BAYESIAN NONPARAMETRIC BIOASSAY ESTIMATION

Bahman Shafii and William J. Price
Statistical Programs
College of Agricultural and Life Sciences
P.O. Box 83844-2337
University of Idaho

Estimation of unknown pesticide levels in experimental samples is an important aspect of many agricultural and environmental studies. Such measurements are often made utilizing a "standard" dose response curve. This methodology compares the biological response of a target organism at known dosages to the response of the same organism exposed to an unknown sample. These "bioassays" are typically more efficient in time and resources than direct chemical assessment of the unknown sample. The form and choice of the standard curve, however, is subjective and can influence the estimation of the unknown dose. Problems may also arise when incomplete or preliminary information is available for determining the standard curve. One means of reducing the effects of these problems is to use a more generalized nonparametric estimation technique. This work will outline an alternative bioassay method based on a Bayesian nonparametric standard curve estimation framework. Empirical results will be demonstrated using data from a trichorpyr herbicide dose-response trial on lettuce germination.

1. Introduction

Dose-response curves are useful for modeling a variety of environmental effects such as exposure to chemical or temperature regimes. Common uses may range from bioassay assessment via calibration curves, estimation of specified quantiles (e.g. Shafii and Price 2006), or the modeling of the time dependent processes such as germination, emergence, and hatching (e.g. Shafii and Price 2001; Shafii, et al. 2009).

For the bioassay problem, an independently observed biological response is compared to a known dose-response curve . The objective is to estimate what dose most likely generated the observed response or, more precisely, what is the probability of a specified dose, given the observed data and the known dose-response (calibration) curve. This type of problem naturally falls into the framework of a Bayesian analysis.

Typically, the underlying bioassay dose-response curve, or tolerance distribution, is assumed to be a known increasing or decreasing sigmoidal function. Common examples would be the probit and logit forms based on normal and logistic probability distributions, respectively. In some situations, however, it may be advantageous to relax the assumptions surrounding the tolerance distribution, allowing for a more flexible shape of the dose-response curve. The need for such conditions could arise where the tolerance distribution is multimodal due to inherent subpopulations in the calibration data or a non-sigmoidal physiological response profile from a

process such as hormesis. In these cases, it would be useful to have a functional form of the dose-response curve which reflected these attributes.

This paper will demonstrate a nonparametric approach to estimating a tolerance curve, as well as, an associated estimation for an unknown dose given an observed response level. Comparative results for standard dose-response estimations using maximum likelihood and Bayesian techniques will also be provided.

2. Methods

Dose-response Curve

The techniques outlined here are applicable to many dose-response forms, however, this paper will concentrate on a logistic model form given by:

$$y_i/N_i = M / (1 + \exp(-\beta (dose_i - \gamma))) \quad (1)$$

where y_i/N_i is the proportion of successes at the i^{th} dose and N_i and y_i are the corresponding total and positive number of responses, respectively. The parameter measures the rate of response change over dose while estimates the dose at which the response equals .5M. M is the estimated theoretical maximum response level ($0 < M \leq 1.0$). Changes in the sign of parameter will describe either an increasing (negative) or decreasing (positive) sigmoidal function.

The model given in (1) can be generalized to the form:

$$y_i/N_i = MC/(C + \exp(-\beta (dose_i - \gamma))) \quad (2)$$

where $C = Q/(M - Q)$ and Q ($0 < Q < M$). In this model form, the parameter estimates the dose required to achieve the response level Q. When C = 1, the model reverts to the common form in (1) where measures the dose necessary to reach 0.5M. The generalized form is also convenient for estimating quantiles other than 50% of the maximum and will be used for the subsequent work.

Estimation: Maximum Likelihood

Given the responses y_i/N_i, a joint likelihood may be defined as:

$$L(\pi_i \mid y_i/N_i) \propto \Pi_i (\pi_i)^{y_i} \cdot (1 - \pi_i)^{N_i - y_i} \quad (3)$$

where $_i$ is given by (2) with the associated parameter vector, = [M, ,] and y_i and N_i are as defined above. The parameter vector is then estimated through maximization of (3) given the data y_i. One means of estimating an unknown dose, $^-$, is carried out through the selection of C in (2) corresponding to the observed response of the unknown dose prior to the dose-response estimation. Inferences on $^-$ are subsequently made assuming $^- \sim N(,)$.

Estimation: Parametric Bayesian

Assuming the parameter vector = [M, ,], the posterior distribution of can be defined as:

$$pr\ (\theta \mid y_i/N_i) \propto pr\ (y_i/N_i \mid \theta) \cdot pr\ (\theta) \qquad (4)$$

where $pr(y_i/N_i \mid\ )$ is the likelihood given in (3) and $pr(\ )$ is a prior distribution for the parameter vector . Estimation of is carried out through computationally intensive methods such as MCMC. Inference on the unknown dose, $^{-}$, can be obtained in a similar manner as that of the maximum likelihood method prior to estimation of the posterior distribution. The marginal posterior distribution of $^{-}$ is then obtained through selective integration over the parameter space of M and in (4) (e.g. Price and Shafii 2005).

Estimation: Non-parametric Bayesian

This methodology was first proposed by Mukhopadhyay (2000) and followed by Kottas et al. (2002). The technique considers the dose-response series as a multinomial process with a parameter vector $P = [p_1, p_2, p_3, \ldots p_k]$. Assuming the responses, y_i, are binomial, a likelihood similar to (3) may be defined as:

$$L\ (P \mid y_i/N_i) \propto \Pi_i\ (p_i)^{y_i} \cdot (1 - p_i)^{N_i - y_i} \qquad (5)$$

with the known dose-response curve, $_i$, replaced by the vector P. If the random segments between true response rates, p_i , are distributed as a Dirichlet Process (DP), a joint prior distribution on the p_i may then be defined by:

$$p(P) \propto \Pi_i\ (\ p_i\ - p_{i-1}\ )^{\lambda_i - 1} \qquad (6)$$

where $_i = \{ F_0(\text{dose}_i) - F_0(\text{dose}_{i-1}) \}$, is a precision parameter , and F_0 is a base response function assumed a priori. The precision parameter, , reflects how closely the final estimation follows the base distribution and may be set as a constant or estimated as part of the process. Low values indicate less correspondence , while larger values indicate a tighter association. The base function itself, $F_0(.)$, defines a family of potential response functions.

A posterior distribution for P can then be defined by combining (5) and (6) as:

$$p(P \mid y_i) \propto \Pi_i\ (\ p_i\ - p_{i-1}\ )^{\lambda_i - 1} \cdot \Pi_i\ (p_i)^{y_i} \cdot (1 - p_i)^{N_i - y_i} \qquad (7)$$

As with the parametric Bayesian method, estimation may be carried out through MCMC procedures. Inference on an unknown dose, $^{-}$, at a known response $\bar{y}_i/N_i$, is obtained through sampling of the posterior given in (7) conditionally, such that $p_i \leq \bar{y}_i/N_i, _i \leq p_{i+1}$, thus deriving the associated distribution of $^{-}$. Specifically, Mukhopadhyay provides a theorem allowing the unknown dose to be estimated by sampling the posterior distribution in (7) and solving for $^{-}$ in:

$$u = \frac{\Gamma(\lambda_\gamma + \lambda'_\gamma)}{\Gamma(\lambda_\gamma)\Gamma(\lambda'_\gamma)} \int_{\gamma_i^*}^{1} \omega^{\lambda_\gamma - 1}(1 - \omega)^{\lambda'_\gamma - 1} \delta\omega \qquad (8)$$

Here u is a random variate drawn from U[0,1], $^*_i = (^- - p_i)/(p_{i+1} - p_i)$, $(F_0(^-) - F_0(\ _i))$, and $'$ F_0 $_{i+1}$ F_0 $^-$ A beta CDF function is used to simplify the computations. Additionally, where $p_{i+1} < p_i$, the difference between these values was set to a small value (0.000001) to avoid an undefined condition when computing logarithms. While this is a modification of Mukhopadhyay, it was necessary for the data at hand, had little effect on the estimations, and accurately captured the condition where $p_{i+1} < p_i$.

3. Demonstration

The data used for demonstration are from a dose-response trial designed to assess the sensitivity of the herbicide trichlorpyr applied to seeds of four crops: Lettuce, Turnip, Radish, and Mustard. For each species, thirty seeds were placed in a petri dish with a specified rate of herbicide. Seven trichlorpyr rates were used ranging from 0 to 0.1 ml/L in concentration and each rate / species combination was replicated three times. Seed germination was recorded after 10 days. Non-germinated seeds at 10 days were considered as seed mortality. For the purposes of demonstration, only the lettuce seed mortality data will be considered here which are depicted in Figure 1. The dose levels have been converted to natural log scale in order to improve the symmetry of the data pattern. This required off setting the dose = 0 level by a small amount (0.03).

Demonstration: Maximum Likelihood

The likelihood given in (3) was maximized utilizing the SAS procedure PROC NLMIXED and the logistic tolerance form given in (2). SAS program codes are provided in Appendix 1. The constant C was set to C = Q/(M-Q) and Q = 0.5, where the variability of the response, proportion of mortality, was expected to be minimal. Subsequent estimation, would, therefore, be analogous to a"best case" scenario. The resulting estimates are shown in Table 1. All parameter estimates were significant with a maximum level of mortality of M = 0.93 and an estimated dose to 50% of = 0.506 ml/L. The fitted curve and associated distribution for the estimate of are shown in Figure 2. While the curve shows a reasonable fit, there is some deviation at the lower and higher dose levels.

Demonstration: Parametric Bayesian

Estimation of (4) was carried out using the SAS procedure PROC MCMC (Appendix 2). assuming similar settings as the maximum likelihood estimation A joint non-informative uniform prior distribution for the parameters was assumed. While this assumption may be varied in

practice, these simple priors were chosen for comparative purposes with the other estimation techniques.

The estimated values and their standard errors are shown in Table 2. As might be expected with the assumptions used, these estimates are very similar to those of the maximum likelihood estimation. Because the estimation procedure relies on the iterative MCMC method, diagnostic plots assessing iterative correlations were also produced (Figure 3). Autocorrelation, in this case, dropped off quickly and posed no problem to the final estimated solution. Figure 4 displays the fitted curve and the associated posterior distribution for the parameter . As with the numeric estimates, the distribution for appears similar to the corresponding maximum likelihood solution.

Demonstration: Non-parametric Bayesian

In order to demonstrate the full affect of the estimation technique, two scenarios are presented here. In both cases the posterior distribution in (7) is estimated via MCMC. Appropriate program codes are provided in Appendix 3.

Linear base dose-response curve, $F_0(.)$

The first example, while not typical in practice, is used for illustration and provides a strong visual assessment of the implications of this estimation technique. Here, the base family for the dose-response curve, F_0, is assumed to be a linear increasing function of dose. In Figure 5, the changing influence of this assumption is illustrated with increasing values of the constant . As increases, the final estimation (solid line) becomes more and more dependent on, and tends to converge to, the assumed linear base tolerance distribution, F_0 (dashed line). The associated estimate and posterior distribution for is also highly influenced by the level of . Determination of in this manner is subjective. Hence, as this example demonstrates, caution should be used in setting its value with lower values having the least amount of influence in the final solution. In practice, the parameter could also be estimated through the addition of an associated prior.

Logistic base dose-response curve, $F_0(.)$

As a second example, a more realistic base function is demonstrated. In this case, a logistic form with parameter values [M, ,] = [0.70, 2.55, and 0.30] is assumed. In practice, these values could be derived from previous trials, a literature review and/or personal expertise. For the example, the curve is close to the data, but deviates somewhat and might represent an earlier estimation from a previous data set.

Following the first example, MCMC estimation was carried out with an increasing series for setting the constant . Table 3 provides the final estimates and credible intervals for the estimated values of at each setting. Estimated fitted curves and the posterior distributions for are shown in Figure 6. As with the linear case, the higher levels of lead towards a convergence of the final solution and the base tolerance distribution. Similarly, the posterior

distribution for is adversely affected becoming narrower and less variable. While the mean estimates for vary by a small amount, it is noted that the fitted curve at the = 50 level follows the actual data pattern well, capturing both the slight initial dip in response levels at low doses, as well as the maximum observed response at the highest dose. Likewise, the lower settings for fail to capture these aspects of the data as well.

Comparison of Methods

Figures 7 and 8 give the overlayed fitted curves and associated distributions for each estimation technique, respectively (the = 10 level was selected for comparison in the non-parametric case). Numeric values for with the appropriate 95% interval estimates are provided in Table 4. As demonstrated earlier, maximum likelihood and parametric Bayesian solutions are very similar given the assumptions used. The strictly sigmoidal tolerance distribution assumed in these models, however, fails to capture the observed response rates at the lower and higher dose rates. The non-parametric Bayesian approach, in contrast, is better able to represent these response areas, given that an appropriate value of is assumed or estimated. Numerically, this is observed in a substantially lower mean estimate of for this method relative to either maximum likelihood or parametric Bayesian approaches. The distributional spread and corresponding interval estimate, however, is wider than the other two methods.

4. Concluding Remarks

Dose-response is an important tool in agricultural and biological research. The bioassay, designed to estimate an unknown dose at an observed response, is a widely used technique within the dose-response framework. While estimation of unknown doses can be problematic for some traditional estimation techniques, it is naturally expressed as a Bayesian problem. Some commonly used dose-response models may not reflect the patterns observed in actual data . These patterns may arise from undefined subpopulations or physiological processes present in the underlying data. The technique outlined here utilizes an assumed base function to define a family of potential tolerance distributions along with a Dirichlet prior assumption to capture some of these data attributes. The method is subjective, however, and must be applied with caution to avoid undue influence of the base tolerance assumption on the final estimation solution. When used properly, this methodology may provide a reasonable solution for a selected class of dose-response problems and give an estimated dose-response fit that accurately reflects the observed data.

Acknowledgments

Research partially funded by USDA-ARS Hatch Project IDA01310, Idaho Agricultural Experiment Station.

Data and technical material provided by Dr. Steven Seefeldt, USDA -ARS, Subarctic Agricultural Research Unit, University of Alaska Fairbanks.

References

Kottas, A., M. D. Branco, and A. E. Gelfand. 2002. A Nonparametric Bayesian Modeling Approach for Cytogenetic Dosimetry. Biometrics 58, 593-600.

Mukhopadhyay, S. 2000. Bayesian Nonparametric Inference on the Dose Level with Specified Response Rate. Biometrics 56, 220-226.

Price, W. J. and B. Shafii. 2005. Bayesian Analysis of Dose-response Calibration Curves. Proceedings of the Seventeenth Annual Kansas State University Conference on Applied Statistics in Agriculture [CDROM], April 25-27, 2005. Manhattan Kansas.

Shafii, B. and W. J. Price. 2001. Estimation of cardinal temperatures in germination data analysis. Journal of Agricultural, Biological and Environmental Statistics. 6(3):356-366.

Shafii, B. and W. J. Price. 2006. Bayesian approaches to dose-response calibration models.Abstract: Proceedings of the XXIII International Biometrics Conference [CDROM], July 16 - 21, 2006. Montreal, Quebec Canada.

Shafii, B., Price, W.J., Barney, D.L. and Lopez, O.A. 2009. Effects of stratification and cold storage on the seed germination characteristics of cascade huckleberry and oval-leaved bilberry. Acta Hort. 810:599-608.

Parameter	Estimate	Std Err	P>F
M	0.9338	0.0152	< 0.0001
β	2.5101	0.3394	< 0.0001
γ	0.5060	0.0383	< 0.0001

Table 1. Parameter estimates, standard errors, and the associated significance levels for the maximum likelihood estimation of the logistic dose-response model.

Parameter	Estimate	Std Dev of Posterior Dist.
M	0.9304	0.0153
β	2.5553	0.3539
γ	0.5132	0.0390

Table 2. Parameter estimates and standard deviations of the marginal posterior distributions for the parametric Bayesian estimation of the logistic dose-response model.

α	Lower Cr. I.	Mean γ	Upper Cr. I.
10	0.2875	0.4155	0.6300
50	0.3525	0.3988	0.5200
100	0.3325	0.3905	0.5050

Table 3. Gamma estimates and 95% credible intervals for the non-parametric Bayesian estimation assuming a logistic base function and three levels of α.

Method	Estimate	Lower*	Upper*
MLE	0.5060	0.4264	0.6856
Parametric Bayesian	0.5132	0.4408	0.5957
Nonparametric Bayesian	0.4155	0.2875	0.6300

Table 4 * Upper and Lower 95% bounds on γ for the three estimation methods. Confidence intervals for MLE, Credible Intervals for Bayesian methods

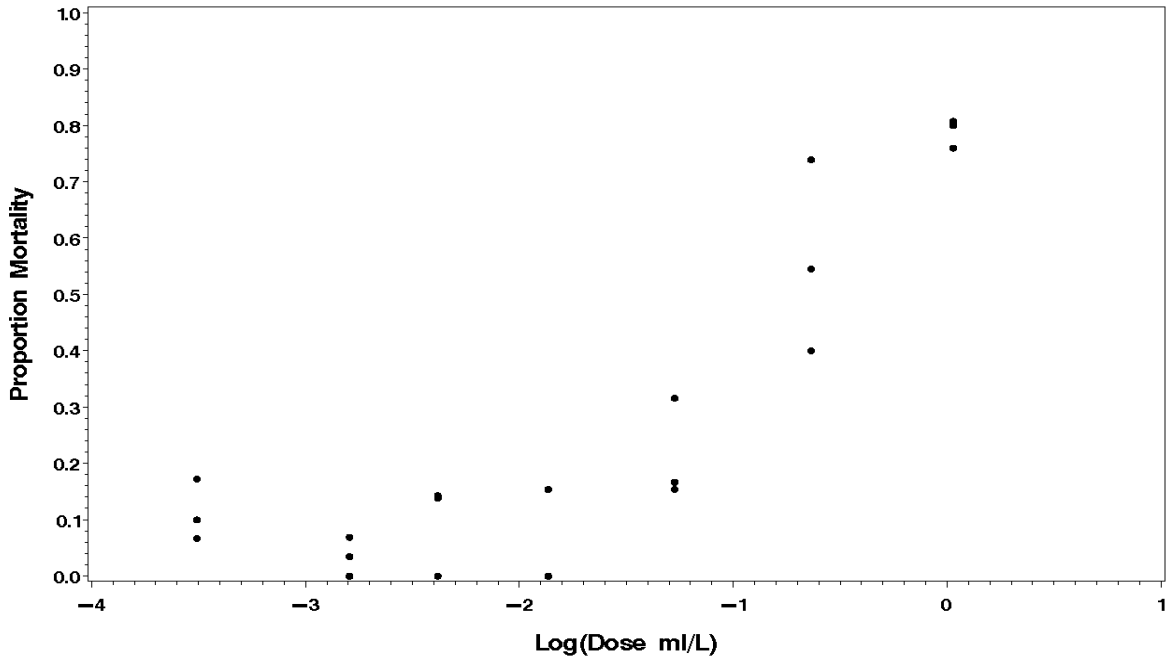

Figure 1. Proportion of lettuce seed mortality in response to trichlorpyr dose.

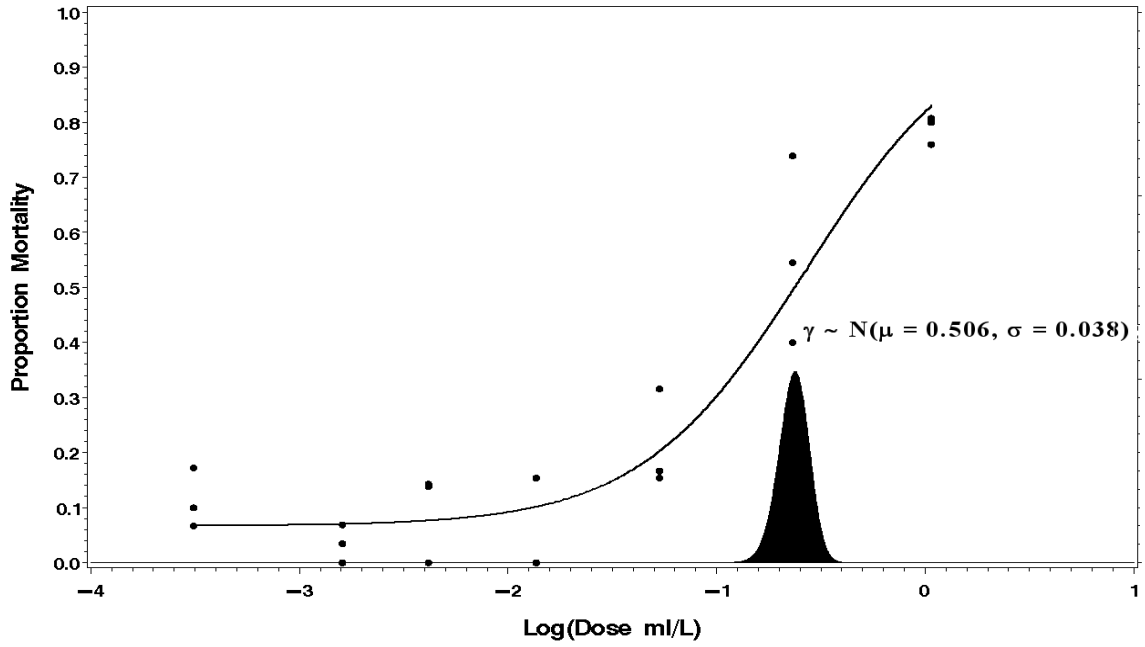

Figure 2. Maximum likelihood estimation of logistic dose-response model (solid line) fitted to lettuce seed mortality (dots). Associated distribution and parameters for the unknown dose at 50% mortality also shown.

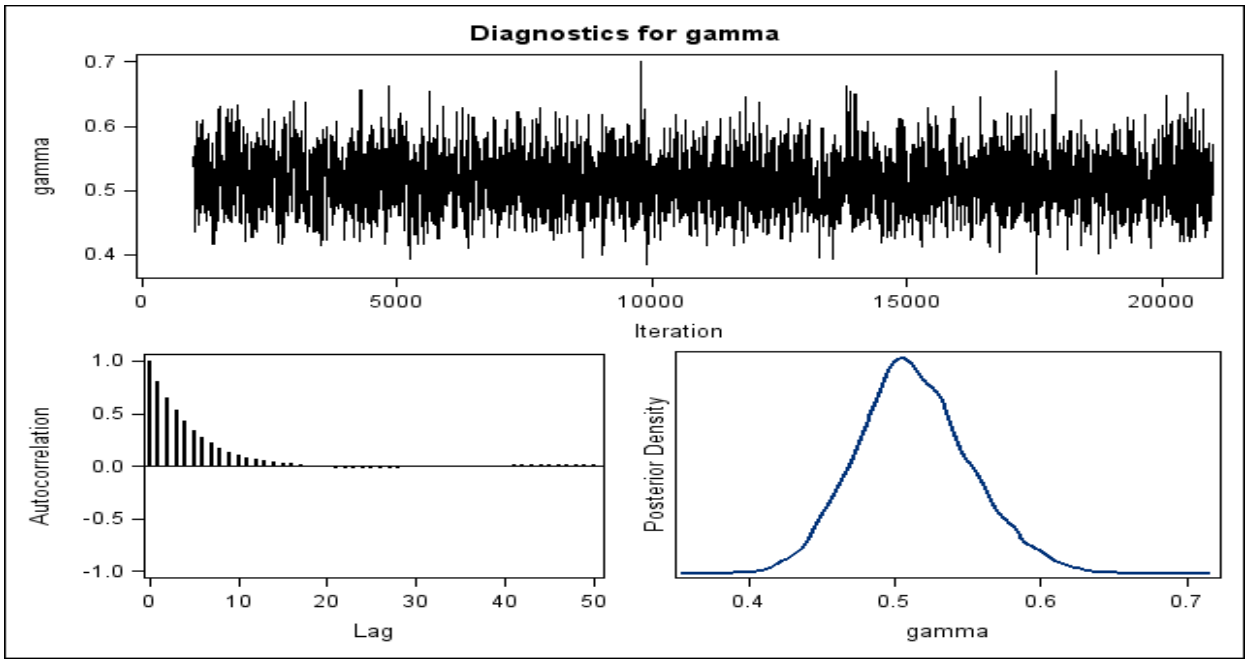

Figure 3. Diagnostic plots for γ in the parametric Bayesian estimation of logistic dose-response model.

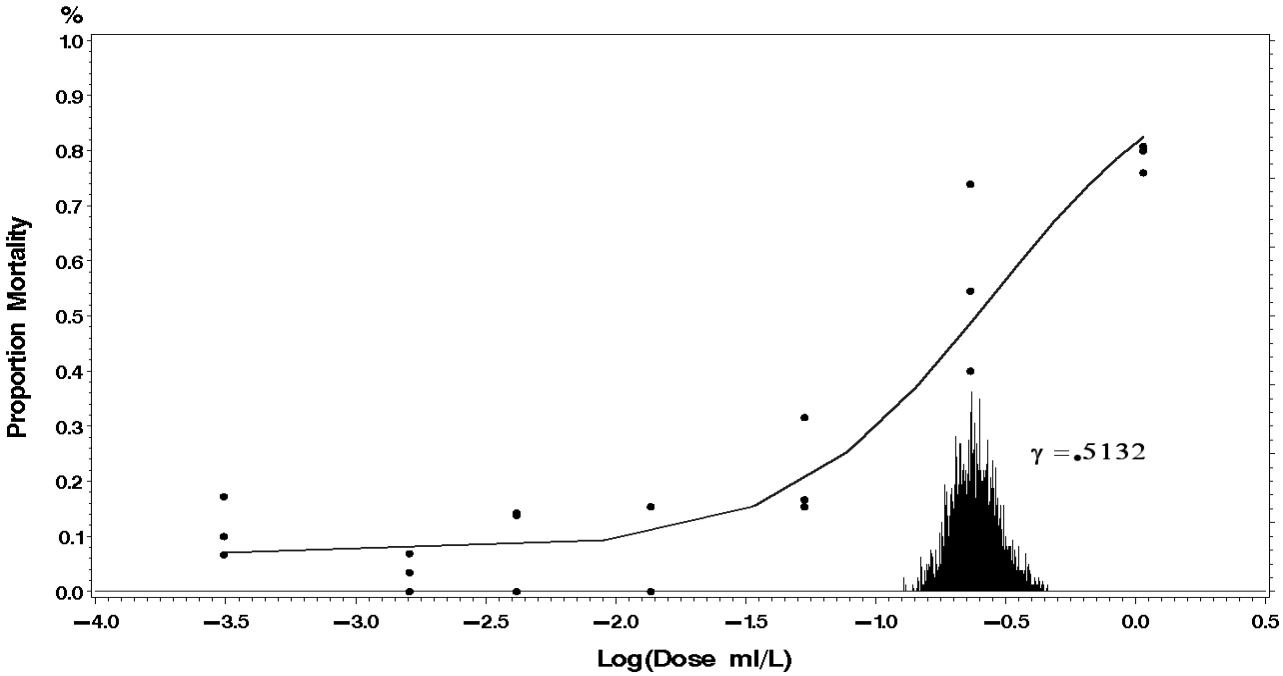

Figure 4. Parametric Bayesian estimation of logistic dose-response model (solid line) fitted to lettuce seed mortality (dots). Associated posterior distribution for the unknown dose at 50% mortality also shown.

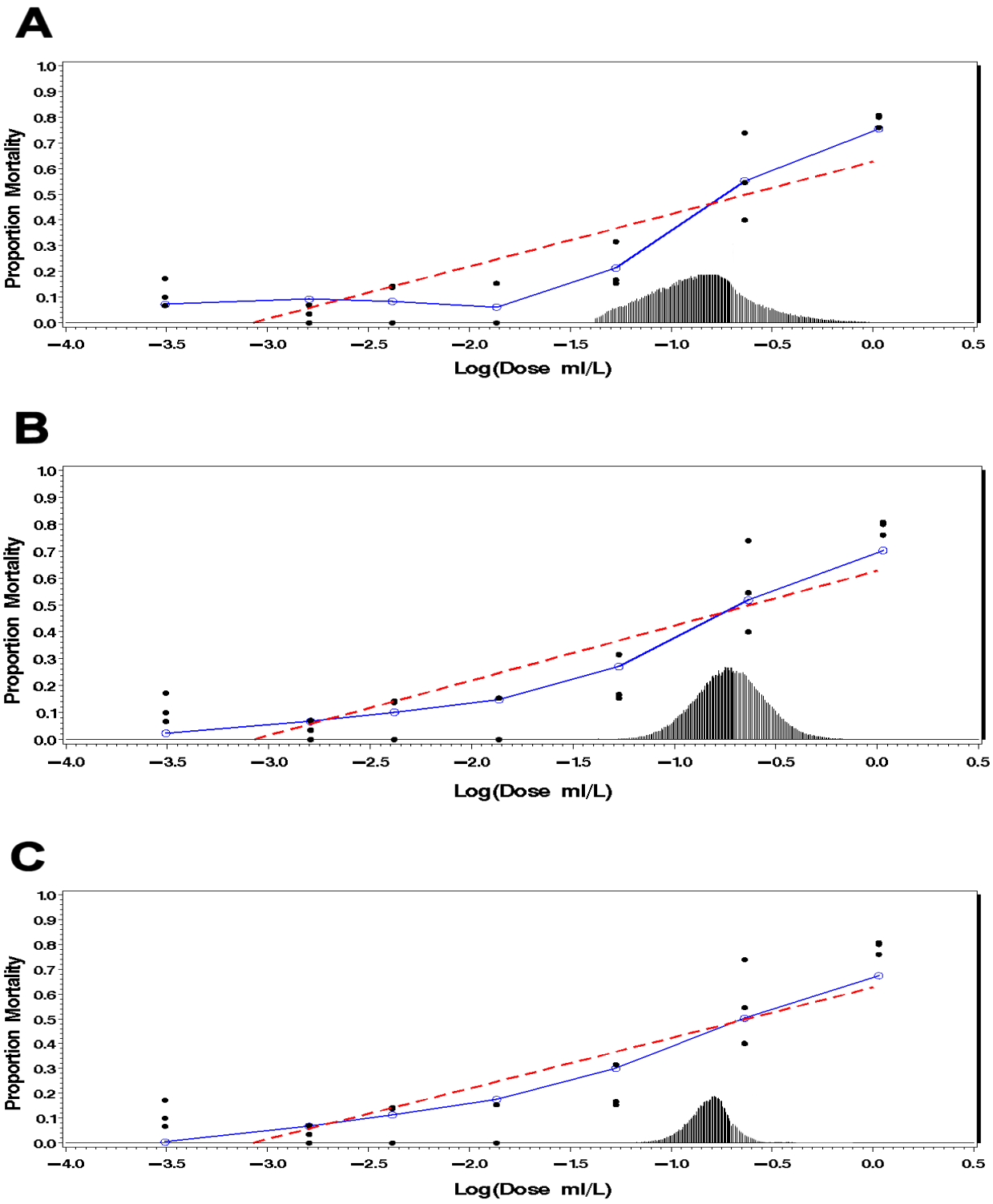

Figure 5. Non-parametric Bayesian estimation of logistic dose-response model (solid line) fitted to lettuce seed mortality (dots). Linear base tolerance function (dashed line) and the estimated posterior distribution for the unknown dose at 50% mortality also shown. Panels reflect changes in the setting of α at A: 10, B: 50, and C: 100.

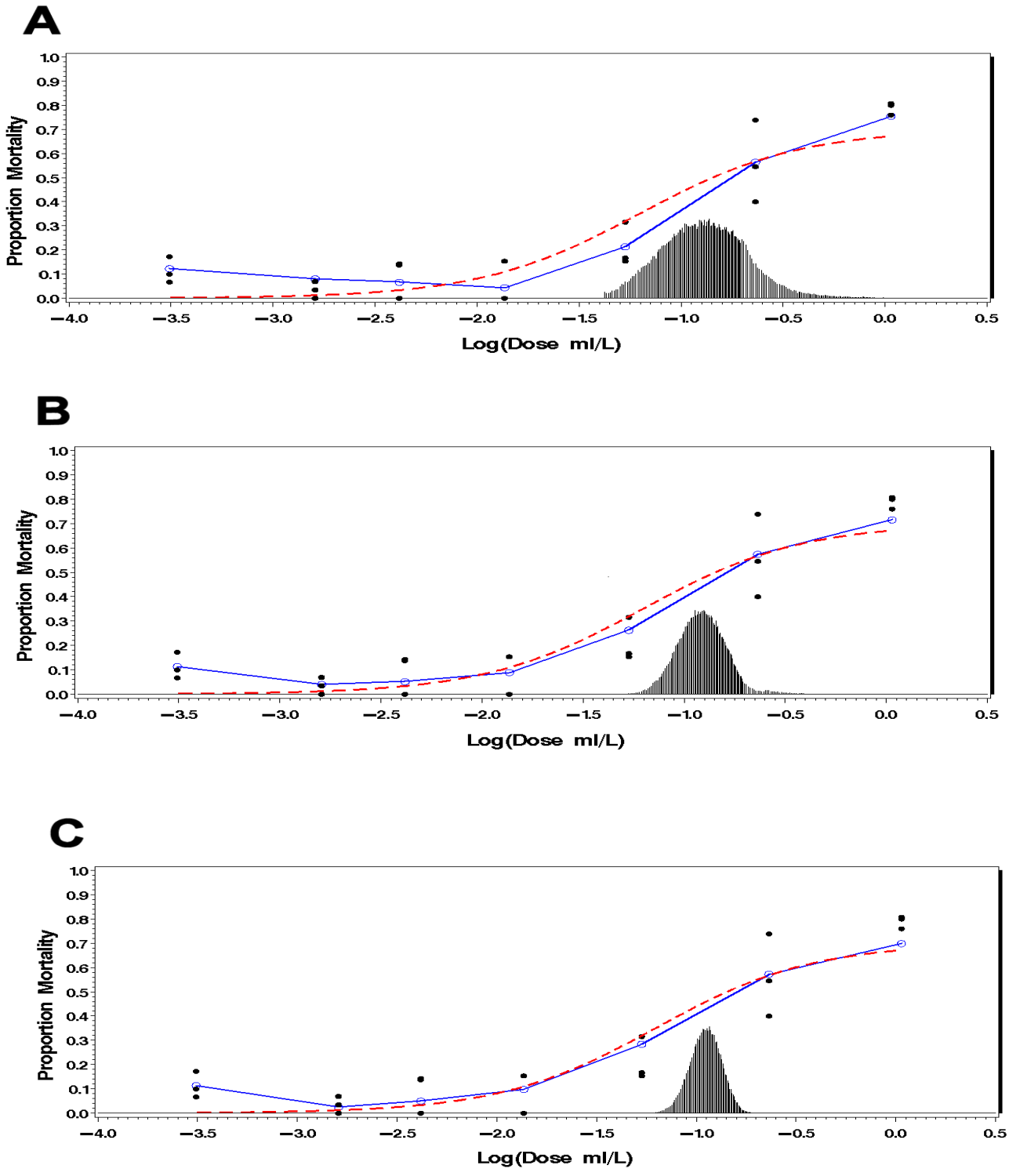

Figure 6. Non-parametric Bayesian estimation of logistic dose-response model (solid line) fitted to lettuce seed mortality (dots). Logistic base tolerance function (dashed line) and the estimated posterior distribution for the unknown dose at 50% mortality also shown. Panels reflect changes in the setting of αat A: 10, B: 50, and C: 100.

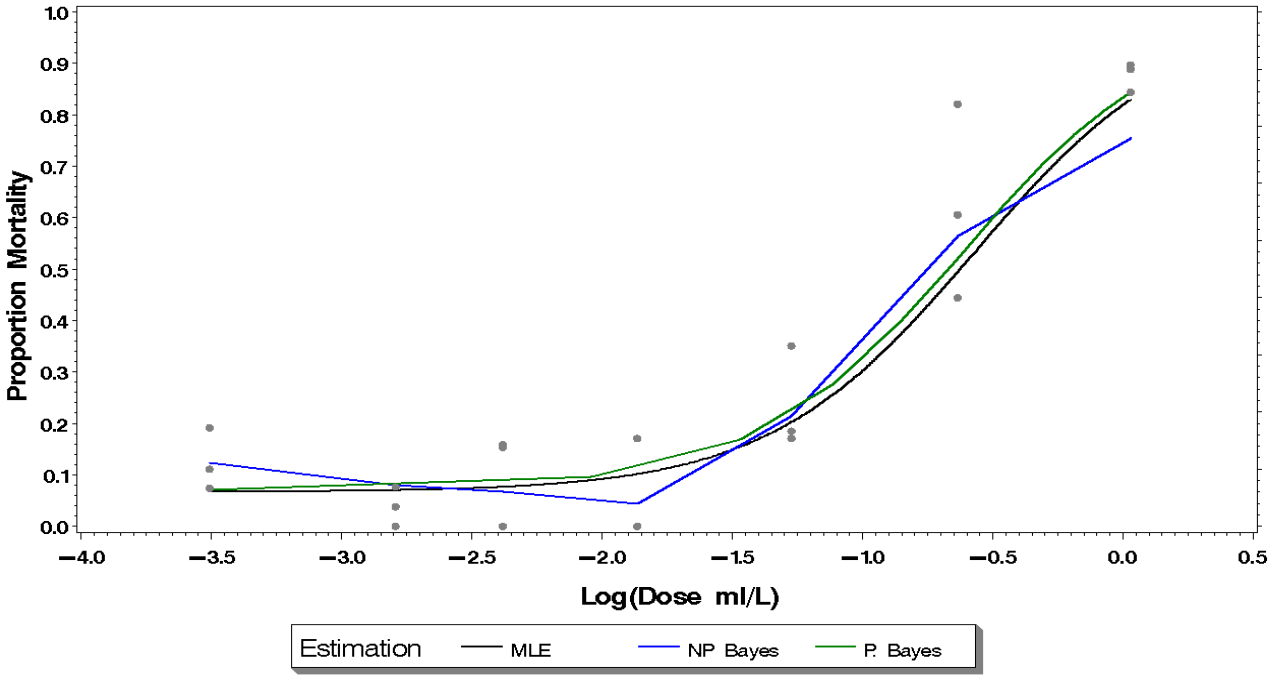

Figure 7. Estimated curves for the maximum likelihood (black), parametric Bayesian (green), and non-parametric Bayesian (blue, α=10) techniques fitted to the lettuce mortality data (dots).

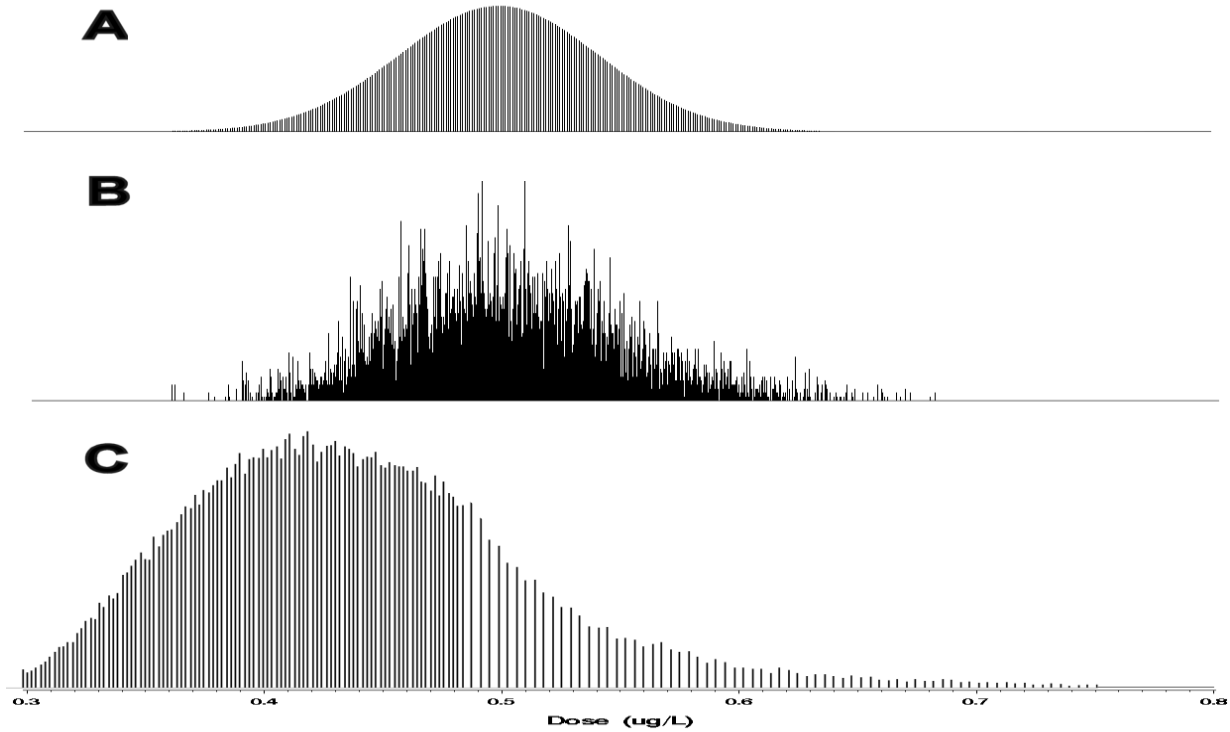

Figure 8. Estimated probability distributions for the maximum likelihood (A), parametric Bayesian (B), and non-parametric Bayesian (C, α=10) techniques fitted to the lettuce mortality data.

Appendix 1. SAS NLMIXED code for maximum likelihood estimation.

```
proc nlmixed data=DR;
        parms m=.9182 B=2.2604 gamma=0.45;
        Q=.5; /** value to estimate yi/Ni=.5 **/
        C = (Q)/(m - Q);
        mu = C*m/(C + exp(-B*(ldose - log(gamma + 0.03))));
        model yes ~ binomial(total, mu);

        predict mu out=pred;
```

Appendix 2. SAS MCMC code for parametric Bayesian estimation.

```
ods graphics on;

proc mcmc data=DR  seed=53197 ntu=1000 nmc=20000 propcov=quanew;
        ods select TADpanel PostSummaries; /** Produce printed & graphic output **/
        ods output PosteriorSample = ests; /**Output estimates **/

        parms M=1 B=2.5 gamma=0.22;
        Q=.5;  /** value to estimate yi/Ni=.5 **/
        C = (Q)/(m - Q);
        mu = M*C/(C + exp(-B*(ldose - log(gamma + 0.03))));
        prior M ~ uniform(0.8, 1.0);
        prior B  ~ uniform(1, 5);
        prior gamma ~ uniform(.15, 1.50);

        model yes ~ binomial(total, mu);
run;
ods graphics off;

data test; /** Compute predicted curve **/
        set ests;
        do dose = 0 to 1 by .1;
                ldose = log(dose+.03);
                Q=.5374;
                C = (Q)/(m - Q);
                mu = m*c/(c + exp(-B*(ldose - log(gamma + 0.03))));
                output;
        end;

proc sort data=test;
        by ldose;

proc univariate noprint data=test; /** Computed some summary stats **/
        var mu;
        by ldose;
        output out = test mean = yhat p5=lower p95=upper;
run;

data test;
        set test steve;

proc freq data=ests noprint;
        tables gamma/out=l_freqs;

data l_freqs;
        set l_freqs;
        dose = (gamma) ;
```

Appendix 3. SAS MCMC code for non-parametric Bayesian estimation.

```
/** Set up base values **/
%LET Q=0.5; /** Percentile to estimate **/
%LET A = 10; /** alpha level.  Comment out to estimate **/
%LET INT = 100; /** number of intervals to use in estimating the dose distribution **/

/** Set up Logistic base function example **/
%LET M = .7;
%LET K = -2.5553;
%LET L = .3;

%MACRO F0(D);
        (&M/(1 + exp(&K*(log(&D + 0.03) - log(&L)))))
%MEND;
/** Set up example linear base function example (uncomment to use) **/
/*
%LET B0 = .62665;
%LET B1 = .20377;
%MACRO F0(D);
        (&B0 + &B1*(log(&D+0.03)))
%MEND;
*/
options mprint;
title1 'Non-Parametric Bayes';

ods graphics on;
proc mcmc data=steve  seed=53197  ntu=1000 nmc=100000
        autocorlag= 500 monitor=(p1 p2 p3 p4 p5 p6 p7) propcov=quanew;
        ods select TADpanel PostSummaries;
        ods output PosteriorSample = ests; /** output sampled posterior **/

        parms p1=.1 p2=.2 p3=.3 p4=.4 p5=.6 p6=.7 p7=.8;* a=50; /** uncomment to estimate **/

         a = &A; /** Comment out to estimate **/

         u1 = max(p1 - 0,.000001); /** Catch condition pi+1 < pi **/
         u2 = max(p2 - p1,.000001);
         u3 = max(p3 - p2,.000001);
         u4 = max(p4 - p3,.000001);
         u5 = max(p5 - p4,.000001);
         u6 = max(p6 - p5,.000001);
         u7 = max(p7 - p6,.000001);
         u8 = max(1 - p7,.000001);


         lam1 = a*((%F0(0)) - (0));
         lam2 = a*((%F0(1/32)) - (%F0(0)));
         lam3 = a*((%F0(1/16)) - (%F0(1/32)));
         lam4 = a*((%F0(1/8)) - (%F0(1/16)));
         lam5 = a*((%F0(1/4)) - (%F0(1/8)));
         lam6 = a*((%F0(1/2)) - (%F0(1/4)));
         lam7 = a*((%F0(1)) - (%F0(1/2)));
```

```
            lam8 = (a)*((1) - (%F0(1)));

            lp1 = (lam1-1)*log(u1) + (1/7)*(lam8 - 1)*log(u8);
            lp2 = (lam2-1)*log(u2) + (1/7)*(lam8 - 1)*log(u8);
            lp3 = (lam3-1)*log(u3) + (1/7)*(lam8 - 1)*log(u8);
            lp4 = (lam4-1)*log(u4) + (1/7)*(lam8 - 1)*log(u8);
            lp5 = (lam5-1)*log(u5) + (1/7)*(lam8 - 1)*log(u8);
            lp6 = (lam6-1)*log(u6) + (1/7)*(lam8 - 1)*log(u8);
            lp7 = (lam7-1)*log(u7) + (1/7)*(lam8 - 1)*log(u8);


            if amount = 0 then mu = p1;
            else if amount = 1/32 then mu = p2;
            else if amount = 1/16 then mu = p3;
            else if amount = 1/8 then mu = p4;
            else if amount = 1/4 then mu = p5;
            else if amount = 1/2 then mu = p6;
            else if amount = 1 then mu = p7;

            prior p1 ~ general(lp1);
            prior p2 ~ general(lp2);
            prior p3 ~ general(lp3);
            prior p4 ~ general(lp4);
            prior p5 ~ general(lp5);
            prior p6 ~ general(lp6);
            prior p7 ~ general(lp7);
          * prior a ~ uniform(45, 55); /** uncomment to estimate **/

          model no ~ binomial(total, mu);

run;
ods graphics off;

proc univariate noprint data=ests;          /** Get average estimates **/
          var p1 p2 p3 p4 p5 p6 p7;

          output out = test mean = p1 p2 p3 p4 p5 p6 p7;
run;

/** Set up macro to compute Theorem 1 from Mukhopadhyay,2000 **/
%MACRO DOSE(P1, P2, D1, D2);
          gam = (&Q - &P1)/(&P2 - &P1);
          inc = ((&D2-&D1)/&INT);
          u = ranuni(0);
          do x = (&D1+INC) to (&D2-INC) by inc;
*         do x = .25 to .5 by (1/&int);
                    lam_x_a = &a*((%F0(x)) - (%F0(&D1)));
                    lam_x1_a = &a*((%F0(&D2)) - (%F0(x)));

                    lam_x_b = &a*((%F0(x+INC)) - (%F0(&D1)));
                    lam_x1_b = &a*((%F0(&D2)) - (%F0(x+INC)));
```

```
                diff_a = abs(u - (1 - probbeta(gam,lam_x_a, lam_x1_a) ));
                diff_b = abs(u - (1 - probbeta(gam,lam_x_b, lam_x1_b) ));
                if diff_b >= diff_a  then leave;
        end;
%MEND;

data dose;
        set ests;
        if p1 < &Q <= p2 then do; %DOSE(p1,p2,0,1/32); output; end;
        else if p2 < &Q <= p3 then do; %DOSE(p2,p3,1/32,1/16); output; end;
        else if p3 < &Q <= p4 then do; %DOSE(p3,p4,1/16,1/8); output; end;
        else if p4 < &Q <= p5 then do; %DOSE(p4,p5,1/8,1/4); output; end;
        else if p5 < &Q <= p6 then do; %DOSE(p5,p6,1/4,1/2); output; end;
        else if p6 < &Q <= p7 then do; %DOSE(p6,p7,1/2,1); output; end;
run;

/** Compute summary on dose distribution **/
data dose;
        set dose;
        ldose = log(x + 0.03);

proc freq data=dose noprint;
        tables ldose/out=l_freqs;

proc univariate data=l_freqs;
        weight count;
        var ldose;
run;
```