# Bayesian Analysis for Food-Safety Risk Assessment: Evaluation of Dose-Response Functions within WinBUGS

**Michael S. Williams**
United States Department
of Agriculture

**Eric D. Ebel**
United States Department
of Agriculture

**Jennifer A. Hoeting**
Colorado State
University

### Abstract

Bayesian methods are becoming increasingly popular in the field of food-safety risk assessment. Risk assessment models often require the integration of a dose-response function over the distribution of all possible doses of a pathogen ingested with a specific food. This requires the evaluation of an integral for every sample for a Markov chain Monte Carlo analysis of a model. While many statistical software packages have functions that allow for the evaluation of the integral, this functionality is lacking in **WinBUGS**. A probabilistic model, that incorporates a novel numerical integration technique, is presented to facilitate the use of **WinBUGS** for food-safety risk assessments. The numerical integration technique is described in the context of a typical food-saftey risk assesment, some theoretical results are given, and a snippet of **WinBUGS** code is provided.

*Keywords*: numerical integration, food safety, **WinBUGS**.

## 1. Introduction

The billions of servings of food and water consumed in developed countries are generally free of pathogenic bacteria, with illness from prepared food occurring at rates often on the order of one illness per 100,000 servings for even the most common pathogens (e.g., Lammerding 2006). The field of microbial food-safety risk assessment deals with the rare event that some number of pathogenic organisms, such as *Salmonella, Campylobacter* and *E. coli* O157:H7, are ingested by an individual and the infection progresses to serious illness. The models used for these risk assessments are typically referred to as "farm-to-table" models. They attempt to model processes mechanistically, such as microbial contamination, growth during transportation and storage, cross contamination, and the inactivation of organisms that occurs during freezing

and cooking.

These models are often complex Monte Carlo simulations that require considerable development effort. Nevertheless, a probabilistic model underlying such a risk assessment is surprisingly simple and consists of only two components. These being:

1. The level of microbial contamination (i.e., dose) $D$ of all servings at the point of consumption is distributed as $D \sim f(\cdot)$ where $f$ is a function of unknown parameters.

2. A dose-response function that estimates the probability of illness occurring given exposure to $D$, denoted $R(D) = P(\text{ill}|D)$.

The average probability of illness is estimated by integrating the product of these two components across the range of possible doses. Monte Carlo models are computationally intensive when second-order modeling is employed to capture both variability in the stochastic system as well uncertainty in the parameters used in the model (Frey and Burmaster 1999; Vose 2008).

Recent research interest has focused on the replacement of Monte Carlo models with a Bayesian approach that uses Markov chain Monte Carlo methods (Hald, Vose, Wegener, and Koupeev 2004; Parsonsa, Ortona, D'Souza, Mooreb, Jonesc, and Dodd 2005; Albert, Grenier, Denis, and Rousseau 2008). Along with this proposed approach comes the inevitable suggestion that the models be built using **WinBUGS** (Vose 2008).

The estimation of the average probability of illness can be a difficult task because the integration of the product of the exposure distribution and dose-response function rarely results in a closed-form solution. Numerical integration is typically used to evaluate the dose-response function, but some computing environments that are designed for implementing Markov chain Monte Carlo methods, such as **WinBUGS** (Lunn, Thomas, Best, and Spiegelhalter 2000), are not well suited to the evaluation of this integral. This article presents a probabilistic framework and a numerical integration technique that can be used for microbial food-safety risk assessment using the **WinBUGS** environment.

## 2. Outline of food-safety risk assessment

A biologically plausible model for microbial contamination assumes that average pathogen counts per serving vary according to a lognormal distribution (Limpert, Stahel, and Abbt 2001). Standard practice in microbiology is to express the mean and variance parameters describing microbial count data in $\log_{10}$, rather than $\log_e$, scale. The conversion between the two different scales is generally necessary for computation and is achieved by division of the $\mu$ and $\sigma$ parameters by $\log_e(10)$. The natural-log scale will be used for the remainder of this presentation.

The lognormal model holds for pathogen levels that are generally measured at the end of the production process (e.g., *Campylobacter* levels on poultry at the end of the slaughter process). Pathogen levels at consumption are generally unknown, so the risk assessment models the change in levels between production and consumption. Average pathogen levels per serving at the point of consumption can be modeled as the product of a large number of independent random variables describing growth and attenuation processes. When the central limit theorem is applied, this product is asymptotically a lognormal distribution (Mitzenmacher 2003).

By making this assumption, the dose distribution is defined by determining the parameters of the lognormal distribution (i.e., $D \sim \text{Lognormal}(\mu, \sigma^2)$) (Haas, Rose, and Gerba 1999). Given the low average pathogen levels per serving at the point of consumption, the $\mu$ parameter is generally much less than zero. The rare instances of high levels of contamination necessitate large $\sigma^2$ values. These parameter values require large sample sizes for accurate Monte Carlo approximation.

A number of different dose-response models have been proposed, with one of the most common models for describing the probability of microbial illness being the beta-Poisson dose-response function (Haas *et al.* 1999). In this case the number of ingested organisms (dose) follows a Poisson distribution with rate parameter $D$. Inter-individual variability in susceptibility is modeled by a beta distribution with parameters $\alpha$ and $\beta$.

Furumoto and Mickey (1967) derive the approximation

$$R(D) = P(\text{ill}|D = d) = 1 - \left(1 + \frac{d}{\beta}\right)^{-\alpha} \tag{1}$$

and show, via a Taylor's series expansion, that this approximation is sufficiently accurate for $\beta >> \alpha$. This model produces a curve that is sigmoidal when plotted in log space and the interpretation of the two parameters is that $\alpha$ controls the slope of the curve while $\beta$ controls the location. The ultimate goal is to determine the probability of an illness occurring across the distribution of possible contamination levels. Averaging the dose-response function across the distribution of all possible doses provides an estimate of the average probability of illness per random serving of the food of interest, i.e.,

$$E[R(D)] = P(\text{ill}) = \int_0^\infty \left(1 - \left(1 + \frac{D}{\beta}\right)^{-\alpha}\right) f(D)dD \tag{2}$$

Calculation of the integral in (2) represents the central component of many food-safety risk assessments. For example, multiplication of $P(\text{ill})$ by the number of servings of a food commodity provides an estimate of the total number of illnesses. The effectiveness of a change in production practices, designed to reduce the level of contamination, will generate a new estimate of the total number of illnesses. A comparison between the two estimates can then be used to predict the marginal benefit of the change in production practice.

## 3. Solution techniques

Solution of the integral in (2) can be approximated using a number of different techniques. A common solution is to draw a large sample of values from the dose distribution and use a Monte Carlo-based approximation. For example, if a sample of size $M$ dose values is drawn from the lognormal distribution and a beta-Poisson dose-response function is appropriate, then

$$\int_0^\infty R(D)f(D)dD \approx \frac{1}{M} \sum_{i=1}^{M} \left(1 - \left(1 + \frac{D_i}{\beta}\right)^{-\alpha}\right). \tag{3}$$

This is a reasonable approach that forms the basis of nearly all Monte Carlo risk assessments. Nevertheless, accurate approximation to the integral with the lognormal distributions typically found in food-safety risk assessments requires values for $M$ on the order of millions of

observations to adequately explore the tails of the distribution, which is generally beyond the capabilities of **WinBUGS**.

A more computationally efficient approach is to replace the integral with a partitioned, discrete equation that provides an analytic solution that can be incorporated into **WinBUGS**. The proposed approximation to (2) is motivated by considering the following: note that

$$P(\text{ill}) = \int_0^\infty R(D)f(D)dD = E[R(D)]. \tag{4}$$

Then partitioning the range of doses yields

$$
\begin{aligned}
\int_0^\infty R(D)f(D)dD &= \int_0^{D_1} R(D)f(D)dD + ... \int_{D_i}^{D_{i+1}} R(D)f(D)dD + ... \int_{D_N}^\infty R(D)f(D)dD \\
&= E[R(D)|0 \le D < D_1]F(D_1)... + \\
&\quad E[R(D)|D_i \le D < D_{i+1}](F(D_{i+1}) - F(D_i))... + \\
&\quad E[R(D)|D_N \le D < \infty](1 - F(D_N)),
\end{aligned}
$$

where $F(D)$ is the cumulative distribution function of $f(D)$.

The proposed integration technique exchanges the order in which the dose-response and expected value are evaluated. That is

$$
\begin{aligned}
\tilde{P}(\text{ill}) &= R(E[D|0 \le D < D_1])F(D_1)... + \\
&\quad R(E[D|D_i \le D < D_{i+1}])(F(D_{i+1}) - F(D_i))... + \\
&\quad R(E[D|D_N \le D < \infty])(1 - F(D_N)).
\end{aligned}
$$

This equation can also be written as:

$$
\begin{aligned}
\tilde{P}(\text{ill}) &\approx R(\bar{D}_0^{D_1})F(D_1) + R(\bar{D}_{D_1}^{D_2})(F(D_2) - F(D_1)) + ... \\
&\quad + R(\bar{D}_{D_N}^\infty)(1 - F(D_N)),
\end{aligned} \tag{5}
$$

where $\bar{D}_{D_i}^{D_{i+1}} = E[D|D_i < D < D_{i+1}]$ is the mean dose in the interval with endpoints $D_i$ and $D_{i+1}$, $F(D_{i+1}) - F(D_i)$ is the proportion of doses in the interval, and $N+1$ is the number of intervals used to partition the range of doses from 0 to $\infty$. This is a type of Newton–Côtes quadrature approximation (Givens and Hoeting 2005, Chapter 5).

The distribution describing $D|D_i < D < D_{i+1}$ is known as a truncated distribution (Johnson, Kotz, and Balakrishnam 1994), where one of the common areas of application for these distributions is in the actuarial sciences (Klugman, Panjer, and Willmot 2004). The expected value of a truncated random variable is given by

$$\bar{D}_{D_i}^{D_{i+1}} = E[D|D_i < D < D_{i+1}] = \frac{1}{F(D_{i+1}) - F(D_i)} \int_{D_i}^{D_{i+1}} Df(D)dD \tag{6}$$

where $F(D)$ is the cumulative density function of $D$.

For the lognormal distribution, the mean in the interval with upper bound $D_i$ is

$$\bar{D}_0^{D_i} = e^{\mu + \sigma^2/2}\Phi\left(\frac{(\log(D_i) - \mu - \sigma^2)}{\sigma}\right) \Bigg/ \Phi\left(\frac{\log(D_i) - \mu}{\sigma}\right), \tag{7}$$

and $\Phi$ is the standard normal cumulative density function (Klugman *et al.* 2004). Equation 7 is used to determine $\bar{D}_0^{D_1}$ while the expected value for the remaining intervals is calculated by the weighted difference of successive intervals using

$$\bar{D}_{D_i}^{D_{i+1}} = \frac{\bar{D}_0^{D_{i+1}}\Phi((\log(D_{i+1}) - \mu)/\sigma) - \bar{D}_0^{D_i}\Phi((\log(D_i) - \mu)/\sigma)}{\Phi((\log(D_{i+1}) - \mu)/\sigma) - \Phi((\log(D_i) - \mu)/\sigma)}. \tag{8}$$

The proposed approximation is beneficial only for probability distributions that have closed-form solutions, or good approximations, to $\bar{D}_{D_i}^{D_{i+1}}$ as well as software that supports the necessary intrinsic functions for the evaluation of both $\bar{D}_{D_i}^{D_{i+1}}$ and $F(D_i)$ (e.g., normal and gamma distributions).

Given Equations 7–8, it is possible to construct an accurate approximation to the integral of the dose-response function using a limited number of terms. Dose intervals need not be equal in size or consistent in pattern. Selection of dose intervals can be informed by considering the uncertainty about the true dose distribution or the shape of the dose-response function. A simplistic, yet reasonable, partitioning of the range of doses is based on the symmetry of the normal distribution. Specifically, the range of log transormed dose can be defined as $(\mu - k\sigma, \mu + k\sigma)$, where $k$ describes the number of standard deviations considered on each side of the mean of the log transformed dose. Choosing $k$ to be from 4 to 6 and setting the width of the partition in the range of 0.05 to 0.1 generally provides estimates of $P(\text{ill})$ that match estimates derived using the adaptive quadrature routine **QUADPACK** in **Netlib** to within the absolute error bounds values reported from **QUADPACK** (Piessens, deDoncker Kapenga, Uberhuber, and Kahaner 1983). In contrast, estimates of $P(\text{ill})$ derived using (3) with $M = 3,000,000$ Monte Carlo samples exhibited discrepancies that were often an order of magnitude larger.

More complex and accurate methods for numerical integration, such as Gaussian quadrature, are available (Givens and Hoeting 2005, Chapter 5). Nevertheless, there are a number of applications in risk assessment where user defined truncation points are important (e.g., a $D_i$ value that represents a maximum safe dose or other food-safety metric (Crouch, Labarre, Golden, Kause, and Dearfield 2009)), while in other applications the mean of the truncated distribution is important for the partitioning of overall risk into categories (Williams, Ebel, and Vose 2011). These applications are natural extensions of the proposed integration technique. Given these practical considerations, the proposed solution represents a good trade-off between simplicity and accuracy and it works well within the limitations inherent in **Win-BUGS**.

# 4. Theoretical properties

For a fixed partition of the dose range, the proposed numerical integration scheme is not an unbiased estimator of the true probability of illness. Bias and convergence properties of the proposed integration technique can be determined using results from basic probability theory.

To assess the bias, consider the relationship between $E[R(D)|D_i \leq D < D_{i+1}]$ and $R(E[D|D_i \leq D < D_{i+1}])$ for an arbitrary interval $i$. The direction of the bias in the estimated probability of illness can be determined by examining the second derivative of the dose-response function and applying Jensen's inequality (Ross 1988). For typical dose-response functions, such as

the beta-Poisson and exponential (Haas *et al.* 1999), the second derivative of $R(D)$ is negative for all $D$, so the function is concave and

$$E[R(D)|D_i \leq D < D_{i+1}] \leq R(E[D|D_i \leq D < D_{i+1}])$$

in every interval. Consequently $\tilde{P}(\text{ill}) > P(\text{ill})$. Though a formal proof is not provided, this same result can be used to show that $\tilde{P}(\text{ill})$ is asymptotically unbiased in the limit as the interval width goes to zero. To demonstrate, note that $E[R(D|D_i \leq D < D_{i+1})] = R(E[D|D_i \leq D < D_{i+1}])$ when $R()$ is linear. All that remains is to show that $R()$ approaches a linear function as the interval width goes to zero.

A closed-form estimate of the bias in each interval, as well as the total bias, can be derived. To facilitate the discussion note that a weighted sum notaion can be use to describe the probability of illness measures as

$$P(\text{ill}) = \sum_{i=0}^{N} E[R(D)|D_i \leq D < D_{i+1})]w_i$$

and

$$\tilde{P}(\text{ill}) = \sum_{i=0}^{N} R(E[D|D_i \leq D < D_{i+1}])w_i,$$

where $D_0 = 0$, $D_{N+1} = \infty$, and the weights are as defined in (5) and represent the proportion of doses in the interval. The magnitude of the bias can be approximated by considering a Taylor's series expansion of each term in the partitioned range of doses. To demonstrate, consider an abitrary dose value $D \in (D_i, D_{i+1})$ from interval $i$. Expanding about $\bar{D}_{D_i}^{Di+1}$ yields the approximation

$$R(D|D_i \leq D < D_{i+1}) \approx R(\bar{D}_{D_i}^{Di+1}) + R'(\bar{D}_{D_i}^{Di+1})(D - \bar{D}_{D_i}^{Di+1}) + \frac{1}{2}R''(\bar{D}_{D_i}^{Di+1})(D - \bar{D}_{D_i}^{Di+1})^2.$$

Taking the expected value of $R(D|D_i \leq D < D_{i+1})$ yields

$$E[R(D|D_i \leq D < D_{i+1})] \approx R(\bar{D}_{D_i}^{Di+1}) + \frac{1}{2}R''(\bar{D}_{D_i}^{Di+1})Var[D_{D_i}^{Di+1}].$$

The contribution to the bias associated with interval $i$ is

$$R(E[D|D_i \leq D < D_{i+1}]) - E[R(D)|D_i \leq D < D_{i+1}] = -\frac{1}{2}R''(\bar{D}_{D_i}^{Di+1})Var[D_{D_i}^{Di+1}]w_i.$$

Expressing the bias across the range of doses leads to the weighted sum

$$\tilde{P}(ill) - P(ill) \approx -\sum_{i=0}^{N} \frac{1}{2}R''(\bar{D}_{D_i}^{Di+1})Var[D_{D_i}^{Di+1}]w_i.$$

For the lognormal distribution, a closed-form solution for $Var[D_{D_i}^{Di+1}]$ is given by:

$$Var[D_{D_i}^{Di+1}] = E[(D_{D_i}^{Di+1})^2] - (\bar{D}_{D_i}^{Di+1})^2,$$

with

$$E[(D_0^{D_1})^2] = \frac{e^{2\mu+2\sigma^2} \times \Phi\left(\left(\log(D_1) - \mu - 2\sigma^2\right)/\sigma\right)}{\Phi\left(\left(\log(D_1) - \mu\right)/\sigma\right)}$$

and

$$E[(D_{D_i}^{D_{i+1}})^2] = \frac{e^{2\mu+2\sigma^2} \times \left(\Phi\left(\left(\log(D_{i+1}) - \mu - 2\sigma^2\right)/\sigma\right) - \Phi\left(\left(\log(D_i) - \mu - 2\sigma^2\right)/\sigma\right)\right)}{\Phi\left((\log(D_{i+1}) - \mu)/\sigma\right) - \Phi\left((\log(D_i) - \mu)/\sigma\right)}.$$

The formulas for the calculation of $\left(\bar{D}_{D_i}^{D_{i+1}}\right)^2$ are provided in Equations 7–8.

The approximation to the bias can be used to select the boundaries of the dose intervals $(D_i, D_{i+1})$ so as to minimize the bias for a given number of intervals. This might, however, be difficult in this application because the uncertainty in the parameters of the exposure distribution and dose-response function causes the optimal grid spacing to vary from one iteration of the model to the next.

# 5. Example

The primary function of the provided code is to implement the proposed integration technique as it might appear as part of a food-safety risk assessment. The level of detail is more simplistic than an actual food safety risk assessment. The primary difference is that many of the parameters are treated as fixed values.

To motivate the example, suppose the public health agency, responsible for the food product, wishes to predict the effect of a risk management option (e.g., an intervention such as mandatory incorporation of a chemical antimicrobial agent) that reduces the mean level of *Campylobacter* contamination at the end of production by one-half $log_e$ unit.

Consider the case where the baseline, or current, distribution of *Campylobacter* per serving at the point of consumption can be modeled as $f(D) \sim \text{Lognormal}(\mu, \sigma)$. It is impossible to collect representative data at the point consumption, so the $\mu$ and $\sigma$ parameters would typically be derived using data collected during the production and processing of the food item and then predictive microbiology models (Haas *et al.* 1999), estimates from previous risk assessments, and possibly expert opinion would be used to determine pathogen levels at the point of consumption, as well as measures of uncertainty in these parameters. For this example, let the parameters and their uncertainties be modeled as $\mu \sim \text{Normal}(-9.5, 1)$ [1] and $\sigma \sim \text{Gamma}(200, 100)$ and assume the $\mu$ and $\sigma$ are independent.

This example assumes a new intervention during the production of carcasses has an estimated proportional reduction of 0.65 (i.e., one-half log reduction or $e^{1/2}$) in the mean level of contamination per carcass. The effectiveness of this hypothetical intervention and its uncertainty, in the log scale, is modeled as $\delta \sim \text{Gamma}(200, 100)$. The distribution of *Campylobacter* per serving at the point of consumption that reflects the reduction in contamination can be modeled as $f_{\text{new}}(D) \sim \text{Lognormal}(\mu - \delta, \sigma)$.

The parameters of the beta-Poisson dose-response function (1) are taken from Black, Levine, Clements, Hughes, and Blaser (1988) and are given by ($\alpha = 0.145, \beta = 7.59$). Fixed values are used for these parameters in the example, but in practice additional uncertainty could be incorporated.

A microbial risk assessment model can reasonably assume that the observed number of sporadic illnesses are distributed as a Poisson random variable (Vose 2008) where the rate pa-

---

[1]Note that the parameterization of the normal distribution in **WinBUGS** is based on precision rather than standard deviation.

rameter is the number of servings multiplied by the probability of illness. Suppose for the population of interest there were $N_{\text{servings}} = 10,000,000$ servings consumed annually of the food product of interest and that a total of $N_{\text{ill}} = 540$ illnesses were attributed to consumption of the product.

In this setting, the likelihood function in the **WinBUGS** environment would be

$$N_{\text{ill}} \sim \text{Poisson}(N_{\text{servings}}P(\text{ill})).$$

In practice $N_{\text{ill}}$ would not be known. Note, however, that in contrast to chemical toxins, estimated total illnesses per annum can be empirically based for microbial pathogens because surveillance systems are in place to detect a portion of such illnesses (Allos, Moore, Griffin, and Tauxe 2004). An additional conversion factor would be required to scale between the observed and the total number of illnesses (e.g., Mead, Slutsker, Dietz, McCaig, Bresee, Shapiro, Griffin, and Tauxe 1999).

A primary objective for the types of risk assessments considered here is the prediction of the number of illnesses that would be avoided ($N_{\text{avoided}}$), assuming the implementation of a new policy or intervention that reduces contamination. This statistic assumes a reduction in pathogens for each serving and models the number of illnesses avoided as $N_{\text{avoided}} \sim \text{Poisson}(N_{\text{servings}}(P(\text{ill}) - P_{\text{new}}(\text{ill})))$, where

$$P_{\text{new}}(\text{ill}) = \int_0^\infty (1 - \left(1 + \frac{D}{\beta}\right)^{-\alpha})f_{\text{new}}(D)dD \tag{9}$$

Figure 1 depicts the resulting posterior distribution for $N_{\text{avoided}}$, with the red lines indicating the boundaries of the 95 % posterior prediction intervals. The number of illnesses avoided can then be used to assess the economic benefit of implementing the policy. These estimated saving can then be compared to the expected cost to industry of implementing the policy to determine if potential benefits exceed the cost of implementation.

The following **WinBUGS** code uses the proposed numerical integration technique for the estimation of the number of human illnesses avoided.

The first line of code inputs the number of intervals (`num.bins`), interval width (`int.width`), number of standard deviations about the mean of the dose distribution (`k`), number of servings consumed (`N.servings`) and number of illnesses attributed to consumption of the product (`N.ill`). The number of intervals used here represents partitioning the interval of approximately 6 standard deviations about the mean into sub-intervals of about 0.1 units, i.e., `num.bins = 2 * k * sigma/interval.width = 2 * 6 * 2/(.1) = 240`.

```
list(num.bins = 240, int.width = 0.1, k = 6,
  N.servings = 10000000, N.ill = 540)
```

In the model step, the parameters of a Beta-Poisson dose-response model for Campylobacter from Black *et al.* (1988) are used (`alpha.dr = 0.145` and `beta.dr = 7.59`). These parameters are fixed values, but assuming prior distributions for these parameters would be even more appropriate in the Bayesian context.

The mu and sigma steps below specify the prior distributions describing the uncertainty in the mean and standard error of the dose distribution at the point of consumption. Recall that the parameterization of the normal distribution in **WinBUGS** is based on precision rather than standard deviation.
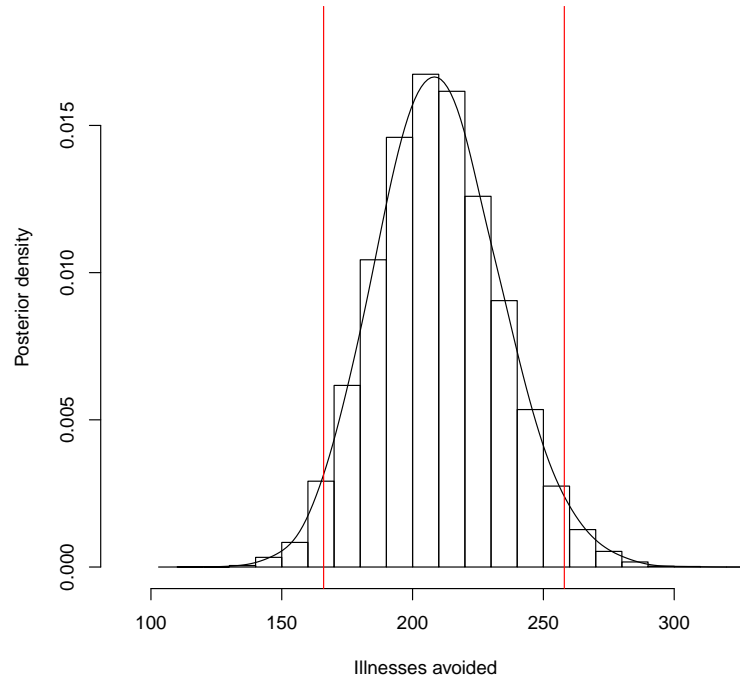
Figure 1: Posterior distribution of $N_{\text{avoided}}$. The vertical line indicates boundaries of the 95 % posterior prediction interval for the number of illnesses avoided as a result implementing the proposed policy. The interval spans the range $(165, 258)$.

The code in the remainder of the manuscript is the model step in the **WinBUGS** code. The delta and mu.new steps specify the anticipated reduction in contamination under the proposed changes to the industry and define the new mean level of contamination.

```
alpha.dr <- 0.145
beta.dr <- 7.59

mu ~ dnorm(-9.5,1)
sigma ~ dgamma(200, 100)

delta ~ dgamma(100, 200)
mu.new <- mu - delta
```

The next steps perform the numerical integration of the dose-response model. First, we set the upper bound, mean dose, and proportion of servings in the first interval for the baseline and reduced contamination scenarios.

```
u[1] <- exp(mu - k * sigma)

D.bar.lo[1] <- 0
```

```
D.bar.up[1] <- exp(mu + 0.5 * pow(sigma, 2)) *
               phi((log(u[1]) - mu - pow(sigma, 2)) /
               sigma) / phi((log(u[1]) - mu) / sigma)

D.bar[1] <- D.bar.up[1]
F.lo.up[1] <- phi((log(u[1]) - mu) /sigma)

u.new[1] <- exp(mu.new - k * sigma)

D.bar.lo.new[1] <- 0
D.bar.up.new[1] <- exp(mu.new + 0.5 * pow(sigma, 2)) *
                   phi((log(u.new[1]) - mu - pow(sigma, 2)) / sigma) /
                   phi((log(u.new[1]) - mu) / sigma)

D.bar.new[1] <- D.bar.up.new[1]
F.lo.up.new[1] <- phi((log(u.new[1]) - mu.new) / sigma)
```

Subsequently, we determine the mean concentration and proportion of servings in all but the last interval.

```
for (i in 2:(num.bins)) {
  u[i]        <- exp(mu - k * sigma + (i-1) * int.width)
  D.bar.lo[i] <- D.bar.up[i-1]
  D.bar.up[i] <- exp(mu + 0.5 * pow(sigma, 2)) *
                 phi((log(u[i]) - mu - pow(sigma, 2)) / sigma) /
                 phi((log(u[i]) - mu)/sigma)
  D.bar[i]    <- (D.bar.up[i] * phi((log(u[i]) - mu) / sigma) -
                 D.bar.lo[i] * phi((log(u[i-1]) - mu) / sigma)) /
                 (phi((log(u[i]) - mu) / sigma) -
                 phi((log(u[i-1]) - mu) / sigma))
  F.lo.up[i]  <- (phi((log(u[i]) - mu) / sigma) - phi((log(u[i-1]) -
                 mu)/sigma))

  u.new[i]         <- exp(mu.new - k * sigma + (i-1) * int.width)
  D.bar.lo.new[i] <- D.bar.up.new[i-1]
  D.bar.up.new[i] <- exp(mu.new + 0.5 * pow(sigma, 2)) *
                     phi((log(u.new[i]) -
                     mu.new - pow(sigma, 2)) / sigma) /
                     phi((log(u.new[i]) - mu.new) / sigma)

  D.bar.new[i] <- (D.bar.up.new[i] *
                  phi((log(u.new[i]) - mu.new) / sigma) -
                  D.bar.lo.new[i] *
                 phi((log(u.new[i-1]) - mu.new) / sigma)) /
                 (phi((log(u.new[i]) - mu.new) / sigma) -
                  phi((log(u.new[i-1]) - mu.new) / sigma))
```

```
    F.lo.up.new[i] <-(phi((log(u.new[i]) - mu.new) / sigma)-
                      phi((log(u.new[i-1]) - mu.new) / sigma))
}
```

Now, determine the mean concentration and proportion of servings in the last interval.

```
D.bar[num.bins+1] <- (exp(mu + 0.5 * pow(sigma, 2)) - D.bar.up[num.bins] *
                      phi((log(u[num.bins]) - mu) / sigma) ) /
                     (1 - phi((log(u[num.bins]) - mu) / sigma))
F.lo.up[num.bins+1] <- 1 - phi((log(u[num.bins]) - mu) / sigma)

D.bar.new[num.bins+1] <- (exp(mu.new + 0.5 * pow(sigma, 2)) -
                          D.bar.up.new[num.bins] *
                          phi((log(u.new[num.bins]) - mu.new) / sigma) ) /
                         (1 - phi((log(u.new[num.bins]) - mu.new) / sigma))
F.lo.up.new[num.bins+1] <- 1 - phi((log(u.new[num.bins]) - mu.new) / sigma)
```

Perform the integration approximation.

```
for (j in 1:(num.bins+1)) {
  tmp.p.ill[j] <- (1.0 - pow((1.0+D.bar[j]/beta.dr),(-1*alpha.dr))) *
                  F.lo.up[j]
  tmp.p.ill.new[j] <- (1.0 - pow((1.0+D.bar.new[j]/beta.dr),(-1*alpha.dr))) *
                  F.lo.up.new[j]
}
```

The probability of a positive results across all production is:

```
p.ill <- sum(tmp.p.ill[])
p.ill.new <- sum(tmp.p.ill.new[])
```

The rate parameters for the total number of illnesses are given by:

```
lambda.ill <- N.servings * p.ill
lambda.ill.new <- N.servings * p.ill.new
```

The Poisson likelihood of observing the total number illnesses per year:

```
N.ill ~ dpois(lambda.ill)
N.ill.new ~ dpois(lambda.ill.new)

lambda.avoid <- (1.0 - p.ill.new/p.ill) * lambda.ill
N.avoided ~ dpois(lambda.avoid)
```

# 6. Discussion

At this point in time the integration technique is presented as a code snippet rather than being integrated directly into **WinBUGS** as a **WBDev** shared component (Lunn 2003). The

motivation for this decision is to provide access to these techniques for users of both **WinBUGS** and **OpenBUGS**, particularly given the current shift of development efforts to **OpenBUGS** (Lunn, Spiegelhalter, Thomas, and Best 2009). Testing of the code snippet under **WinBUGS** and **OpenBUGS** produced essentially equivalent inferences.

The provided code snippet is specific to the lognormal distribution. Nevertheless, the technique relies only on the existence of a closed-form solution to the expected value of a truncated random variable. Additional distributions that satisfy this condition can be found in Johnson *et al.* (1994) and Klugman *et al.* (2004). Examples include the logistic, normal, and gamma distributions. The adaptation of this numerical integration technique to these distributions may facilitate the use of the **WinBUGS** environment in other applications.

# References

Albert I, Grenier E, Denis JB, Rousseau J (2008). "Quantitative Risk Assessment from Farm to Fork and Beyond: A Global Bayesian Approach Concerning Food-Borne Diseases." *Risk Analysis*, **28**, 557–571.

Allos BM, Moore MR, Griffin PM, Tauxe RV (2004). "Surveillance for Sporadic Foodborne Disease in the 21st Century: The FoodNet Perspective." *Clinical Infectious Diseases*, **38**, 115–120.

Black RE, Levine MM, Clements ML, Hughes TP, Blaser MJ (1988). "Experimental *Campylobacter Jejuni* Infection in Humans." *Journal of Infectious Diseases*, **157**, 472–479.

Crouch EA, Labarre D, Golden NJ, Kause JR, Dearfield KL (2009). "Application of Quantitative Microbial Risk Assessments for Estimation of Risk Management Metrics: *Clostridium perfringens* in Ready-to-Eat and Partially Cooked Meat and Poultry Products as an Example." *Journal of Food Protection*, **72**, 2151–2161.

Frey HC, Burmaster D (1999). "Methods for Characterizing Variability and Uncertainty: Comparison of Bootstrap Simulation and Likelihood-Based Approaches." *Risk Analysis*, **19**, 119–130.

Furumoto WA, Mickey R (1967). "A Mathematical Model for the Infectivity-Dilution Curve of Tobacco Mosaic Virus: Theoretical Considerations." *Virology*, **32**, 216–223.

Givens G, Hoeting JA (2005). *Computational Statistics*. John Wiley & Sons, Hoboken.

Haas CN, Rose JB, Gerba CP (1999). *Quantitative Microbial Risk Assessment*. John Wiley & Sons, New York.

Hald T, Vose D, Wegener HC, Koupeev T (2004). "A Bayesian Approach to Quantify the Contribution of Animal-Food Sources to Human Salmonellosis." *Risk Analysis*, **24**, 255–269.

Johnson NL, Kotz S, Balakrishnam N (1994). *Continuous Univariate Distributions*. 2nd edition. John Wiley & Sons, New York.

Klugman SA, Panjer HH, Willmot GE (2004). *Loss Models: From Data to Decisions*. 2nd edition. John Wiley & Sons, Hoboken.

Lammerding AM (2006). "Modelling and Risk Assessment for *Salmonella* in Meat and Poultry." *Journal of AOAC International*, **89**(2), 543–552.

Limpert E, Stahel WA, Abbt M (2001). "Log-Normal Distributions across the Sciences: Keys and Clues." *Bioscience*, **51**, 341–352.

Lunn D (2003). "**WinBUGS** Development Interface (**WBDev**)." *ISBA Bulletin*, **10**(3), 10–11.

Lunn D, Spiegelhalter DJ, Thomas A, Best NG (2009). "The **BUGS** Project: Evolution, Critique and Future Directions." *Statistics in Medicine*, **28**, 3049–3082.

Lunn D, Thomas A, Best NG, Spiegelhalter DJ (2000). "**WinBUGS** – A Bayesian Modelling Framework: Concepts, Structure, and Extensibility." *Statistics and Computing*, **10**, 325–337.

Mead PS, Slutsker L, Dietz V, McCaig LF, Bresee JS, Shapiro C, Griffin PM, Tauxe RV (1999). "Food-Related Illness and Death in the United States." *Emerging Infectious Diseases*, **5**, 607–625.

Mitzenmacher M (2003). "A Brief History of Generative Models for Power Law and Lognormal Distributions." *Internet Mathematics*, **1**, 226–251.

Parsonsa DJ, Ortona TG, D'Souza J, Mooreb A, Jonesc R, Dodd CER (2005). "A Comparison of Three Modelling Approaches for Quantitative Risk Assessment Using the Case Study of *Salmonella* spp. in Poultry Meat." *International Journal of Food Microbiology*, **98**, 35–51.

Piessens R, deDoncker Kapenga E, Uberhuber C, Kahaner D (1983). *QUADPACK: A Subroutine Package for Automatic Integration*. Springer-Verlag, New York.

Ross S (1988). *A First Course in Probabilty*. 3rd edition. Macmillan Publishing, New York.

Vose D (2008). *Risk Analysis: A Quantitative Guide*. 3rd edition. John Wiley & Sons, West Sussex.

Williams MS, Ebel ED, Vose D (2011). "Methodology for Determining the Appropriateness of a Linear Dose-Response Function." *Risk Analysis*, **31**, 345–350.

**Affiliation:**

Michael S. Williams, Eric D. Ebel
Risk Assessment Division
Food Safety and Inspection Service
United States Department of Agriculture
2150 Centre Avenue, Building B
Fort Collins, Colorado 80526, United States of America
E-mail: mike.williams@fsis.usda.gov

Jennifer A. Hoeting
Department of Statistics
Colorado State University
Fort Collins, Colorado 80526, United States of America
E-mail: jah@colostate.edu