CentER

Discussion paper

No. 2002-58

STATISTICAL ANALYSIS OF RANDOM SIMULATIONS: BOOTSTRAP TUTORIAL

By David Deflandre, Jack P.C. Kleijnen

June 2002

ISSN 0924-7815

File: bootMM1CDP.doc

Written: 3 June 2002

Printed: 03-06-02 16:51

Statistical analysis of random simulations: bootstrap tutorial

David Deflandre [1] and Jack P.C. Kleijnen [2]

[1] Laboratoire Inter-établissement d'Informatique, de Modélisation et d'Optimisation des Systèmes (LIMOS)

Equipe de Recherche en Systèmes de Production de l'Institut Français de Mécanique Avancée (IFMA)

B.P. 265, F-63175 Aubière Cedex, France

[2] Corresponding author

Department of Information Management/Center for Economic Research (CentER)

Tilburg University (KUB), Postbox 90153, 5000 LE Tilburg, The Netherlands

Phone: +31 13 4662029; Fax: +31 13 4663377; E-mail: kleijnen@kub.nl

Web: http://www.tilburguniversity.nl/faculties/few/im/staff/kleijnen/

Abstract

The bootstrap is a simple but versatile technique for the statistical analysis of random simulations. This tutorial explains the basics of that technique, and applies it to the well-known M/M/1 queuing simulation. In that numerical example, different responses are studied. For some responses, bootstrapping indeed gives better statistical results than parametric statistical techniques do.

Keywords: Bootstrapping; normality; robustness; queuing simulation; statistics

JEL: C1

1. Introduction

This paper is a tutorial that explains the basics of the statistical analysis technique known as the bootstrap, and illustrates the application of the bootstrap through the derivation of confidence intervals for various responses of an M/M/1 queuing simulation. The M/M/1 is a well-known building block in discrete-event simulation; see Law and Kelton [6].

Bootstrapping implies resampling - with replacement – of a given sample. In our numerical illustration, this sample consists of the responses of (say) m simulation runs or replicates. For example, the response is the average waiting time per simulation run, and the sample consists of this average response observed for ten runs that use ten different pseudorandom number (PRN) streams - but the same traffic rate. In practice, much computer time (for example, five hours) is often needed to obtain the response for a single simulation run. However, once these data are obtained, bootstrapping is a fast analysis technique, which requires only seconds to compute statistically sound conclusions. Bootstrapping does not assume a specific distribution – such as the normal (Gaussian) distribution - for the response of interest.

Conceptually, the bootstrap may be explained as follows. Suppose that a sample of size *m* is available (for example, *m* average waiting times per simulation run). Now suppose that by chance one of these data elements gets lost. To keep the sample size constant at *m*, another data element is then counted twice. Obviously, the value of the sample average now changes. By repeating this chance experiment many times, the bootstrap gives many different average values – all computed from the same original sample. We shall define and illustrate the bootstrap more precisely, in Section 3.2.

Our *main conclusion* will be that bootstrapping can give valid statistical results even if the standard statistical assumption of normality does not hold. So, the bootstrap is a simple non-parametric (distribution-free) technique. Moreover, the statistic to be studied may be more complicated than the mean and variance, which we focus on in the illustrations; for example, Kleijnen and Van Groenendaal [5] use bootstrapping to classify journals into distinct quality classes.

This tutorial is written because the bootstrap technique is simple and versatile, but is not well known among simulation practitioners and theorists. A few recent discussions of bootstrapping in simulation are Demirel and Willemain [1], Friedman and Friedman [3], and Kleijnen, Cheng, and Bettonvil [4].

The remainder of this paper is organized as follows. Section 2 presents a simulation of the M/M/1 queuing system, using the Arena software and the C language respectively. Section 3 considers M/M/1 simulation outputs that are normally distributed; this section analyzes these responses through both the parametric Student *t* test and the bootstrap, which is explained in some detail. Section 4 analyzes other M/M/1 responses, including the means and variances of the responses in the transient state. Section 5 presents conclusions.

2. M/M/1 queuing simulation

By *definition*, M/M/1 assumes that the interarrival times of customers are independently and exponentially distributed with a constant arrival rate (say) λ; likewise, the customer service times are independently exponentially distributed with constant service rate μ; arrival and service times are also independent of each other. The symbol M in the notation M/M/1 refers to the Markov properties of the arrival and service times in this model; the symbol 1 means that there is a single server. Implicit in this notation are the assumptions of an unlimited capacity of the waiting room, and a First-In-First-Out (FIFO) priority rule (queue discipline). The traffic rate (utilization factor, load) ρ equals λ / μ. The M/M/1 reaches a steady state provided $\rho < 1$. In our examples we assume that the steady state is indeed reached if we select $\rho = 1/3$ and simulate $n = 10^7$ customers per run. However, when we simulate only ten customers per run, the M/M/1 shows transient behavior (see Section 4.2).

In the *steady state*, the M/M/1 has analytically known means (so it is easy to compare simulation results with the ‘true’ results) for the following responses: the number of customers in the system (say) L , the number of customers in queue (excluding the customer being served) *Lq*, the waiting time in the system *W*, and the waiting time in the queue (excluding the customer being served) *Wq*:

$$L = \frac{\lambda}{\mu - \lambda} \tag{1}$$

$$Lq = \frac{\lambda^2}{\mu(\mu - \lambda)} \tag{2}$$

$$W = \frac{1}{\mu - \lambda} \tag{3}$$

$$Wq = \frac{\lambda}{\mu(\mu - \lambda)} \tag{4}$$

At the start of each M/M/1 simulation run, we make the server idle and the queue empty. We program the simulation in the Arena simulation package and in the faster C language. In Arena we use its standard PRN generator. In C, we use L'Ecuyer's generator taken from Law & Kelton [6] (432-435). This generator will also be used to implement the bootstrap.

3. M/M/1 example with Gaussian simulation responses

In the first example we analyze the M/M/1 when we conjecture that its simulation gives responses that are normally distributed. A parametric statistical technique - such as the Student *t* test - should then give a *correct coverage probability*: the 1- α confidence interval should cover the true value with a 1- α probability. This true value is given by (1) through (4) if the simulation has indeed reached the steady state. Therefore, we simulate ten million (10^7) customers per run. For each run we estimate the four responses corresponding with (1) through (4) through their averages (say) $\overline{Y}_{i;\,j}$ with i = 1, … , 4 and j = 1, … , m. These m averages are independent and identically distributed (IID) because they result from the same M/M/1 simulation program with the same input value for ρ and non-overlapping PRN streams. This IID assumption is crucial for both the parametric and the bootstrap techniques. Moreover, we select a 'large' m value so that the Central Limit Theorem (CLT) applies; for example, we select an m of 80.

Let us first consider only one of the four responses, and let $\overline{\overline{Y}}$ denote the average of the m IID $\overline{Y}_j$. Further, let η denote the mean of these $\overline{Y}_j$, and let η_0 denote the true value following from (1) through (4). Then H_0 in (5) is the null hypothesis, whereas H_1 is the alternative hypothesis:

$$H_0: \eta = \eta_0; \; H_1: \eta \neq \eta_0. \tag{5}$$

To test each of the four null hypotheses, we use a type-I error probability of $\alpha = 0.05$. First we apply the parametric t test (Section 3.1); then the bootstrap (Section 3.2).

3.1 Student *t* test

To test the normality assumption implied by the t test, we construct the empirical probability distribution $\hat{F}$ that has a probability of $1/m$ at each element $\overline{Y}_j$ of the sample. Figure 1 gives this $\hat{F}$ and the corresponding estimated density function $\hat{f}$ for one of the four responses, namely $\overline{W}_q$. The chi-square goodness-of-fit test accepts $\hat{F}$ as a Gaussian distribution.

Table 1 displays the results of the t test for the four outputs. This table shows very small standard errors S, so – on hindsight – $m = 80$ is a high value. The t test does not reject H_0, as we expected from the start.

3.2 Bootstrap test

Based on Efron and Tibshirani [2] (45-53, 170-173) - and also Mooney and Duval [7] (10-11, 36-40) - we bootstrap the original sample of m IID observations $\overline{Y}_j$ (j = 1,…, m), as follows.

1. From the original sample, we draw a random sample of the same size m - *with replacement* – denoted by $\{\overline{Y}_1^*, \ldots, \overline{Y}_j^*, \ldots, \overline{Y}_m^*\}$ (j = 1, .., m). Figure 2 gives an example of the resulting bootstrapped distribution function $\hat{F}^*$ and density function $\hat{f}^*$, which resembles Figure 1 but is not identical to that figure. This bootstrap sample gives the bootstrap estimator $\overline{\overline{Y}}^* = \sum_{j=1}^{m} \overline{Y}_j^* / m$, which has zero probability of being identical to the original value of $\overline{\overline{Y}}$.

2. We *replicate* Step 1 (say) B times, where replicate b gives $\overline{\overline{Y}}_b^* = \sum_{j=1}^{m} \overline{Y}_{b;\,j}^* / m$ (b = 1, …, B). We take B = 1,000 .

3. We *sort* the B bootstrap observations $\overline{\overline{Y}}_b^*$, from the smallest observation – denoted by $\overline{\overline{Y}}_{(1)}^*$ - to the largest one – denoted by $\overline{\overline{Y}}_{(B)}^*$. (The sorted observations $\overline{\overline{Y}}_{(b)}^*$ are the so-called order statistics.) Figure 3a shows an example of the resulting empirical distribution function, which has a probability mass of 1/B at each point. A bootstrap 1 - α confidence interval is then $[\overline{\overline{Y}}_{(B\alpha/2)}^*,\ \overline{\overline{Y}}_{(B[1-\alpha/2])}^*]$. For example, if B = 1,000 and α = 0.05, then the lower limit is the 25th ordered value of the bootstrapped observations, and the upper limit is the 975th value. If non-integer values result for the particular B and α values, then we round to the next integer.

In Figure 3, the solid vertical line is at $\overline{Wq}$ = 1.250191; the square-dotted lines at 1.249817 and 1.250586, which are the 2.5% and 97.5% percentiles of the histogram; the dashed line is at η_0 = Wq_0 = 1.25. So this figure implies that the bootstrap interval does cover the true value η_0. Table 2 shows the bootstrap confidence interval for all our four responses.

4. M/M/1 example: more simulation responses

The preceding section gave normally distributed M/M/1 simulation responses with correct results for both the parametric *t* test and the bootstrap. This result agrees with the statistics literature showing that the *t* statistic is not very sensitive to non-normality. Obviously, this sensitivity decreases as the sample size *m* increases. We therefore investigate the effect of the number of simulation runs, *m* (Section 4.1).

Moreover, the simulation literature shows that the average of a long run is asymptotically normally distributed – even though the individual observations are non-normal and auto-correlated. We therefore investigate the effect of the number of customers per simulation run, *n* (Section 4.2).

Finally, the statistics literature shows that the χ^2 and the *F* statistics are more sensitive to non-normality than the *t* statistic. We therefore analyze the variances – instead of the means - of the simulation responses (Section 4.3).

4.1 Mean responses of run with $n = 10^7$ customers and varying number of runs *m*

Table 3 shows the *t* and the bootstrap intervals for five values of *m*, namely 2, 5, 10, 25, 50. This table shows that for *m* = 2, the bootstrap does not include the true value E(*Wq*) = 1.25. In fact, in this example the simulation provides only two numbers – namely 1.250349 and 1.250928, which both exceed E(*Wq*); so each bootstrap sample average exceeds the true value. In general, we recommend that m = 2 should not be used in bootstrapping.

4.2 Mean responses of run with n = 10 customers and varying run numbers m

With only n = 10 customers per run, the simulation remains in the *transient* state so we should not apply (1) through (4). In this academic example we can afford to estimate the true mean η through a big number of runs, namely 10^6; see Table 4. This table shows very small standard errors. So we use the averages in this table as the true values η to decide whether the confidence intervals cover the true mean. Figure 4 illustrates that the density function (estimated from m = 10,000 observations) does not look Gaussian.

Table 5 shows that neither the parametric interval nor the bootstrap interval ever misses η for $m \geq 50$. The t statistic turns out to be insensitive to the non-normality shown in Figure 4. In general, we conclude that bootstrapping is not useful when the simulation response is a run average.

4.3 Response variances of run with n = 10 customers and varying run numbers m

As the simulation response of interest we now consider *variances* instead of means, so (5) is replaced by

$$H_0\colon \sigma^2 = \sigma_0^2;\ H_1\colon \sigma^2 \neq \sigma_0^2. \qquad (6)$$

To obtain the 'true' value σ_0^2, we again use Table 4: we multiply the numbers in the last column by 10^3 (= $\sqrt{m}$). We consider different run numbers: m = 5, 10, 25, 40, 50, 80, 100.

The parametric 90% confidence interval for σ^2 is

$$\frac{(m-1)S^2(\overline{Y})}{\chi^2_{0.05;\,m-1}} \le \sigma^2 \le \frac{(m-1)S^2(\overline{Y})}{\chi^2_{0.95;\,m-1}} \tag{7}$$

where $\chi^2_{0.05;\,m-1}$ and $\chi^2_{0.95;\,m-1}$ are the 5% and the 95% quantiles of the χ^2 distribution with m – 1 degrees of freedom.

For the bootstrap intervals, we resample the m run averages, and re-estimate the variance from these m bootstrap observations, etc.

Table 6 shows that for $m \ge 50$ the bootstrap intervals do cover σ_0^2, whereas the parametric intervals do not.

Finally, we do not study the individual variance magnitudes, but *compare* the variances of two independent random samples. The first sample consists of m averages of n = 10 customers each, obtained through PRN stream 1 of L'Ecuyer's generator defined in Law & Kelton [6] (433-434); the second sample is obtained with stream 10. We test

$$\mathrm{H}_0\!: \sigma_1^2 = \sigma_{10}^2;\ \mathrm{H}_1\!: \sigma_1^2 \neq \sigma_{10}^2 \tag{8}$$

where σ_1^2 and σ_{10}^2 are the variances obtained through streams 1 and 10 respectively, so we know that this null-hypothesis is true. We consider m = 10, 50, 100 runs.

The parametric 90% confidence interval for the ratio σ_1^2/σ_{10}^2 is given by

$$\frac{S_1^2}{S_{10}^2}F_{0.05;\,m-1,m-1} \le \frac{\sigma_1^2}{\sigma_{10}^2} \le \frac{S_1^2}{S_{10}^2}F_{0.95;\,m-1,m-1} \tag{9}$$

where $F_{0.05;\ m\text{-}1,m\text{-}1}$ and $F_{0.95;\ m\text{-}1,m\text{-}1}$ are the 5% and the 95% quantiles of the F distribution with m – 1 degrees of freedom for both the numerator and the denominator.

Next, we bootstrap the m run averages, estimate the variance from these m IID. averages, etc. Table 7 shows that for $m \geq 50$ the bootstrap intervals do cover the ratio $\sigma_{1_0}^2/\sigma_{10_0}^2$ = 1, whereas the parametric intervals do not; that is, the F statistic is sensitive to non-normality.

5. Conclusion

We used a basic simulation model – namely the M/M/1 queue – to compare parametric and bootstrap tests. In case of normally distributed responses both methods give correct results; that is, the procedures give confidence intervals that cover the true value with a probability of 1 - α. In case of 'serious' non-normality, however, only the bootstrap gives good confidence intervals; such non-normality occurs if other responses than means are of interest, for example, variances.

References

- [1] O.F. Demirel and T.R. Willemain, Generation of simulation input scenarios using bootstrap methods, Journal Operational Research Society 53 (2002), 69-78.

- [2] B. Efron and R.J. Tibshirani, An introduction to the bootstrap, Chapman & Hall (New York, 1993).
- [3] L.W. Friedman and H.H. Friedman, Analyzing simulation output using the bootstrap method, Simulation 64, Vol. 2 (1995) 95-100.
- [4] J.P.C. Kleijnen, R.C.H. Cheng, and B. Bettonvil, Validation of trace-driven simulation models: bootstrapped tests, Management Science 47, Vol. 11 (2001) pp. 1533-1538.
- [5] J.P.C. Kleijnen and W. Van Groenendaal, Measuring the quality of publications: new methodology and case study. Information Processing & Management 36 (2000) 551-570.
- [6] A.M. Law and W.D. Kelton, Simulation modeling and analysis, 3rd edition, McGraw-Hill (New York, 2000).
- [7] C.Z. Mooney and R.D. Duval, Bootstrapping: a nonparametric approach to statistical inference, Sage Publications, Series: Quantitative Applications in the Social Sciences, no. 07-095, Newbury Park, 1993.

Table 1: Student t test for mean responses of M/M/1 simulation with m = 80 runs and n = 10^7 customers per run; critical value $t_{0.05;\,79}$ =1.99045

	η_0	$\bar{\bar{Y}}$	S	$\lvert t_0 \rvert$
L	0.50000	0.50005	0.00037	1.24215
Lq	0.16667	0.16670	0.00025	1.07155
W	3.75000	3.75044	0.00237	1.67355
Wq	1.25000	1.25019	0.00175	0.97651

Table 2: Bootstrap confidence intervals with B = 1000 bootstrap samples, for the M/M/1 responses corresponding with Table 1

	η_0	Confidence Interval
L	0.50000	[0.499972,0.500134]
Lq	0.16667	[0.166644,0.166751]
W	3.75000	[3.749963,3.750972]
Wq	1.25000	[1.249817,1.250586]

Table 3: Parametric and bootstrap confidence intervals for varying *m*;

bold face denotes type-I error; remaining symbols defined in Tables 1 and 2

		Parametric *t* test	Bootstrap
m = 50	*L*	[0.499850 ; 0.500080]	[0.499849 ; 0.500061]
	Lq	[0.166564 ; 0.166730]	[0.166565 ; 0.166716]
	W	[3.749144 ; 3.750531]	[3.749158 ; 3.750409]
	Wq	[1.249317 ; 1.250454]	[1.249335 ; 1.250340]
m = 25	*L*	[0.499876 ; 0.500194]	[0.499883 ; 0.500175]
	Lq	[0.166587 ; 0.166813]	[0.166592 ; 0.166799]
	W	[3.749380 ; 3.751203]	[3.749435 ; 3.751153]
	Wq	[1.249497 ; 1.251015]	[1.249537 ; 1.250928]
m = 10	*L*	[0.499890 ; 0.500270]	[0.499920 ; 0.500222]
	Lq	[0.166573 ; 0.166867]	[0.166595 ; 0.166831]
	W	[3.749097 ; 3.751223]	[3.749268 ; 3.750967]
	Wq	[1.249278 ; 1.251225]	[1.249455 ; 1.250981]
m = 5	*L*	[0.499792 ; 0.500419]	[0.499911 ; 0.500288]
	Lq	[0.166504 ; 0.166963]	[0.166590 ; 0.166876]
	W	[3.747991 ; 3.752172]	[3.748745 ; 3.751368]
	Wq	[1.248638 ; 1.251893]	[1.249265 ; 1.251351]
m = 2	*L*	[0.497573 ; 0.502909]	[**0.500031** ; 0.500451]
	Lq	[0.164953 ; 0.168739]	[**0.166697** ; 0.166995]
	W	[3.739419 ; 3.762676]	[**3.750593** ; 3.752502]
	Wq	[1.239679 ; 1.262842]	[**1.250349** ; 1.250928]

Table 4: Mean responses of simulation run with only 10 customers each, estimated from 10^6 observations

	Average	Standard error
L	0.487292	0.000382
Lq	0.164769	0.000266
W	3.469899	0.001828
Wq	0.970643	0.001280

Table 5: Confidence intervals for only n = 10 customers per simulation run for varying m; see Tables 3 and 4

		Parametric (t-test)	Bootstrap
m = 50	*L*	[0.378961 ; 0.678679]	[0.410282 ; 0.691098]
	Lq	[0.081722 ; 0.326162]	[0.110272 ; 0.333543]
	W	[3.274720 ; 4.391356]	[3.320852 ; 4.418864]
	Wq	[0.703133 ; 1.632256]	[0.761513 ; 1.651382]
m = 25	*L*	[0.337200 ; **0.474021**]	[0.343763 ; **0.471647**]
	Lq	[0.063101 ; **0.133002**]	[0.068224 ; **0.132341**]
	W	[2.853836 ; 3.684307]	[2.882404 ; 3.666295]
	Wq	[0.473647 ; **0.928672**]	[0.507649 ; **0.925314**]
m = 10	*L*	[0.341220 ; **0.474677**]	[0.320291 ; 0.503626]
	Lq	[0.038709 ; **0.134443**]	[0.050122 ; **0.126516**]
	W	[2.641867 ; 3.546088]	[2.755872 ; **3.443103**]
	Wq	[0.331028 ; **0.836934**]	[0.385879 ; **0.790322**]
m = 5	*L*	[0.205214 ; 0.676804]	[0.300132 ; 0.581887]
	Lq	[-0.000873 ; 0.199383]	[0.039158 ; **0.161557**]
	W	[2.451091 ; 4.164820]	[2.817221 ; 3.831998]
	Wq	[0.145700 ; 1.194277]	[0.346806 ; 0.993171]
m = 2	*L*	[-1.615904 ; 2.576032]	[0.315107 ; 0.645021]
	Lq	[-0.625336 ; 0.832008]	[0.045988 ; **0.160684**]
	W	[-3.287186 ; 10.746990]	[3.177641 ; 4.282159]
	Wq	[-3.312499 ; 4.796973]	[0.423121 ; 1.061353]

Table 6: Response variance σ^2 of four responses

$\sigma_0^2(L) = 0.1459$, $\sigma_0^2(Lq) = 0.0708$, $\sigma_0^2(W) = 3.3423$, $\sigma_0^2(Wq) = 1.6379$

		Parametric (χ^2-test)	Bootstrap
m = 100	*L*	[**0.1361** ; 0.2177]	[0.0726 ; 0.2841]
	Lq	[0.0767 ; 0.1227]	[0.0310 ; 0.1781]
	W	[2.5686 ; 4.1081]	[2.0092 ; 4.2601]
	Wq	[1.4207 ; 2.2722]	[0.0197 ; 3.1053]
m = 80	*L*	[**0.1517** ; 0.2568]	[0.0943 ; 0.2971]
	Lq	[**0.0770** ; 0.1304]	[0.0440 ; 0.1590]
	W	[**3.5473** ; 6.0043]	[2.2412 ; 7.0636]
	Wq	[**2.0240** ; 3.4259]	[1.1925 ; 4.3765]
m = 50	*L*	[**0.2054** ; 0.4015]	[0.0453 ; 0.6153]
	Lq	[**0.1366** ; 0.2671]	[0.0179 ; 0.4165]
	W	[2.8507 ; 5.5736]	[1.7493 ; 6.5558]
	Wq	[**1.9737** ; 3.8588]	[0.8229 ; 5.1368]
m = 40	*L*	[0.0653 ; **0.1387**]	[0.0393 ; 0.1508]
	Lq	[0.0225 ; **0.0479**]	[0.0122 ; **0.0515**]
	W	[1.7187 ; 3.6501]	[1.3694 ; **3.3209**]
	Wq	[0.6514 ; **1.3834**]	[0.4649 ; **1.3700**]
m = 25	*L*	[0.0181 ; **0.0476**]	[0.0170 ; **0.0363**]
	Lq	[0.0047 ; **0.0124**]	[0.0065 ; **0.0110**]
	W	[0.6669 ; **1.7537**]	[0.5799 ; **1.3972**]
	Wq	[0.2002 ; **0.5265**]	[0.1103 ; **0.4756**]

Table 7: Response variance ratios $\sigma_1^2 / \sigma_{10}^2 = 1$

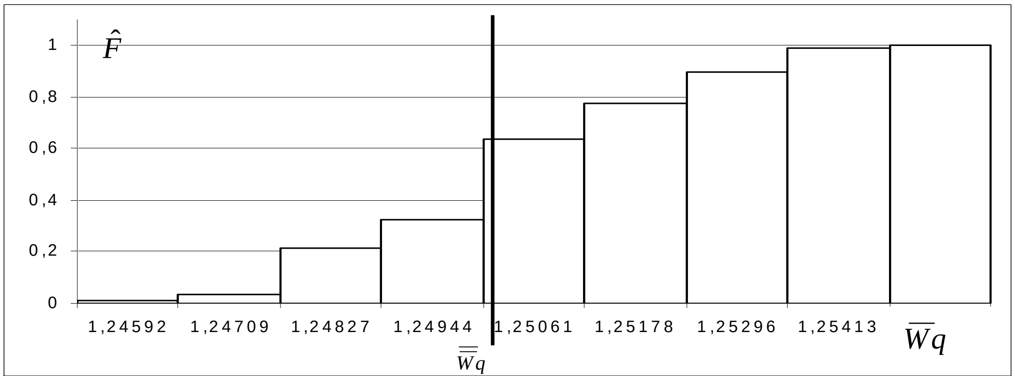
		Parametric *F* test	Bootstrap
m = 100	*L*	[0.73807 ; 1.43436]	[0.73914 ; 2.88741]
	Lq	[0.69675 ; 1.35407]	[0.31897 ; 4.17641]
	W	[**1.07688** ; 2.09282]	[0.41237 ; 3.01069]
	Wq	[**1.15986** ; 2.25409]	[0.65561 ; 1.57299]
m = 50	*L*	[0.93755 ; 2.42206]	[0.47687 ; 4.89792]
	Lq	[**1.02148** ; 2.63887]	[0.41887 ; 7.24482]
	W	[**1.29542** ; 3.34656]	[0.83133 ; 4.37372]
	Wq	[**1.60051** ; 4.13474]	[0.78944 ; 6.79824]
m = 10	*L*	[0.27827 ; 2.81200]	[0.02318 ; 3.21575]
	Lq	[0.21617 ; 2.18444]	[0.00545 ; 3.39574]
	W	[0.07635 ; **0.77159**]	[0.09051 ; **0.64896**]
	Wq	[0.05964 ; **0.60270**]	[0.00632 ; 1.02862]

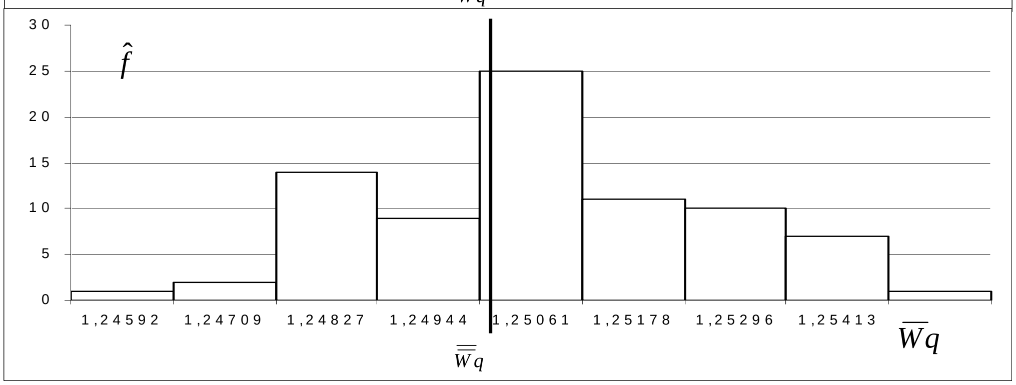
Figure 1: (a) Empirical Distribution $\hat{F}$ (b) Empirical Density Function $\hat{f}$ from the sample $\overline{Wq}$ (m = 80)

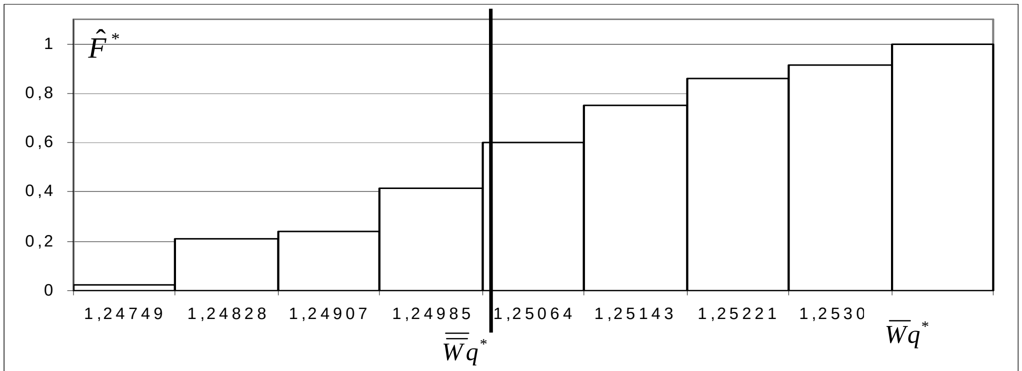
Figure 2: (a) Empirical Distribution, (b) Empirical Density Function, from the bootstrapped $\overline{Wq}^*$ (m = 80)

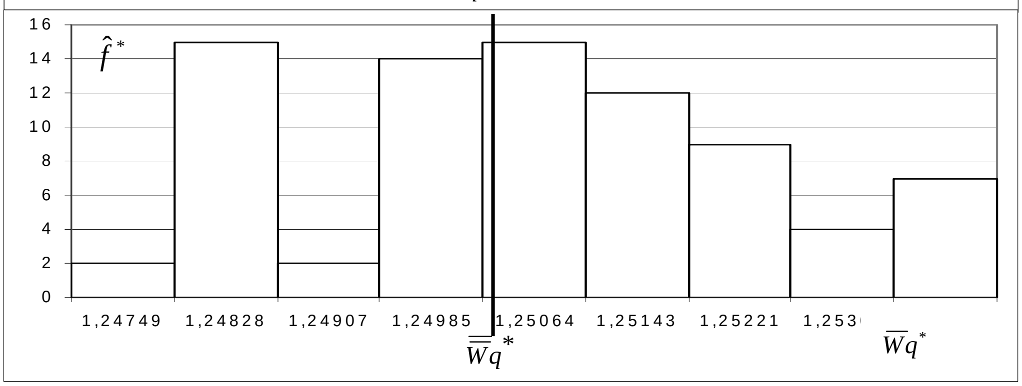
Figure 3: (a) Probability Distribution from the B = 1000 sample variables $\overline{\overline{Wq}}^*$, (b) Density Function from the B = 1000 sample variables $\overline{\overline{Wq}}^*_{(b)}$

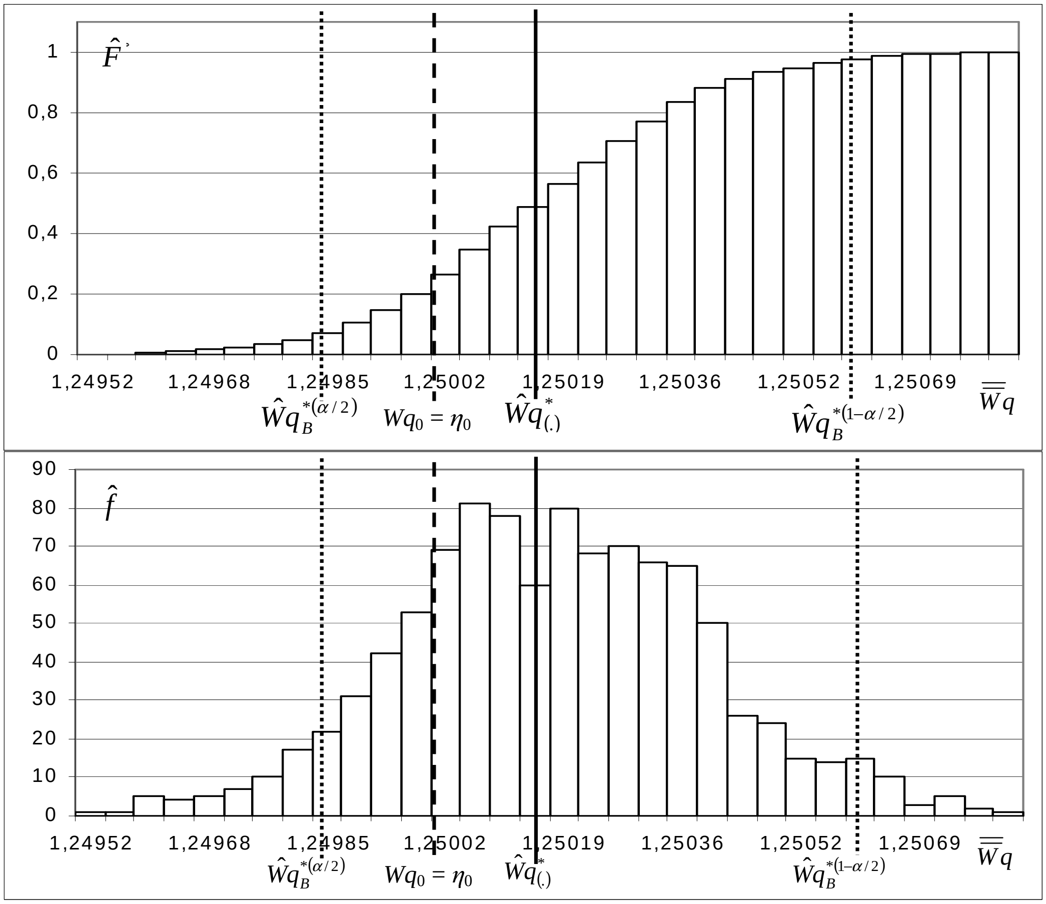
Figure 4: Empirical Density Function for m = 10,000 runs, each run simulating only n = 10 customers

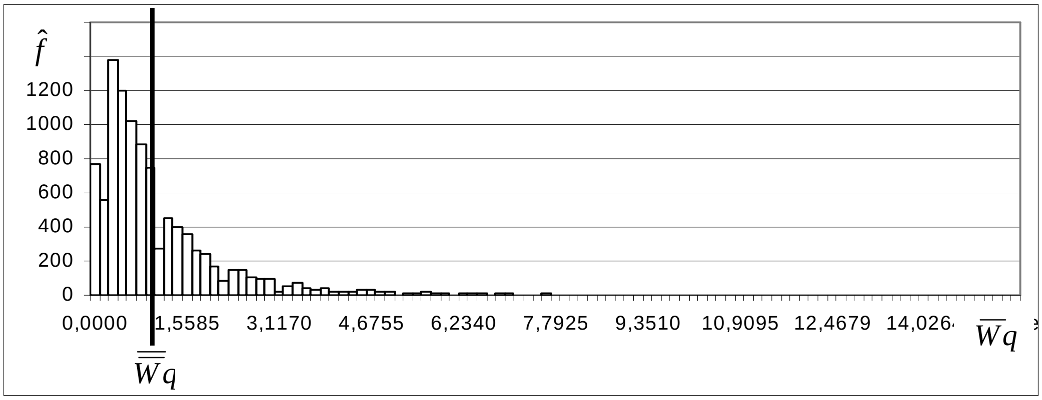

$\hat{F}$
1
0,8
0,6
0,4
0,2
0
1,24592
1,24709
1,24827
1,24944
1,25061
1,25178
1,25296
1,25413
$\overline{Wq}$
$\overline{\overline{Wq}}$


$\hat{f}$
30
25
20
15
10
5
0
1,24592
1,24709
1,24827
1,24944
1,25061
1,25178
1,25296
1,25413
$\overline{Wq}$
$\overline{\overline{Wq}}$


$\hat{F}^*$
1
0,8
0,6
0,4
0,2
0
1,24749
1,24828
1,24907
1,24985
1,25064
1,25143
1,25221
1,2530
$\overline{\overline{W}}q^*$
$\overline{W}q^*$


$\hat{f}^*$
16
14
12
10
8
6
4
2
0
1,24749
1,24828
1,24907
1,24985
1,25064
1,25143
1,25221
$\overline{\overline{W}}q^*$
$\overline{W}q^*$


$\hat{F}$
1
0,8
0,6
0,4
0,2
0
1,24952
1,24968
1,24985
1,25002
1,25019
1,25036
1,25052
1,25069
$\overline{\overline{W}}q$
$\hat{W}q_B^{*(\alpha/2)}$
$Wq_0 = \eta_0$
$\hat{W}q^*_{(.)}$
$\hat{W}q_B^{*(1-\alpha/2)}$
$\hat{f}$
90
80
70
60
50
40
30
20
10
0
1,24952
1,24968
1,24985
1,25002
1,25019
1,25036
1,25052
1,25069
$\overline{\overline{W}}q$
$\hat{W}q_B^{*(\alpha/2)}$
$Wq_0 = \eta_0$
$\hat{W}q^*_{(.)}$
$\hat{W}q_B^{*(1-\alpha/2)}$


$\hat{f}$
1200
1000
800
600
400
200
0
0,0000
1,5585
3,1170
4,6755
6,2340
7,7925
9,3510
10,9095
12,4679
$\overline{W}q$
$\overline{\overline{W}}q$