
90	7	15	38	77	388
95	8	18	46	93	473
99	11	24	64	130	662

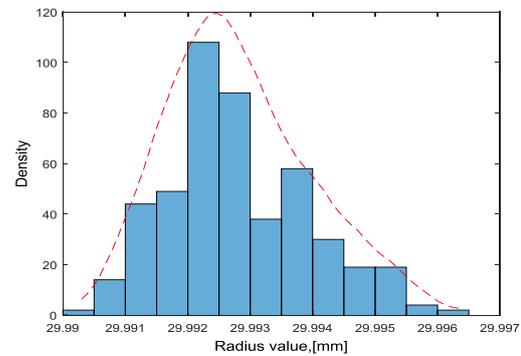


Figure 1. A histogram and kernel density estimate of $f(r)$ for Section A with sample size 473 points

Regarding to the above description for the two-sided statistical tolerance interval with $p = 0.99$ and $100 \cdot (1-\alpha)\% = 95\%$ (the actual confidence level is 95.02%), the minimum sample size is 473 (Table 2). That is the total number of measuring points we used in our inspection of the cylinder sections.

2.2 Kernel density estimation

In practice, the data distribution is often unknown and/or may contain outliers. Hence, it is reasonable to estimate the tolerance intervals based on more general assumptions when it is impossible to describe sample data with any known standard distribution functions. A possible solution in such case is an estimation of the pdf directly from the measured data sample. A non-parametric statistic may be used in this way. One possibility to do that is the well-known histogram technique. However, the histogram suggests a distribution

interpretation only in form of bins and it is less useful for further appliance due to lack of continuity. Meanwhile a limited number of known pdf $f(r)$ are available to describe a continuous-valued random variable (logarithmic, exponential and so on). To avoid such restrictive assumptions about the form of $f(r)$ the KDE may be applied [16]. Further, using the kernel estimator based on original measured data, an opportunity appears to generate any different random data samples regarding to the initial data distribution.

2.3 Kernels and weighting function

Similar to the histogram we need an estimator of $f(r)$. The probability that random variable is within of the interval $r \pm b$ can be written as following:

$$P(r-b < R < r+b) = \int_{r-b}^{r+b} f(\delta)d\delta \approx 2bf(r), \quad (3)$$

and hence

$$f(r) \approx \frac{1}{2b} P(r-b < R < r+b). \quad (4)$$

Alternatively, the frequency for the given interval could be estimated by the equation:

$$\hat{f}(r) = \frac{1}{n} \sum_{i=1}^n w(r-r_i, b), \quad -\infty < r < \infty \quad (5)$$

where the estimator $\hat{f}(r)$ has the properties of a pdf, i.e. positive for any r and an integral area equal to 1. Then a weighting function $w(\delta, b)$ can be generalized in this way:

$$w(\delta, b) = \frac{1}{b} K\left(\frac{\delta}{b}\right), \quad (6)$$

where b is the *bandwidth* or *smoothing* constant of weighting function and K is the standardized weighting function (with $b = 1$), which is the *kernel*.

2.4 Kernels' parameters

The degree of smoothing of $\hat{f}(r)$ depends on parameters of $w(\delta, b)$ such as the kernel K and the bandwidth b , which determine a shape and a width of the weighting function respectively. The proper choice of K and b is a subject of an optimization problem.

The accuracy of the kernel density estimator can be evaluated by mean squared error (MSE), mean integrated squared error (MISE) and asymptotic mean integrated squared error (AMISE).

According to previous research [17] *Epanechnikov* function was defined as the optimal kernel in respect to $MISE(\hat{f})$:

$$K(\delta) = \begin{cases} \frac{3}{4\sqrt{5}} \left(1 - \frac{1}{5}\delta^2\right) & \text{for } |\delta| < \sqrt{5} \\ 0 & \text{otherwise.} \end{cases} \quad (7)$$

The bandwidth b depends on different factors e.g. unknown pdf $f(r)$, kernel type, number of observations in the sample and so on. There are a number of methods available to optimize the bandwidth parameter such like

bias cross-validation (BCV), unbiased cross-validation (UCV), direct plug-in rule (DPI) and others. The methods can have a different performance dependently on the estimation function $\hat{f}(r)$ used and the pdf $f(r)$ estimated. Thus, we use *Epanechnikov* kernel and default MATLAB bandwidth estimation in this study.

2.5 Date estimation by kernel function

The radius variable r_i used in simulation was computed by (1) with assumption of a unique circle centre, which was obtained from LSC based on 473 measurements points. The rounding of the values by one decimal place ($1 \cdot 10^{-4+1}$ mm), on the one hand allows considering the cylinder form tendency and possible outliers, and on the other hand do not take into account unnecessary accuracy requirements to the data estimated circle centre coordinates.

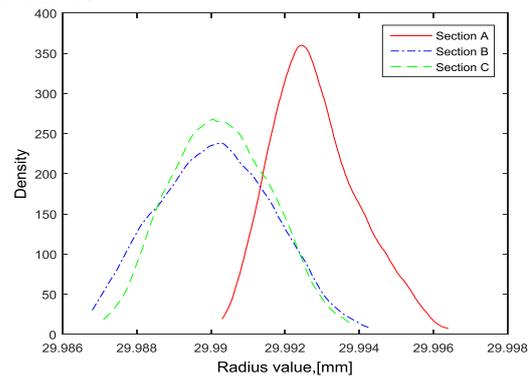


Figure 2. Kernel estimates $\hat{f}_A(r)$, $\hat{f}_B(r)$, $\hat{f}_C(r)$ for three sections based on 473 points sample size machine.

Estimates of pdf $f(r)$ for the three cylinder sections based on Epanechnikov kernel and the sample size 473 measuring points are shown on Fig. 2. Observing the curves one can notice a distinction of the distribution parameters such as the mean values, the variations and the data spread for the cross-sections, which belong to the same cylinder.

All these mentioned facts together with rounding of the sample point numbers give us the reason to presume that the small centre coordinate offsets can be neglected. The robustness of the simulation model is discussed in the next sections.

2.6 Estimation of an optimal sample size

In order to discover an optimal sample strategy for the inspection of the part, a statistical simulation was carried out in MATLAB, by using the KDE $\hat{f}_A(r)$, $\hat{f}_B(r)$, $\hat{f}_C(r)$, estimated from the CMM measurements of workpiece, see Fig. 2.

The eight initially predefined different sample sizes $n = \{5; 10; 15; 30; 60; 90; 93; 95\}$ were simulated with 10^5 iterations for each sample size n_i . The maximum r_{max} and the minimum r_{min} values were detected for every new generated sample. The population content p for the each iteration was evaluated as a difference of the

Table 3. The statistical test simulation with 10^5 iterations for each n_i sample size

<i>Section A</i>								
Sample size n	5	10	15	30	60	90	93	95
Probability P%, ($p \geq 0.95$) ^a	2.2	8.5	16.8	44.9	81.2	94.5	95.3	95.5
<i>Section B</i>								
Sample size n	5	10	15	30	60	90	93	95
Probability P%, ($p \geq 0.95$)	2.1	8.5	16.4	44.3	80.9	94	95	95.4
<i>Section C</i>								
Sample size n	5	10	15	30	60	90	93	95
Probability P%, ($p \geq 0.95$)	1.9	8.5	18	45.1	81.3	93.6	94.9	95.5

a. Probability P of covering a range with population content p

cumulative distribution functions (cdf) of maximum and minimum random variable r_i , based on KDE. Then conditions of equality/exciding 0.95 of total radius variation range was tested by:

$$p(r_{\min} < r < r_{\max}) = F_R(r_{\max}) - F_R(r_{\min}) \geq 0.95, \quad (8)$$

where $F_R(r)$ is cdf of a real-valued (either maximum or minimum) random variable r_i , calculated with the kernel pdf estimator $\hat{f}(r)$. The number of successful iterations was assigned as 1 (or 0 – otherwise) and summed up as Sum_N . The final probability P% for each sample size n_i was calculated as a rate of Sum_N / M , where M is the total iteration number.

In spite of the predefined initial sample size (473 points) taken from the Table 2 with given $p = 0.99$ and $100(1 - \alpha)\% = 95\%$, the total area under the estimated kernel function is equal to 1 (from pdf properties). That gives as an opportunity to generate any size of data sample even larger than the initial sample size. This in turn leads to independency of our simulated results presented in Table 3 on the initial parameters of the model (2).

3 Discussion of results

Analysis of measurements obtained from cylinder sections shows that parts produced by a turning operation has unknown non-normal distribution of radius variables. The parameters (e.g. mean, variance, skewness) of the distributions in difference sections of the cylinder are apart from each other. However, the sample strategy according to table 3 for the sections is finally the same. Thus, the applied simulation model based on the experimental data confirms the robustness property of the method proposed by (2) in section 1. For example, Table 3 shows that the optimal sample size is about 93 points for all sections, which agrees with the data in Table 2 (for $100(1 - \alpha)\% = 95\%$, $p = 0.950$). Thereby the simulation based on distributions estimated by kernel function confirms Wilks model given by equation (2). We can also notice that the probability to estimate at least 0.95

fraction of radius variation range is only about 2% in the case of 5 points sample.

This fact gives us the reason to expect that further research of cylindrical parts with larger radius values, from other machine operations and different materials of workpiece, most likely will provide the similar results.

In the simulated model, the circle centre is assumed the same for any data samples. For different sample points the centre point would vary, but maximum possible range between r_{\min} and r_{\max} remains a similar. In addition, the minimum number of points n_{\min} (Table 2) was rounded to the nearest upper integer, thus it makes negligible the influence of the centre coordinates. Again, a good compliance of the simulation results with the distribution free model given in formula (2) demonstrate the insignificant influence of the assumptions about the centre of coordinates. That also confirm that the simulation model itself employs a robust principle. Namely, the identical optimal number of points (Table 3) for different sections with observably diverse distributions improves this statement.

4 Conclusion

The innovation of this work is to show the possibility to use the distribution free model (2) to predict the sample size, its least content and confidence level for GD&T inspection with CMM before any measurements are performed.

In addition, the simulation model for robust estimation of the optimal sample size based on the experimental measuring date has been developed. The provided simulation procedure allows evaluating the sample sizes and their confidence levels for real cylindrical components in industry independently of their dimensions and machining process accuracy. Moreover, the finite sample sizes, which often used in industry, were evaluated. The obtained results demonstrate the particular low confidence level especially for the sample sizes from 5 to 30 measuring points.

The inspection sample size in production is often defined with cost and time-consumption in mind, and thereby it could be too small. The applied technique provide the demanded guidance criteria based on the confidence level and the real data distribution for

choosing the proper measuring sample strategy for GD&T inspection with CMM in manufacture.

In solving of a practical problem, it is recommended to evaluate the distribution of the initial data in the very beginning. If the data distribution is close to the normal distribution then the standardized procedure (ISO16269-6) can be used to estimate the tolerance interval limits. Otherwise, the original distribution based on the CMM measurements with predefined confidence level $1-\alpha$, the variation proportion p and the minimum sample size n_{\min} should be estimated by (2). Then the smallest and the largest order statistics of the sample should be used as the tolerance limits.

References

1. Chiffre, L.D., *Geometrical Metrology and Machine Testing*, DTU Mechanical Engineering, (2011)
2. Jiang, B. and S.-D. Chiu, *Form tolerance-based measurement points determination with CMM*, Journal of Intelligent Manufacturing, **13**(2): p. 101-108, (2002)
3. Hong-Tznong Yau, C.-H.M., *An automated dimensional inspection environment*, (1992)
4. N. Cappetti, A.N., F. Vilecco, *Fuzzy approach to measures correction on Coordinate Measuring Machines: The case of hole-diameter verification*, Measurement, **93** (Elsevier Ltd.): p. 41-47, (2016)
5. Changcai Cui, S.F., Fugui Huang, *Research on the uncertainties from different form error evaluation methods by CMM sampling*, Int J Adv Manuf Technol, **43**: p. 136-145, (2008)
6. K. D. Summerhays, R.P. Henke, J. M. Baldwin, R. M. Cassou and C. W. Brown, *Optimizing discrete point sample patterns and measurement data analysis on internal cylindrical surfaces with systematic form deviations*, Precision Engineering, **26**(1): p. 105-121, (2002)
7. G. Barbato, E.M.B., P. Pedone, D. Romano, G. Vicario, *Sampling point sequential determination by kriging for tolerance verification with CMM*, in Proceedings of the 9th Biennial ASME Conference on Engineering Systems Design and Analysis, ESDA08, ASME: Israel. p. 10, (2008)
8. S. Ruffa, G.D. Panciani, F. Ricci, and G. Vicario, *Assessing measurement uncertainty in CMM measurements: comparison of different approaches*, Int.J.Metrol. Qual. Eng. **4**, p. 163-168, (2013)
9. A.Weckenmann, M. Knauer, H. Kunzmann, *The influence of measurement strategy on the uncertainty of CMM measurements*, **47**: p. 451-454, (1998)
10. ISO16269-6, *Determination of statistical tolerance intervals*, in Statistical interpretation of data, (2014)
11. I. Janiga, I. Garaj, V. Witkovský, *On Exact Two-Sided Statistical Tolerance Intervals for normal distributions with unknown means and unknown common variability*, Journal of Mathematics and Technology, **3**: p. 25-32, (2012)
12. Janiga, I., Garaj, I., *On exact two-sided statistical tolerance intervals for normal distributions with unknown means and unknown common variability*, In 2009 Quality and Productivity Research Conference, Yorktown Heights, New York: IBM Thomas J. Watson Research Center, (2009)
13. Royston, P., *An extension of Shapiro and Wilk's W test for normality to large samples*, Applied Statistics, **31**: p. 151-124, (1982)
14. Wilks S.S., *Determination of Sample Sizes for Setting Tolerance Limits*, AMS Ann. Math. Statist., **12**: p. 91-96, (1941)
15. Wilks S.S., *Statistical Prediction with Special Reference to the Problem of Tolerance Limits*, AMS Ann. Math. Statist., **13**: p. 400-409, (1942)
16. Alexandre B.T., *Introduction to Nonparametric Estimation*, New York: Springer-Verlag, (2009)
17. Wand M.P., Jones M.C., *Kernel Smoothing*, Monographs on statistics and applied probability, Boca Raton, FlaChapman&Hall/CRC, (1995)