11-18-2010

A New Statistical Procedure for Validation of Simulation and Stochastic Models

Robert G. Sargent
Syracuse University, Department of Electrical Engineering and Computer Science, rsargent@syr.edu

Recommended Citation
SYR-EECS-2010-06

Department of Electrical Engineering and Computer Science

Technical Report

SYR-EECS-2010-06 Nov. 18, 2010

A New Statistical Procedure for Validation of Simulation and Stochastic Models

Robert G. Sargent rsargent@syr.edu

ABSTRACT: We describe a new statistical procedure for validating simulation and analytic stochastic models using hypothesis testing when the amount of model accuracy is specified. This procedure provides for the model to be accepted if the difference between the system and the model outputs are within the specified ranges of accuracy. The system must be observable to allow data to be collected for validation.

KEYWORDS: Validation of Simulation Models, Stochastic Model Validation, Hypothesis Tests of Intervals, Operating Characteristic Curve, Risk Curve

Syracuse University - Department of EECS,
4-206 CST, Syracuse, NY 13244
(P) 315.443.2652 (F) 315.443.2583
http://eecs.syr.edu

A New Statistical Procedure for Validation of Simulation and Stochastic Models

Robert G. Sargent

Department of Electrical Engineering and Computer Science
Syracuse University
Syracuse, NY 13244
rsargent@syr.edu

ABSTRACT

We describe a new statistical procedure for validating simulation and analytic stochastic models using hypothesis testing when the amount of model accuracy is specified. This procedure provides for the model to be accepted if the difference between the system and the model outputs are within the specified ranges of accuracy. The system must be observable to allow data to be collected for validation.

1. Introduction

Simulation and analytic stochastic (e.g., queueing) models are often used to solve problems and to aid in decision-making. The developers and users of these models, the decision makers using information obtained from the results of these models, and the individuals affected by decisions based on such models are all rightly concerned with whether a model and its results are “correct”. This concern is addressed through model validation. Model validation is usually defined to mean “substantiation that a model within its domain of applicability possesses a satisfactory range of accuracy consistent with the intended application of the model” and is the definition used here. There is a considerable literature on validation of simulation models (see, e.g., Law and Kelton 2000, Sargent 2009) and little literature on validation of analytic stochastic models. In this paper we describe a new statistical procedure for validating models where the amount of model accuracy is specified.

A model should be developed for a specific purpose (or application) and its validity determined with respect to that purpose. If the purpose of a model is to answer a variety of questions, the validity of the model needs to be determined with respect to each question. Numerous sets of experimental conditions are usually required to define the domain of a model’s intended applicability. A model may be valid for one set of experimental conditions and invalid for another. A model is considered valid for a set of experimental conditions if the model’s accuracy is within its acceptable range of accuracy, which is the amount of accuracy required for the model’s intended purpose. This usually requires identifying the model’s output variables of interest (i.e., the model variables used in answering the questions that the model is being developed to answer) and specifying the required acceptable range of accuracy for each variable. The acceptable range of accuracy is typically specified for each model variable of interest as the range that the difference between that model variable and the corresponding system variable can have for the model to be valid. The amount of accuracy required should be specified prior to starting the development of the model or very early in the model development process. If the variables of interest are random variables, then properties and functions of the random variables such as means and variances are often what is of primary interest and are the quantities that are used in determining model validity.

As an example, consider a single-server queueing model where the mean waiting time is of interest. Then validity should determine if the model's mean waiting time has sufficient accuracy for the model's intended purpose within its domain of applicability. The acceptable range of accuracy would be the difference between the model and system mean waiting times with, e.g., a range of -0.1 to 0.1 hour. The domain of applicability would be the range that the model variables of interest in the domain of applicability can take; e.g., the mean arrival rate and the mean service rate. A single value for each variable in the domain of applicability would give the values for a set of experimental conditions; e.g., a mean arrival rate of 2 per hour and a mean service rate of 3 per hour.

A new statistical procedure is described in this paper for validating models using hypothesis testing when the amount of model accuracy is specified as a range and the system being modeled is observable for the experimental condition under test. This procedure would be used in performing operational validity, where we determine whether the model's output behavior has the accuracy required for the model's intended purpose over the domain of the model's intended applicability. (See, e.g., Sargent 2009 for a discussion of operational validity.) A system is observable if data can be collected on the system for model validation. Current statistical procedures using hypothesis tests for model validation use only a single point (see, e. g., Banks et al. 2010) instead of a range, or they consider ranges indirectly (see Balci and Sargent (1981, 1982a, 1982b, 1983)). We first develop the procedure for validation of simulation models and then discuss its use for validation of analytic stochastic models.

2. Validation of Simulation Models

This new statistical procedure uses hypothesis testing. The first step in hypothesis testing is to state the hypotheses to be tested:

H_0: Model is valid for the acceptable range of accuracy under the set of experimental conditions.

H_1: Model is invalid for the acceptable range of accuracy under the set of experimental conditions.

Two types of errors are possible in testing hypotheses. The first, or type I error, is that of rejecting the validity of a valid model and the second, or type II error, is that of accepting the validity of an invalid model. The probability of a type I error, α, is called *model builder's risk*, and the probability of type II error, β, is called *model user's risk* (Balci and Sargent 1981). In model validation, the *model user's risk* is extremely important and must be kept small. Thus both type I and type II errors must be carefully considered when using hypothesis testing for model validation.

This statistical procedure is going to be developed to determine if the mean of a simulation model output random variable has the desired range of accuracy; e.g., the mean waiting time of a queueing system. This determination will be made by comparing data from the simulation model and from the real system. Classical hypothesis testing is typically concerned with a point (or a value). Here we are interested in conducting hypothesis testing for an interval because we want to consider the amount of model accuracy. This type of statistical problem commonly uses the t-distribution, as we will do here. The procedure requires the data to be used for analysis from both the simulation model and the system to be approximately Normally Independent and Identically Distributed (NIID). The specific test statistic used with the t-distribution will depend on whether the variances are known or unknown and assumed equal or unequal.

There are general two types of simulations with respect to output analysis: terminating and nonterminating (Law and Kelton 2000). Terminating simulations are a function of the initial conditions and a terminating event, which is either a time event or a state event. An example of a terminating simulation where time is the terminating event is that of simulating a store that is open from 9 am until 5 pm. A system performance measure of "how much" during a fixed time interval is usually of interest in this type of terminating simulation, e.g., the amount of sales from 9 am until 5 pm. An example of a terminating simulation where a state event is the terminating event is the length of time a new system operates until its first failure occurs. The "amount of time" until the terminating event occurs is usually of interest in this type of terminating simulation. Terminating simulations are generally analyzed statistically by the method of replications. In this method n independent replications are made with the same initially conditions and terminating event to give n IID data observations which can be used for analysis. These n IID data observations may or may not be normally distributed. If these observations are not normally distributed then n can usually be selected large enough such that the usual asymptotic statistical methods can be used; e.g., n of size 20 – 30 is usually sufficient for the use of the t-distribution to handle non-normal distributions.

Nonterminating simulations, often referred to as steady-state simulations, are where some aspect of the steady-state behavior is usually of interest. Simulation runs start with some initial condition, the initial transient behavior is truncated (deleted) using some truncation point, and then steady-state behavior data are collected for analysis. We note that the steady-state observations are usually correlated. There are two major methods of conducting steady-state analysis. One method is the method of replication, where n independent replications are made to produce n IID data values. The data value for each replication is the average of, say, m steady-state observations in that replication. This set of IID data values will be approximately normally distributed if m is made large enough. Thus the method of replication gives n approximately NIID data values for steady-state analysis. The other major method of steady-state analysis is the method of batch means. In this method of analysis, one long simulation run is made starting with some initial condition, the initial transient behavior is truncated, and then the presumedly steady-state observations are divided into n nonoverlapping batches of equal length of size m. The observations in each batched are averaged to give n data values. With m sufficiently large, these n data values are approximately NIID.

We have just summarized the standard methods used in output analysis of simulations to obtain n NIID data values. (For details of these method see, e.g., Law and Kelton 2000 or Banks et al 2010.) These same methods can be used for collecting and analyzing data on systems to obtain NIID data values. We will use a superscript of m to identify simulation model data and a superscript of s to identify system data; e.g., n^m will indicate the number of model NIID data values and n^s will indicate the number of system NIID data values.

2.1 Statistical Procedure

We are interested in determining if the difference between the means of a model, μ^m, and system, μ^s, performance measure is within the acceptable range of accuracy for the performance measure of interest. Specifically we want to determine if $D = \mu^m - \mu^s$ is contained in the acceptable range of accuracy. Let the acceptable range of accuracy for D be given by L for the lower limit and U for the upper limit. This interval (L, U) will include the value of zero and often U = -L. We can state the statistical form of the hypothesis to be tested as:

H_0: $L \le D \le U$

H_1: $D < L$ or $D > U$

We will collect data on both the model and system as discussed above for a set of experimental conditions to obtain approximately NIID data for the model and for the system. The t-distribution with the appropriate test statistic will be used for testing the means of our NIID data. As mentioned above, both type I and type II errors are important in model validation. The probabilities of both of these errors are characterized by the operating characteristic (OC) curve, which is defined as $P_A(D)$ = probability of accepting the null hypothesis when event D prevails. (We assume that the interval (L, U) is larger than the range of the sampling distribution unless stated otherwise.) The probability of type I error, $\alpha(D)$, is 1

- $P_A(D)$ when D has a value where the null hypothesis is true and the probability of type II error, $\beta(D)$, is $P_A(D)$ when D has a value where the alternative hypothesis is true. (See, e.g., Johnson, Miller, and Freund 2010 or Hines, et al. 2003 for a detailed discussion on the use of the t-distribution for testing means , type I and type II errors, and OC curves.)

We first consider the case of fixed sample sizes and then the case of selecting the sample sizes. There is a direct tradeoff between type I error, the *model builder's risk*, and type II error, the *model user's risk* when the sample sizes remain fixed. For *Methods 1, 2, and 3* below, the appropriate hypothesis test to be performed on D is the two-sample t-test with unknown and equal variances. (The sample sizes can be equal or unequal, depending on what is selected.)

Method 1: The first method that we will discuss is the simplest and is for the case of specifying the acceptable range of accuracy for D by (L, U) and a value for $\alpha/2 < 0.5$ for the *model builder's risk* at L and at U, which gives an overall maximum *model builder's risk* of α. Consider Figure 1. Figure 1(a) shows (L, U) on D with the sampling distribution at both L and U. Using a *model builder's risk* of $\alpha_1(D = L) = \alpha_1(D = U) = 0.05$ and the model and system data, calculate the critical value for the sampling distribution at L, say a_1, and the critical value for the sampling distribution at U, say b_1. These critical values give the acceptance region for the null hypothesis H_0 as the closed interval $a_1 \leq D \leq b_1$ and the rejection regions for the null hypothesis as $D < a_1$ and $D > b_1$. (See Figure 1(a).) The OC curve is calculated for these two critical values and is shown in Figure 1(b) by the curve that goes through L and U at 0.95 and is labeled (1). Figure 1(c) shows the risk curves which come from the OC curve. (Note: In Figure 1(c) risk curves $\alpha_i(D)$ and $\beta_j(D)$ are labeled α_i and β_j, respectively.) Risk curve $\alpha_1(D)$ is the *model builder's risk* and is for $L \leq D \leq U$. ($\alpha_1(D) = 1 – OC(D)$ for $L \leq D \leq U$, and equals zero elsewhere.) Risk curve $\beta_1(D)$ is the *model user's risk* and is for $D < L$ and $D > U$. ($\beta_1(D) = OC(D)$ for $D < L$ and $D > U$ and equals zero elsewhere.) To obtain the results of a specific hypothesis test, the calculated test statistic from the model and system data is used to determine if the null hypothesis is rejected or accepted.

Suppose a model user looking at the risk curve $\beta_1(D)$ believes the risk of having a model accepted as valid when in fact it is invalid is too high. Further, suppose that after some discussion and negotiation it is decided to set the *model builder's risk* at L and U to 0.4 instead of 0.05, i.e., have $\alpha_2(D = L) = \alpha_2(D = U) = 0.4$. (Recall there is a direct trade-off between type I and type II errors.) Suppose the new critical values are, say, a_2 and b_2. The acceptance region for the null hypothesis H_0 is now the closed interval $a_2 \leq D \leq b_2$ and the rejection regions for the null hypothesis are $D < a_2$ and $D > b_2$. We note that the new acceptance region is shorter, as can be seen in Figure 1(a). Figure 1(b) shows the new OC curve which goes through L and U at 0.6 and is labeled (2). Note that the shape of this new OC curve is identical to the previous OC curve except for the two locations on D where the OC curve value changes between 0 and 1 at L and at U. Figure 1(c) show the (new) risk curves $\alpha_2(D)$ and $\beta_2(D)$. The decrease in the *model user's risk* and the increase in the *model builder's risk* can readily be seen in Figure 1(c); these changes resulted from the increase in the *model builder's risk* to α_2 from α_1 at L and at U.

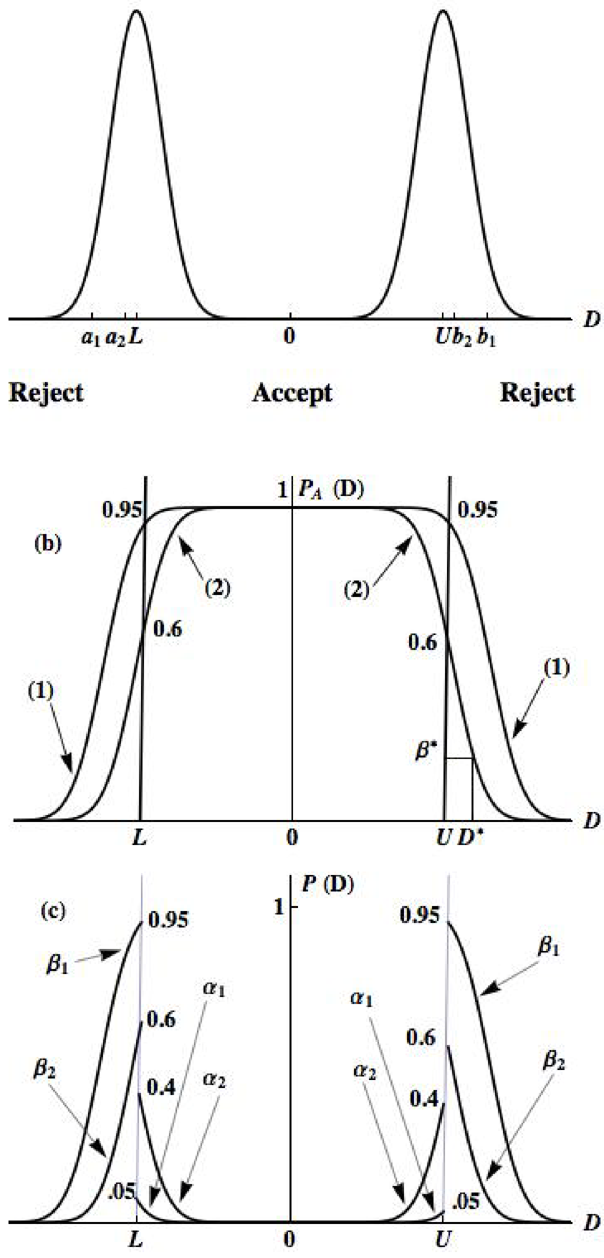

Figure 1. Fixed Sample Sizes.

There is no reason that the *model builder's risk* at L and U must be the same in Method 1 as was used in the two examples; but most often in applications the values chosen would be the same. If they are chosen to be different, the sum of the two must be less than one. The critical values at L and U would be calculated using the appropriate value for the *model builder's risk* at L and also at U.

Method 2: The second method also uses fixed sample sizes and is for the case of specifying the acceptable range of accuracy for D by (L, U) and the *model user's risk*, β, at locations (L – δ) and (U + δ) . Recall from the discussion of Figure 1(a) in Method 1 that the OC curves for the different values of *model builder's risk* were identical except for the locations where the OC curves values changed between 0 and 1 at L and at U. This is the situation when the same sample sizes and the same test statistic are used. In Method 1, the location of the OC curve is determined by the specification of the *model builder's risk* for D = L and for D = U along with the values of D for L and U. In Method 2, the location of the OC curve is determined by the specification of L, U, δ, and the *model user's risk*, β, for D = (L – δ) and D = (U + δ). (Look at Figure 1(b) to see how β^* and D^* are related to an OC curve.) The *model builder's risk* for D = L and for D = U will be determined in Method 2. An OC curve must first be calculated and then the portions of the OC curve that change between 0 and 1 must be shifted at both L and U to have proper locations. Different ways can be used to calculate the OC curve. We suggest calculating critical values using sample distributions at L and U as was done in Method 1 using some arbitrary chosen value for α(D = L) = α(D = U), say, 0.05. Using these values, an OC curve is determined. Shift the location of the portion of the OC curve that changes between 0 and 1 going through L such that the OC curve has the specified value of β at D = (L – δ). We next shift the location of the portion of the OC curve that changes between 0 and 1 going through U such that the OC curve has the specified value of β at D = (U + δ). Now find the value of α(D = L) (or α(D = U)) from where the OC curve intersects L (or U). One now can proceed as in Method 1 to calculate the critical values at L and U to obtain the acceptance and rejection regions for the null hypothesis, and then obtain the risk curves for the *model builder's risk* and the *model user's risk* from the OC curve.

The same values of δ and β have been used at L and at U in the presentation of Method 2. They do not need to be equal but they usual are in practice. If different values are used for δ, β, or both δ and β, the process described for Method 2 is still used.

The above discussion kept the sample sizes of observations on the model and system constant with the result that there is a direct tradeoff between type I and type II errors. Changing the sample sizes n^m and n^s causes the α and β risks to change. Increasing (decreasing) n^m or n^s or both n^m and n^s decreases (increases) the standard deviation of the sampling distribution, which decreases (increases) both α and β simultaneously. Consider Figure 2. (We will only discuss the behaviors at U since the behaviors at L and U are similar.) Figure 2(a) shows two sampling distributions at U; distribution (1) has smaller sample sizes than distribution (2). Note that the standard deviation of distribution (1) is larger than that of distribution (2). We set the *model builder's risk* at U to 0.3 for both distributions; and thus α_1(D = U) = α_2(D = U) = 0.3. Using the t-distribution to calculate the critical values of both distributions we obtain, say, b_1 and b_2, respectively, for distributions (1) and (2). Note that the critical value b_1 for distribution (1) will give a larger acceptance region for the null hypothesis than the critical value b_2 does for distribution (2). Figure 2(b) contains the OC curves for the distributions (1) and (2). Note that the OC curve is much steeper for distribution (2) and is preferred over the OC curve for distribution (1). However, that comes with the requirement of larger sample sizes for sampling distribution (2). Figure 2(c) contains the risk curves for both of these distributions. (Note: In Figure 2(c) risk curves $\alpha_i(D)$ and $\beta_j(D)$ are labeled α_i and β_j, respectively.) We see that both risk curves $\alpha_2(D)$ and $\beta_2(D)$ are smaller than risk curves $\alpha_1(D)$ and $\beta_1(D)$; i.e., both the *model builder's risk* and the *model user's risk* for sampling distribution (2) are smaller than those of sampling distribution (1). We note that sample sizes are important in the values of the risk curves. Thus one should consider sample sizes as well *model builder's risk* and *model user's risk*.

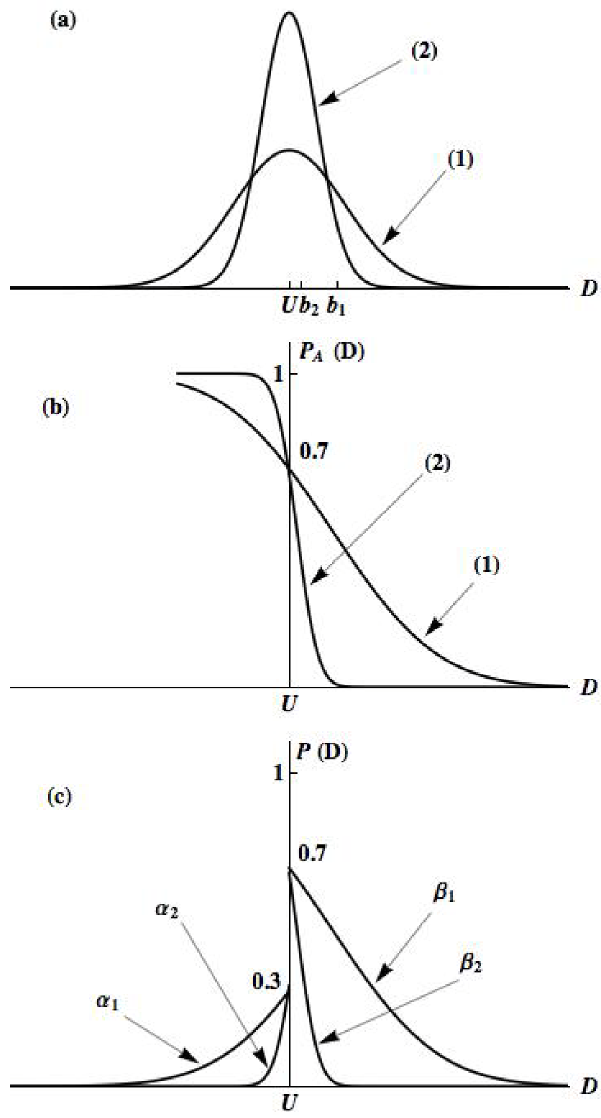

Figure 2. Unequal Sample Sizes.

Method 3: Methods 1 and 2 used fixed sample sizes and specified the *model builder's risk* and the *model user's risk*, respectively. In Method 3 both *model builder's risk* and *model user's risk* are specified along with L, U, and δ, and then the sample sizes are determined. Some statistic books discuss the determination of sample sizes when type I and type II errors are specified and include tables for equal sample sizes for some specific values of α and β; e.g. Johnson, Miller, and Freund 2010 and Hines, et al. 2003. After the sample sizes have been determined, Method 1 can be used to determine the critical values, the OC curves, and the risk curves.

Method 4: Trace-drive simulation models (Balci and Sargent 1983, Law and Kelton 2000) can be validated using Methods 1, 2, and 3 by replacing the two-sample t-test with the paired-sample t-test. The application of the statistical procedure is simpler for trace-driven simulations because there is only one sample to analyze instead of two samples. Also, more tables and graphs exists for the one-sample case for investigating type I and type II errors and sample sizes than for the two-sample case.

2.2 Comments on the Statistical Procedure

We now discuss various aspects of this new statistical procedure. Regarding the observable of systems to have system data: (a) some systems such as computer systems, communication systems, and call centers can have large amounts of data available, (b) some systems such as hospitals or manufacturers can have moderate amounts of data available, (c) some systems such as fire departments or systems having failures can have only small amounts of data available, and (d) other systems may not be observable to have any data available. Thus the number of data values on observable systems can vary from a few to large. We next discuss sample sizes, which relates to the availability of data. As shown in Figure 2, the sample sizes used for testing have a major effect on the risks associated with the hypothesis test. The larger the sample sizes, the less the risk. The sample size used for the system depends on what may be available on the system and then how much effort one wants to use to obtain the data values to use in testing the model validity. Regarding the sample size for the simulation model, any number of observations is possible if one wants to pay the price for them. Thus a judgment must be made on how much risk should be taken versus the cost and effort. Method 3 describes the determination of the sample sizes for specified risks; however, the determination may be difficult in some cases. An alternative to using Method 3 is to iterate the samples sizes using Methods 1 or 2 to achieve the desired risks; and with the use of computers today that may not take much effort.

For Methods 1, 2, and 3, we used the two-sample t-test with unknown and equal variances (with the sample sizes being equal or unequal depending on what is desired). The OC curve does not exist for the two-sample t-test with unknown and unequal variances (2003 Hines, et al.) and therefore this statistical procedure cannot be performed for this case. Assuming equal variances is reasonable since the same output process of the simulation model and of the system are being compared. Furthermore, Johnson, Miller, and Freund (2010) states that the t-test is not sensitive to small differences between the variances. If there is concern that the variances may have large differences, they can be statistically be tested to determine if they are indeed different (see, e.g., Johnson, Miller, and Freund 2010). In the rare case that the variances are known, then the appropriate t-test can be used in Methods 1, 2, and 3 instead of the t-test with unknown variances.

This new statistical procedure was illustrated by the testing of means using the t-test. The same statistical procedure can also be used for the testing of variances using the F-test. (For a discussion of using the F-test for variances see, e.g., Hines, et al. (2003), which also contains Tables of OC curves for the F-test.)

The usual procedure for validation is to specify only the acceptable range of accuracy and not to specify anything about the statistical analysis such as type I and type II errors for hypothesis testing (or confidence levels for use with confidence intervals). The selection of the values for type I and type II errors is often left to the people conducting model validation. The discussion of this statistical procedure above clearly illustrates that the interested parties (model builders, model users, models sponsors, etc.) should be involved in determining the values used in the statistical analysis and examining the risk curves. Their involvement would help with model credibility (Law and Kelton 2000, Sargent 2009).

It is possible that this statistical procedure will not work if the range of the sampling distribution is close to the length of the acceptable range of accuracy. This situation can occur if the length of the acceptable range of accuracy being specified is very small or the width of the sampling distribution is large due to having only a few system observations, or the variance of the performance measure of interest being large, or some combination of the previous three possible causes. If this situation occurs, then the sample sizes should be increased to reduce the standard deviation of the sampling distribution. Hopefully, this will solve the issue. This, however, may not provide a solution if there are only a few data observations available for the system and the system performance measure does not have a small enough variance. In this case, a formal statistical test may not be possible and thus a graphical approach will need to be used to determine validity (See Sargent 1996, 2001).

3. Validation of Analytic Stochastic Models

This new statistical procedure can be used for validation of analytic stochastic models in a straight forward manner. The NIID data values for the system for estimating the mean of the performance of interest and for use in the statistical procedure are obtained as they were for validation of simulation models. The mean of the performance measure for the analytic model is calculated from the analytic model and thus there is no statistical estimation of the mean. The statistical procedure now uses the one-sample t-test in Methods 1, 2, and 3 instead of the two-sample t-test used in the validation of simulation models. The use of this statistical procedure will be simpler here since there is only one sample now instead of two samples. Furthermore, there are more graphs and tables available for use with these Methods for one-sample cases than for two-sample cases; especially for Method 3. The comments made in Subsection 2.2 also apply for the validation of analytic stochastic models.

4. Conclusions

A new statistical procedure that uses hypothesis testing for validation of simulation and analytic stochastic models when an acceptable range of accuracy is specified was presented. Different methods were discussed that allow different ways of handling the *model builder's risk* and the *model user's risk*. These Methods are based on the use of the OC curve and they provide the risk curves for the *model builder's risk* and the *model user's risk*. It is common practice in some fields to use OC and risk curves such as those for acceptance sampling in control quality (Johnson, Miller, and Freund 2010). This new statistical procedure should also aid in model credibility if the model users are involved in its use.

Acknowledgment

The author thanks David M. Goldsman for his suggestions on the presentation of this work.

References:

Balci, O. and R. G. Sargent. 1981. A methodology for cost-risk analysis in the statistical validation of simulation models. *Comm. of the ACM* 24 (6): 190-197.

Balci, O. and R. G. Sargent. 1982a. Validation of multivariate response simulation models by using Hotelling's two-sample T^2 test. *Simulation* 39(6): 185-192.

Balci, O. and R. G. Sargent. 1982b. Some examples of simulation model validation using hypothesis testing. In *Proc. 1982 Winter Simulation Conf.*, ed. H. J. Highland, Y. W. Chao, and O. S. Madrigal, 620-629. Piscataway, New Jersey: IEEE.

Balci, O. and R. G. Sargent. 1983. Validation of multivariate response trace-driven simulation models. In *Performance 83*, ed. Agrawada and Tripathi, 309-323. North Holland.

Banks, J., J. S. Carson II, B. L. Nelson, and D. M. Nicol. 2010. Discrete-Event System Simulation, 5th Edition, Prentice Hall.

Hines, W. W., D. C. Montgomery, D. M. Goldsman, and C. M. Borror. 2003. Probability and Statistics in Engineering, 4th Edition, John Wiley.

Johnson, R. A., I. Miller, and J.E. Freund. 2010. Miller & Freund's Probability and Statistics for Engineers, 8th Edition, Prentice Hall.

Law, A. M. and W. D. Kelton. 2000. Simulation Modeling and Analysis, 3rd Edition, McGraw Hill.

Sargent, R. G. 1996. Some subjective validation methods using graphical displays of data. In *Proc. 1996 Winter Simulation Conf.*, ed. J. M. Charnes, D. J. Morrice, D. T. Brunner, and J. J. Swain, 345-351. Piscataway, New Jersey: IEEE.

Sargent, R. G. 2001. Graphical displays of simulation model data as statistical references. In *Simulation 2001: Proc. of the 4th St. Petersburg Workshop on Simulation*, ed. S. M. Ermakor, Yu. N. Kashtanov, and V. B. Melas, 109-118. Publisher: Chemistry Research Institute of St. Petersburg University.

Sargent, R. G. 2009. Verification and Validation of Simulation Models, In *Proceedings 2009 Winter Simulation Conference*, ed. M. D. Rossetti, R. R. Hill, B. Johansson, A. Dunkin, and R. G. Ingalls, 162-176. Piscataway, New Jersey: IEEE, 2009.