DOE/NV/25946--2133

Statistical Distributions from Lens Manufacturing Data

Morris I. Kaufman,[*,a] Brandon B. Light,[b] Robert M. Malone,[a] Michael K. Gregory[b]
[a]National Security Technologies, LLC, Los Alamos Operations; [b]Optimax Systems

ABSTRACT

Optical designers assume a mathematically derived statistical distribution of the relevant design parameters for their Monte Carlo tolerancing simulation. Presented are measured distributions using lens manufacturing data to better inform the decision-making process.

optical design, statistics, optomechanical, manufacturing, distribution, nonparametric, tolerance

1. INTRODUCTION

Generating optical tolerances in lens design software such as Code V or Zemax begins by modeling a perturbed optical system to capture the effect of simultaneous changes in lens radius, center thickness, and other critical parameters in amounts matching the lens manufacturer's stated capabilities. The analyst would assume a statistical distribution of the relevant parameters for the Monte Carlo simulation. In some cases, the assumed distributions do not resemble the distributions that occur in lens manufacturing. For example, optical designers assume a center thickness tolerance that is symmetric about the nominal value, but lens manufacturing brings parts to tolerance, not nominal. The best option would use actual manufacturing data to build a statistical distribution for the Monte Carlo simulation. Understanding true distributions will have other beneficial consequences. The designer or manufacturer could easily predict the cost of changing a particular tolerance (as a result of lost yield) for a given manufacturing process.

Code V[1] offers the designer a relatively conservative default choice of uniform distribution for the *TOLMONTE* analysis (i.e., Monte Carlo perturbations of the system to derive tolerances), although other choices, such as a truncated Gaussian distribution, are available. However, all of the readily available options use distributions symmetric about the nominal value. The *TOR* option (i.e., wavefront differential tolerancing) also permits an *END* option that only allows for worst-case values of ±T where T is the maximum permissible tolerance.

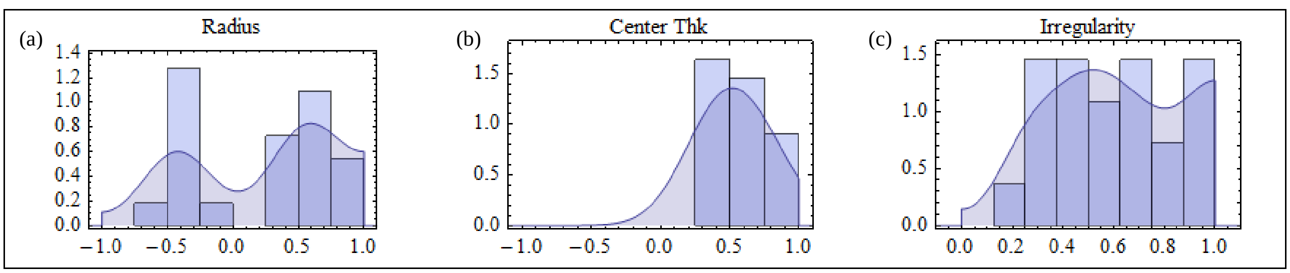

Figure 1. Binned and smooth histograms from a typical lens batch with 22 lenses. The vertical axes of all binned histograms are normalized.

Measured lens data do not fit neatly into any mathematically derived format (Fig. 1). Often, measured distributions are bimodal, indicating more than one process is at work, and distributions generally do not drop to zero at the edges. Taken together, these two facts suggest some of the lenses are polished a second time to pass inspection. The factors that influence the shape of the distribution include shape (convex/concave), softness of the glass, the precision grade, the R-number (radius to diameter ratio), and the size of the batch (how many lenses being produced in the run). Manufacturers typically leave center thickness larger than the nominal value in case rework is needed. Manufacturing can only remove material, so radii are ground and polished starting larger (i.e., flatter) than nominal for a convex surface and smaller than nominal for a concave surface. Because of the rework phenomenon, the combined mixture distribution for

[*] kaufmami@nv.doe.gov; phone: 1 505-663-2034; fax: 1 505-663-2003

radius looks very much like a slightly biased uniform distribution. The center thickness data on the other hand, show a strong preference to be thicker than the nominal value.

The fact that the mean value of center thickness is offset from the intended nominal value would create systematic deviations from the intended overall metric, such as wavefront error. This fact can be exacerbated because the optomechanical mounting assumes that the nominal value is also the expected value. The end result is a systematic error in the distance-to-next value between optical surfaces.

As the lens manufacturer ramps up from prototyping to large-scale production, defined here as a >20x increase in batch size, the nature of these errors and the resulting measurement distribution shape may change. Fabrication choices depend on order quantity to address daily output and often price targets.

In order to compare lens data for a variety of conditions, center thickness and radius data were normalized from ±T to ±1. The surface irregularity and wedge data were normalized to (0, 1). The normalization also has the effect of removing the possibility of revealing proprietary data (not a small consideration). One would expect that as the absolute value of these tolerances gets tighter, the normalized standard deviation would increase compared to the interval (and in some cases this is true). One would also expect that as the batch size increases, the distribution shape would more closely resemble a truncated Gaussian distribution, but this is shown not to be true.

Our study examined seven lens batches. The word "batch" in this context refers to an order of nominally identical lens elements. In this study, the word "aggregate" refers to the statistics of all of the batches or lens elements mixed together. The manufacturing process for the batches was not disclosed. The lens data were supplied by Optimax Systems, Inc. All lenses were ground from flat cylindrical stock. Batches ranged in size from 13 to 48 lenses, totaling 182 lenses; of the 14 surfaces produced, eight were convex and six were concave. The R-number ranged from 0.8 to 15. The glass hardness is hypothesized to be an important factor in the distribution shape, but the present study did not include glass hardness as a parameter.

The subject of optical tolerancing has a rich history[2], but there are no available studies that connect the parameters of lens grinding and polishing probability density function (PDF) data to the tolerancing practices of optical engineers. Juergens[3] states that the engineers at Raytheon routinely look at lens production data from their suppliers and feed the resulting PDF estimations back into their Code V Monte Carlo tolerancing analysis. But these data are not available to the public. In fact, design houses vigilantly guard their optical designs, and fabrication houses are not anxious to give competitors insight into the true capabilities of their manufacturing practices.

2. ANALYSIS METHODS

Two general approaches exist for analyzing this type of data: parametric and nonparametric methods. Any method needs to consider the facts that the data are, by nature, bounded on some interval and may be bimodal.

2.1 Parametric methods

In a parametric approach the analyst selects a distribution, such as the Gaussian or Pearson, and then estimates the parameters for the distribution. There are special distributions for bounded data sets, such as the Beta distribution, but this distribution does not appear to be applicable for the data under consideration. Another way of analyzing bounded data is to truncate a non-bounded distribution. Generally speaking, the truncation of a distribution f(x) restricts the support to the interval (x_{min}, x_{max}). The new truncated version of f(x) is given by $f(x)/(F(x_{max}) - F(x_{min}))$, where F(x) is the cumulative distribution function of f(x).

To clear up a confusing point, must mention that the standard deviation parameter one would use as an input into a Gaussian distribution is not the same as the standard distribution of the truncated Gaussian distribution. As you might imagine, the truncated Gaussian distribution starts resembling the uniform distribution (Fig. 2) as the standard deviation parameter goes to infinity. Our intuition regarding the meaning of the standard deviation is confounded when a distribution is truncated.

Bimodal data can be modeled using a mixture of two distributions. A well-known test case of bimodal data is the wait-times between eruptions of the Old Faithful Geyser.[4–7] If we use a mixture of two Gaussian distributions to analyze the data, we need to estimate five parameters: two means, two standard deviations, and a weighting parameter. Several methods exist for estimating these parameters by using the function *EstimatedDistribution* in the Mathematica software package,[8] including the "maximum likelihood" and "method of moments" algorithms. Similar functions exist in other software.[7] Mixture distributions are implemented in Mathematica using the function *MixtureDistribution*. As the number of parameters increases, numerical problems also increase and the parameter values would have more variance.[9] Having said that, if there are reasons to suspect that the underlying data distributions are mixtures of Gaussian or non-Gaussian functions, these algorithms represent a reasonable approach. In the case of the batch lens data, it was decided that truncated mixture distributions did not describe the data shape very well so this approach was abandoned. A wise colleague commented: "Forget about math for a minute and listen to what the data is saying. The distributions for any particular lens run cannot be fit into any neat category; it is inherently messy."

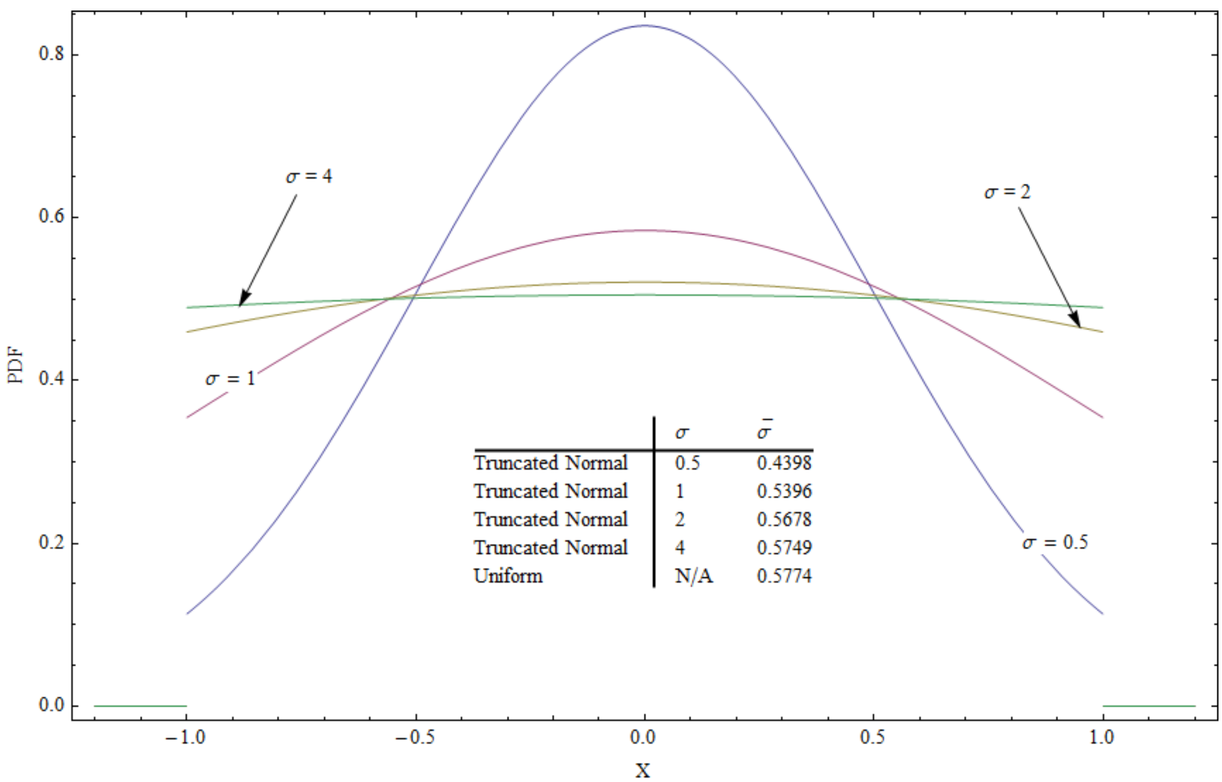

Figure 2. Gaussian distributions, truncated at (–1, 1). Note that σ is the standard deviation parameter whereas $\overline{\sigma}$ is the standard deviation of the truncated distribution.

2.2 Nonparametric methods

Nonparametric methods avoid assumptions regarding the shape of the PDF that describes the data in question. But as we shall see, a few assumptions inevitably creep in. Binned histograms and nonparametric methods are often used in exploratory data analysis. The kernel density estimate (KDE) is an established nonparametric method for finding the shape of the PDF for a data set with unknown properties,

$$f(x) = \frac{1}{n \cdot h} \sum_{i=1}^{n} K\left(\frac{x - x_i}{h}\right), \tag{1}$$

where $f(x)$ is the PDF estimate, K is the kernel function, x_i is abscissa value of the i^{th} data point, n is the number of data points, and h is a smoothing parameter that is usually called the bandwidth. The popular Gaussian kernel was used for this analysis,

$$K(u) = \frac{1}{\sqrt{2 \cdot \pi}} e^{-\frac{1}{2}u^2}. \tag{2}$$

The Silverman rule[4] is often used to calculate the optimal bandwidth h,

$$h_{optimal} = \overline{\sigma} \cdot (\frac{4}{3 \cdot n})^{1/5} \approx 1.06 \cdot \overline{\sigma} \cdot n^{-1/5}, \tag{3}$$

where $\overline{\sigma}$ is the calculated standard deviation of the data set and n is the number of data points. There are other important methods for estimating the optimal bandwidth, for example the Sheather-Jones method.[10,11] In a sense, the KDE algorithm behaves like a low pass filter that attenuates features smaller than the bandwidth *h*. If *h* is far below the Silverman optimum, the resulting PDF will have increased roughness and reflect the randomness within the data rather than its intrinsic structure. For a given bandwidth *h*, more data points will reduce the roughness of the estimated PDF. An interesting consequence of the Silverman rule is that reducing *h* to resolve small features in the data would require a very large increase in the number of data points. Roughness is a statistical parameter that is calculated as

$$R(f) = \int (\frac{\partial^2 f(x)}{\partial x^2})^2 \cdot dx. \tag{4}$$

What are the hidden assumptions in the KDE method and the Silverman rule? The KDE method assumes that the data set is either unbounded, or if it is bounded that the PDF drops to zero at the edges. The KDE method makes no assumptions about the shape of the distribution (i.e., uni-modal or multi-modal). The Silverman rule assumes that the data are close to a Gaussian shape (therefore unbounded and uni-modal). Interestingly enough, Bernard Silverman was also captivated by the bimodal "Old Faithful" problem.[4] Our problem involves bounded data that does not drop to zero at the edges and may be bimodal. But all is not lost; there is a body of knowledge to guide us through these difficulties.[7,12,13]

Fig. 3 shows PDF approximations generated by the KDE method (also called smooth histograms) of a uniform distribution with 5000 random data points over the interval (–1, 1). These plots were created in Mathematica using the *Histogram* and *SmoothHistogram* functions, and the *SmoothKernelDistribution* function to generate the PDF estimates. The truncation in the middle plot was done using the *TruncationDistribution* function. The bounded PDF plot on the right was generated using an undocumented feature in Mathematica 9: *SmoothKernelDistribution* [data, 0.25, {"Bounded", {–1, 1}, "Gaussian"}]; the value 0.25 defines the bandwidth *h* and the "Gaussian" option defines the kernel function as being Gaussian. The region in Fig 3a outside the (–1, 1) interval is called spillover, and it constitutes about 10% of the overall PDF. The truncation algorithm was used in Fig. 3b to cut away the spillover. The underestimated region to the right and left in Fig. 3b demonstrates the "edge bias" effect. This effect and ways to fix it are a subject of ongoing research.[12,13] The "Bounded" option on Fig. 3c removes the edge bias, but Wolfram Research will not discuss how this option is implemented in Mathematica.

Part of the problem with edge bias has to do with the fact that a data point in the middle has an influence that goes in two directions whereas a data point on the edges has an influence that goes in only one direction. Again we see that our intuition is confounded by edges.

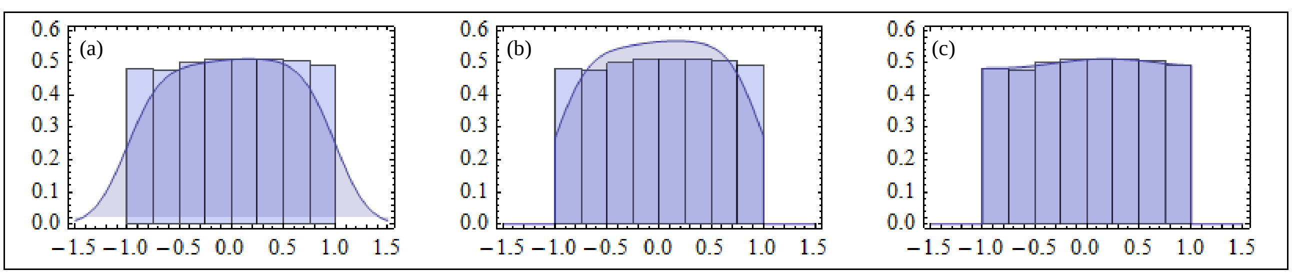

Figure 3. Left to right: standard KDE algorithm, truncated KDE algorithm, bounded KDE algorithm; shown with binned histograms (ordinate dimension normalized to the PDF function) for comparison.

It is interesting that much of the research concerning bounded and mixture distributions[7,13,14] focuses on the multivariate case. We have been concerned with the univariate data. Whereas the realities of boundedness and bimodality are merely annoying in the univariate case, they are positively vexing in the multivariate case.

How do we pick a reasonable bandwidth given that our data violate the assumptions underlying the Silverman rule? It seems natural to think of the bandwidth as a fraction of the bounded data interval. Fig. 4 shows the results of a data simulation. Data was randomly generated using a mixture of two Gaussian distributions with means centered on 0.66 and –0.6; standard deviations of 0.5 and 0.25; and weights equal to 1 and 0.25. The shape of this distribution is intended to mimic actual lens data. The Silverman optimal h was calculated to be 0.26; the standard deviation of the data was calculated to be 0.539; and 50 data points were used. The original PDF distribution is plotted along with KDE estimates using h = 0.5, 0.25, and 0.125. This plot shows several interesting features. Because the data are randomly generated, the histogram plot does not exactly match the original PDF mixture distribution. The KDE with the smallest h (0.125) has the most roughness and intersects seven of the eight histogram bin tops. The KDE with the largest h (0.5) shows the strongest attenuation, least roughness, and only intersects two of the bin tops. The KDE with the middle h value (0.25) is close to the Silverman optimum; it shows some attenuation, but captures the flavor of the underlying data as it is close to most of the bin tops. It was decided that an h value equal to $1/8^{th}$ of the interval (0.25 for radius and center thickness, 0.125 for wedge and irregularity) is a reasonable choice for the lens data in this analysis, and it appears to be close to the Silverman optimum in most of the cases.

It could be argued that the decision to use $1/8^{th}$ of the interval for an h-value is subjective. At this point we are less concerned with mathematical objectivity and more concerned with gathering insight into the data. For that reason it was thought that having the same h-value for all the batches would make comparisons easier. In the end we will be using parametric methods on the aggregate data as we develop recommendations for optical designers.

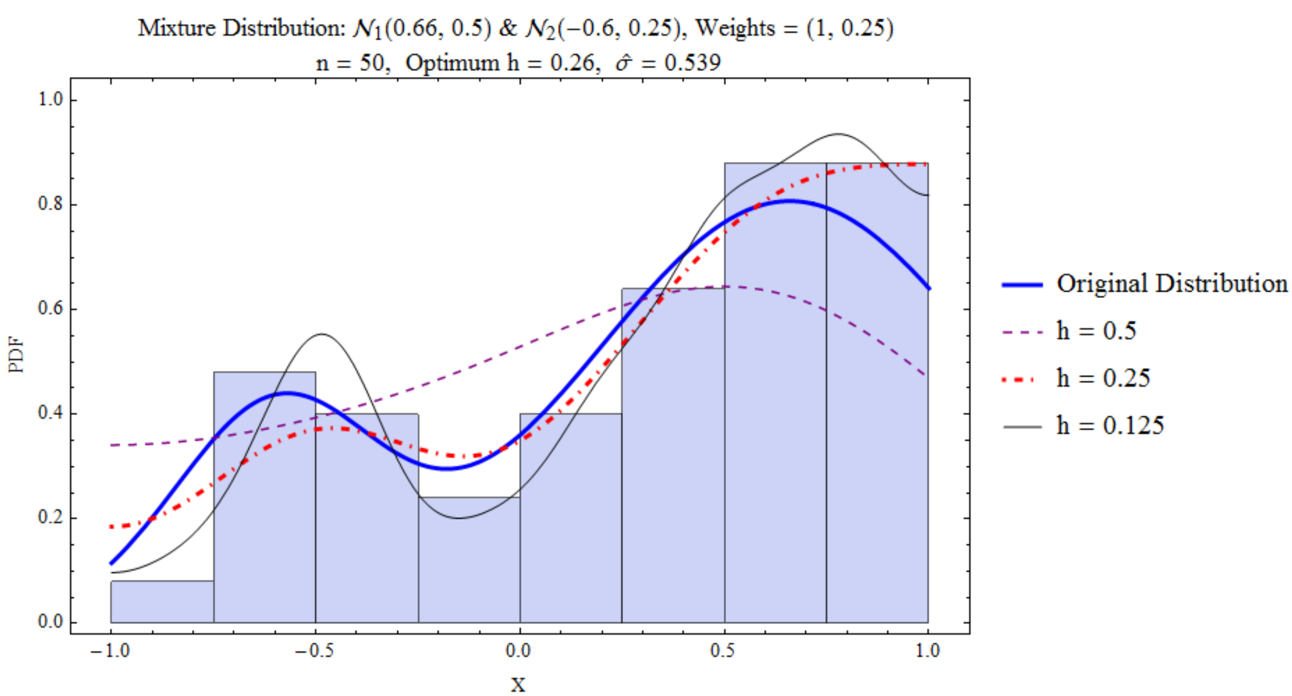

Figure 4. Simulated data showing data histograms, the original distribution, and several KDE estimations.

3. ANALYSIS RESULTS

Because the number of lenses in a batch could influence the shape of the distribution, it was initially decided to equally weight large and small batches. As the data are normalized, it is possible to superimpose multiple smooth histograms to get a sense of an aggregate distribution. A smooth histogram was made from an equally weighted mixture distribution (Fig. 5) from all of the batch results.

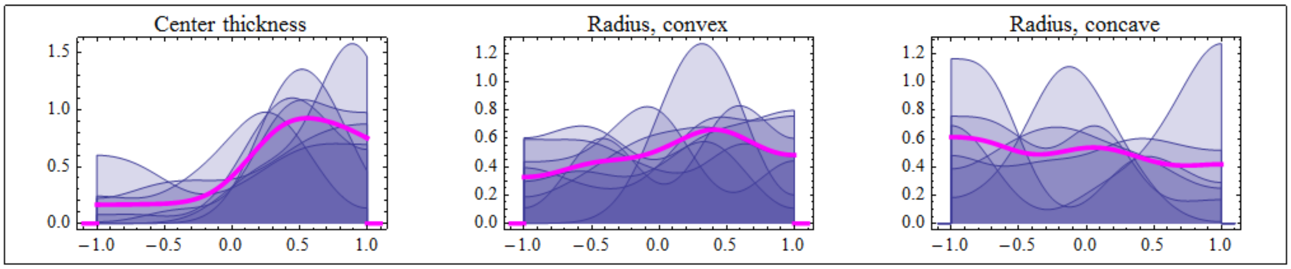

Figure 5. Individual (thin line) and mixture (thick line) smooth histograms for center thickness, convex, and concave radii.

But the analysis in Fig. 5 begs the question: What would happen if we had simply mixed all the lens data together (equal lens weighting), and how would that be different from an equal weighting of the batches? As we see in Fig. 6, if we equally weight each lens or equally weight each batch, the aggregate distributions are almost identical. This would indicate that the aggregate distribution is not a strong function of batch size for a batch size of fewer than 50 lenses. However, it is possible that the batch distribution would assume a characteristic shape as the batch gets large enough, and the manufacturing process becomes mature enough.

Notice that the messiness of the individual lens data settles down considerably when the aggregate distribution is considered. The convex and concave radius data show a weak asymmetry, but the center thickness shows a strong asymmetry. We know anecdotally that convex radii tend to be larger and concave radii tend to be smaller when the glass is ground from flat stock. But this effect is much smaller than we had initially imagined. One might be tempted to think of the center thickness as a right-shifted truncated normal distribution, but this wouldn't work because of the strong left tail, which is strong because fabricators occasionally make mistakes that require more grinding than desired.

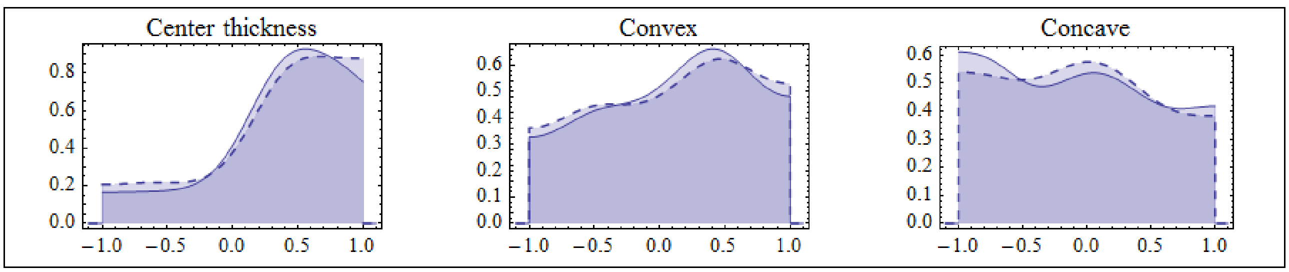

Figure 6. Equal batch weighting (solid) versus equal lens weighting (dashed) of aggregate PDF's for center thickness, convex, and concave radii.

Because the number of lenses in this study is rather small, it was not possible to tease out the factors that influence the distribution shapes in any great detail. We do see a slight difference in distribution shape between convex and concave radii. We also see that that the irregularity data fall into groups according to the precision grade (Figs. 7 and 8). In this case, the precision grade is described by a root mean squared (RMS) specification. As stated earlier, the irregularity and wedge data was normalized to a (0, 1) interval. Table 1 summarizes the findings of Figs. 5, 7, and 8.

Why do the irregularity data separate so neatly according to precision grade? To some extent, you are seeing the result of a process, not a tolerance. The fabricator has a small number of discrete finishing processes to choose from, and each has a "best expected outcome" that the process is capable of yielding. Magneto-Rheological finishing for example is expected to yield irregularity under 0.07 λ RMS, and if parts with a looser tolerance are run using these processes the measured irregularity fall well below expectation. In effect, the customer gets a better result than his optics code accounts for. To take advantage of this effect, the optical engineer would need an intimate understanding of the particular supplier process advantages and limitations.

National Security Technologies, LLC (NSTec) procurement is governed by the U.S. Department of Energy (DOE) rules. We have to be very careful not to violate these rules as we try to understand our suppliers. Any conversation between an engineer and supplier during a procurement cycle could potentially give that supplier an advantage during a bid process, and would be a violation. But by knowing the supplier process statistics, you can estimate a cost factor due to reduced yield given that some of the lenses will be rejected if you pick a tolerance that the chosen process cannot fully support. Of course, this idea presupposes that the supplier has a good statistical understanding of his own process statistics, and this might not be the case. In general, process advantages and limitations is part of what gives a particular supplier competitive advantage, so the optics industry has incentives that prohibit the dissemination of this information.

Supplier/consumer relationships are very complex in the optics industry. The vendor that supplies metrology equipment to a fabricator like Optimax may also buy lenses from Optimax and fabricate lenses that compete with Optimax. NSTec has a very simple relationship with Optimax—they sell lenses, we buy lenses. This is partially why we were able to co-author this study (procurement rules aside). If the optics industry is to collectively move to a deeper understanding of how fabrication and design are related, we will have to overcome structural problems in our industry.

Fig. 8 shows plots of edge thickness deviation (ETD). In general, optical engineers are less concerned with how wedge is measured and are more concerned with the angular boresighting effect. A wedge measurement is often a mechanical indicator runout measured on one optical surface near the clear aperture with the other optical surface properly mounted on a spindle and a zero-runout condition on the edge diameter. Hence the term "edge thickness deviation" since the runout measured is the change of edge thickness per revolution. Special care should be taken to make sure the quantity that is toleranced in the design program is measurable and then properly specified on lens drawings.

Statistics is not necessarily a comfortable subject for the precision engineer. Engineers prefer to view the world as a Newtonian, deterministic place, actively engaged in creating methodologies to force the world to conform to this view, and what better place to showcase these impressive accomplishments than optical fabrication? Yet, even in this sacred garden we see randomness if we know where to look, and not just any randomness, but bimodal—shocking! Let's suppose for a moment the thick line in Fig. 9 is known to be the master PDF and individual lens batches were just random expressions of the master. This supposition is embodied in the plot on the right in Fig. 9, but the plot on the left shows smooth histograms generated from actual lens data. Note the concave data plots in Figs. 5 and 9 are identical. So to what extent is lens manufacturing an exercise in pure randomness, and to what extent is it predictable and deterministic?

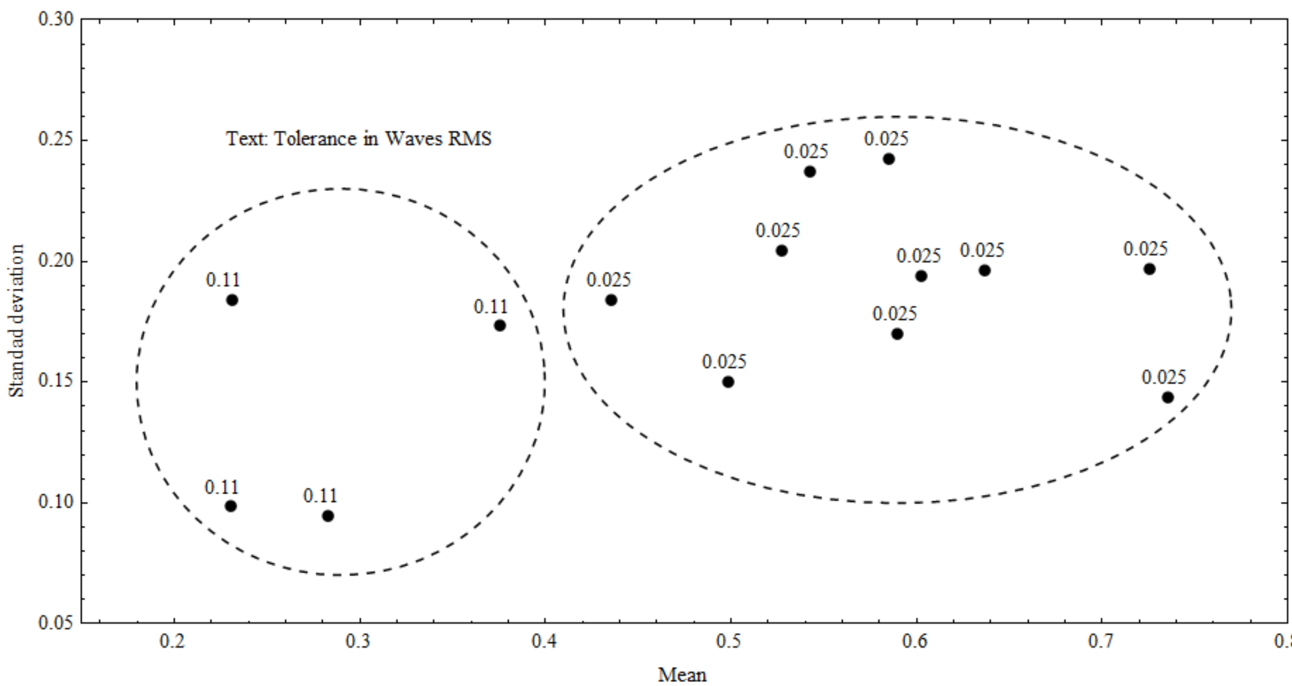

Figure 7. Irregularity mean versus standard deviation. Notice the high precision tolerances separate nicely from the low precision tolerances.

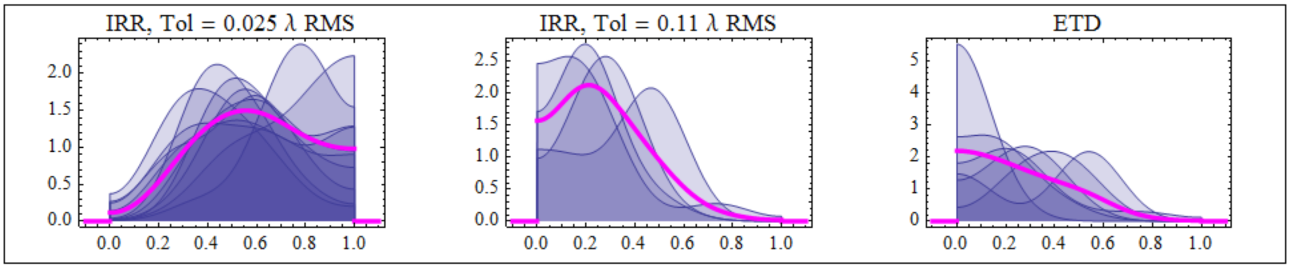

Figure 8. Superimposed smooth histograms for two precision grades of irregularity and edge thickness difference. The wavelength used for measuring the irregularity is λ, usually 632.8 nm. Batch smooth histograms are shown with thin lines, equally weighted mixture distributions shown with a thick line.

Table 1. Mean and standard deviation data for radii, center thickness, irregularity, and wedge

	Mean	Standard Deviation	Comments
Convex radius	0.0898	0.5569	Normalized to (–1, 1)
Concave radius	–0.0559	0.5559	Normalized to (–1, 1)
Center thickness	0.3367	0.5046	Normalized to (–1, 1)
Irregularity	0.5748	0.2168	Tol = 0.025 λ RMS, data normalized to (0, 1)
Irregularity	0.2722	0.1463	Tol = 0.11 λ RMS, data normalized to (0, 1)
Wedge	0.2497	0.1818	Measured using ETD, data normalized to (0, 1)

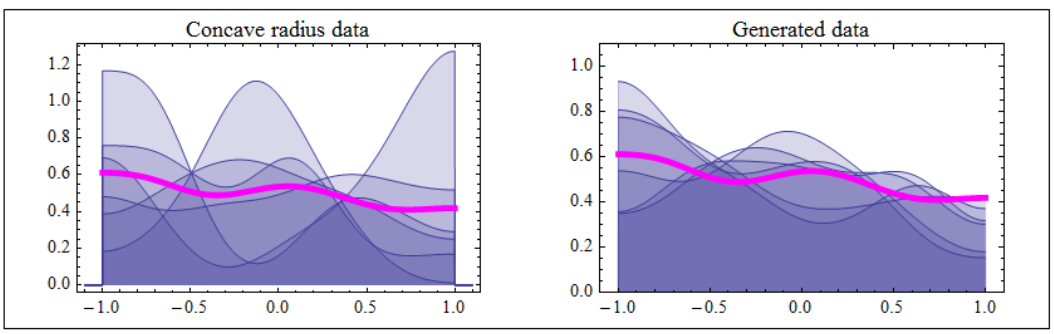

Figure 9. Measured (left) and generated (right) data. The superimposed smooth histograms on the left come from individual lens batches; the aggregate mixture PDF is the thick line. On the right, the individual batches with 30 data points (a typical lens batch size) were randomly generated from the aggregate PDF distribution.

4. CONCLUSION

This modest pilot study seeks to document the shop statistics that could help optical engineers more properly tolerance systems. In the process, we worked with statistical methods and tools not well known in the optics community. We have been forced to come to grips with economic realities that have prevented prior studies of this type being attempted. And we have had to confront deterministic engineering paradigms that have prevented us from thinking about the very statistical methods that could help us the most.

There was a time, before computer design codes were invented, when optical designers and fabricators worked together under the same roof, or might have even been the same person, but we don't have to go back to the time of Galileo or Spinoza. The separation of fabrication from design is really a modern phenomenon. Yes, computer codes have surely helped us, but we have also separated the craftsman and the designer. Perhaps we should think about taking a step towards re-establishing that collaboration in spite of the structural problems that exist in our industry.

Lens fabrication is the result of highly skilled labor, but glass is a rather truculent material, and modern optical designs can be very demanding. We have something to learn from quirky physical processes like the classic bimodal "Old Faithful" data. But the individual batch statistics for radius are 36% unimodal. Batch statistics sometimes have a clumpy character. A careful inspection of the sample batch plotted in Fig. 1a shows that the smaller mode has very little variance. Opticians are learning about how the particular glass reacts as they rework the material. We see in Fig. 5 that individual batch distributions occasionally have a character that seems very different from that of the aggregate distribution. We also see in Fig. 9 that the differences between batch statistics and aggregate statistics are much larger than one would predict from a Monte Carlo simulation. Hence, batch statistics are a fascinating mixture of deterministic lens polishing and unexpected randomness.

Despite the small size of this study, certain basic trends can be noted. "Your lenses are always fat." It sounds like one of those horrible subjects we avoid discussing in mixed company during dinner parties. Center thickness does not seem to be the most important optical parameter, but when errors occur systematically they can have a powerful effect. Center thickness affects the optomechanical distance-to-next parameter that governs lens spacing, and systematic center thickness errors lead to unpredicted focus and spherical aberration.

We can see the effects of process on statistical outcomes, particularly irregularity and radius. With irregularity, we see the precision class determines the fabrication process, which, in turn, affects the statistical outcome. With radius, we see a weak correlation between shape (convex/concave) and statistical outcome due to how the lenses are ground. We suspect glass hardness or softness is an important statistical parameter but do not have the data to support this assertion. This could be an area of further work.

We have shown uniform distribution is actually a reasonable approximation for radius tolerances. We can see a weak asymmetry in the data, and the asymmetry is supported anecdotally by an understanding of the grinding process. But the numerical value for the asymmetry could easily be within an error bar.

Specific recommendations follow in the Appendix A. We hope that this pilot study encourages more lens suppliers to share fabrication data.

Appendix A

In order to apply the data presented here, the aggregate statistical distributions need to be parameterized in a way that can be used by an optical tolerancing program. The parameters for the concave and convex radii distributions were estimated using the Mathematica function *"EstimatedDistribution"* with the *"MaximumLikelihood"* option without problems. The *"MethodOfMoments"* option was tried without success. The *"EstimatedDistribution"* function is actually a local optimization that benefits from a judiciously chosen starting point. The results of this analysis are shown in Fig. A-1.

The convex and concave radii data were fit to a normal distribution, $N[\mu, \sigma]$ truncated to the (–1, 1) interval, where μ is the mean parameter, and σ is the standard deviation parameter. The concave and convex parameterized distributions shown in Fig. A-1 are $N[-0.3118, 1.3101]$ and $N[0.5907, 1.4312]$ respectively.

The center thickness distribution is difficult to fit and in most cases failed to converge using the *"EstimatedDistribution"* function. The center thickness data was eventually fit by maximizing the "*LogLikelihood*" function with the global optimization search function "*NMaximize*", subject to the appropriate constraints. The equations (A-1) and (A-2) are two successively more complex approximations with improvements in the log-likelihood metric also shown. Equation (A-1) is a truncated normal function; equation (A-2) is a weighted mixture with two components: a truncated normal (N) and a uniform distribution (U) spanning the interval (–1, 1).

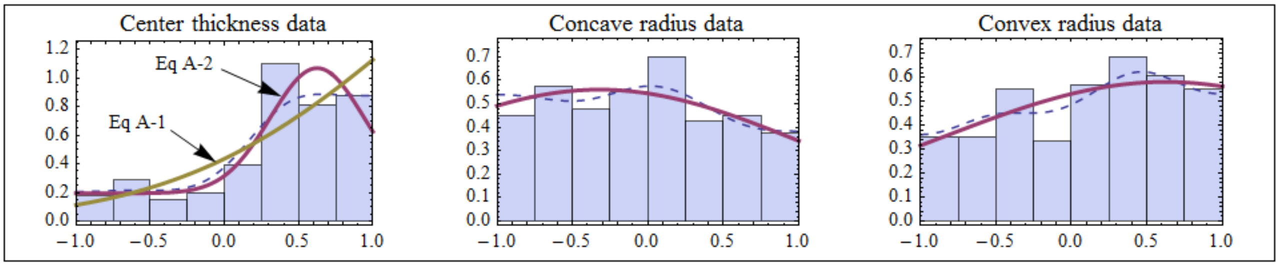

Figure A-1. Center thickness, concave, and convex radii parameterized distributions fit to the raw aggregate data. The parametric approximation is thick, the nonparametric approximation is dashed.

The center thickness distribution parameterized distribution is

$$\mathrm{PDF}_{CT} = N\,[2.908,\ 1.3101] \qquad \text{(log-likelihood = –94)} \qquad \text{(A-1)}$$

or

$$\mathrm{PDF}_{CT} = 0.639\times N\,[0.625,\ 0.315] + (1–0.639)\times U\,[–1,\ 1] \qquad \text{(log-likelihood = –87)} \qquad \text{(A-2)}$$

What happens if we look at the radius data without regard to figure (convex/concave)? These data are displayed in Fig. A-2. The mean and standard deviation for the aggregate data are 0.026 and 0.56, respectively. The data was fit to N [0.180, 1.483], truncated to the (–1, 1) interval using the maximum-likelihood method.

The wedge (measured using ETD) and irregularity data were all fit to a normal distribution, N [μ, σ] truncated to the (0, 1) interval. These distributions are N [0.0607, 0.2843] for ETD; N [0.2497, 0.1651] for irregularity (tolerance = 0.11 λ RMS); and N [0.5990, 0.2522] for irregularity (tolerance = 0.025 λ RMS). Comparisons of the results together with binned histograms are shown in Fig. A-3.

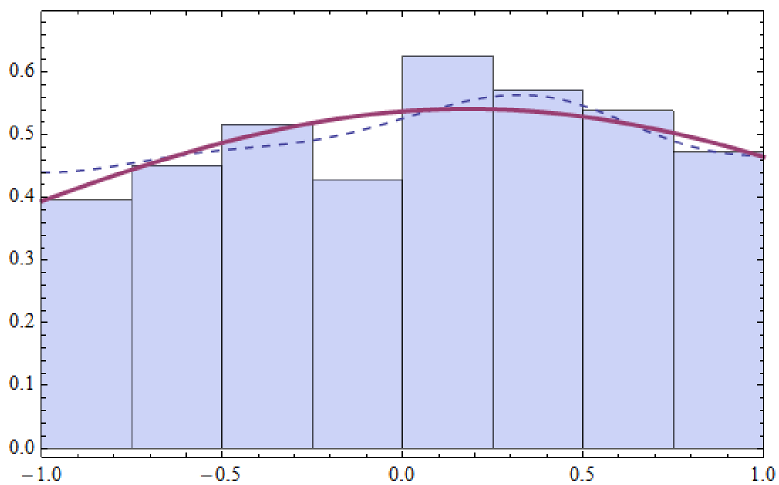

Figure A-2. Parametric (thick) and nonparametric (dashed) approximations for radius (convex and concave combined) data.

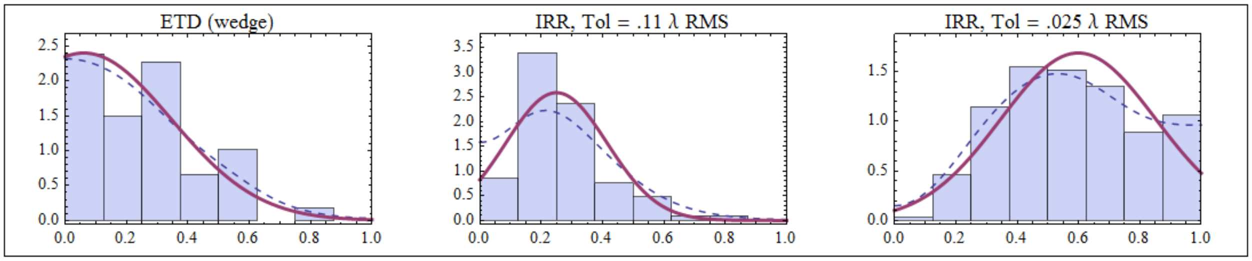

Figure A-3. Wedge and irregularity plots. The thick lines are parametric; the dashed lines are nonparametric approximations.

It is interesting to note that the parametric and nonparametric approximations are noticeably different in some cases, particularly the middle irregularity data of Fig. A-3. How can this be? The nonparametric distributions for irregularity and wedge were created using an *h*-bandwidth of 0.125. An examination of the standard deviation parameter generated by the maximum likelihood estimate shows that they are relatively close to the bandwidth. If you think of the standard deviation parameter as a feature size, the nonparametric estimate becomes less accurate if the feature size is too close to the bandwidth. So what should we believe? The choice of $h = 0.125$ was a subjective decision, and one could argue that a better choice exists. But reducing the bandwidth *h* also has the consequence of increasing roughness. We could say that that the Silverman rule provides an objective choice of bandwidth, but we have already shown that the Silverman estimate uses assumptions that are violated by the type of data we are examining. We could use the bandwidth estimate method of Bouezmarni,[14] but this method has not necessarily been universally accepted. The maximum likelihood estimate was able to resolve a relatively small standard deviation, but the choice of using the normal distribution two-parameter system is also a subjective decision. We are left with the uncomfortable realization that modeling choices cannot always be objectively supported. In reality, optimal bandwidth estimation methods like Silverman,[4] Sheather-Jones,[10] Bouezmarni,[13] and parametric estimation methods like maximum-likelihood, are tools that help us to make those subjective choices, and we have to decide which is best.

Recommendations

Because the aggregate radius data are so close to the uniform distribution, for simplicity we recommend that the uniform distribution is a reasonable option for radius tolerancing. Note that the uniform distribution is already a common option in many optical design programs. However, the center thickness distribution is strongly asymmetrical with a mean value of 0.337 T (Table 1) for the lenses measured in this study. This asymmetry is supported anecdotally as well. Lenses have traditionally been toleranced in a symmetric manner with the nominal value ±T where T is the tolerance value. The issue here is that the nominal value is known a priori to be smaller from the expected value (if the tolerance zone is symmetric), and this is a problem. Another approach would be to tolerance the center thickness with nominal +0.663 T and –1.337 T where T in this case is half of the total tolerance zone and the nominal value is assumed to be equal to the expected mean. The optical designer would also need to use an asymmetric PDF in his tolerancing program to support the asymmetric tolerance zone. Admittedly, this is a fairly radical change from the established practice, and the implications for the various optical design programs would have to be investigated. But in the end we would have lenses that are closer to the nominal value in most cases, although in extreme cases some of the lenses would be further away from the nominal.

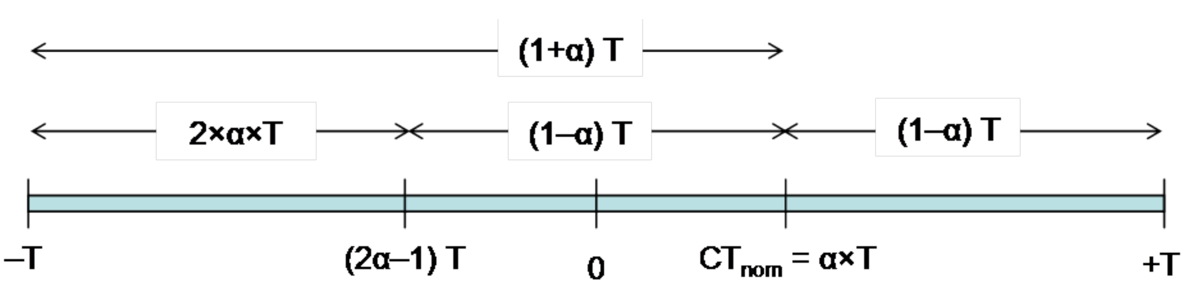

Figure A-4. The structure of an asymmetric center thickness tolerance zone. The smaller symmetric zone would be CT_{nom}± (1–α) T.

The most common approach for tolerancing optics in Code V is the *TOR* option (wavefront differential tolerancing). The *TOLMONTE* option is also available, but being a true Monte Carlo method, it takes longer, so it is often used as a final check. There is no way to change the predefined tolerance distributions in the TOR option. However, one could apply a user-defined tolerance distribution in *TOLMONTE* through the *UTOCHNG.seq* file.[15]

Using this approach (asymmetrical tolerancing), let us assume that the reader would normally tolerance the desired center thickness as CT_{nom} ±T, the desired tolerance zone being 2×T. Instead, applying the *TOR* option with a uniform probability distribution, use *TOR* with the smaller symmetrical tolerance zone of $CT_{nom} \pm (1–\alpha)$ T. The data here suggests that a reasonable value for α is 0.337 (Table 1 and Fig. A-4). As a final check, apply T*OLMONTE* with the *UTOCHNG.seq* and equations (A-1) or (A-2). The final engineering drawings would need to be reformulated for the asymmetrical tolerance zone of $CT_{nom} +(1–\alpha)$ T and $–(1+\alpha)$ T.

We find that the probability of a lens center thickness in this study falling inside of the smaller symmetric zone is 86%. Mathematically this would be equivalent to Pr [(2α–1) T < x < T] for the center thickness probability distribution shown in Fig. A-1. Now we have a nominal value that is equal to the expected value but we have a 14% probability that the lens falls outside of the smaller symmetric tolerance zone. One could think of the asymmetrical part (2×α×T) as a safety bonus for the supplier. The intent is that, knowing the structure of the manufacturing probability distributions, the designer could loosen tolerances a bit and have a nominal optical design that is much closer to the expected as-built values for downstream optomechanical design work.

There is a caveat, however. The preceding analysis is true for all of the lenses taken together. But the aggregate statistics are not necessarily reflected in the individual lens batches. In five of the seven batches, there are almost no examples of individual lens center thicknesses in the region from (–1, –0.5). But in two of the seven batches there were a significant number of lenses in this region; this is what caused the fat left tail in the aggregate distribution. We don't know if there is a way of predicting if extra grinding/polishing is likely (i.e., thin lenses). We suspect that it is related to a combination of soft glass and stringent irregularity tolerances.

In the case of wedge and irregularity, the desired nominal value is zero, but some deviation is allowed. It is possible that using the PDFs supplied here, tolerances could conceivably be loosened slightly, but this would depend on strong communication between the supplier and designer. This is particularly true for the less restrictive 0.11 λ RMS irregularity tolerance class.

The reader is encouraged to work with his (or her) suppliers to see if these conclusions are applicable to specific situations.

References

[1] Code V, v10.6, Reference Manual, Synopsys Optical Solutions, Pasadena (2013).
[2] Weise, G. E., [Selected Papers on Optical Tolerancing], SPIE Optical Engineering Press, Bellingham (1991).
[3] Juergens, R. C. and Wood, H. W., "Random thoughts on Monte Carlo tolerancing," Proc. SPIE 6676, 667605 (2007).
[4] Silverman, B. W., [Density Estimation for Statistics and Data Analysis], Chapman and Hall, London, 1–22 (1986).
[5] Dekking, M. A., [A Modern Introduction to Probability and Statistics: Understanding Why and How], Springer, London, 206–230 (2005).
[6] Shaughnessy, J. and Pfannkuch, M., "How faithful is Old Faithful? Statistical thinking: A story of variation and prediction," Math. Teacher, 95(4), 252–259 (2002).
[7] Benaglia, T., Chauveau, D., Hunter, D. R., and Young D., "Mixtools: An R package for analyzing finite mixture models," J. Statistical Software, 32(6), 1–29 (2009).
[8] Wolfram Mathematica documentation, Wolfram Mathematica 9 Documentation Center, Champagne, (2014).
[9] MacKenzie, D. I., Nichols, J. D., Royle, J. A., Pollack, K. H., Bailey, L. L., and Hines, J. E., [Occupancy Estimation and Modeling: Inferring Patterns and Dynamics of Species Occurrence], Academic Press, New York, 77–79 (2005).
[10] Sheather, S. J. and Jones, M. C., "A reliable data-based bandwidth selection method for kernel density estimation," J. Royal Statistical Soc. B, 53(3), 683–690 (1991).

[11] Jones, M. C., Marron, J. S., and Sheather, S. J., "A brief survey of bandwidth selection for density estimation," J. Amer. Statistical Assoc., 91(433), 401–407 (1996).
[12] Albers, M. G., "Boundary estimation of densities with bounded support," Swiss Federal Institute of Technology, Zurich, 19–44 (2011).
[13] Bouezmarni, T. and Rombouts, J. V. K., "Nonparametric density estimation for multivariate bounded data," J. Statistical Planning and Interference, 140(1), 139–152 (2010).
[14] Benaglia, T. A., "Nonparametric estimation in multivariate finite mixture models," UMI Dissertation Publishing, Ann Arbor, 1–35 (2011).
[15] Shi, W., personal communication, Synopsys Optical Solutions, (May 8, 2014).

Acknowledgments

This manuscript has been authored by National Security Technologies, LLC, under Contract No. DE-AC52-06NA25946 with the U.S. Department of Energy. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government purposes.