VIRGINIA COMMONWEALTH UNIVERSITY

Virginia Commonwealth University
VCU Scholars Compass

Theses and Dissertations | Graduate School

1997

Estimation of the Ld100p When p Is Small

Daniel Bettendorf
danielbettendorf@yahoo.com

Follow this and additional works at: http://scholarscompass.vcu.edu/etd

 Part of the Life Sciences Commons

© The Author

Downloaded from
http://scholarscompass.vcu.edu/etd/4373

This Thesis is brought to you for free and open access by the Graduate School at VCU Scholars Compass. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of VCU Scholars Compass. For more information, please contact libcompass@vcu.edu.

Virginia Commonwealth University
School of Medicine

This is to certify that the thesis prepared by Daniel Bettendorf entitled "Estimation of the LD100p when p is Small" has been approved by his committee as satisfactory completion of the thesis requirement for the degree of Master of Science.

[redacted]

Sung C. Choi, Ph.D., Director of Thesis

[redacted]

Alvin M. Best, Ph.D., School of Medicine

[redacted]

Patricia Pepple Williamson, Ph.D., School of Humanities & Sciences

[redacted]

Walter H. Carter, Jr., Ph.D., Department Chair

[redacted]

Hermes A. Kontos, M.D., Ph.D., Vice-President for Health Sciences and Dean, School of Medicine

[redacted]

Jack L. Haar, Ph.D., Dean, School of Graduate Studies

July 31, 1997

Date

Estimation of the LD100p when p is Small

A thesis submitted in partial fulfillment of the requirements of the degree of Master of Science at Virginia Commonwealth University.

By

Daniel M. Bettendorf
B.A., Washington and Lee University, 1991
Lexington, Virginia

Director: Sung C. Choi, Ph.D.
Professor
Department of Biostatistics

Virginia Commonwealth University
Richmond, Virginia
August, 1997

Acknowledgments

Whatever merits this thesis has are chiefly the result of the guidance and direction of my advisor, Dr. Sung Choi. He suggested the problem, indicated the path and supervised the progress.

I would also like to acknowledge the helpful comments and suggestions of the other members of my committee, Dr. Al Best and Dr. Patti Williamson, as well as those of my undergraduate academic advisor, Professor Robert Johnson of Washington and Lee University.

Finally, I owe a special debt of gratitude to my friend and colleague Roger Gibb for endless hours of patient assistance and kind encouragement. His technical prowess in computer matters and related issues made the completion of this project possible.

Ad Maiorem Dei Gloriam

Table of Contents

Page

List of Tables

Table **Page**

List of Figures

Figure **Page**

Abstract

ESTIMATION OF THE LD_{100p} WHEN P IS SMALL

Daniel M. Bettendorf, M.S.

A thesis submitted in partial fulfillment of the requirements of the degree of Master of Science at Virginia Commonwealth University.

Virginia Commonwealth University, 1997

Director: Sung C. Choi, Ph.D.
Professor, Department of Biostatistics

This thesis concerns the estimation of extreme quantiles on a dose-response curve. It focuses on the Robbins-Monro and up-and-down procedures. Simulation studies run in search of the LD_{15} and LD_{30} using a variety of methods suggest that the Robbins-Monro procedure is optimal in terms of Monte Carlo MSE and bias. The up-and-down procedure's performance differs in many cases only slightly from that of the Robbins-Monro process, therefore indicating its value as a practical alternative to the Robbins-Monro process for extreme-quantile estimation.

CHAPTER 1: INTRODUCTION

1.1 The Tolerance Distribution

This thesis concerns the level of stimulus at which a certain proportion of responses can be expected. It is assumed that all experiments discussed have a quantal response. That is, the response to stimulus is binary rather than continuous. For example, the response might be death and the stimulus is some dose (or log dose) of a toxic substance.

The stimulus-response relationship is described by the tolerance distribution, which gives the probability of response at given levels of the stimulus or dose. In the case of subjects exposed to a toxin, we expect that at increasing levels of the toxin the probability of death will increase. Figure 1 illustrates this concept most generally.

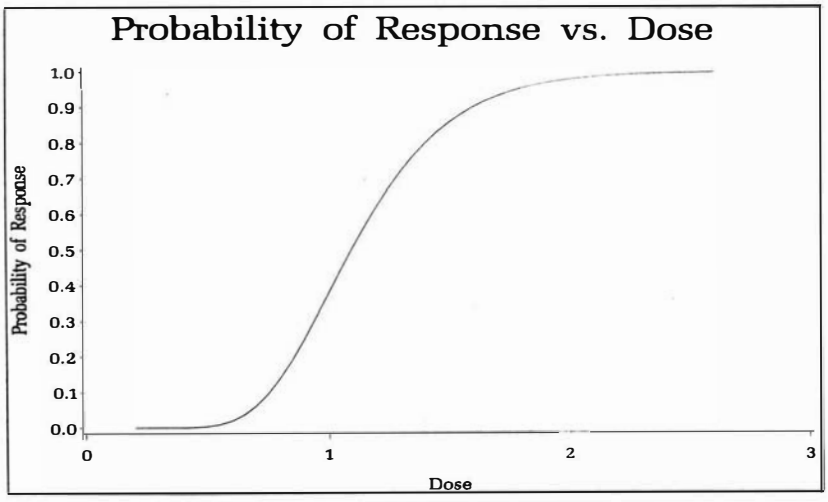

Figure 1 Tolerance Distribution

We might arrange these initial concepts in a more rigorous fashion in the following manner. Quantal response is a situation where stimulus (e.g., dose of a drug) is applied to n experimental units (e.g., animals) and r respond and $n - r$ do not respond. Furthermore, our main assumption is that for any given individual there is an individual effective dose (IED) and the tolerance distribution is a distribution of these IED's across the population. In practice, we often assume this tolerance distribution to be normal or logistic; failing that, a transformation of the dose might make that assumption more plausible.

An important problem is to find the level of stimulus or dose where a certain proportion, say p, of the population can be expected to respond. We wish to estimate the quantile q_p such that

$$\int_{-\infty}^{q_p} dF(x) = p$$

for a given tolerance distribution F. When $p = 0.5$ and the response is death, the dose level is called the Lethal Dose-50 or LD_{50}; alternatively, such a dose, regardless of response, is referred to as the Effective Dose-50 or the ED_{50}. In general, for the quantile where $100p$ % of the population can be expected to respond, the terms are LD_{100p} and ED_{100p}.

1.2 Examples

There are a variety of circumstances under which one might be interested in the estimation of such a quantile. In toxicity studies, one is commonly searching for thresholds, i.e., levels after which a certain unacceptable proportion of responses (such as death in laboratory animals) might be expected. Several methods for estimating the LD_{50} have been proposed (Hamilton, 1979). In this study we are interested mainly in those quantiles where p is relatively small (e.g., 0.15 or 0.3); relatively few studies have focused on methods for finding the quantiles where p is other than 0.5 (Wu, 1985).

Toxicity studies are not the only application of this methodology. The approaches discussed below are appropriate whenever the outcome is binary and the search is for the point or place where a certain proportion of outcomes may be expected. We might easily imagine an agricultural firm that wishes to discover how to maximize use of fertilizer while containing the amount of crop damage due to overfertilization. While fertilizers naturally help plants to grow, they are in certain quantities toxic to plants as well. Each plant, or group of plants, would have a certain tolerance level to the amount of fertilizer; that is, each plant would have an Individual Effective Dose (IED).

The researchers might therefore design an experiment in which they applied various amounts of fertilizer to different rows of a particular crop. Then each crop has to be assessed after a certain time period as either seriously damaged by fertilizer or not. If the case of damaged crops is considered as a response, the researchers might have a certain amount of crop damage in mind that would be considered acceptable given their

desire to maximize fertilizer use; let us say 10%. In that case, conducting the experiment as described, the researchers would be searching for the LD_{10} on the response curve as a function of amount of fertilizer.

Other examples of the application of these procedures are numerous. There might be a certain metal strip under testing which has a breaking point distribution; the researcher might be then interested in finding the maximum tension where a 5% failure rate is expected (see, for example, Wu, 1985). In short, in any situation where there is some kind of binary response and the above assumptions can reasonably be made, the investigator may be interested in estimating a given quantile.

Finally, a compelling example of estimating a quantile other than the LD_{50} is the case of the maximum tolerable dose (MTD) in cancer studies; the MTD is the LD_{33} on the tolerance distribution of the drug under study. Other methods besides those discussed here have also been explored for this application (Storer, 1989).

1.3 Purpose of Study

The sequential procedures such as the Robbins-Monro process or the up-and-down method are alternatives to a fixed sample, non-sequential approach such as the Spearman-Karber estimator. In a variety of fixed-sample situations, the trimmed Spearman-Karber estimator has been found to be optimal (Hamilton et. al, 1979).

Investigation of the sequential procedures for the LD_{50} has been extensive; Davis conducted a particularly germane simulation study involving most of the methods discussed below. His results indicated that the Robbins-Monro (RM; Robbins, 1951)

method had a remarkably good performance when compared to the other methods, including the non-sequential Spearman-Karber estimator.

The problem is, however, that there has not been sufficient investigation into the relative advantages of these sequential procedures when the quantile to be estimated is small (or large). This oversight is pertinent because, for example, it is known that difficulties arise in the Robbins-Monro process when the quantile sought after is extreme in this sense. The main problem is that a single positive response late in the estimation sequence can cause a big jump in dose while it will take the process a long time to recover from this jump; that is, in later iterations the step sizes are quite small and a large jump past the quantile may take too many iterations to correct.

This difficulty alone is enough to warrant some comparative simulation studies of the most popular sequential procedures in clinical trials. In order to see how the procedures compare in the case of extreme quantiles, we have conducted the investigation that follows.

1.4 Literature Review

The most important results relative to this inquiry are given by Davis (1971). As we mentioned above, Davis' simulation results indicated the overall superiority of the Robbins-Monro method for the LD_{50}, at least when the sample size is relatively small.

Much later Wu (1985) conducted similar simulation studies on the estimation of other quantiles and achieved similar results. While he considered the up-and-down method for estimating the LD_{50}, he did not report his results because he claimed that the

method was consistently the worst of those he tried. Much of the paper is devoted to an exploration of an adaptation of the Robbins-Monro procedure as well as a parametric method he himself proposed.

The Adaptive Robbins-Monro (ARM) involves an effort to use regression to estimate the slope of the response curve and to incorporate that into the sequential procedure. As a result of its relevance to our present aims we have included the ARM in our investigation. Wu's own contribution, however, seemed too far afield for our consideration. In short, the procedure involves making some general assumptions about the parametric form of the tolerance distribution and using the data to obtain maximum-likelihood estimates of the parameters to approximate the distribution and draw the next dose level from what would be the quantile on that approximated curve. While the idea is clearly clever, it involves too many assumptions about the form of the tolerance distribution and it is far more complicated than either RM or up-and-down.

Simplicity is, in fact, the main virtue of the up-and-down method. While it might be inferior to RM, it is possible that some level of outperformance is less attractive than the clinical simplicity offered by the up-and-down procedure. One cannot overemphasize the importance of having a procedure that can be easily explained and correctly executed.

Despite its simplicity and efficiency relative to fixed-sample methods, the literature on the up-and-down procedure is scarce. A review of the last 25 years under the heading "up-and-down" in the *Current Index of Statistics* will show only a handful of entries. Those relevant to this study are two papers by Little (1974a, 1974b).

Little was concerned with simulation studies of the different estimators generally available. He described another method of obtaining estimators, the "minimum chi-square" analysis, and compared it to the maximum-likelihood estimators of Dixon-Mood and Brownlee. His studies showed that there was very little difference among these estimators for either normal or logistic tolerance distributions; in the latter paper he confirmed that the methods based on these symmetric distributions are relatively robust by simulating results from an extreme-value distribution.

Hsi (1969) is the main resource for a multiple up-and-down design. This design uses several experimental units per trial rather than the classic single-unit trials of the traditional up-and-down method. Most of the details of this paper relevant to this undertaking are explored below. Durham and Flournoy (1994) provided the main outline for how to modify the up-and-down design to accommodate quantiles other than the LD_{50}.

McLeish and Tosh (1983) focused on estimation of extreme quantiles (such as the LD_{05}). Their concern centered around experiments where the experimental units were precious or highly valuable, such as primates. They therefore explored an estimation procedure that began at very low doses and continued increasing by small increments until a response was recorded. This process is repeated. The distribution of ending points of these dose sequences is then approximated to locate the LD_{05}. It is an ingenious approach when there is the added constraint that a response is to be avoided as much as

possible, but it is rather complicated, especially when compared to the two methods investigated in this study.

Because Hamilton has shown the superiority of the trimmed Spearman-Karber estimator for the LD_{50} with symmetric background distributions (Hamilton, 1979), we thought it worthwhile to search for similar estimators for the general LD_{100p}. Wu pointed out that the Spearman-Karber can be easily modified to a sequential procedure by taking the estimate at each stage and using that to decide the next dose level; however, as no modified estimator for the LD_{100p} could be found, this avenue is available only in the case of the LD_{50}.

CHAPTER 2: DISCUSSION OF METHODS OF ESTIMATION

2.1 The Robbins-Monro Process

The Robbins-Monro process is a simple sequential allocation scheme that easily accommodates various group sizes at different trial stages. Let x_0 be the initial esimate of the LD_{100p}. The design can be summarized neatly by writing the next design level or dose level at stage $n + 1$ as

$$x_{n+1} = x_n - \frac{\Delta}{n}\left(\frac{r_n}{t_n} - p\right),$$

where Δ is a constant, $n = 1,2, ..., r_n$ is the number of responders in a group of (usually) fixed size $t_n = k$, and p is the quantile whose estimate is sought. This formula applies, of course, when we wish to test only one unit at a time, rendering the proportion of responders invariably either 0 or 1 (i.e., $r_n = 0$ or 1).

The asymptotic variance of the estimate x_n takes a minimum when the step constant is related to the slope of the response curve at the desired quantile; that is,

$$\beta = \left\{\frac{dF(x)}{dx}\right\}_{x=q_p}$$

where q_p is the quantile as defined in Section 1.1. Given the slope β and response curve $F(x)$, we set the step constant $\Delta = 1/\beta$ to attain the lower bound for the asymptotic variance (Wetherill, 1986).

That the RM estimator is consistent (i.e., $\lim_{n\to\infty} = q_p$) follows directly from the following theorem due to Robbins (Robbins, 1951).

THEOREM Suppose that the cumulative distribution function F(x) of the tolerance distribution as described above has the following properties:

(1) $F(x)$ is nondecreasing;

(2) $F(q_p) = p$;

(3) $F'(q_p) > 0$,

then $\lim_{n\to\infty} \mathrm{E}(x_{n+1} - q_p)^2 = 0$.

Naturally, it is seldom the case that sufficient prior knowledge is available about the underlying tolerance distribution to warrant hazarding a guess at the slope of the curve at the proposed quantile. While there are methods, discussed below, for obtaining reasonable estimates of this value, it might be safer merely to postulate a value for the step constant Δ that seems suitable for the case at hand. A simple rule of thumb might be to make an estimate of the standard deviation of the underlying distribution and use that figure; that is, if the distribution is fairly spread out we might expect the slope of the response curve to be small and its reciprocal correspondingly large. In any case, in all practical circumstances, unless one wishes to incorporate "adaptive" methods for the process, some sort of guess will have to be made.

One possible adaptive method is to estimate the slope of the response curve at the quantile by applying, somewhat naively, ordinary regression estimation to the responses.

That is, we might estimate the slope β by setting it equal, after each iteration ($n > 1$), to the following:

$$\hat{\beta}_n = \sum y_i (x_i - \bar{x}_n) / \sum (x_i - \bar{x}_n)^2$$

This leads to the "adaptive" Robbins-Monro process:

$$x_{n+1} = x_n - \frac{1}{n\hat{\beta}_n}\left(\frac{r_n}{t_n} - p\right)$$

It has been proved that this process, under certain regularity conditions, together with proper truncation of $\tilde{\beta}_n$, has the same asymptotic distribution as the optimal nonadaptive procedure above with $\Delta = 1/\beta$ (Wu, 1985). The truncation is necessary to avoid a situation where the slope of the response curve is too close to 0, for example, when the doses tested are being drawn from the tails of the distribution. Therefore, a truncation of the form

$$\max[\min(\hat{\beta}_n^{-1}, d), \delta], \; d > \delta > 0,$$

instead of $\hat{\beta}_n^{-1}$, is advisable to maintain good performance of the adaptive procedure while preserving the asymptotic optimality mentioned above (Wu, 1985).

It is noteworthy that the Robbins-Monro process is not limited to quantal-response data; on the contrary, the procedure works equally well for continuous-response data. No adaptation is required. Furthermore, for the adaptive procedure, continuous-response data are more suitable because the regression approximations are more reasonable. Regression is generally not well-suited for binary-response data.

2.2 Up-and-Down Method

The up-and-down method is a simple variation of the Robbins-Monro technique in which the intervals between experimental levels are fixed. That is, we begin testing at an initial estimate and decide to move up one step or down one step depending on the outcome of the experiment. The best way to illustrate this procedure is to consider first the method for estimating the LD_{50}.

In order to estimate the LD_{50}, we first choose an initial dose, say 0. Then we decide on a suitable step size Δ based on whatever prior experience may indicate, for example, $\Delta = 1$. (Often, one standard deviation of the tolerance distribution is deemed the best choice for step size.) Then we begin by testing the first experimental unit at dose 0; an outcome is observed. If it is a response, then the next dose is one unit down; otherwise, the next dose is one unit up. Figure 2 contains an illustration of these steps.

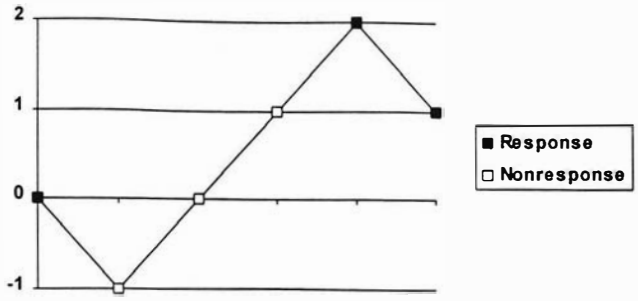

Figure 2 Up-and-Down Design for LD50

There are two relatively easy ways to estimate the LD_{50} using this method. The first is to take the mode of the dose levels, so that for the above example the estimate would be 0 or 1. (The fact that there are two choices indicates an obvious weakness with this estimate.) The second is to consider an average of dose levels, namely

$$LD_{50} = \sum_{i=0}^{n} \frac{x_i}{n+1}$$

While the first method is fairly crude, it is correspondingly simple. The second estimate intuitively seems more trustworthy.

Both of these estimates, however, admit to the likelihood of bias. If the initial estimate of the LD_{50} is inaccurate (and it most likely is, else we would not be conducting the experiment), then the choice will clearly influence the location of the dose levels included in the calculation of the estimate. Two clear alternatives arise: (1) strike the first

dose level and begin at the second, or (2) begin including dose levels after a change in direction has occurred. This latter alternative is clearly appropriate only for the LD_{50}.

It is also noteworthy that in this method there is a "free" extra dose level without requiring an experiment at that point; that is, if at the terminal dose level there is a response (or nonresponse) then it is known that the next dose level will be one step down (or up). Using this fact and the first adjustment above (1), we have

$$LD_{50} = \sum_{i=1}^{n+1} \frac{x_i}{n+1}$$

Another estimator for the LD_{50} was proposed by Wetherill; this estimator only counts the mean values between "turning points" in the up-and-down diagram. Such an estimator is clearly not generalizable to the search for quantiles other than the LD_{50}, because it assumes one is interested in the place on the curve where responses alternate evenly.

2.2.1 Biased Coin Method

It is evident that the up-and-down method described above is no longer appropriate when the quantile of interest is not the median. An alternative algorithm for determining dose levels is required. One way to alter this method for other quantiles is the biased-coin method following Durham and Flournoy (1994). The method is based on the up-and-down design described in the previous section; the modification is as follows.

Let us assume the quantile is below 0.5 (i.e., $p < 0.5$ for LD_{100p})[1]. We then construct a biased coin with the probability of obtaining a head given by

$$P(H) = \frac{p}{1-p}$$

Then if the outcome is a failure or nonresponse, we flip the biased coin to decide whether to go up a step ("heads" means go up; "tails" means stay at the same level). For a success or response, we automatically drop down a step. This procedure is illustrated in Figure 3 below, where T and H denote "tails" and "heads," respectively. It should be noted that the procedure reduces to the standard up-and-down design when $p = 0.5$.

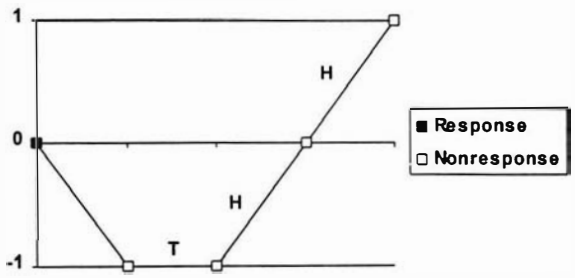

Figure 3 Up-and-Down Design: Biased-Coin Technique

In the case of the biased-coin technique, the estimator used in Section 2.2 for the LD_{50} is no longer adequate; rather, an adjustment has to be made to avoid a certain amount of bias. The new estimator is given by

[1] For $p > 0.5$, we will adjust the procedure slightly; that is, $P(H) = (1-p)/p$ and we flip only at success to decide whether to drop.

$$LD_{100p} = \sum_{x=1}^{n+1} \frac{x_i}{n+1} - \Delta(p - 0.5)$$

where Δ is the size of the step. It has been shown that this procedure will converge at the estimate of the LD_{100p} (Durham, 1994).

2.2.2 Multiple Up-and-Down Process (MUD)

Typically, the up-and-down procedure involves one experimental unit per trial. However, there are methods for dealing with multiple units per trial. In general the advantage of multiple units is one of economy of experimental time and effort (Hsi, 1969). Moreover, Hsi has found that under certain conditions MUD can be nearly as efficient (in terms of bias and mean-squared error) as the single-unit up-and-down procedure.

The procedure available to us follows the general pattern described here (Hsi, 1969):

(1) a series of doses is chosen;

(2) *n* trials of *k* subjects are performed; after each trial, the following decision is made: increase by one dosage level if there are *s* or fewer responses; decrease if there are *r* or more; and remain the same otherwise;

(3) after *n* trials the experiment is terminated and the estimator (as given above) is calculated.

The obvious difficulty with this procedure in regard to our aims is how to decide what values to use for *s* and *r* to obtain good estimates of the LD_{15} and LD_{30}.

Furthermore, the question arises whether some form of the biased-coin technique should be incorporated into the scheme to insure better estimates or greater efficiency (with respect to statistical error). While it might be expected that a straightforward way to achieve the ED_{10}, for example, would be to use 10 units per trial and stay at the same level for one response and go down for two or more, that method does not work as well as choosing $s = 0$ and $r = 1$ (Hsi, 1969).

It should be noted that for determining the LD_{50}, Hsi found that it is best to find an s and an r such that $s = k - r$. Thus, for example, if the trial has five subjects and we decide to increase if only one or no response, we should decrease for four or more responses.

For our purposes we would like to be able to formulate a general guideline for estimation of the LD_{100p}. We would like to propose as an alternative a combination of the biased-coin technique and the multiple up-and-down design.

2.2.3 Biased-Coin Multiple Up-and-Down Technique

The idea behind this approach is to combine the multiple-trial method (i.e., MUD) for the LD_{50} with the biased-coin technique in the following manner. For simplicity we will consider only two cases here, namely trials with 2 or 5 subjects ($k = 2$ or 5). In the first case ($k = 2$), the dose will be decreased if any responses are observed, and if no response is observed the biased coin will be flipped, where the odds[2] are determined as before as $p/(1-p)$. With 5 subjects per trial ($k = 5$), the dose level will decrease if there is

[2] Again, these odds apply only to the case $p < 0.5$.

more than one response; otherwise, the coin will be flipped with the biased odds as before.

This procedure consists of proceeding as though it were the LD_{50} under consideration, but before the decision to increase a dose level is made, the biased coin is flipped. This avenue avoids the question of the search for suitable s and r for every situation other than the LD_{50}.

One of the solutions given by Hsi (1969) to the problems when searching for LD_{100p} with $p < 0.5$ is to use a method of varying step sizes; however, that suggestion seems to us to violate the entire advantage of the up-and-down design, which surely rests in its simplicity. If we were to consider adjusting step sizes, the up-and-down design reverts back to the general Robbins-Monro process in terms of allocation.

CHAPTER 3: SIMULATION STUDIES

The simulations for the assessment of the procedures below were conducted in SAS using both the standard normal and gamma ($\alpha = 2$, $\beta = 1$) distributions as background or tolerance distributions. The normal distribution represents the case of a symmetric distribution; the gamma (quite skewed with these parameters) represents an asymmetric case.

Dixon and Mood suggested a step size about equal to the underlying standard deviation; Brownlee showed that the up-and-down design is most efficient when the step size is between 2/3 and 3/2 of the standard deviation. We have therefore chosen 1/2, 1 and 3/2 of the underlying standard deviation as a suitable step size for both methods.

The simulation size for every simulation in this study is 500, and sample sizes (N) of both 20 and 30 were used. Performance of the estimate is measured in terms of Monte Carlo mean-squared error (MSE) and bias.

The programs for two of the simulations that follow can be found in Appendix A; the other programs differ only slightly from those given as examples in the appendix.

3.1 Robbins-Monro Process

The first question related to this process specifically was how to choose the values of the step constant for consideration. As we will see below, one possibility is to employ an adaptive procedure where the step constant is generated by the prior data within the experiment. However, commonly we wish to set the step constant at the outset, so in this

first experiment we have done so. Our choices for this constant are 1/2, 1 and 3/2 for the standard normal; these step sizes represent, if desired, those multiples of the standard deviation of the distribution. Of course, for our choice of shape parameter in the gamma distribution, the standard deviation would be $\sqrt{2}$ or, roughly, 1.41. Consequently, if we wish to continue with this guideline for step constant for the gamma distribution, we must choose values of 0.7, 1.4 and 2.1.

Another consideration is the choice of initial doses. For the standard normal case, we have chosen -1, 0, and 1, representing, very roughly, good, mediocre and poor choices for initial dose with response to the 15% and 30% quantiles. The analogous choices for the gamma case might be 1, 2 and 3.

A frequent criticism of the Robbins-Monro procedure is that the step sizes decrease too rapidly if the initial estimate is far from the true quantile; this problem is exacerbated by large "one-way" jumps in the procedure for extreme quantiles when there is only one subject per trial. In order to ameliorate this deficiency, various delay mechanisms are often recommended. For example, in the search for the median one might recommend that the steps ought not begin to decrease until both a response and a nonresponse have been observed (Davis, 1971). A somewhat more clever approach advanced by Kesten (1985) is to decrease the step size if the last two responses immediately prior to the present trial are opposite and to leave it static otherwise.

Unfortunately, neither mechanism is appropriate for the present case, because each applies only to the search for the median. That is, it makes little sense to wait for

both a response and a nonresponse if we are indeed searching for a quantile where few responses are expected (e.g., LD_{15}). Therefore, we have followed a different course.

Our interest is in the estimation of LD_{15} and LD_{30} as representative of extreme quantiles; therefore, our delay mechanism might reasonably take into account places where we are expecting a greater number of nonresponses than responses. Consequently, we have decided to begin the decrease of the step sizes when 3 nonresponses have been observed. After that stage we begin decreasing by the usual increment; that is,

$$a_t = \frac{\Delta}{(t - t^* + 2)}$$

where a_t is the step constant, t is the index of trials and t^* is the stage at which we have observed the third nonresponse. This adjustment has the effect of decreasing c by consecutive factors only after the delay criterion has been met. All simulations in this section followed this pattern.

The results of the simulation are summarized in Table 1 for $p = 0.15$ and Table 2 for $p = 0.3$. Note that the "design" refers both to the total sample size and the background distribution assumed for simulation purposes.

Table 1 MSEX10^2 and BiasX10^2 for Robbins-Monro Process when p=15%

MSE

		k=1 First Dose			k=2 First Dose			k=5 First Dose		
DESIGN	**STEP**	**-1**	**0**	**1**	**-1**	**0**	**1**	**-1**	**0**	**1**
	1/2	05	22	40	02	35	73	0	21	152
normal	**1**	13	16	18	07	17	24	01	14	69
N = 20	**3/2**	21	22	20	11	14	16	03	17	35
	1/2	06	21	31	02	29	58	0	13	137
normal	**1**	11	14	17	05	13	18	01	07	54
N = 30	**3/2**	18	17	17	09	12	11	02	06	24

		First Dose			First Dose			First Dose		
		1	**2**	**3**	**1**	**2**	**3**	**1**	**2**	**3**
	1/2	08	17	29	05	29	61	06	67	167
gamma	**1**	11	10	13	07	12	21	05	29	62
N = 20	**3/2**	07	26	07	10	11	12	06	15	28
	1/2	07	16	24	05	24	49	05	58	137
gamma	**1**	07	07	08	05	09	14	04	21	45
N = 30	**3/2**	05	17	05	07	08	07	04	11	15

BIAS

		k=1 First Dose			k=2 First Dose			k=5 First Dose		
DESIGN	**STEP**	**-1**	**0**	**1**	**-1**	**0**	**1**	**-1**	**0**	**1**
	1/2	-2	35	48	0	53	77	2	76	123
normal	**1**	-8	10	11	-1	27	35	2	54	78
N = 20	**3/2**	-14	-5	-9	-5	11	11	1	37	46
	1/2	-1	30	46	1	49	73	2	71	113
normal	**1**	-7	7	10	0	24	30	2	51	66
N = 30	**3/2**	-14	-11	-8	-2	11	8	-1	32	40

		First Dose			First Dose			First Dose		
		1	**2**	**3**	**1**	**2**	**3**	**1**	**2**	**3**
	1/2	7	23	35	13	47	64	21	78	127
gamma	**1**	-3	-1	1	3	15	19	13	44	68
N = 20	**3/2**	-9	-18	-8	-6	2	2	6	23	33
	1/2	4	19	29	11	43	58	19	72	116
gamma	**1**	-1	-2	-1	1	9	15	11	37	56
N = 30	**3/2**	-8	-15	-5	-5	0	-1	5	18	24

Table 2 MSEX10^2 and BiasX10^2 for Robbins-Monro Process when p=30%

MSE

		k=1 First Dose			k=2 First Dose			k=5 First Dose		
DESIGN	**STEP**	**-1**	**0**	**1**	**-1**	**0**	**1**	**-1**	**0**	1
	1/2	09	10	20	11	10	36	16	14	79
normal	**1**	12	11	13	08	09	14	12	09	32
N = 20	**3/2**	13	13	14	10	09	13	08	08	17
	1/2	09	11	15	10	09	29	14	13	67
normal	**1**	10	09	10	07	07	10	09	07	25
N = 30	**3/2**	10	10	11	07	08	07	07	06	12

		First Dose			First Dose			First Dose		
		1	**2**	**3**	**1**	**2**	**3**	**1**	**2**	**3**
	1/2	07	16	25	04	18	43	01	33	121
gamma	**1**	10	12	14	06	12	17	03	18	47
N = 20	**3/2**	10	10	10	08	10	13	04	12	24
	1/2	07	15	19	03	15	38	02	29	105
gamma	**1**	07	09	10	05	08	12	02	13	37
N = 30	**3/2**	06	07	06	06	07	09	03	09	17

BIAS

		k=1 First Dose			k=2 First Dose			k=5 First Dose		
DESIGN	**STEP**	**-1**	**0**	**1**	**-1**	**0**	**1**	**-1**	**0**	**1**
	1/2	-24	12	25	-30	21	49	-39	36	87
normal	**1**	-12	2	2	-19	8	18	-31	23	49
N = 20	**3/2**	-14	-7	-6	-12	2	1	-26	14	28
	1/2	-22	10	26	-28	22	43	-38	33	79
normal	**1**	-12	-2	2	-18	7	14	-29	20	44
N = 30	**3/2**	-10	-7	-5	-8	-1	5	-22	14	20

		First Dose			First Dose			First Dose		
		1	**2**	**3**	**1**	**2**	**3**	**1**	**2**	**3**
	1/2	-4	15	25	-5	29	54	-9	54	108
gamma	**1**	-4	2	6	-1	8	16	-6	30	57
N = 20	**3/2**	-5	0	0	-1	2	9	-4	17	29
	1/2	-3	15	24	-5	26	50	-8	49	96
gamma	**1**	-1	2	5	0	8	15	-5	26	50
N = 30	**3/2**	-2	0	1	-3	3	5	-3	13	26

There are several noteworthy results above. First, observe that the patterns discernible in these results are very similar for the LD_{15} and LD_{30}; this fact bodes well for the generalizability of our recommendations.

There are some clear trends evident in the results in terms of Monte Carlo MSE. The best results when the initial estimate is close to the true value are achieved by using a larger trial size; that is, the best estimates for both the symmetric (normal) and asymmetric (gamma) background distributions are achieved with trial sizes of 5. For the symmetric distribution, that trend is more pronounced (i.e., some of the 'second-best' estimates are found in trial sizes 1 or 2 for the asymmetric).

We note, however, that if the initial estimate is very poor then using a large trial size is equally so: the results for trial size 5 and the 'bad' initial guess are uniformly disappointing. This outcome is not surprising: if the initial guess is far removed from the true quantile, it will take a larger number of total iterations to approximate it. Similarly unsurprising is the fact that a larger step size has a hugely salutary effect when the initial guess is poor and the trial sizes are large. All of these observations seem to apply equally well to a total sample size of 20 or 30.

It seems that, unless the researcher has a very good idea of where the quantile is, it is best to avoid using large trial sizes. If time is important, however, there is very little difference between using trial sizes of 1 or 2, except when the step size is small. Consequently, one might recommend in general using one or two experimental units per

trial with a reasonably large step size, unless the investigator is certain that the quantile is within a very small range indeed.

Finally, we note that the biases reported do not contradict these conclusions, though the patterns are not precisely the same. In those cases where bias would lead to a slightly different choice, the discrepancies of bias are minuscule.

3.2 Adaptive Robbins-Monro Process

As previously discussed, the adaptive Robbins-Monro Process (ARM) is one in which the slope of the response curve at the quantile under consideration is estimated via ordinary linear regression. To this end we arrive at an estimate for the slope given by

$$\hat{\beta}_n = \sum y_i (x_i - \bar{x}_n) / \sum (x_i - \bar{x}_n)^2$$

To achieve the step coefficient, a_t, we then take

$$a_t = \frac{\Delta}{n\hat{\beta}}$$

This substitution can take place only after both nonresponses and responses have been observed; otherwise, we would have an estimate of 0 for β and be unable to take its reciprocal. Also, there cannot be an equal number of responses and nonresponses, else the regression line would be horizontal; in that case, we merely use the former step coefficient as if the procedure were nonadaptive.

Furthermore, to avoid unduly large estimates of β^{-1}, we may use the truncation mentioned above. In our simulations we followed the ordinary RM pattern from the

previous section until the criterion for both responses and nonresponses was met, and then we applied the truncation

$$\max[\min(\hat{\beta}_n^{-1}, d), \delta]$$

instead of $\hat{\beta}_n^{-1}$ with d = 50 and $\delta = 1$. Clearly, if $\hat{\beta}_n^{-1}$ is very large our truncation will cut it at 50, while if it is very small it will truncate at 1. Also, of course, the estimate cannot be obtained until two trials at two different doses have been observed; in that case, we also merely continue with the procedure as before. The results of the simulation are summarized in Table 3 for $p = 0.15$ and Table 4 for $p = 0.3$.

Table 3 $MSEX10^2$ and $BiasX10^2$ for Adaptive Robbins-Monro Process when p=15%

MSE

		k=1 First Dose			k=2 First Dose			k=5 First Dose		
DESIGN	**STEP**	**-1**	**0**	**1**	**-1**	**0**	**1**	**-1**	**0**	**1**
	1/2	15	51	126	34	81	322	22	79	232
normal	**1**	14	42	44	20	55	65	26	47	156
N = 20	**3/2**	23	46	25	23	48	38	40	63	68
	1/2	13	28	70	21	40	209	28	71	282
normal	**1**	11	23	23	19	30	37	32	38	190
N = 30	**3/2**	17	21	17	15	27	33	34	65	66

		First Dose			First Dose			First Dose		
		1	**2**	**3**	**1**	**2**	**3**	**1**	**2**	**3**
	1/2	17	56	210	34	403	661	90	615	2180
gamma	**1**	15	28	47	20	242	314	62	484	1131
N = 20	**3/2**	37	29	57	28	199	338	66	204	744
	1/2	12	38	177	16	275	506	86	630	1909
gamma	**1**	13	17	46	13	177	242	51	335	853
N = 30	**3/2**	36	11	35	24	146	202	76	291	667

BIAS

		k=1 First Dose			k=2 First Dose			k=5 First Dose		
DESIGN	**STEP**	**-1**	**0**	**1**	**-1**	**0**	**1**	**-1**	**0**	**1**
	1/2	7	1	7	8	3	14	23	39	33
normal	**1**	9	-3	22	13	-4	25	24	22	46
N = 20	**3/2**	7	-4	6	14	0	14	27	10	38
	1/2	8	-6	-1	15	-8	0	32	12	21
normal	**1**	9	-8	14	13	-5	20	26	12	17
N = 30	**3/2**	6	-2	6	12	-8	17	32	-8	35

		First Dose			First Dose			First Dose		
		1	**2**	**3**	**1**	**2**	**3**	**1**	**2**	**3**
	1/2	-32	-12	-17	-27	-89	-90	-35	-99	-197
gamma	**1**	-32	-16	3	-26	-53	-61	-29	-84	-93
N = 20	**3/2**	-43	-25	-21	-32	-51	-71	-30	-51	-66
	1/2	-28	-9	-25	-21	-76	-85	-41	-119	-198
gamma	**1**	-31	-14	-6	-25	-52	-56	-35	-71	-82
N = 30	**3/2**	-43	-18	-15	-29	-49	-55	-41	-78	-74

Table 4 $MSEX10^2$ and $BiasX10^2$ for Adaptive Robbins-Monro Process when p=30%

MSE

		k=1			k=2			k=5		
		First Dose			First Dose			First Dose		
DESIGN	**STEP**	**-1**	**0**	**1**	**-1**	**0**	**1**	**-1**	**0**	**1**
	1/2	13	21	52	37	32	253	53	13	239
normal	**1**	12	13	23	20	23	36	52	26	178
N = 20	**3/2**	11	12	18	26	23	38	38	12	86
	1/2	11	12	23	26	17	104	63	29	181
normal	**1**	11	8	15	16	19	23	57	15	135
N = 30	**3/2**	9	10	11	17	13	20	43	15	57

		First Dose			First Dose			First Dose		
		1	**2**	**3**	**1**	**2**	**3**	**1**	**2**	**3**
	1/2	26	76	165	54	179	267	36	418	1422
gamma	**1**	39	61	103	37	208	203	48	247	904
N = 20	**3/2**	36	74	87	39	157	189	47	124	733
	1/2	16	28	123	34	168	264	28	407	1334
gamma	**1**	18	39	65	27	166	159	28	221	882
N = 30	**3/2**	26	50	51	20	129	184	33	155	677

BIAS

		k=1			k=2			k=5		
		First Dose			First Dose			First Dose		
DESIGN	**STEP**	**-1**	**0**	**1**	**-1**	**0**	**1**	**-1**	**0**	**1**
	1/2	-2	6	6	4	5	-10	16	21	-29
normal	**1**	3	5	1	5	5	14	14	13	14
N = 20	**3/2**	4	5	-2	15	4	5	14	6	10
	1/2	0	1	1	2	2	-8	29	8	-18
normal	**1**	1	6	7	7	4	10	26	5	-6
N = 30	**3/2**	3	6	-1	11	2	0	19	5	8

		First Dose			First Dose			First Dose		
		1	**2**	**3**	**1**	**2**	**3**	**1**	**2**	**3**
	1/2	-34	-27	-28	-35	-65	-48	-24	-86	-155
gamma	**1**	-37	-33	-24	-28	-67	-48	-27	-67	-92
N = 20	**3/2**	-39	-43	-29	-33	-59	-62	-28	-52	-100
	1/2	-23	-11	-28	-31	-69	-59	-29	-103	-197
gamma	**1**	-23	-27	-16	-27	-71	-49	-30	-63	-123
N = 30	**3/2**	-34	-37	-24	-24	-58	-68	-31	-63	-100

This procedure appears to be uniformly worse than the nonadaptive Robbins-Monro for quantal-response data. It is noteworthy, however, that here the recommendation is much clearer if the ARM is to be used: sticking with one experimental unit per trial seems to be always the best approach, whether in terms of MSE or bias. On the whole, however, there is a clear advantage in avoiding the procedure altogether.

3.3 Up-and-Down Design: Biased-Coin Technique

The simulations for both the single-unit trials and the MUD were conducted via the same program; for the specific case at hand as described above, the biased-coin MUD with $n = 1$ is the same as the single-unit biased-coin technique.

The delay mechanism for this procedure was the same as in the RM simulations, namely the calculation of the estimate never involved any observations before three nonresponses had been recorded. The estimate from the up-and-down design often excludes the initial observation to remove bias (following Brownlee), but we see no reason why this principle should not be extended for small p cases as was done in the case of Robbins-Monro. Thus the estimate for the LD_{100p} was calculated as

$$LD_{100p} = \sum_{i=x^*}^{n+1} x_i / (n+1-x^*) - \Delta(p-0.5)$$

where x^* is the first dose level after at least three nonresponses have been observed and Δ is the step size.

Note that the biases listed below would be exaggerated by a factor of $\Delta(p - 0.5)$ were it not for the adjustment. Thus, for our $p = 0.15$ and steps 1/2, 1 and 3/2, the biases listed below would be less by the amounts 17, 35 and 52. Similarly exaggerated biases would apply to the other results below.

On a very few occasions, when starting at an initial point far from the true quantile, the experiment never led to at least three nonresponses, and therefore the above estimate could not be calculated. (Such was the case on at most 5 out of the 500 simulations per individual design.) The estimator is determined before an experiment is conducted; therefore, if the estimate cannot be calculated then there is no estimate. This strategy is consistent with one's overall aims, namely to get an accurate estimate of the quantile. Indeed, if there were not even three nonresponses, then clearly the doses tried were almost all far away from the true LD_{100p}.

Of course, in the case of the up-and-down there is no further adjustment made to the step within the simulation: once the step is chosen for the experiment the increment is fixed throughout, following the up-and-down design methods.

The results from the simulations for the up-and-down design are presented in Table 5 for $p = 0.15$ and Table 6 for $p = 0.3$.

Table 5 MSEX10^2 and BiasX10^2 for Biased-Coin MUD when p=15%

MSE

DESIGN	STEP	k=1 First Dose -1	0	1	k=2 First Dose -1	0	1	k=5 First Dose -1	0	1
	1/2	14	22	25	11	16	27	6	26	115
normal	1	21	27	26	22	21	24	17	17	36
N = 20	3/2	35	35	33	36	41	34	28	29	27
	1/2	12	16	20	13	10	13	11	10	50
normal	1	17	19	19	21	19	19	23	15	15
N = 30	3/2	23	27	24	27	34	29	33	33	21

		First Dose 1	2	3	First Dose 1	2	3	First Dose 1	2	3
	1/2	17	22	24	26	46	68	29	113	245
gamma	1	25	31	32	38	56	56	63	122	160
N = 20	3/2	39	36	47	65	56	78	117	128	148
	1/2	11	14	15	17	31	41	29	89	172
gamma	1	15	25	26	28	39	39	53	98	113
N = 30	3/2	25	35	27	46	42	53	88	92	126

BIAS

DESIGN	STEP	k=1 First Dose -1	0	1	k=2 First Dose -1	0	1	k=5 First Dose -1	0	1
	1/2	10	29	35	-9	19	36	-10	46	106
normal	1	9	21	20	-10	0	8	-12	16	49
N = 20	3/2	15	16	14	-15	-10	-5	-24	4	21
	1/2	9	21	29	-14	4	18	-20	22	68
normal	1	11	15	15	-17	-6	-9	-27	-6	11
N = 30	3/2	12	13	15	-20	-20	-15	-33	-17	-8

		First Dose 1	2	3	First Dose 1	2	3	First Dose 1	2	3
	1/2	22	29	32	32	50	67	43	97	147
gamma	1	20	23	39	23	40	50	54	86	105
N = 20	3/2	8	50	11	17	55	30	63	76	84
	1/2	18	23	26	23	40	50	39	83	121
gamma	1	16	22	39	21	27	44	44	71	84
N = 30	3/2	3	52	6	12	52	18	44	73	68

Table 6 MSEX10^2 and BiasX10^2 for Biased-Coin MUD when p=30%

MSE

		k=1 First Dose			k=2 First Dose			k=5 First Dose		
DESIGN	**STEP**	**-1**	**0**	**1**	**-1**	**0**	**1**	**-1**	**0**	**1**
	1/2	11	16	17	34	17	14	41	6	25
normal	**1**	17	18	20	36	32	37	60	36	19
N = 20	**3/2**	22	22	24	45	45	46	81	64	45
	1/2	8	9	12	28	17	15	56	16	6
normal	**1**	11	12	13	34	28	28	66	48	36
N = 30	**3/2**	16	14	15	39	35	36	79	73	59

		First Dose			First Dose			First Dose		
		1	**2**	**3**	**1**	**2**	**3**	**1**	**2**	**3**
	1/2	16	18	22	20	28	48	16	61	128
gamma	**1**	27	27	26	43	44	46	56	85	101
N = 20	**3/2**	39	32	43	68	48	72	104	99	125
	1/2	11	14	15	15	24	30	16	47	93
gamma	**1**	17	20	19	34	34	33	42	69	81
N = 30	**3/2**	25	25	32	48	36	52	79	70	97

BIAS

		k=1 First Dose			k=2 First Dose			k=5 First Dose		
DESIGN	**STEP**	**-1**	**0**	**1**	**-1**	**0**	**1**	**-1**	**0**	**1**
	1/2	-1	10	14	-44	-21	-3	-58	-11	47
normal	**1**	3	5	6	-39	-32	-33	-64	-44	-17
N = 20	**3/2**	5	6	2	-39	-34	-36	-71	-55	-46
	1/2	-1	6	9	-43	-27	-20	-69	-31	11
normal	**1**	4	7	3	-42	-37	-34	-71	-56	-44
N = 30	**3/2**	5	6	2	-40	-38	-39	-74	-67	-61

		First Dose			First Dose			First Dose		
		1	**2**	**3**	**1**	**2**	**3**	**1**	**2**	**3**
	1/2	13	17	25	10	26	43	8	60	99
gamma	**1**	21	24	27	14	24	36	24	50	60
N = 20	**3/2**	14	34	22	7	31	25	30	38	48
	1/2	15	20	21	11	25	32	11	47	80
gamma	**1**	18	23	27	18	23	32	17	44	56
N = 30	**3/2**	17	35	17	21	34	16	26	36	40

There is a clear correlation between increased step size and poverty of the estimate in terms of MSE and bias. This observation is highly intuitive, as the larger step sizes limit the number of dose choices. The other quite remarkable trend here is that it does not seem to matter much how good the initial guess is; that is an extremely valuable point to keep in mind. When compared to the Robbins-Monro procedure, however, we note that the MSE is almost always greater here; however, when the step size is small the difference is correspondingly insignificant.

We note that Hsi has already anticipated a poor performance for the MUD if the initial guess is far from the desired quantile (Hsi, 1969). This expectation is certainly confirmed by the above results.

Second, we note that for $p = 15\%$ and the symmetric case, the MUD is actually comparable in performance to the single-trial method. This comparability collapses for $p = 30\%$ or for the asymmetric case. Aside from the advantage of speed, the MUD seems to have little to offer. In every other instance it is consistently outperformed by the single-trial up-and-down design.

3.4 Simultaneous Trials

The entire purpose of the MUD procedure is to speed up the experiment. The same end can be accomplished, however, without the loss of statistical accuracy; such a compromise is possible if we allow several one-unit-per-trial up-and-down designs to take place at the same time. We might well expect an improvement over the MUD in terms of both MSE and bias.

To this end we conducted a simulation that took the simultaneous trials approach with the up-and-down design. The approach was to run simultaneous trials with $k = 2$ and 5, find their estimates in the ordinary fashion, and then average those estimates. (Trying to obtain an overall estimate from all the data points would fail to take into account the delay mechanism described earlier.) The results of the simulations studies are presented in Table 7 for $p = 0.15$ and Table 8 for $p = 0.3$.

Table 7MSEX10^2 and BiasX10^2 for Simultaneous Biased-Coin MUD when p=15%

MSE

		k=2			k=5		
		First Dose			First Dose		
DESIGN	**STEP**	**-1**	**0**	**1**	**-1**	**0**	**1**
	1/2	8	47	83	4	73	195
normal	**1**	18	36	49	14	64	108
N = 20	**3/2**	26	35	41	27	63	93
	1/2	7	26	52	5	57	131
normal	**1**	14	24	31	13	46	77
N = 30	**3/2**	19	29	26	22	44	60

		First Dose			First Dose		
		1	**2**	**3**	**1**	**2**	**3**
	1/2	131	19	17	189	20	18
gamma	**1**	65	4	4	122	5	5
N = 20	**3/2**	89	1	20	95	1	24
	1/2	100	19	18	166	19	17
gamma	**1**	41	4	4	90	4	4
N = 30	**3/2**	87	1	19	90	1	21

BIAS

		k=2			k=5		
		First Dose			First Dose		
DESIGN	**STEP**	**-1**	**0**	**1**	**-1**	**0**	**1**
	1/2	15	64	86	18	84	138
normal	**1**	25	47	58	31	75	99
N = 20	**3/2**	27	41	47	42	70	88
	1/2	13	45	66	17	74	112
normal	**1**	17	37	44	28	63	83
N = 30	**3/2**	21	34	35	37	60	70

		First Dose			First Dose		
		1	**2**	**3**	**1**	**2**	**3**
	1/2	-113	-43	-41	-137	-45	-41
gamma	**1**	-76	-19	-19	-110	-21	-20
N = 20	**3/2**	-93	6	-43	-97	2	-48
	1/2	-98	-43	-42	-128	-44	-41
gamma	**1**	-60	-19	-18	-93	-20	-20
N = 30	**3/2**	-92	6	-43	-94	4	-44

Table 8 MSEX10^2 and BiasX10^2 for Simultaneous Biased-Coin MUD when p=30%

MSE

		k=2			k=5		
		First Dose			First Dose		
DESIGN	**STEP**	**-1**	**0**	**1**	**-1**	**0**	**1**
	1/2	9	14	31	9	16	78
normal	**1**	12	15	22	7	16	37
N = 20	**3/2**	16	18	23	10	20	28
	1/2	7	10	18	7	13	50
normal	**1**	9	11	14	6	12	25
N = 30	**3/2**	11	13	16	9	15	21

		First Dose			First Dose		
		1	**2**	**3**	**1**	**2**	**3**
	1/2	159	68	62	282	72	60
gamma	**1**	82	35	34	178	40	39
N = 20	**3/2**	166	14	68	187	18	77
	1/2	123	66	62	223	70	60
gamma	**1**	64	34	33	123	37	35
N = 30	**3/2**	159	13	65	172	15	70

BIAS

		k=2			k=5		
		First Dose			First Dose		
DESIGN	**STEP**	**-1**	**0**	**1**	**-1**	**0**	**1**
	1/2	-14	25	48	-27	35	85
normal	**1**	0	20	27	-12	30	49
N = 20	**3/2**	3	14	23	-3	26	34
	1/2	-11	19	33	-21	31	67
normal	**1**	-2	14	18	-4	23	40
N = 30	**3/2**	0	13	14	3	21	29

		First Dose			First Dose		
		1	**2**	**3**	**1**	**2**	**3**
	1/2	-125	-82	-78	-168	-85	-77
gamma	**1**	-89	-59	-57	-132	-63	-61
N = 20	**3/2**	-128	-36	-81	-136	-41	-86
	1/2	-110	-81	-78	-149	-83	-77
gamma	**1**	-79	-58	-57	-110	-61	-59
N = 30	**3/2**	-125	-34	-80	-131	-38	-82

For the symmetric distributions, the simultaneous-trials approach (SMUD) is often more efficient in terms of MSE when the initial guess is relatively good. This advantage is not as widespread or as consistent as we might have hoped; there appears to be little advantage, if a symmetric tolerance distribution is supposed, in using the SMUD.

The simultaneous trials outperform the standard MUD in the asymmetric case (both in terms of MSE and bias) for $p = 0.15$ when the initial guess is not near the true quantile. Hence, if the researcher has reason to believe the tolerance distribution is asymmetric and not much is known about the location of the true (extreme) quantile, it may be worthwhile to use the simultaneous-trials approach.

3.5 Comparison of Methods

In order to summarize the relative efficiency of the various methods, we have chosen four representative scenarios. In the graphs in Figure 4 through Figure 7, the patterns already described in Sections 3.1 to 3.4 can be seen. (Note that those values so large as to be unworthy of comparison are not recorded on the graphs.)

In the case of one subject per trial (Figure 4), the relative parity of the up-and-down design and the RM procedure is clear. The uniform inferiority of the ARM is also evident, and it is the more exaggerated the worse the initial estimate is.

Similar results are evident in the case of two subjects per trial with an asymmetric tolerance distribution(Figure 5); however, here we have the additional observation that

the simultaneous MUD is less efficient than the MUD for a good initial estimate but more efficient if the initial estimate is very poor.

In Figure 6 the RM and the SMUD clearly outperform the other methods for a good initial estimate; furthermore, for less precise initial doses the three non-ARM methods appear roughly equally efficient. It is noteworthy that in this case (with a symmetric tolerance distribution) the SMUD underperforms the MUD for a very poor initial choice.

Finally, in Figure 7 we see a scenario where there appears to be a clear hierarchy of methods: RM, SMUD, ARM, MUD. It should be emphasized that here we observe an unususally good performance in the ARM.

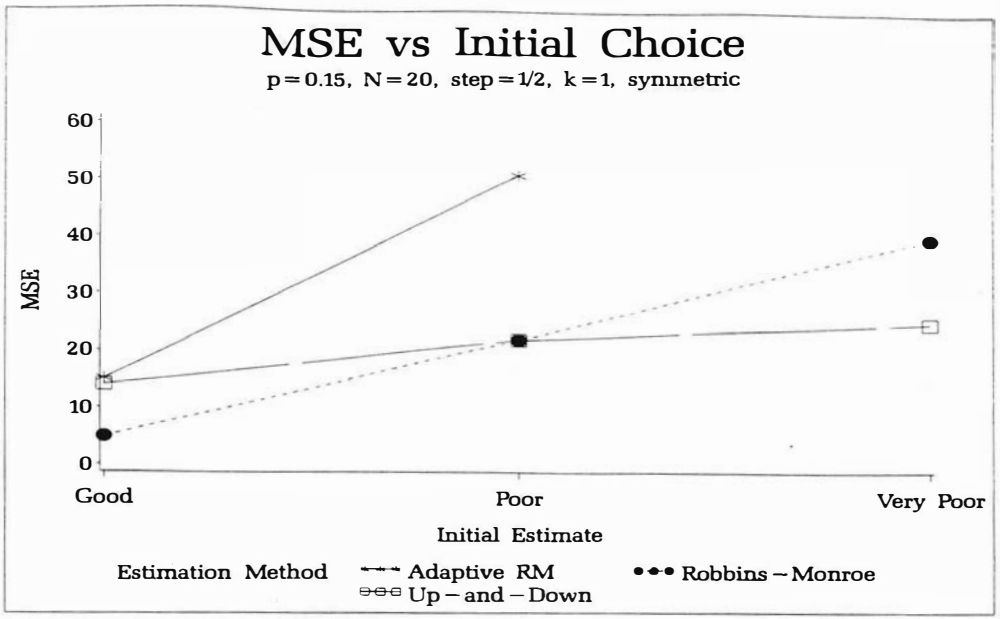

Figure 4 Comparison of Methods: p=0.15, N=20, k=1, step=1/2, symmetric

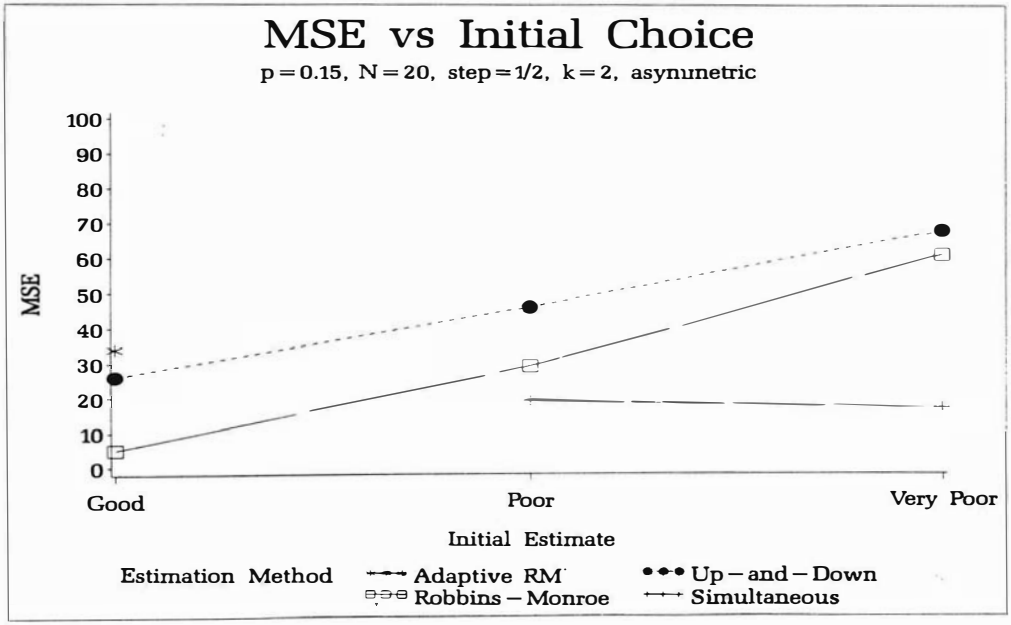

Figure 5 Comparison of Methods: p=0.15, N=20, k=2, step=1/2, asymmetric

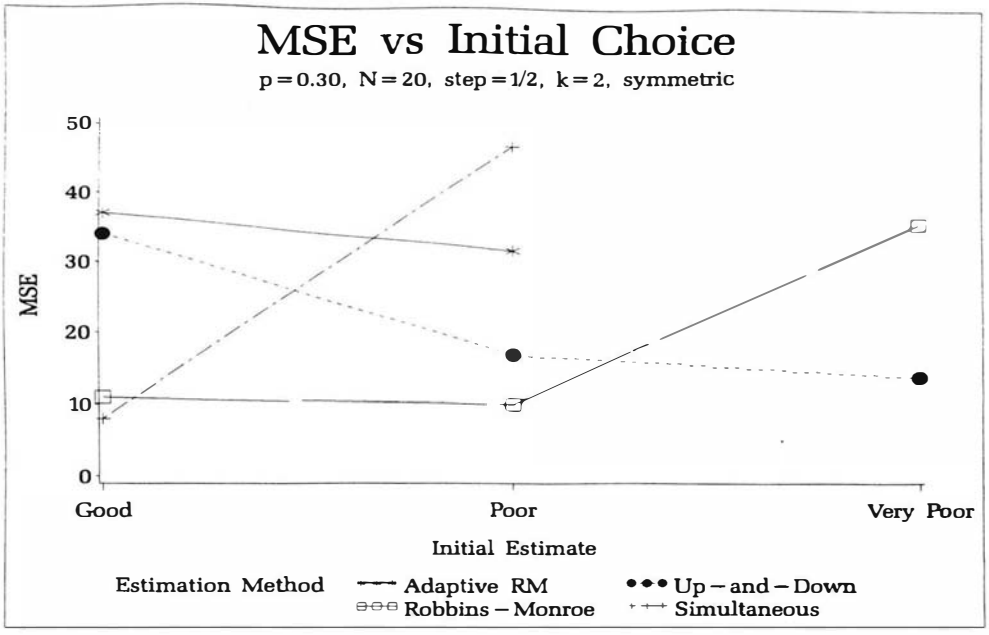

Figure 6 Comparison of Methods: p=0.30, N=20, k=2, step=1/2, symmetric

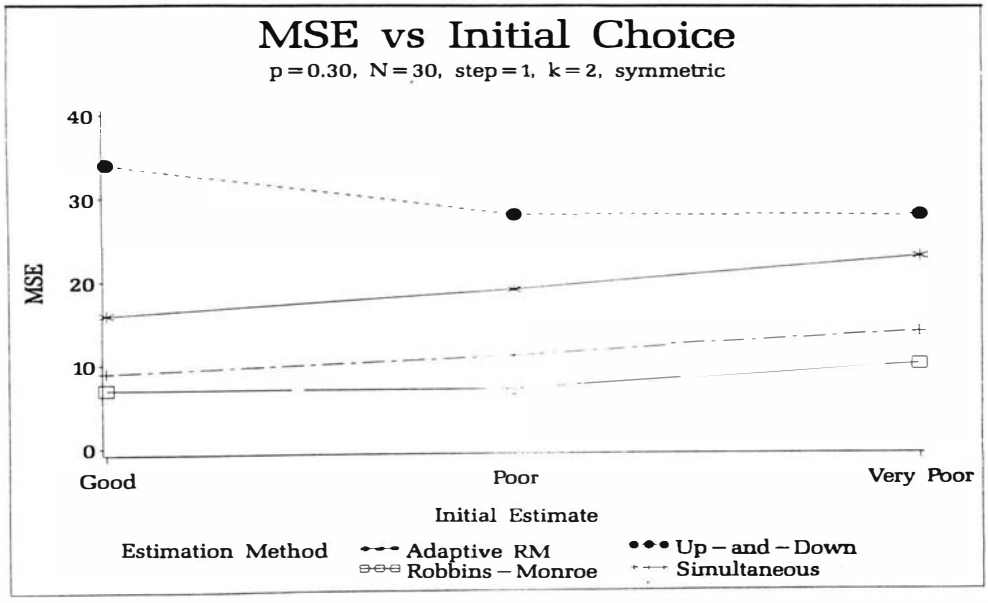

Figure 7 Comparison of Methods: p=0.30, N=30, k=2, step=1, symmetric

CHAPTER 4: CONCLUSION

4.1 Summary and Recommendations

This thesis is concerned with sequential estimation of the LD_{100p} when $p \neq 0.5$. Although we have focused our attention on cases where p is relatively small (i.e., $p = 0.15$ or 0.3), the conclusions are likely to be generalized to other relatively small and large values of p. Noteworthy, however, is that overall the MSE and bias are often lower when $p = 0.3$ than when $p = 0.15$. Such a result is consistent with expectations, because more extreme quantiles are more difficult to estimate.

In terms of both MSE and bias, the clear overall recommendation is for the Robbins-Monro procedure (RM) in virtually every case. Such a result is consistent with the published results for the LD_{50}. The only possible exceptions to this general rule would be the following two cases: (1) if the researcher is interested in conducting multiple experiments per trial and has reason to believe that the tolerance distribution is asymmetric, the simultaneous-trials biased-coin MUD procedure may be optimal; (2) if the step sizes are small (e.g., one-half standard deviation of tolerance distribution), the initial estimate is relatively far from the true quantile, and there is reason to believe the tolerance distribution is symmetric, the up-and-down procedure is often superior or at least comparable to RM in terms of MSE.

There is often lower bias generally found for the adaptive RM (ARM) in the symmetric case, especially for the larger quantile ($p = 0.30$); however, this discrepancy is negligible.

It is not surprising that the RM process works so well, given its variable step sizes and its proven record with the LD_{50}. On the other hand, it is important to remember that the delay mechanisms we have introduced may play a role in this apparent superiority revealed in our simulations. In any event, using the procedure as we have advised appears to work very well.

The performance of the MUD for $k = 5$ can be better than that of RM in the symmetric case with poor initial estimates, so it may be advisable to use the MUD if such trial sizes are necessary given the small overall samples (N = 20 or 30). Note, however, that SMUD appears to have little advantage over the MUD except for the asymmetric case. We would continue to recommend this biased-coin MUD because it is easy to use in general, instead of searching for the right s and r as required by the unaltered MUD presented by Hsi (Hsi, 1969).

Finally, we repeat that there appears to be no reason to use the ARM process when analyzing quantal-response data. The RM process appears to be the most efficient, while the up-and-down procedure is often comparable enough to warrant its use and is sometimes even superior.

4.2 Recommendations for Further Study

The methods in this study have been modified slightly in some ways from their form given by the original proponents (e.g., delay mechanisms and the biased-coin element in the MUD); therefore, it may be necessary to compare via simulation these methods with their unaltered versions to see if the modifications are indeed improvements. Such a comparison would involve only slight changes in the programs used to run the simulations in this study (see Appendix A).

Furthermore, in most procedures in sequential analysis, the sample size is a random variable determined by the outcomes of the experiments. In the case of the procedures describes in this study, however, there is no stopping rule. Although work has been done in this area, a satisfactory, well-tested method for deciding how to stop a test using either RM or up-and-down has not been devised (see Pflug, 1988). Further development in this arena is needed.

Finally, there are also some methods for determining confidence intervals for the estimates, but they often involve modifications of the original estimation procedures (Ghosh, 1991). Further development for constructing confidence-intervals is necessary.

BIBLIOGRAPHY

BIBLIOGRAPHY

Brownlee, K.A., Hodges, J.L., and Rosenblatt, M. (1953), "The Up-and-Down Method with Small Samples," *Journal of the American Statistical Association*, 48, 262-277.

Choi, S.C. (1990), "Interval Estimation of the LD_{50} Based on an Up-and-Down Experiment'" *Biometrics*, 46, 485-492.

Davis, M. (1971), "Comparison of Sequential Bioassays in Small Samples," *Journal of the Royal Statistical Society*, B 33, 78-87.

Dixon, W.J. (1965), "The Up-and-Down Design for Small Samples," *Journal of the American Statistical Association*, 60, 967-978.

Durham, S. D., Flournoy, N. (1994), "Random Walks for Quantile Estimation," *Statistcal Decision Theory and Related Topics V*, 467-476, New York: Springer-Verlag.

Ghosh, B.K. and Sen, P.K. (1991), *Handbook of Sequential Analysis*, New York: M. Dekker.

Hamilton, M.A. (1979), "Robust Estimates of the ED50," *Journal of the American Statistical Association*, 74, 344-354.

Hsi, B.P. (1969), "The Multiple Sample Up-and-Down Method in Bioassay," *Journal of the American Statistical Association*, 64, 147-162.

Hubert, J.J. (1992), *Bioassay*, (Third Edition), Dubuque: Kendall/Hunt.

Kesten, H. (1958), "Accelerated Stochastic Approximation," *Annals of Mathematical Statistics*, 29, 41-59.

Little, R.E. (1974a), "A mean square error comparison of certain median response estimates for the up-and-down design," *Journal of the American Statistical Association*, 69, 202-206.

----------. (1974b), "The up-and-down method for small Samples with Extreme Value Response Distributions," 69, 803-806.

McLeish, D. and Tosh, D. (1983), "The estimation of extreme quantiles in logit bioassay," *Biometrika*, 70, 625-632.

Pflug, G. C. (1988), "Stepsize Rules, Stopping Times and Their Implementation in Stochastic Quasigradient Algorithms," *Numerical Techniques for Stochastic Optimization* , ed. Y. Emolieve and R.J. Wets, New York: Springer-Verlag, 353-372.

Robbins, H. annd Monro, S. (1951), "A Stochastic Approximation Method," *Annals of Mathematical Statistics*, 22, 400-407.

Storer, B.E. (1989), "Design and Analysis of Phase I Clinical Trials," *Biometrics*, 45, 925-937.

Wetherill, G.B., Glazebrook, K.D. (1986), *Sequential Methods in Statistics* , (Third Edition), New York: Chapman and Hall.

Wu, C. F. Jeff. (1985), "Efficient Sequential Designs with Binary Data," *Journal of the American Statistical Association*, 80, 974-984.

APPENDIX A:SAS PROGRAMS FOR SIMULATIONS

Appendix A-1 Program for Robbins-Monro Design

```
*********************************************************************
* This program runs simulations for the Robbins-Monroe technique*
*  It assumes a background NORMAL distribution (i.e., N(0,1)).
* Author:  Dan Bettendorf
*********************************************************************;
Title 'Robbins-Monroe Macro';
Footnote 'UD$BHS:[dbettendorf.thesis.simulations]robbins_monroe.sas';

*********************************************************************
* VARIABLES:
* INIT is the inital value for the algorithm
* DIVIDE is the divisor of 20 to create the group SIZE
* STEP is the step coefficient in the algorithm
* P is the p as in LD100p
*Then we can discuss the ORDINARY VARIABLES:
*   RESP counts the number of responses within group
*   TOTAL counts the number of responses within experiment
*   QUANT is the estimate of the LD100p
*   ISTAR is the iteration at which 3 nonresponders have been reached
*           for the purposes of the delay mechanism
*   EXP is the index variable for the experiments
*   RESPOND is an indicator whether the random variate is a response
*   MARGIN is the count of nonresponders
*************************************************************************
;
options ls=80 ps=54;
data driver;
 do q=1 to 3;
  do w=1 to 6;
    do e=1 to 3;
      do r=1 to 2;

        init=q-2;

        if w=1 then do; divide=20;
```

```
                size=1;
              end;
        if w=2 then do; divide=10 ;
                size=2 ;
              end;
      if w=3 then do; divide= 4;
                size=5;
              end;
      if w=4 then do; divide= 30 ;
                size=1 ;
              end;
      if w=5 then do; divide= 15 ;
                size=2 ;
              end;
      if w=6 then do; divide= 6 ;
                size=5 ;
              end;

      step=e*0.5;
      p= r*(0.15);
      truth= probit(p);
     output;
     end;
   end;
  end;
end;

*proc print data=driver;

data collect(keep=step p init size sam_siz quant row sqbias);
 set driver; row=_N_;
 retain resp total quant istar;
 do exp=1 to 500;
    quant=init; total=0;  istar=0;
       do i=1 to divide;
         resp=0;
           do j=1 to size;
             z=normal(-1);
             respond= (z le quant);
             resp=resp+respond;
           end;
```

```
        total=resp+total;
        margin=size*i-total;
        if margin ge 3 then do;
                        if istar=0 then do;
                                        istar=i;
                                end;
                         a=step/(i-istar+2);
                       end;
        else a=step;
        quant=quant+a*(p-resp/size);
      end;
   sqbias=(quant-truth)**2;
   sam_siz=divide*size;
  output collect;
end;

proc univariate data=collect noprint;
   var sqbias;
   by row;
output out=info
mean=mse;


data final(drop=quant);
  merge collect
        info;
   by row;
  if first.row;
proc print data=final(drop=row sqbias);
* END OF PROGRAM
```

APPENDIX A-2 Program for Biased-Coin Multiple Up-and-Down Design (MUD)

```
*****************************************************************
* This program runs simulations for the Up-and-Down technique   *
* using the NORMAL distribution. It was composed by Dan         *
* Bettendorf,                                                   *
*****************************************************************;
Title 'UP-AND-DOWN DESIGN/NORMAL';
```

```
Footnote 'UD$BHS:[dbettendorf.thesis.simulations]upanddown_normal.sas';

*********************************************************************
* INIT is the inital value for the algorithm
* DIVIDE is the divisor of 20
* STEP is the step coefficient in the algorithm
* P is the p as in LD100p
*Then we can discuss the ORDINARY VARIABLES:
*   RESP counts the number of responses within group
*   TOTAL counts the number of responses within experiment
*   QUANT is the estimate of the LD100p
*
*   EXP is the index variable for the experiments
*   RESPOND is an indicator whether the random variate is a response
*   MARGIN is the count of nonresponders
*********************************************************************
;
options ls=80 ps=54;
data driver;
 do q=1 to 3;
  do w=1 to 6;
   do e=1 to 3;
    do r=1 to 2;

      init=q-2;
        if w=1 then do; divide=20;
                  size=1;
                end;
       if w=2 then do; divide=10 ;
                  size=2 ;
                end;
      if w=3 then do; divide= 4;
                  size=5;
                end;
      if w=4 then do; divide= 30 ;
                  size=1 ;
                end;
      if w=5 then do; divide= 15 ;
                  size=2 ;
                end;
      if w=6 then do; divide= 6 ;
                  size=5 ;
```

```
            end;


      step=e*0.5;
      p= r*(0.15);
      truth= probit(p);
     output;
     end;
   end;
  end;
 end;
run;

***Now that the driver dataset is ready we move on to the kill;
*data collect check;
data collect(keep=step p init sam_siz size est row bias sqbias);
 set driver; row=_N_;
 retain resp total quant count numer;
 do exp=1 to 500;
    quant=init; total=0;  count=0; numer=0;
           do i=1 to divide;
        resp=0;
         do j=1 to size;
           z=normal(-1);
           respond= (z le quant);
           resp=resp+respond;
          end;

        total=resp+total;
        margin=size*i-total;
           if margin ge 3 then do;
                       numer=numer+quant;
                       count=count+1;
                        end;
        if resp ge 1 then do;
                      quant=quant-step;
                   end;
        else do;
              try=ranuni(-1);
               if try le p/(1-p) then quant=quant+step;
            end;
*output check;
```

```
        end;
     est=(numer/count) -step*(p-0.5);
     bias=(est-truth);
     sqbias=(est-truth)**2;
     sam_siz=size*i;
   output collect;
 end;
run;

proc univariate data=collect noprint;
    var sqbias bias;
    by row;
 output out=info
 mean=mse bias;
run;

data almost;
   merge collect
         info;
    by row;
   if first.row;
run;

proc sort data=almost;
 by p sam_siz size  init  step;
  run;

/*Writing the data to a suitable format*/
/* Make sure you open a blank sheet in Excel FIRST */
**Also, you need to put a put statement in your data step;

filename random dde 'excel|sheet1!r1c1:r140c8';
run;


data final;
  set almost;
  file random;
 bias=round(bias*100,1);
 mse=round(mse*100,1);
 put p sam_siz size init  step mse bias;
```

```
run;

proc print data=final(drop=row sqbias);
run;
```

Vita