UNIVERSITY OF CALIFORNIA
SANTA CRUZ

A NEW FRAMEWORK FOR BAYESIAN ANALYSIS OF DOSE-RESPONSE STUDIES THROUGH DEPENDENT NONPARAMETRIC MODELING FOR CATEGORICAL RESPONSES

A dissertation submitted in partial satisfaction of the
requirements for the degree of

DOCTOR OF PHILOSOPHY

in

STATISTICS AND STOCHASTIC MODELING

by

Kassandra Marie Fronczyk

September 2011

The Dissertation of Kassandra Marie Fronczyk is approved:

Professor Athanasios Kottas, Chair and Advisor

Professor David Draper

Professor Marc Mangel

Professor Abel Rodriguez

Tyrus Miller
Vice Provost and Dean of Graduate Studies

Table of Contents

List of Figures

List of Tables

Abstract

A new framework for Bayesian analysis of dose-response studies through dependent nonparametric modeling for categorical responses

by

Kassandra Marie Fronczyk

We develop a Bayesian nonparametric mixture modeling framework for replicated count responses in dose-response settings. Data from dose-response experiments typically involve features that can not be captured by standard parametric approaches. The proposed mixture models are built from dependent Dirichlet process priors to provide flexibility in the functional form of both the response distribution and the probability of positive response. The dependence of the mixing distributions is governed by the dose level. We explore this methodology with applications in developmental toxicity studies and quantal bioassay settings. In particular, we propose a collection of dependent Dirichlet process mixture models for developmental toxicology data, for which the primary objective is to determine the relationship between the level of exposure to a toxic chemical and the probability of a physiological or biochemical response, or death. These models have increasing levels of complexity to account for the different data structures encountered in developmental toxicity studies, including multicategory discrete outcomes, clustered binary and continuous outcomes, and responses from studies with pre-implantation exposure. For standard bioassay settings, we formulate a more structured version of the dependent nonparametric mixture model that incorporates monotonicity for the dose-response curve. Here, particular emphasis is placed on the key risk assessment goal of inference for the dose level that corresponds

to a specified response, including an application to the area of cytogenetic dosimetry. The proposed modeling framework yields highly flexible inference for the response distribution, for the dose-response relationship, and thus for risk assessment. The approach is extended to incorporate data with multiple, ordered classifications. The methodology is illustrated with simulated and real data, and also compared with parametric and semiparametric Bayesian approaches.

To my family, without whom none of this would be possible.

Acknowledgments

Through my years in graduate school, my family and friends have given me unconditional support - be it via phone calls, emails, and care packages, or Friday afternoon activities, nights out on the town, trips to various parts of the country, and watching sports. I owe you all, big time.

There are many professors who have shared life stories, advice, and an immense amount of knowledge that I will keep with me for years to come. You have been incredibly helpful on many levels. I fully appreciate everything that I've learned from you and all that you've done for me. I will do my best to pay it forward in my future endeavors.

Finally, I am eternally indebted to my advisor, Thanasis Kottas. Without your guidance, dedication, and constant attention to all Greek words and the dreaded sphinx, I would not be where I am today. You have pushed me to grow and excel, while keeping me grounded and aware of what lies ahead. You have given me opportunities and opened doors I never thought possible, I am lucky to have you as my mentor and my friend. With the number of projects left unfinished, untouched, and unformulated, we still have many, many, many years of collaborative work ahead.

Chapter 1

Introduction

1.1 Background and motivation

Dose-response studies consist of the exposure of humans or animals to a substance which is thought to have some effect depending on the dose level. The response of the individual is conventionally binary, either affected or not; although multicategory classifications or continuous responses are also recorded in some experiments. A number of individuals are exposed at several levels of the substance, giving a single, aggregate count at each dose level.

A generalization of the standard bioassay setting is found in the area of developmental toxicology, in which birth defects induced by toxic chemicals are investigated. In the most common experiment, at each experimental dose level, a number of pregnant laboratory animals (or dams) are exposed to the toxin after implantation and the number of prenatal deaths, the number of live pups, and the number of live malformed pups from each dam are typically recorded. Additional outcomes measured on each of the live pups may include body weight and length. A more general setting for developmental toxicity studies involves exposure before implantation

in an effort to capture the potential early effects on reproduction and development which are otherwise not manifested.

The main objective of developmental toxicity studies and bioassay studies is to examine the relationship between the level of exposure to the toxin (dose level) and the probability of response. The dose-response curve is defined by the probability of an outcome across the dose levels. Many of the traditional dose-response analyses assume that the probability of response increases as dose increases; that is, the dose-response curve is a non-decreasing function. Other inferential objectives are "inversion", where interest lies in estimation of the dose level that corresponds to a specified response rate, as well as the closely related "calibration" problem, where the aim is to infer about the dose level associated with a specified response vector. Also of interest is quantitative risk assessment, which evaluates the probability that adverse effects may occur as a result of the exposure to the substance. There are a number of probabilities and/or functions that are examined for risk assessment, including conditional probabilities of an outcome given specific conditions.

While the objectives of the studies are clear, the resulting data are a veritable gold mine of statistical challenges. The common assumption of a monotonically increasing dose-response curve, specified without unreasonable restrictions, is one of many hurdles in the bioassay setting. A further challenge in the context of developmental toxicity experiments arises from the inability of simple parametric discrete distributions to capture the non-standard features typically suggested by the data. This difficulty mainly emanates from the inherent heterogeneity in the data due to the clustering of individuals within a group and the variability of the reaction of the individuals to the toxin. Another challenging feature of developmental toxicology data is associated with the multiple related outcomes, both continuous (e.g., body weight) and discrete (e.g.,

number of malformations).

In the next section, we present data sets coinciding to the specific types of dose-response studies considered in this dissertation. In the process, more details are provided for these studies, and the proposed Bayesian nonparametric framework is motivated.

1.2 Data examples

In a Segment II developmental toxicity experiment, n_i pregnant dams are exposed to dose level, $x_i,\ i = 1, \ldots, N$. Dam $j = 1, \ldots, n_i$ at dose x_i has m_{ij} implants, of which the number of resorptions and prenatal deaths, as well as the number of live pups at birth with a certain defect are typically recorded. The prevailing data structure found in the statistical literature appears to be of the first type, where the random variables involved are the number of implants and the sum of all negative outcomes. Under this setting, each triplet of data (m, y, x), comprising the number of implants, number of negative outcomes, and dose level, defines a particular dam.

Two data sets commonly considered in the statistical literature for developmental toxicity experiments are shown in Figure 1.1. The left panel plots a data set from a toxicity study regarding the effects of the herbicide 2,4,5-trichlorophenoxiacetic (2,4,5-T) acid (Holson et al., 1991). We work with the version of the data given in Table 3 of Bowman and George (1995), where the number of combined endpoints consists of the number of resorptions and prenatal deaths, and the number of fetuses with cleft palate malformation. The experiment considers $N = 6$ doses, one control and 5 active dose groups. The number of animals per dose level ranges from 25 to 97 dams. The number of implants ranges from 1 to 21 across all dams and all dose levels, with 25th, 50th and 75th percentiles given by 10, 12 and 13, respectively. Based

on exploratory analysis, the data suggest varying departures from the Binomial model across the dose levels, indicating the need for a flexible model to capture the evolution of the response distributions over the range of dose levels.

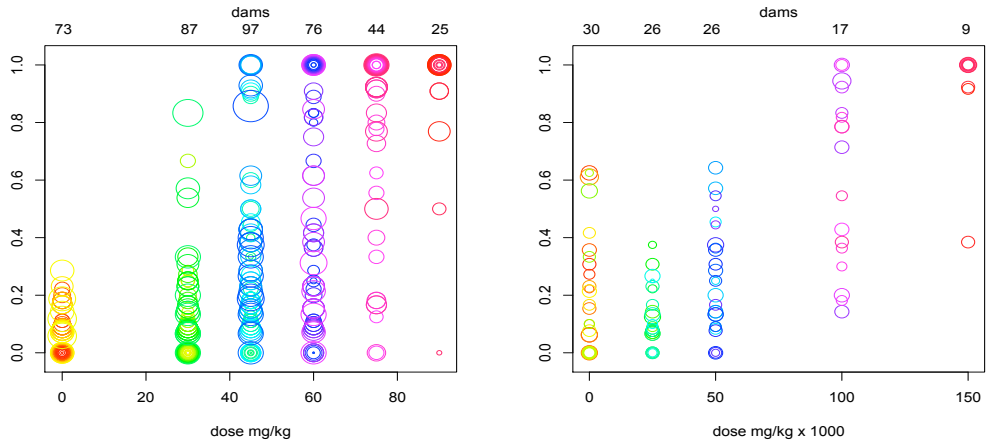

Figure 1.1: Plots of the 2,4,5-T data (left panel) and the DEHP data (right panel). Each circle corresponds to a particular dam, the size of the circle is proportional to the number of implants, and the coordinates of the circle are the dose level and the proportion of combined negative outcomes. Shown at the top of each panel is the number of animals per dose level.

The second data set (Figure 1.1, right panel) is from an experiment that explored the effects of diethylhexalphthalate (DEHP), a commonly used plasticizing agent. It is known that these plasticizers may leak in small quantities from plastic containers with various solvents such as food or milk. The possibility of toxic effects from these agents have been recognized and tested in developmental toxicity studies, such as the one described in Tyl et al. (1983). The DEHP study is also discussed by Molenberghs and Ryan (1999), although they consider a different version of the data set than the one available from the database of the National Toxicology Program (which is the version we work with). Here, the combined endpoints include resorption, prenatal death, and malformation of a live fetus (external, visceral, or skeletal malformation).

The number of dams per dose level is about a third of those found in the 2,4,5-T data; the number of implants across all dams and dose levels ranges from 4 to 18, with 25th, 50th, and 75th percentiles equal to 11, 13, and 14, respectively.

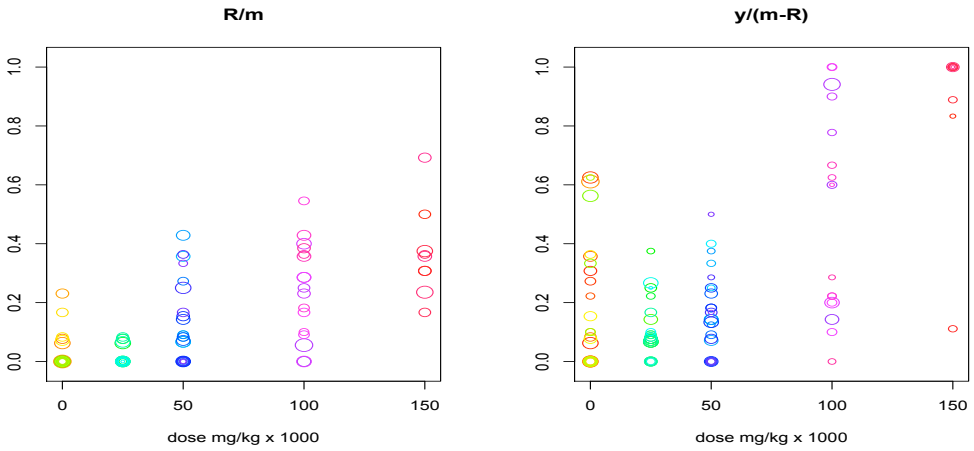

Figure 1.2: DEHP data expanded to investigate multiple endpoints. In the left panel, the proportion of non-viable fetuses of the implants for each dam at each dose level, and, on the right, the proportion of malformed pups of the viable fetuses for each dam at each dose level.

Particularly noteworthy in the DEHP data is the drop in the proportions of combined negative outcomes from dose 0 to 25 mg/kg $\times 1000$, which may indicate a hormetic dose-response relationship. Hormesis refers to a dose-response phenomenon characterized by favorable biological responses to low exposures to toxins, and thus by opposite effects in small and large doses. For endpoints involving disease incidence (e.g., mutation, birth defects, cancer), hormesis results in a J-shaped dose-response curve. Although the possibility of different low dose effects is accepted, the suggestion of positive low dose effect is debated, hence, hormesis is a controversial concept in the toxicological sciences (e.g., Calabrese, 2005). Notwithstanding the ultimate scientific conclusions, to be able to uncover non-standard dose-response relationships, we seek a

modeling framework for the dose-dependent response distributions which enables flexible inference for the implied, possibly non-monotonic, dose-response curve.

One extension of this data structure to be explored is the expansion of the collapsed category of "negative outcomes," i.e., the practically important setting with multicategory responses. Figure 1.2 gives for the DEHP data, the proportions of the implants which are considered non-viable (left panel), and among the viable fetuses, the proportion of malformed pups (right panel). Note the proportion of prenatal deaths (a dose-response curve of interest) suggests an increasing trend, while the proportion of malformed, viable pups (another dose-response curve) has a non-monotonic structure at the smaller dose levels.

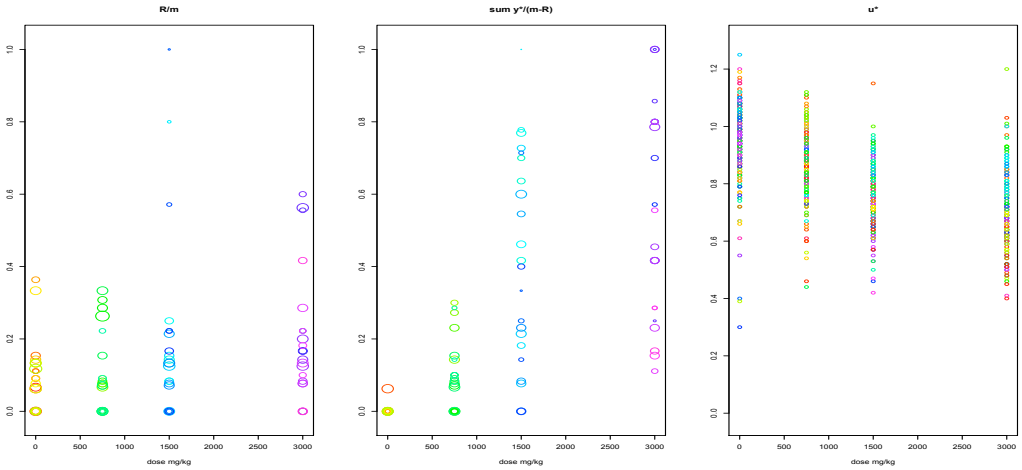

Figure 1.3: EG data. In the left panel, the proportion of non-viable fetuses of the implants for each dam at each dose level, in the middle the proportion of malformed pups of the viable fetuses for each dam at each dose level, and on the right the birth weights of the viable pups at each dose level.

Another scientifically relevant extension includes a continuous outcome for each of the viable pups, such as birth weight. In a study found at the National Toxicology Program database, ethylene glycol (EG), a widely used industrial chemical, is evaluated

for toxic effects in pregnant mice (Price et al., 1985). Recorded from each dam are the number of implants, the number of prenatal deaths, the number of live, malformed pups, and the birth weight of the viable pups (see Figure 1.3). There are four dose levels, three active (750, 1500, and 3000 mg/kg) and one control, with approximately the same number of animals exposed at each level of the toxin, ranging from 22 to 25. The number of implants across all dams and dose levels spans from 3 to 19, with 25th, 50th, and 75th percentiles of 11, 13, and 15, respectively.

We also investigate modeling for studies in which the dams are exposed to the toxin before implantation, and therefore, the early reproductive processes are expected to be interrupted. In particular, we explore a dominant lethal assay experiment (Lüning et al., 1966) where the number of dose levels is three (the control and two active dose levels). The dose levels are evenly spaced (0, 300, and 600 Rads) and there are about six times the number of animals at each dose level compared to the previous studies, ranging from 486 to 683 dams.

Turning to the bioassay setting, n_i subjects are exposed to dose level x_i, $i = 1, \ldots, N$, and y_i subjects respond in some fashion. These observations may be reported as quantal, as ordinal, or as continuous data, describing some response of the subject such as enzyme activity, heart rate, or death. We will mainly focus on quantal data, illustrating the methods through examples found in the literature. For instance, an interesting example is the trypanosome data (Ashford and Walker, 1972); see the left panel of Figure 1.4. The plot suggests skewness and multiple modes in the tolerance distribution, where traditional parametric models will fail.

This setting also lends itself to cytogenetic dosimetry, the particular area of dose-response modeling concerned with the relationship between exposure to radiation and some measure of genetic aberration. For example, in an experiment, a portion

of the full set found in Madruga et al. (1996), blood samples from individuals were exposed in vitro to ^{60}Co radiation with doses of 0, 20, 50, 100, 200, 300, 400, and 500 cGy (centograms). The cultures were analyzed for the presence of binucleated cells with none, one, or two or more micronuclei (MN). For the purposes of this illustration, we collapse these groups to none or one or more MN. The right panel of Figure 1.4 gives the proportions of one or more micronuclei from those blood samples from the older healthy subjects, though further illustration with the trinary responses on the ordinal scale is provided in Chapter 4. As seen in Figure 1.4, an important distinction of cytogenetic dosimetry data is that only a portion of the dose-response curve is observed, rendering extrapolation beyond observed dose levels a primary inferential target. Also essential is calibration, inference for unknown exposures given observed responses.

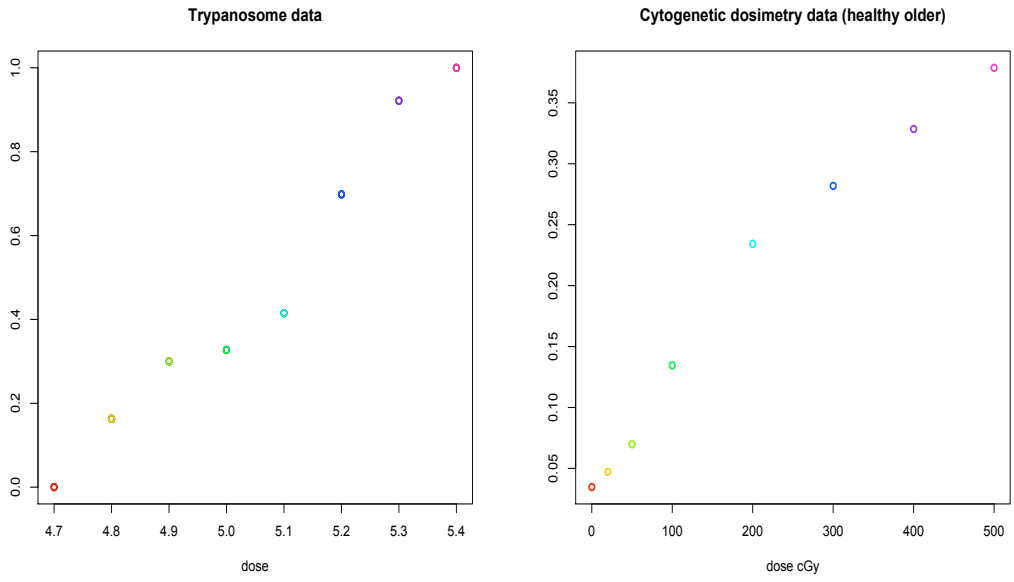

Figure 1.4: Quantal response setting: Trypanosome bioassay data (left) and the cytogenetic dosimetry data (right). Each data point represents the proportion of responses at each dose level.

1.3 Objectives and contributions

Experiments falling under developmental toxicology and the bioassay setting contain intricate processes for which parametric models may be too restrictive and will likely provide unreasonable inferences. Regarding modeling for the dose-dependent response distributions, most commonly found in the toxicology literature is the beta-binomial model, used to compensate for the extra variability in the data. With the continuous nature of the mixture model, each animal has its own probability of response, which requires immense variation to counteract the restriction of the binomial response distribution. This, in turn, produces a dose-response curve which has very wide uncertainty bands and renders the inference virtually untenable (see Section 3.1.3 for more details and an illustration).

In lieu of the insufficiency of standard parametric models, the statistics literature is riddled with approaches to flexibly model, specifically, the dose-response curve. In terms of developmental toxicity studies, many authors assume a parametric response distribution and focus on innovative determination of the dose-response curve. However, due to the various sources of heterogeneity, data from many studies indicate vast departures from parametric models. A different line of research has focused on classical semiparametric or likelihood estimation for the joint distribution of the vector of binary responses associated with each dam under the assumption of exchangeability. Although such approaches provide more general modeling for the response distribution than traditional parametric models, dose-response relationships are still introduced through parametric forms. Moreover, inferential challenges include interpolation at unobserved dose levels (a key objective for risk assessment) as well as uncertainty quantification for point estimates.

By comparison to likelihood and classical semiparametric approaches, Bayesian methods have not been widely used for the analysis of developmental toxicity studies. Examples of parametric Bayesian hierarchical models, generally comprising joint discrete-continuous outcomes, are found sparingly throughout the literature. In terms of semiparametric approaches for developmental toxicity studies, the only model available is presented in Dominici and Parmigiani (2001) (to be discussed in Section 3.1.4). The approach does attempt to capture the complexity in the response distributions, however the induced dose-response curve produces arguably unreasonable uncertainty quantification. Bayesian nonparametric methods have been applied extensively to the quantal bioassay setting. In fact, many of the pioneering works in the area were driven by experiments of this type. As motivated by the data examples and the discussion thereafter, we espouse the use of Bayesian nonparametric methods, to be elucidated in Chapter 2.

For developmental toxicity studies, we introduce a new approach to inference and risk assessment based on nonparametric modeling for collections of dose-dependent response distributions, which allows functionals of interest to inherit the flexibility of the distributions. We provide a comprehensive framework built upon nonparametric mixture models, resulting in flexibility in both the collection of response distributions as well as the probability of response at any given dose level. The dependence of the distributions is governed by the dose level, implying that distributions corresponding to nearby dose levels are more closely related than those far apart. The mixture models provide a means to appropriately quantify the uncertainty and variability in the response distributions, which carries over to the dose-response curve. The assumptions of the mixture models bestow a foundation for interpolation and extrapolation of the dose-response curve at unobserved dose levels. Note also that the dams are labeled and

recorded in ascending numerical order across dose levels; that is, the smallest ID number corresponds to data from the first dam at the first dose level, the first dam at the second dose level has the next ID number, and so on. Therefore, the animals can be linked as a response vector across the dose levels, and the replicated response vectors are considered to be exchangeable. Presented in Chapter 3 are the proposed models for the different types of developmental toxicology data structures, including methodological details, posterior inference techniques, simulation studies, comparison, and illustrative data examples.

The approach to the analysis of the bioassay setting is also built upon modeling dose-dependent response distributions. However, because of the simplified structure of the data, this is equivalent to modeling the dose-response curve with a mixture model. Contrary to the models for developmental toxicity studies, we adopt a more structured version of the prior mixture model, which begets a monotonicity restriction for the dose-response curve. Another advantage of the restricted mixture model is the facilitation of full exchangeability for the binary responses. That is, each binary response, an expansion of the aggregated count, is regarded as exchangeable both across and within each dose level. Also of interest is a framework for experiments with a single response from each subject measured on an ordinal scale, as in the cytogenetic dosimetry application. Modeling details, discussion, and examples are found in Chapter 4.

Chapter 2

Bayesian nonparametric mixture models

Nonparametric Bayesian mixture priors offer more flexible modeling tools that can capture complexity inherent in data. These models can be viewed as extensions of finite mixture or continuous mixture models, where the random mixing distribution is not defined with a particular parametric family of distributions. Countable mixture models built from Dirichlet process (DP) priors (Ferguson, 1973) for the mixing distribution are the most widely used alternative to finite mixture models.

2.1 Dirichlet process mixture models

The DP is a random probability measure which defines a prior for random distributions, or equivalently, random distribution functions. We will use $\mathrm{DP}(\alpha, G_0)$ to denote the DP prior for random mixing distribution G, defined in terms of a parametric centering (base) distribution G_0 (thus, $\mathrm{E}(G) = G_0$), and precision parameter $\alpha > 0$, which controls the variability of G about G_0 with larger values of α resulting in realizations G that are closer to G_0. Using its constructive definition (Sethuraman, 1994), the DP prior generates countable mixtures of point masses with locations drawn

from the base distribution and weights defined by a stick-breaking process. Specifically, a random distribution, G, drawn from $\mathrm{DP}(\alpha, G_0)$ has an almost sure representation as

$$G(\cdot) = \sum_{l=1}^{\infty} \omega_l \delta_{\eta_l}(\cdot), \tag{2.1}$$

where δ_a denotes a point mass at a, the η_l are i.i.d. from G_0, and $\omega_1 = \zeta_1$, $\omega_l = \zeta_l \prod_{r=1}^{l-1}(1-\zeta_r)$ for $l \geq 2$, with ζ_l i.i.d. from a Beta(1,α) distribution (independently of the η_l).

The DP generates flexible, albeit discrete distributions. While it can be useful in categorical data modeling (e.g., Cifarelli and Regazzini, 1978; Carota and Parmigiani, 2002) or specialized modeling of cumulative distribution functions as in dose response modeling (e.g., Antoniak, 1974; Bhattacharya, 1981; Disch, 1981; Kuo, 1983, 1988; Gelfand and Kuo, 1991; Mukhopadhyay, 2000; Kottas et al., 2002), it shines as a prior in mixture models.

Mixture models are robust alternatives to standard parametric models. Continuous mixtures, like the beta-binomial or poisson-gamma models, increase heterogeneity but are generally limited to symmetry and unimodal distributions. While finite mixtures are more flexible, intuitive, and, feasible to implement, they include large numbers of parameters and require prior specification of the number of components. Assuming a random mixing distribution with, say, a DP prior, the model is not restricted to a specific parametric family, and therefore symmetric, unimodal distributions, and the number of components that are necessary for a given data set are decided in a more automatic fashion.

The Dirichlet process mixture density or probability mass function is given as $f(\cdot; G) = \int k(\cdot; \boldsymbol{\theta}) \mathrm{d}G(\boldsymbol{\theta})$, where $k(\cdot; \boldsymbol{\theta})$ is a parametric family of density (probability

mass) functions indexed by parameters $\boldsymbol{\theta}$. The mixture density $f(\cdot; G)$ is random, due to the fact that G is random, and can be discrete or continuous, univariate or multivariate, depending on $k(\cdot; \boldsymbol{\theta})$. The discreteness of G, implied by its DP prior, is a key feature as, given the data, it enables flexible shapes for the mixture density $f(\cdot; G)$ through data-driven clustering of the mixing parameters associated with each observation. These mixtures are valuable not only for density estimation, but in inference for regression through DP mixture modeling for the joint-response-covariate distribution (e.g., Müller et al., 1996; Rodriguez et al., 2009; Taddy and Kottas, 2010; Fronczyk et al., 2011), as well as in more general nonparametric models for collections of distributions related through covariates, time, or space.

2.2 Dependent Dirichlet process priors

Consider data indexed by a continuous covariate, $x \in \mathcal{X}$, which is represented by the dose level in the models developed in this dissertation. This situation gives rise to the need to extend the DP prior to a prior which relates distributions across dose level to varying degrees. Many models restrict the collection of distributions to either be independent realizations of the DP or equal to a single realization of the DP, possibly with some hyperparameters that change across the covariate space, e.g., $\mathcal{X} \subseteq \mathcal{R}^+$ or $\mathcal{R}$ for dose-response settings. The idea is to model dependent, but not identical groups of distributions. One solution to this modeling problem is the dependent Dirichlet process (DDP) prior. Reported in MacEachern (1999, 2000), the DDP prior is built upon the constructive definition of the DP in (2.1), where the locations of the DP are replaced by a sample path from a stochastic process, denoted by $\eta_{l\mathcal{X}} = \{\eta_l(x) : x \in \mathcal{X}\}$, and the underlying beta draws driving the weights are also replaced by a process, $\zeta_{l\mathcal{X}}$. We

assume $\eta_{l\mathcal{X}} = h(\theta_{l\mathcal{X}})$ with $\theta_{l\mathcal{X}}$ a realization from a Gaussian process (GP) and $h(\cdot)$ an appropriate transformation, and that the $\zeta_{l\mathcal{X}}$ are defined by a realization from an independent GP with an inverse Beta$(1, \alpha_x)$ CDF transform. That is, the DDP$(\alpha, G_{0\mathcal{X}})$ is a prior given by $G_{\mathcal{X}}(\cdot) = \sum_{l=1}^{\infty} \omega_{l\mathcal{X}} \delta_{\eta_{l\mathcal{X}}}(\cdot)$, where the $\omega_{l\mathcal{X}}$ arise from the stick-breaking process using the $\zeta_{l\mathcal{X}}$ and the atoms are realizations from a possibly transformed GP.

Shown in MacEachern (2000), the DDP prior has full support and a useful continuity property which allows smooth evolution of the distributions G_x across the covariate space. This property is explored in Chapter 3 in the context of the specific models. Also found within the DDP manuscript are various properties assuming the weights do not change across the covariate, a version of the DDP prior to be discussed in further detail below.

We turn to the most natural simplification of the DDP construction, where only the atoms vary across the dose levels with common weights at all covariate levels. That is, the constructive definition becomes

$$G_{\mathcal{X}}(\cdot) = \sum_{l=1}^{\infty} \omega_l \delta_{\eta_{l\mathcal{X}}}(\cdot), \tag{2.2}$$

where the w_l arise from the standard DP stick-breaking process and the $\eta_{l\mathcal{X}}$ are sample paths from a (transformed) GP. The restriction implies the model can be viewed as a countable mixture of stochastic processes, with weights matching those from a single DP model. This structure allows the application of computational strategies for Dirichlet process mixture models, while preserving the great flexibility of the DDP. A key feature of this DDP prior is that for any finite collection of covariate levels $(x_1, ..., x_k)$ it induces a multivariate DP prior for the corresponding collection of mixing distributions $(G_{x_1}, ..., G_{x_k})$. This prior structure is referred to as the "single-p" DDP

prior (MacEachern, 2000; DeIorio et al., 2004; Gelfand et al., 2005; Rodriguez and ter Horst, 2008; Kottas and Krnjajić, 2009).

Contrarily, we may choose to simplify the DDP prior by allowing only the weights to evolve with x and assume constant atoms. In the general definition, the stick-breaking weights are driven by a Gaussian process with the Beta inverse transform. This can be difficult to implement, and therefore authors have resorted to defining the weights by retaining the stick-breaking weights and multiplying by some function of the covariate to create changes across the space (see Reich and Fuentes, 2007; Dunson and Park, 2008). Several other constructions have been proposed in the recent literature, including Fuentes-García et al. (2009), Taddy (2010), and Rodriguez and Dunson (2011). In these constructions, the marginal DP property is lost. Also, for kernel stick-breaking priors, the weights depend heavily on the assumed function of the covariate and the specific choice is not always apparent.

When the collections of random distributions are constructed with weights that change across the covariate and the atoms are common to all covariate levels, we call this the "single-θ" DDP; that is, $G_{\mathcal{X}} = \sum_{l=1}^{\infty} w_{l\mathcal{X}} \delta_{\theta_l}$. This formulation presents a formidable complication with regard to the main inferential objectives for dose-response experiments. The dose-response relationship is an important goal and, in general, it is assumed that the probability of a response increases as the dose level increases. As discussed in Chapter 3, the proposed DDP mixture models may not force a monotonically increasing relationship, however, we can incorporate monotonicity in prior expectation. This is imperative in terms of prediction at unobserved dose levels to anchor the inference with an increasing trend. In the case of the single-θ DDP prior, there is no means to force such a trend across the dose levels for any functional. As the weights and atoms are independent of each other, the expectation of a functional of

the mixture becomes the infinite sum of the expectation of the dose-dependent weights multiplied by the functional evaluated at each atom. By definition, the sum of the expectation of the weights is equal to one. What remains is the sum of the functional evaluated at the atoms, and therefore the expectation of the functional is free of the dose level.

While the DDP prior in full generality can be applied in models for a vast array of experimental designs and has desirable properties, it can be complicated to implement and may require large data sets to sufficiently learn about parameters. In our nonparametric mixture model formulations, the choice of the single-p version of the DDP prior strikes a good balance between model flexibility and computational feasibility. Data from the experiments of interest have on the order of 5-10 dose levels and are not likely to indicate drastic changes in distributional shapes between nearby dose levels. The small number of dose levels may be problematic for learning about the parameters in more general models where the DDP prior weights are also dose dependent.

Ergo, the general DDP prior, including covariate dependent weights and atoms, may struggle in these settings due to the generally limited data. The inability to induce a trend *a priori* in the dose-response curve terminates the single-θ version. As the methodological details and data examples will demonstrate, the single-p DDP prior mixture model is sufficiently flexible to capture the dependence structure of the distributions across dose levels, while remaining interpretable and manageable to implement.

In fact, the single-p DDP mixture framework can be simplified even further to accommodate special requirements for experimental data. This is fitting for the types of bioassay experiments studied in Chapter 4, where monotonicity for the dose-

response curves is a standard assumption. Under particular choices for the mixture kernel, we may choose to reduce the centering GP of the prior to a linear function of the dose levels and mix over the regression coefficients. Specifically, setting $\eta_l(x) = \gamma_{0l} + \gamma_{1l}x$ in (2.2) yields the linear DDP prior (e.g., DeIorio et al., 2009). Here, the $(\gamma_{0l}, \gamma_{1l})$ are i.i.d. from a centering distribution given hyperparameters, typically, with independent components. The lessening of the link between the mixing distributions supports implementation details which are straightforward. With respect to inference for bioassay studies, provided a restriction on the γ_{1l}, the linear-DDP models presented in Chapter 4 imbue, with probability 1, a non-decreasing characteristic on the dose-response curves.

The following section showcases the practicality of the single-p DDP prior in a generic mixture model setting. Specifically, the details within the section build the foundation for the Markov chain Monte Carlo (MCMC) framework utilized for the models in Chapter 3.

2.3 Implementation for single-p DDP mixture models

For demonstration of the implementation details, consider estimation of a collection of densities indexed by a single covariate, x. For all the models disclosed in this dissertation, the covariate is dose level. Given response y, we assume a prior model of

$$f(y \mid G_{\mathcal{X}}) = \int k(y;\theta))\mathrm{d}G_{\mathcal{X}}(\theta), \quad G_{\mathcal{X}} \sim \mathrm{DDP}(\alpha, G_{0\mathcal{X}}).$$

The base stochastic process, $G_{0\mathcal{X}}$, is a univariate Gaussian process for illustrative purposes. To model more complex data structures, we may need multiple mixing parameters and, consequently, $G_{0\mathcal{X}}$ will increase in dimension accordingly.

There are multiple avenues to traverse by way of implementation of DP mixture models. One popular route, especially in the early literature, takes advantage of the DP Pólya urn representation (e.g., Escobar and West, 1995; Bush and MacEachern, 1996; MacEachern and Müller, 1998; Neal, 2000). Alternatively, we choose to truncate the infinite dimensional DP from the outset and use the blocked Gibbs sampler (Ishwaran and Zarepour, 2000; Ishwaran and James, 2001) to exploit the ready implementation and the ease with which it can handle unbalanced replicates.

The responses, y, are observed at N distinct levels of the covariate with n_i replicates at each level. To connect with the data structures considered in Chapter 3, we assume these response replicates arise as vectors across the range of observed x. That is, let $\boldsymbol{y}_j = (y_{1j}, \ldots, y_{Nj})$ be the j-th response vector, for $j = 1, \ldots, n$. To account for possible unbalanced designs, we introduce missing value indicators, s_{ij}, such that $s_{ij} = 1$ if the jth replicate at dose level i is present and $s_{ij} = 0$ otherwise. We denote the mixing parameters for the jth replicate as $\boldsymbol{\theta}_j = (\theta_j(x_1), \ldots, \theta_j(x_N))$. Hence, the first stage of the hierarchical model for data can be expressed as

$$\{y_{ij}\} \mid \{\boldsymbol{\theta}_j\} \sim \prod_{j=1}^{n} \prod_{i=1}^{N} \{k\left(y_{ij}; \theta_j(x_i)\right)\}^{s_{ij}},$$

where the $\boldsymbol{\theta}_j$, given $G_{\boldsymbol{x}}$, are i.i.d. $G_{\boldsymbol{x}}$ for $j = 1, \ldots, n$, and $G_{\boldsymbol{x}}$ has a $\mathrm{DP}(\alpha, G_{0\boldsymbol{x}})$ prior induced by the DDP prior for $G_{\mathcal{X}}$. Resultantly, $G_{0\boldsymbol{x}}$ is the N-variate normal distribution implied by the GP, with mean $\boldsymbol{\mu}$ and covariance matrix Σ. We will use ψ to denote the vector of $G_{0\mathcal{X}}$ hyperparameters.

Note that the hierarchical model for the data is a DP mixture model induced by the DDP mixture prior. For MCMC posterior simulation, we will use blocked Gibbs sampling which is based on truncation of $G_{\boldsymbol{x}}$, induced by a finite truncation

approximation to $G_{\mathcal{X}}$. Truncating $G_{\mathcal{X}}$ at a sufficiently large level L, the model includes $G_{\mathcal{X}}^L = \sum_{l=1}^L p_l \delta_{Z_{l\mathcal{X}}}$, where the $Z_{l\mathcal{X}} = \{Z_l(x) : x \in \mathcal{X}\}$ are independent realizations, given ψ, from $G_{0\mathcal{X}}$, and the weights p_l arise from a truncated version of the stick-breaking construction: $p_1 = V_1$, $p_l = V_l \prod_{r=1}^{l-1}(1 - V_r)$, $l = 2, \ldots, L-1$, and $p_L = 1 - \sum_{l=1}^{L-1} p_l$, with the V_l independent, given α, from Beta$(1, \alpha)$. The joint distribution for the set of L random weights is given by a special case of the generalized Dirichlet distribution (Connor and Mosimann, 1969),

$$f(\boldsymbol{p} \mid \alpha) = \alpha^{L-1} p_L^{\alpha-1}(1 - p_1)^{-1}(1 - (p_1 + p_2))^{-1} \times \cdots \times (1 - \sum_{l=1}^{L-2} p_l)^{-1}.$$

Moreover, $Z_l(\boldsymbol{x}) = (Z_l(x_1), \ldots, Z_l(x_N)) \equiv \boldsymbol{Z}_l$, with the $\boldsymbol{Z}_l$ i.i.d., given ψ, from $G_{0\boldsymbol{x}}$, for $l = 1, ..., L$. Introducing configuration variables $\boldsymbol{w} = (w_1, \ldots, w_n)$, where each w_j takes a value in $\{1, \ldots, L\}$, we have $w_j = l$ if and only if $\boldsymbol{\theta}_j = \boldsymbol{Z}_l$, for $j = 1, \ldots, n$ and $l = 1, \ldots, L$. Therefore, the hierarchical model is given as follows

$$\begin{aligned}
\{y_{ij}\} \mid \{\boldsymbol{Z}_l\}, \boldsymbol{w} &\sim \prod_{j=1}^{n} \prod_{i=1}^{N} \{k(y_{ij}; Z_{w_j}(x_i))\}^{s_{ij}} \\
\{w_j\} \mid \boldsymbol{p} &\sim \prod_{j=1}^{n} \sum_{l=1}^{L} p_l \delta_l(w_j) \\
\boldsymbol{p} \mid \alpha &\sim f(\boldsymbol{p} \mid \alpha) \\
\{\boldsymbol{Z}_l\} \mid \psi &\sim \prod_{l=1}^{L} G_{0\boldsymbol{x}}(\boldsymbol{Z}_l; \psi)
\end{aligned} \tag{2.3}$$

and is completed with hyperpriors for α and ψ.

2.3.1 Prior specification

While prior specification is model specific, there are some commonalities; namely, the approach for the prior given to α and the value of the truncation level are related and discussed here. The DDP prior precision parameter, α, controls the number, $n^* \leq n$, of distinct mixture components (e.g., Antoniak, 1974; Escobar and West, 1995). In particular, for moderate to large sample sizes, a useful approximation to the prior expectation $\mathrm{E}(n^* \mid \alpha)$ is given by $\alpha \log\{(\alpha+n)/\alpha\}$. This expression can be averaged over the gamma(a_α, b_α) prior to be placed on α to obtain $\mathrm{E}(n^*)$, thus selecting a_α and b_α to agree with a prior guess at the expected number of distinct mixture components.

Additionally, the truncation level L for the DDP prior approximation can be chosen using standard distributional properties for the weights arising from the stick-breaking structure. For instance, $\mathrm{E}(\sum_{l=1}^{L} \omega_l \mid \alpha) = 1 - \{\alpha/(\alpha+1)\}^L$, which can be averaged over the prior for α to estimate $\mathrm{E}(\sum_{l=1}^{L} \omega_l)$. Given a specified tolerance level for the approximation, this expression is solved numerically to obtain the corresponding value L.

2.3.2 Implementation and posterior inference

Next, we provide an outline of the methods for MCMC posterior simulation and predictive inference. For notational purposes, we denote the n^* distinct values of vector $\boldsymbol{w}$ by $w_1^*, \ldots, w_{n^*}^*$, $M_k^* = |\{j : w_j = w_k^*\}|$, $k = 1, \ldots, n^*$, and $M_l = |\{w_j : w_j = l\}|$, $l = 1, \ldots, L$.

Updating the mixing parameters, $\{\boldsymbol{Z}_l\}$, depends on the value of l. If $l \notin \{w_k^* : k = 1, \ldots, n^*\}$, then $\boldsymbol{Z}_l$ arises from the prior multivariate normal distribution induced

by the GP, $G_{0\boldsymbol{x}} = \mathrm{N}_N(\boldsymbol{\mu}, \Sigma)$. If $l \in \{w_k^* : k = 1, \ldots, n^*\}$, then $\boldsymbol{Z}_l$ is drawn from

$$\boldsymbol{Z}_{w_k^*} \mid \psi, \text{data} \propto \mathrm{N}_N(\boldsymbol{Z}_{w_k^*}; \boldsymbol{\mu}, \Sigma)) \prod_{\{j: w_j = w_k^*\}} \prod_{i=1}^{N} \{k(y_{ij}; Z_{w_k^*}(x_i))\}^{s_{ij}}.$$

This requires a Metropolis-Hastings step for updating, unless $k(\cdot)$ is the normal distribution and the data have balanced replicates. Further attention is given to these updates in Chapter 3, where the proposed mixture models are built from discrete or combined discrete-continuous kernel distributions.

The new draws for w_j come from a discrete distribution given by $\sum_{l=1}^{L} \tilde{p}_{lj} \delta_l(\cdot)$, where $\tilde{p}_{lj}$ is proportional to $p_l \prod_{i=1}^{N} \{k(y_{ij}; Z_l(x_i))\}^{s_{ij}}$.

The conditional posterior of $\boldsymbol{p}$ is given by a generalized Dirichlet distribution with parameters $(M_1 + 1, \ldots, M_{L-1} + 1)$ and $(\alpha + \sum_{k=2}^{L} M_k, \ldots, \alpha + M_L)$ (Ishwaran and James, 2001). This distribution is not easy to work with directly, but vector $\boldsymbol{p}$ can be generated through latent $V_1^*, \ldots, V_{L-1}^*$ by setting $p_1 = V_1^*$, $p_l = V_l^* \prod_{r=1}^{l-1}(1 - V_r^*)$, $l = 2, \ldots, L-1$, and $p_L = 1 - \sum_{l=1}^{L-1} p_l$, where the V_l^* are independent from $\mathrm{Beta}(M_l + 1, \alpha + \sum_{k=l+1}^{L} M_k)$. The full conditional for α is a gamma distribution with shape parameter $L + a_\alpha - 1$ and rate parameter $b_\alpha - \log p_L = b_\alpha - \sum_{l=1}^{L-1} \log(1 - V_l^*)$. Finally, the GP mean and covariance function drive the updates for ψ. Specific details are provided in the context of the models of Chapter 3.

Regarding posterior predictive inference and inference for functionals of the mixture, consider new responses y_0 corresponding to a generic x_0. The approach to estimating predictive densities or mixture functionals depends on whether x_0 is among the observed covariate levels or a new level. Inference for observed x_0 uses the samples of $\boldsymbol{Z}$ and $\boldsymbol{p}$, which define the mixing distribution at all the observed x values. To

interpolate and extrapolate at unobserved covariate levels, $\tilde{\boldsymbol{x}}_0 = (\tilde{x}_{01}, \ldots, \tilde{x}_{0M})$, we need the predicted sample paths of the Gaussian process realizations. The mixing parameter vector is extended to $(\boldsymbol{Z}_l, \tilde{\boldsymbol{Z}}_l)$, for $l = 1, \ldots, L$, where given model parameters and the $\boldsymbol{Z}_l$, the new $\tilde{\boldsymbol{Z}}_l$ are obtained through standard conditioning under multivariate normal distributions. Again, the form of the mean and covariance matrix of the conditional normal distributions relies on the assumptions of the Gaussian process mean and covariance function.

We have described the main details of the MCMC algorithm to obtain samples from the posterior distribution of a single-p DDP mixture model. Each specific model in Chapter 3 and 4 describes briefly the distinguishing updates, as well as the inference methods for key risk assessment functionals of the mixture models.

2.4 Discussion

Bayesian nonparametrics and the Dirichlet process provide flexible priors for mixture models. The resulting models can capture both standard and non-standard densities and functionals, while remaining easy to implement and interpret. Based on these attributes, we move forward with the development of the modeling frameworks for the various types of dose-response data studied in this dissertation.

Chapter 3

Bayesian nonparametric framework for developmental toxicity studies

This chapter provides details for the different developmental toxicology settings: the Segment II studies in Section 3.1 for combined negative outcomes and the multicategory responses in Section 3.2; studies including a continuous response for the live pups in Section 3.3; experiments where the animals are exposed to the toxin prior to implantation in Section 3.4.

3.1 Segment II studies: combined negative outcomes

A variety of approaches for the analysis of Segment II developmental toxicity studies have been suggested in the statistical literature. Modeling approaches based on standard parametric response distributions and/or customary parametric forms for dose-response curves include Chen et al. (1991), Catalano and Ryan (1992), Ryan (1992), Zhu et al. (1994), and Regan and Catalano (1999). However, due to the various sources of heterogeneity, data from many studies indicate vast departures from

parametric models. A different line of research has focused on classical semiparametric or likelihood estimation for the joint distribution of the vector of binary responses associated with each dam under the assumption of exchangeability (e.g., Bowman and George, 1995; George and Bowman, 1995; Kuk, 2004; Pang and Kuk, 2005). Although such approaches provide more general modeling for the response distribution than traditional parametric models, dose-response relationships are still introduced through parametric forms. Moreover, inferential challenges include interpolation at unobserved dose levels (a key objective for risk assessment) as well as uncertainty quantification for point estimates.

By comparison to likelihood and classical semiparametric approaches, Bayesian methods have not been widely used for the analysis of developmental toxicity studies. To our knowledge, the only Bayesian semiparametric model is presented by Dominici and Parmigiani (2001), using a product of mixtures of Dirichlet process prior structure.

To overcome the limitations of parametric approaches, and at the same time retain a fully inferential framework, we develop a Bayesian nonparametric mixture model that provides flexibility in both the response distribution and the dose-response relationship. We seek mixture modeling for response distributions that are related across doses with the level of dependence driven by the distance between the dose values. To this end, we consider a dependent Dirichlet process (DDP) prior for the dose-dependent mixing distributions, in particular, the single-p DDP prior structure, as introduced and motivated in Chapter 2. Inference and prediction under DDP priors require replication, which arises through the number of dams observed at each dose level. The replicated count responses in conjunction with a nonparametric mixture of Binomial distributions, induced at each dose value by the DDP mixture model, enable flexible inference for the response distribution at any observed dose level. And,

importantly, the dependence of the DDP prior across dose levels allows data-driven prediction for collections of response distributions, as well as inference for the implied dose-response relationship, through interpolation (and extrapolation) over any range of dose values of interest. We develop properties of the DDP model that are key for the application to developmental toxicity studies. We discuss various forms of inference that are available under the model, as well as the MCMC details that deviate from those presented in Chapter 2. Traditional parametric dose-response models are shown to be special (limiting) cases of the nonparametric DDP mixture model, which, using simulated data sets, is also compared with simpler semiparametric Bayesian methods. In particular, in the context of the simulation study, we provide comparison of the semiparametric model from Dominici and Parmigiani (2001) with the proposed DDP model.

As exhibited in Section 1.2, models for the combined negative outcomes are the prevalent structure found in the statistical literature. We focus on this case inasmuch as it provides a conveniently simple setting for notation, model details, and comparison with other models. More general models which are applicable to more biologically realistic settings are established in later sections.

3.1.1 Methods

Under the Segment II toxicity study design, exposure occurs after implantation. Thus, following standard arguments from the literature (e.g., Zhu et al., 1994), we treat the number of implants, m, as a random quantity containing no information about the dose-response relationship. That is, we assume $m \mid \boldsymbol{\kappa} \sim f(m; \boldsymbol{\kappa})$, where $\boldsymbol{\kappa}$ are parameters of the implant distribution which do not depend on x. Here, we take $f(m; \boldsymbol{\kappa}) = \text{Pois}(m; \lambda)$, where we have shifted the Poisson distribution to

have support $m \geq 1$; however, more flexible distributions can be readily utilized. We focus on the data structure that involves the number of implants, m_{ij}, and the corresponding number of combined negative outcomes, y_{ij}, for dam $j = 1, \ldots, n_i$ at dose level $x_i,\ i = 1, \ldots, N$. With the assumption of an implant distribution that does not depend on dose level, the modeling for the number of implants and the number of negative outcomes is decomposed to $f(m, y) = f(m)f(y \mid m)$. Therefore, inference for the parameters of the implant distribution is carried out separately from inference for the parameters of the model for $f(y \mid m)$.

To relax the potentially restrictive assumptions imposed by standard parametric models, we propose a nonparametric mixture modeling framework for the dose-dependent conditional distribution of the number of negative outcomes given the number of implants. In other words, we seek modeling for the response distribution that allows nonparametric dependence structure across dose levels. We achieve such modeling by representing $f(y \mid m)$ as a mixture of Binomial distributions with dose-dependent mixing distribution. Placing a DDP prior on the collection of mixing distributions indexed by dose level yields the desired nonparametric prior model for the collection of dose-dependent response distributions.

More formally, we propose the following DDP prior mixture model

$$f(y \mid m; G_{\mathcal{X}}) = \int \text{Bin}\left(y; m, \frac{\exp(\theta)}{1+\exp(\theta)}\right) \mathrm{d}G_{\mathcal{X}}(\theta), \qquad G_{\mathcal{X}} \mid \alpha, \psi \sim \text{DDP}(\alpha, G_{0\mathcal{X}}) \tag{3.1}$$

where $\text{DDP}(\alpha, G_{0\mathcal{X}})$ denotes the single-p DDP prior for $G_{\mathcal{X}} = \sum_{l=1}^{\infty} \omega_l \delta_{\eta_{l\mathcal{X}}}$ with precision parameter α and base stochastic process $G_{0\mathcal{X}}$ that depends on parameters ψ; the model is implemented with hyperpriors on α and ψ. We take $G_{0\mathcal{X}}$ to be a GP with a linear mean function, constant variance, and isotropic power exponential correlation

function. Hence, for all l, $\mathrm{E}(\eta_l(x) \mid \beta_0, \beta_1) = \beta_0 + \beta_1 x$, $\mathrm{Var}(\eta_l(x) \mid \sigma^2) = \sigma^2$, and $\mathrm{Corr}(\eta_l(x), \eta_l(x') \mid \phi) = \exp(-\phi|x - x'|^d)$, with $\phi > 0$ and (fixed) $d \in [1, 2]$ (and thus $\psi = (\beta_0, \beta_1, \sigma^2, \phi)$). The DDP Binomial mixture model is completed with (independent) priors for the DDP hyperparameters, in particular, we place a gamma prior on the precision parameter α, normal prior on β_0, an inverse gamma prior on σ^2, and a uniform prior on ϕ over $(0, b_\phi)$. Moreover, a gamma prior is taken for β_1 to incorporate a non-decreasing trend in prior expectation for the dose-response curve, which is a key consideration for the modeling approach. The linear mean function enables connections with standard parametric dose-response models, which arise as limiting cases of the DDP mixture model, as discussed in the next section which also includes development of further properties of the DDP Binomial mixture model.

Model properties

Hereinafter, $\pi(u) = \exp(u)/(1 + \exp(u))$, $u \in \mathbb{R}$, will be used to denote the logistic function.

The DDP Binomial mixture model in (3.1) includes both the hierarchical Binomial-logistic-normal model and the standard Binomial-logit model as special (limiting) cases. As $\alpha \to \infty$, each response replicate has a distinct mixing parameter. If we also assume the GP for $G_{0\mathcal{X}}$ is a white noise process, (i.e., $\phi \to \infty$) we obtain the hierarchical Binomial-logistic-normal model, $y_{ij} \mid m_{ij}, \theta_{ij} \overset{ind.}{\sim} \prod_{i=1}^{N} \prod_{j=1}^{n_i} \mathrm{Bin}(y_{ij}; m_{ij}, \pi(\theta_{ij}))$, with $\theta_{ij} \mid \beta_0, \beta_1, \sigma^2 \overset{ind.}{\sim} \mathrm{N}(\beta_0 + \beta_1 x_i, \sigma^2)$. If we let $\sigma^2 \to 0^+$, we arrive at the standard Binomial-logit model as a further special limiting case. In the other extreme, as $\alpha \to 0^+$, all the response replicates are assigned to a single mixture component. This, along with the white noise process assumption, yields a Binomial-logistic normal model with a common mean for each animal within a dose

level. In other words, as $\alpha \to 0^+$, $y_{ij} \mid m_{ij}, \theta_i \overset{ind.}{\sim} \prod_{i=1}^{N} \prod_{j=1}^{n_i} \text{Bin}(y_{ij}; m_{ij}, \pi(\theta_i))$, with $\theta_i \mid \beta_0, \beta_1, \sigma^2 \overset{ind.}{\sim} \text{N}(\beta_0 + \beta_1 x_i, \sigma^2)$. Again, with the additional restriction of $\sigma^2 \to 0^+$, we obtain the standard Binomial-logit model.

The DDP prior for $G_{\mathcal{X}} = \sum_{l=1}^{\infty} \omega_l \delta_{\eta_{l\mathcal{X}}}$ allows for a flexible response distribution at each fixed level of x through a DP mixture model. Consider a realization $\theta_{\mathcal{X}} = \{\theta(x) : x \in \mathcal{X}\}$, which, given $G_{\mathcal{X}}$, arises from $G_{\mathcal{X}}$. Then, for any x, $x' \in \mathcal{X}$, $\text{Cov}(\theta(x), \theta(x') \mid G_{\mathcal{X}}) = \sum \omega_l \eta_l(x) \eta_l(x') - \{\sum \omega_l \eta_l(x)\}\{\sum \omega_l \eta_l(x')\}$. Therefore, although $G_{\mathcal{X}}$ is centered around a stationary GP, it generates nonstationary realizations with non-Gaussian finite dimensional distributions. Moreover, if G_x and $G_{x'}$ denote the marginal distributions of $\theta(x)$ and $\theta(x')$ under $G_{\mathcal{X}}$, then the continuity of the $\eta_{l\mathcal{X}}$ implies that, as the distance between x and x' gets smaller, the difference between G_x and $G_{x'}$ gets smaller; formally, for any $\varepsilon > 0$, $\lim_{|x-x'| \to 0} \Pr(\mathcal{L}(G_x, G_{x'}) < \varepsilon) = 1$, where $\mathcal{L}$ is the Lévy distance (MacEachern, 2000). Hence, the level of dependence between G_x and $G_{x'}$, and thus between $f(y \mid m; G_x)$ and $f(y \mid m; G_{x'})$, is driven by the distance of the dose levels. The practical implication is that in prediction for the inferential objectives, we learn more from covariate levels x' nearby x than from more distant levels, a desirable property for distributions that are expected to evolve relatively smoothly with the dose level.

Next, we discuss a useful connection of the DDP mixture model in (3.1), which is built from the Binomial kernel for the number of combined negative outcomes within a dam, with a DDP mixture model based on a product of Bernoullis kernel for the set of binary responses for all implants corresponding to that dam. This connection is requisite for the study of risk assessment quantities such as the dose-response curve and intra-cluster correlations. The mixture model using the underlying vector of binary responses (denoted by $\boldsymbol{y}^* = (y_1^*, \ldots, y_m^*)$) for a generic dam with number of implants m

at dose level x is given by

$$f^*(\boldsymbol{y}^* \mid m; G_{\mathcal{X}}) = \int \prod_{k=1}^{m} \text{Bern}(y_k^*; \pi(\theta)) \text{d}G_{\mathcal{X}}(\theta), \tag{3.2}$$

where the same DDP prior as before would be assigned to $G_{\mathcal{X}}$. Note that the model formulation involves a common mixing parameter for all binary outcomes associated with the same dam. Below, we show that mixture models (3.1) and (3.2) are equivalent in the sense that the moment generating function for the number of negative outcomes under (3.1) is equal to the moment generating function for the sum of binary responses under (3.2).

Connection between the DDP Binomial and DDP product of Bernoullis mixture models: The moment generating function for the number of negative outcomes under model (3.1) is given by

$$\begin{aligned}
\text{E}(e^{ty} \mid m; G_{\mathcal{X}}) &= \int \sum_{y=0}^{m} e^{ty} \, \text{Bin}(y; m, \pi(\theta)) \, \text{d}G_{\mathcal{X}}(\theta) = \int \left(\frac{1 + \exp(\theta + t)}{1 + \exp(\theta)} \right)^m \text{d}G_{\mathcal{X}}(\theta) \\
&= \int \prod_{k=1}^{m} \left(\frac{1 + \exp(\theta + t)}{1 + \exp(\theta)} \right) \text{d}G_{\mathcal{X}}(\theta) = \int \prod_{k=1}^{m} \sum_{y_k^*=0,1} e^{ty_k^*} \, \text{Bern}(y_k^*; \pi(\theta)) \, \text{d}G_{\mathcal{X}}(\theta) \\
&= \sum_{y_1^*=0,1} \cdots \sum_{y_m^*=0,1} e^{t \sum_{k=1}^{m} y_k^*} \int \prod_{k=1}^{m} \text{Bern}(y_k^*; \pi(\theta)) \text{d}G_{\mathcal{X}}(\theta) = \text{E}(e^{t \sum_{k=1}^{m} y_k^*} \mid m; G_{\mathcal{X}})
\end{aligned}$$

i.e., the moment generating function for the sum of binary responses under model (3.2).

Under the truncated mixing distributions, $G_x^L = \sum_{l=1}^{L} p_l \delta_{Z_l(x)}$, we can also determine the correlation between two binary responses within the same dam at a generic dose level x, i.e., $\text{Corr}(y_k^*, y_{k'}^*; G_x^L)$; we will refer to this as the intracluster correlation (where the dam serves as the *cluster*). The expectations needed to

obtain the intracluster correlation are $\mathrm{E}(y_k^*; G_x^L) = \mathrm{E}(y_{k'}^*; G_x^L) = \sum_{l=1}^L p_l \pi(Z_l(x))$, and $\mathrm{E}(y_k^* y_{k'}^*; G_x^L) = \sum_{l=1}^L p_l (\pi(Z_l(x)))^2$. Therefore, $\mathrm{Var}(y_k^*; G_x^L) = \mathrm{Var}(y_{k'}^*; G_x^L) = \{\sum_{l=1}^L p_l \pi(Z_l(x))\} - \{\sum_{l=1}^L p_l \pi(Z_l(x))\}^2$, and

$$\mathrm{Corr}(y_k^*, y_{k'}^*; G_x^L) = \frac{\{\sum_{l=1}^L p_l (\pi(Z_l(x)))^2\} - \{\sum_{l=1}^L p_l \pi(Z_l(x))\}^2}{\{\mathrm{Var}(y_k^*; G_x^L)\mathrm{Var}(y_{k'}^*; G_x^L)\}^{1/2}}. \tag{3.3}$$

Developmental toxicity studies typically give rise to overdispersed binary responses, that is, $\mathrm{Corr}(y_k^*, y_{k'}^*; G_x^L) > 0$. Capitalizing on results for mixtures from exponential families, it can be shown that the DDP mixture model supports positive intracluster correlations.

Positive intracluster correlation result: Consider the vector of binary responses, $\boldsymbol{y}^* = (y_1^*, \ldots, y_m^*)$, for a generic dam with m (≥ 2) implants at dose level x. Denote by $\pi^* = \Pr(y_k^* = 1; G_x) = \int \pi(\theta) \mathrm{d}G_x(\theta)$, $k = 1, ..., m$, the probability of a negative outcome, and by $\gamma = \mathrm{Corr}(y_k^*, y_{k'}^*; G_x)$, the correlation between any pair of binary outcomes within the same dam; γ is given by (3.3) under the DDP truncation approximation.

Under the implicit assumption of common π^* and γ for all binary responses within the same dam, the variance for the number of combined negative outcomes, $y = \sum_{k=1}^m y_k^*$, is given by $\mathrm{Var}(y \mid m; G_x) = m\pi^*(1-\pi^*)\{1 + (m-1)\gamma\}$. (Note that this result does not rely on the specific form of the mixture model for y in (3.1) or the equivalent model for $\mathbf{y}^*$ in (3.2).) Now, consider a random variable u, which has a Binomial distribution with the same mean as y arising from $f(y \mid m; G_x) = \int \mathrm{Bin}(y; m, \pi(\theta)) \mathrm{d}G_x(\theta)$, that is, $u \sim \mathrm{Bin}(m, \pi^*)$. Then, using overdispersion results for mixtures from exponential families (e.g., Shaked, 1980), we have $\mathrm{Var}(y \mid m; G_x) \geq \mathrm{Var}(u) = m\pi^*(1-\pi^*)$, which yields $\gamma \geq 0$. The case of $\gamma = 0$ arises only under the limiting case of the mixture model, $\alpha \to 0^+$, where the mixture reduces to the Binomial

kernel with a GP driving the probability of response.

Using four different combinations of fixed values for the DDP prior hyperparameters, Figure 3.1 gives the resulting correlation structures across dose levels. While always positive, the smooth evolution of correlations across dose level can have parabolic, oscillatory, or linear patterns.

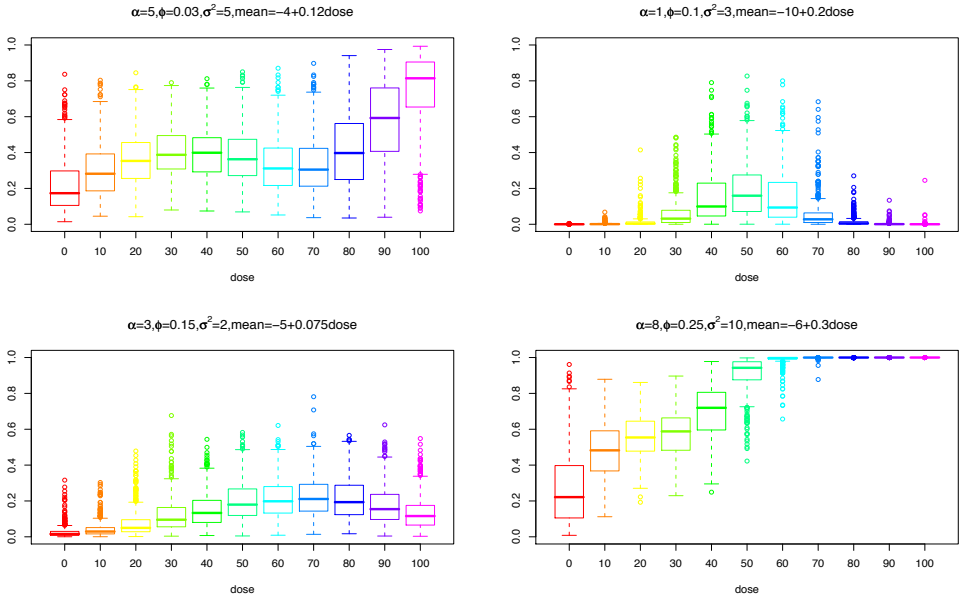

Figure 3.1: Boxplots of prior realizations for the intracluster correlation across dose levels under four different combinations of fixed values for the DDP prior hyperparameters.

Dose-response relationship

Using the mixture formulation in (3.2) for the underlying binary outcomes, we define the dose-response curve through the probability of a negative outcome for a generic implant expressed as a function of dose level. Therefore, under the DDP

truncation approximation, the dose-response curve is given by

$$\Pr(y^* = 1; G_{\mathcal{X}}^L) = \int \pi(\theta) \mathrm{d}G_{\mathcal{X}}^L(\theta) = \sum_{l=1}^{L} p_l \pi(Z_{l\mathcal{X}}). \tag{3.4}$$

Note that, although this is a conditional probability (given $m = 1$), we suppress this implicit conditioning in the notation.

Smoothness properties of prior realizations for the dose-response function emerge directly from properties of prior realizations $Z_{l\mathcal{X}}$ under the centering GP $G_{0\mathcal{X}}$. In particular, for choices of $d \in [1, 2)$ ($d = 2$) for the GP correlation function, the continuity (differentiability) of the $Z_{l\mathcal{X}}$ yields continuous (differentiable) dose-response curves under the DDP Bernoulli mixture model.

A key aspect of the model is that it does not force a monotonicity restriction to the dose-response function, which is an assumption for standard parametric dose-response models. However, the prior expectation $\mathrm{E}(\Pr(y^* = 1; G_{\mathcal{X}}^L))$ is non-decreasing in x provided $\beta_1 > 0$, as shown next.

Monotonicity of the prior expectation for the dose-response curve: Denote by $D(x)$, $x \in \mathcal{X}$, the prior expectation for the dose-response curve.

Under the DDP truncation approximation, $D(x) = \mathrm{E}(\Pr(y^* = 1; G_x^L)) = \sum_{l=1}^{L} \mathrm{E}(p_l)\mathrm{E}(\pi(Z_l(x)))$, since $\{Z_l(x) : l = 1, ..., L\}$ is independent of $\{V_l : l = 1, ..., L-1\}$, which is the collection of i.i.d. Beta$(1, \alpha)$ variables that define the p_l through stick-breaking. Therefore, for any $x < x'$, $D(x) - D(x') = \sum_{l=1}^{L} \mathrm{E}(p_l)\{\mathrm{E}(\pi(Z_l(x))) - \mathrm{E}(\pi(Z_l(x')))\}$. Now, for any $l = 1, ..., L$, random variables $Z_l(x)$ and $Z_l(x')$ follow $\mathrm{N}(\beta_0 + \beta_1 x, \sigma^2)$ and $\mathrm{N}(\beta_0 + \beta_1 x', \sigma^2)$ distributions, respectively. Hence, if $\beta_1 > 0$, $Z_l(x)$ is stochastically smaller than $Z_l(x')$, for each $l = 1, ..., L$. This in turn implies

$\mathrm{E}(\pi(Z_l(x))) \leq \mathrm{E}(\pi(Z_l(x')))$, for all $l = 1, ..., L$ (since $\pi(u)$ is an increasing function), and thus $D(x) \leq D(x')$.

The above argument establishes the result for the form in (3.4), which is the one we work with to obtain inference for the dose-response relationship. The result can also be obtained without the truncation approximation. In this case, we have $D(x) = \mathrm{E}(\Pr(y^* = 1; G_x)) = \mathrm{E}\{\int \pi(\theta) \mathrm{d}G_x(\theta)\} = \int \pi(\theta) \mathrm{d}G_{0x}(\theta)$, where $G_{0x} = \mathrm{N}(\beta_0 + \beta_1 x, \sigma^2)$. Therefore, $D(x)$ is the expectation of the (increasing) logistic function with respect to G_{0x}, which is stochastically ordered in x provided $\beta_1 > 0$, and thus $D(x)$ is a non-decreasing function of x.

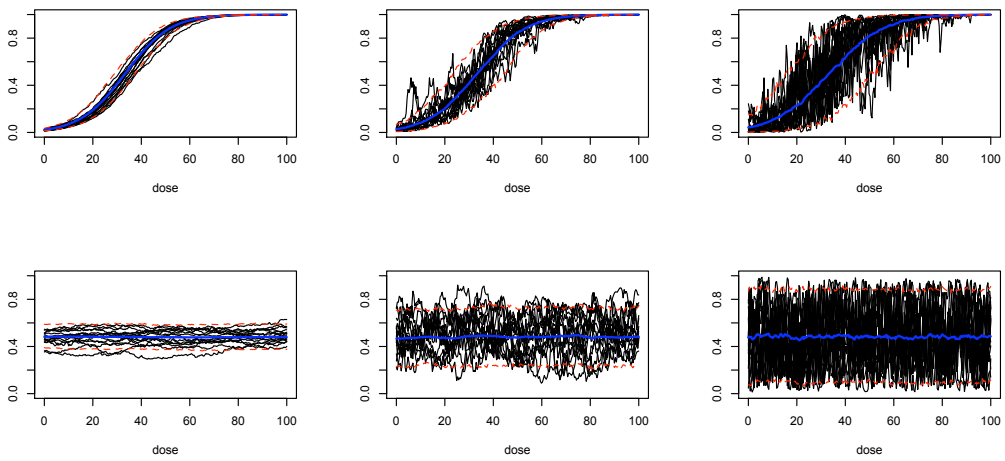

Figure 3.2: Prior realizations with mean (blue) and 90% probability bands (red) for the dose-response curve given $\beta_1 > 0$ (top row) and $\beta_1 = 0$ (bottom row). The remaining hyperparameters are fixed to values that increase the variability of the realizations moving from the left to the right column.

Including an increasing trend in prior expectation is crucial for practicable posterior inference. In particular, if the model is applied using a constant mean function for the DDP prior centering GP (i.e., setting $\beta_1 = 0$), there is little hope to obtain

meaningful interpolation and extrapolation results for the dose-response curve. Figure 3.2 shows prior realizations of the dose-response curve, evaluated using (3.4) over a grid of dose values, along with the mean and 90% interval bands. The top row has β_1 fixed at 0.12, whereas in the bottom row β_1 is fixed at zero. In the far left column, the hyperparameters are chosen such that the error bands are narrow and the individual draws are relatively smooth. Moving to the right, the hyperparameters are designated such that the variability of the realizations increases. In all cases, the top row depicts the increasing trend in prior expectation, though not necessarily for any given realization, whereas the bottom row does not show this trend.

Even though we insist on the non-decreasing trend in prior expectation, prior (and thus posterior) realizations for the dose-response curve are not structurally restricted to be non-decreasing. An illustration of this model feature is provided in Section 3.1.3, where the shape of the estimated dose-response curve for the DEHP data is indicative of a possible hormetic relationship (see discussion in Chapter 1). The fact that the DDP mixture model allows non-monotonic dose-response relationships to be uncovered is an asset of the proposed modeling approach, and, arguably, a practical advance relative to existing methods.

Prior specification

Regarding the choice of d, we have experimented with both exponential and Gaussian correlation functions ($d = 1$ and $d = 2$, respectively). For all data examples of Sections 3.1.3 and 3.1.4, inferences were largely unaffected by the particular choice. However, as numerical instabilities may arise assuming a Gaussian correlation function, we take the exponential for this analysis and the remaining analyses within the chapter. To specify b_ϕ, we consider the limiting case of the DDP mixture model with $\alpha \rightarrow$

0^+, which corresponds to a Binomial response distribution with a GP prior for the dose-response function on the logistic scale. Then, under the exponential correlation function, $3/\phi$ is the *range of dependence*, i.e., the distance between dose levels that yields correlation 0.05. The range is usually assumed to be a fraction of the maximum interpoint distance over the index space. Let $D_{\max}$ be the maximum distance between observed doses. Since $3/b_\phi < 3/\phi$, we specify b_ϕ such that $3/b_\phi = rD_{\max}$ for a small r. This approach to prior specification for ϕ is noninformative, in particular, for all data analyses considered, the posterior distribution for ϕ is concentrated on values substantially smaller than b_ϕ.

We set the means of the normal priors for the centering GP linear mean parameter, β_0, to 0, β_1 to $1/b_\phi$, and the shape parameter of the inverse gamma prior for the GP variance, σ^2, to 2 (implying infinite prior variance). The prior variances for β_0 and β_1 and the prior mean for σ^2 are chosen by studying the induced prior distribution for the dose-response curve for which prior realizations can be readily sampled using the definition in (3.4). Specifically, under the prior choice discussed below, the prior mean for $\Pr(y^* = 1; G^L_{\mathcal{X}})$ begins around 0.5 with a slight increasing trend, and the corresponding 95% interval bands are essentially spanning the $(0, 1)$ interval.

Note that this prior specification approach is fairly automatic as it only requires a range of dose values along with a reasonable prior for α (chose as discussed in Chapter 2). In particular, since the range is comparable for all data examples in Sections 3.1.3 and 3.1.4, we used the same prior setting for all analyses: a normal prior for β_0 with mean 0 and variance 10; an exponential prior with mean 0.1 for β_1; an inverse gamma prior for σ^2 with shape parameter 2 and mean 25; a uniform prior for ϕ over $(0, 10)$; and a gamma$(2, 1)$ prior for α. Prior sensitivity analysis revealed robust posterior inference under less and more dispersed priors. Finally, we set the truncation level to $L = 50$,

which results in $\mathrm{E}(\sum_{l=1}^{L} w_l) \approx 0.9999593$ averaging over the given prior on α.

3.1.2 Posterior inference

MCMC simulation

As noted in Chapter 1, we observe that for the DEHP data, discussed in Section 1.2, the dams are labeled and recorded in ascending numerical order across dose levels. (This is also the case for other data sets available from the database of the National Toxicology Program.) Therefore, the animals can be linked as a response vector across the dose levels with the conditional independence assumption built for the replicated response vectors. Hence, the data structure and corresponding hierarchical model is along the lines of the spatial DP (Gelfand et al., 2005) rather than, for instance, the ANOVA DDP (DeIorio et al., 2004). The latter would be appropriate for more traditional quantal response bioassay settings where exchangeability both across and within dose levels is the more natural assumption.

More specifically, let $\boldsymbol{y}_j = (y_{1j}, \ldots, y_{Nj})$ be the jth response replicate with number of implants vector $\boldsymbol{m}_j = (m_{1j}, \ldots, m_{Nj})$, for $j = 1, \ldots, n$, and $\boldsymbol{\theta}_j \equiv \boldsymbol{\theta}_j(\boldsymbol{x}) = (\theta_j(x_1), \ldots, \theta_j(x_N))$ be the latent mixing vector for $\boldsymbol{y}_j$ and $\boldsymbol{x} = (x_1, \ldots, x_N)$. Then, the first stage of the hierarchical model for the data is written as

$$(\{m_{ij}\}, \{y_{ij}\}) \mid \{\boldsymbol{\theta}_j\} \sim \prod_{j=1}^{n} \prod_{i=1}^{N} \{\mathrm{Bin}\,(y_{ij}; m_{ij}, \pi(\theta_j(x_i)))\}^{s_{ij}} \prod_{i=1}^{N} \prod_{j=1}^{n_i} \mathrm{Pois}(m_{ij}; \lambda)$$

where $G_{0\boldsymbol{x}} = \mathrm{N}_N(\beta_0 \boldsymbol{j}_N + \beta_1 \boldsymbol{x}, \Sigma)$, where $\boldsymbol{j}_N$ is an $N \times 1$ vector of ones and Σ is induced by the GP covariance function, that is, $\Sigma = \sigma^2 H(\phi)$ with $H_{ij}(\phi) = \exp(-\phi|x_i - x_j|^d)$.

Note that inference for the implant distribution, using a gamma prior for λ, is implemented independent of the DDP mixture, and is not discussed further except

with regard to posterior predictive calculations.

Given the truncation approximation and configuration variables as discussed in Chapter 2 (see model (2.3)), we can write the hierarchical model for the y_{ij} as

$$\begin{aligned}
\{y_{ij}\} \mid \{m_{ij}\}, \boldsymbol{w}, \boldsymbol{Z} &\sim \prod_{j=1}^{n} \prod_{i=1}^{N} \left\{ \text{Bin}\left(y_{ij}; m_{ij}, \pi(Z_{w_j}(x_i))\right) \right\}^{s_{ij}} \\
w_j \mid \boldsymbol{p} &\overset{i.i.d.}{\sim} \sum_{l=1}^{L} p_l \delta_l(w_j), \ j = 1, \ldots, n \\
(\boldsymbol{p}, \boldsymbol{Z}) \mid \alpha, \psi &\sim f(\boldsymbol{p} \mid \alpha) \times \prod_{l=1}^{L} G_{0\boldsymbol{x}}(\boldsymbol{Z}_l \mid \psi)
\end{aligned} \tag{3.5}$$

with priors on α and ψ. Prior specification is discussed in the section following the dose-response inferences.

Under model (3.5), updating the mixing parameters $\boldsymbol{Z}_l$ efficiently is vital. In light of this, we discuss here options for these Metropolis-Hastings steps. Specifically, the full conditional for $\boldsymbol{Z}_l$ depends on whether l corresponds to one of the distinct components. If $l \notin \{w_k^* : k = 1, \ldots, n^*\}$, then $\boldsymbol{Z}_l$ is drawn from the $\text{N}_N(\beta_0 \boldsymbol{j}_N + \beta_1 \boldsymbol{x}, \Sigma)$ distribution. For $l \in \{w_k^* : k = 1, \ldots, n^*\}$,

$$\boldsymbol{Z}_{w_k^*} \mid \boldsymbol{w}, \psi, \text{data} \propto \text{N}_N(\boldsymbol{Z}_{w_k^*}; \beta_0 \boldsymbol{j}_N + \beta_1 \boldsymbol{x}, \Sigma) \prod_{\{j: w_j = w_k^*\}} \prod_{i=1}^{N} \left\{ \text{Bin}\left(y_{ij}; m_{ij}, \pi(Z_{w_k^*}(x_i))\right) \right\}^{s_{ij}}$$

which is sampled with a Metropolis-Hastings step. This step was approached in many ways including slice sampling and random-walk updates with various covariance matrices. One course included scaled identity covariance matrices, another used the output from the previous model to dynamically estimate the proposal covariance matrix. Furthermore, component-specific covariance matrices were estimated from initial runs. Mixing and acceptance rates were optimal given a covariance matrix for

the proposal distribution that is of the same form as the GP prior, with (i, j)-th element $a \exp(-b|x_i - x_j|)$, where a and b are tuning parameters. For instance the 2,4,5-T data set (Section 3.1.3), $a = 1.12$ and $b = 0.1$ provided acceptance rates between 0.15 and 0.20.

The conditional posterior for each w_j, $j = 1, ..., n$, is a discrete distribution, $\sum_{l=1}^{L} \tilde{p}_{lj} \delta_l(w_j)$, where $\tilde{p}_{lj} \propto p_l \prod_{i=1}^{N} \{\text{Bin}\,(y_{ij}; m_{ij}, \pi(Z_l(x_i)))\}^{s_{ij}}$, for $l = 1, ..., L$.

Given the GP assumptions, the joint full conditional for the hyperparameters of $G_{0\mathcal{X}}$ is shown by

$$p(\beta_0, \beta_1, \sigma^2, \phi \mid \boldsymbol{Z}, \boldsymbol{w}, \text{data}) \propto p(\beta_0)p(\beta_1)p(\sigma^2)p(\phi) \prod_{k=1}^{n^*} \text{d}G_{0\boldsymbol{x}}(\boldsymbol{Z}_{w_k^*} \mid \beta_0, \beta_1, \sigma^2, \phi),$$

where $p(\cdot)$ denotes the respective priors. Specifically, β_0 is given a normal prior, $\text{N}(m_0, s_0^2)$, β_1 has an exponential prior, and σ^2 is assigned an inverse-gamma prior with shape parameter $a_\sigma > 1$ and mean $b_\sigma/(a_\sigma - 1)$. Then, β_0 has a normal posterior full conditional distribution with mean $\left(m_0 s_0^{-2} + \boldsymbol{j}_N' \Sigma^{-1} \sum_{k=1}^{n^*} (\boldsymbol{Z}_{w_k^*} - \beta_1 \boldsymbol{x})\right) / \left(s_0^{-2} + n^* \boldsymbol{j}_N' \Sigma^{-1} \boldsymbol{j}_N\right)$ and variance $\left(s_0^{-2} + n^* \boldsymbol{j}_N' \Sigma^{-1} \boldsymbol{j}_N\right)^{-1}$. We sample β_1 through a random-walk Metropolis-Hastings step with a normal proposal on the log scale. Moreover, σ^2 has an inverse gamma full conditional with shape parameter $a_\sigma + 0.5 n^* N$ and rate parameter $b_\sigma +$ $0.5 \sum_{k=1}^{n^*} (\boldsymbol{Z}_{w_k^*} - \beta_0 \boldsymbol{j}_N - \beta_1 \boldsymbol{x})' H^{-1}(\phi) (\boldsymbol{Z}_{w_k^*} - \beta_0 \boldsymbol{j}_N - \beta_1 \boldsymbol{x})$. Under the Unif(0,$b_\phi$) prior for ϕ, its full conditional is proportional to

$$|H(\phi)|^{-n^*/2} \exp\left(-0.5\sigma^{-2} \sum_{k=1}^{n^*} (\boldsymbol{Z}_{w_k^*} - \beta_0 \boldsymbol{j}_N - \beta_1 \boldsymbol{x})' H^{-1}(\phi) (\boldsymbol{Z}_{w_k^*} - \beta_0 \boldsymbol{j}_N - \beta_1 \boldsymbol{x})\right),$$

for $0 < \phi < b_\phi$. While this distribution is not available in closed form, it can be sampled

using a Metropolis-Hastings step or by discretizing its support.

Inference for the dose-response relationship and risk assessment

Consider new responses (m_0, y_0) corresponding to a generic dose level x_0. Under the assumed formulation for the joint distribution, $f(m, y; \lambda, G_{\mathcal{X}}) = \text{Pois}(m; \lambda) f(y \mid m; G_{\mathcal{X}})$, the posterior predictive distribution for (m_0, y_0) can be separated into the conditional predictive for y_0 and the marginal predictive for m_0. That is, $p(m_0, y_0 \mid x_0, \text{data}) = \int \text{Pois}(m_0; \lambda) p(\lambda \mid \text{data}) \text{d}\lambda \times p(y_0 \mid m_0, x_0, \text{data})$. The expression for $p(y_0 \mid m_0, x_0, \text{data})$ depends on whether x_0 is among the observed doses or a new dose level, and can be obtained as a special case of the general form below. Therefore, to obtain inference for the joint posterior predictive distribution, at each iteration of the MCMC algorithm, we draw m_0 from a shifted Poisson with mean λ, then given m_0 and the current values of the DDP parameters, obtain a predictive draw for y_0.

Moreover, each posterior sample for $(\boldsymbol{p}, \boldsymbol{Z})$ provides a posterior realization for $G^L_{\boldsymbol{x}}$ directly through its definition, $\sum_{l=1}^{L} p_l \delta_{Z_l(\boldsymbol{x})}$. Next, given the predictive draw for the number of implants m_0, for any vector $\boldsymbol{y}_0 = (y_{10}, \ldots, y_{N0})$, $f(\boldsymbol{y}_0 \mid m_0; G^L_{\boldsymbol{x}}) = \sum_{l=1}^{L} p_l \prod_{i=1}^{N} \text{Bin}\left(y_{i0}; m_0, \pi(Z_l(x_i))\right)$ is a posterior realization from the conditional response distribution at the observed doses.

To extend the inference beyond the N observed dose levels, we predict across M new doses, $\tilde{\boldsymbol{x}} = (\tilde{x}_1, \ldots, \tilde{x}_M)$, which may include values outside the range of the observed doses. To predict a new vector of responses at all dose levels, $(\boldsymbol{y}_0, \tilde{\boldsymbol{y}}_0) = (y_{10}, \ldots, y_{N0}, \tilde{y}_{10}, \ldots, \tilde{y}_{M0})$, given the corresponding number of implants m_0, the mixing parameter vector is extended to $(\boldsymbol{Z}_l, \tilde{\boldsymbol{Z}}_l)$, for $l = 1, \ldots, L$, as described in Chapter 2. Denoting $\tilde{\boldsymbol{Z}} = \{\tilde{\boldsymbol{Z}}_l : l = 1, ..., L\}$, the conditional posterior predictive distribution for

$(\boldsymbol{y}_0, \tilde{\boldsymbol{y}}_0)$ given m_0 is given by

$$p((\boldsymbol{y}_0, \tilde{\boldsymbol{y}}_0) \mid m_0, \tilde{\boldsymbol{x}}, \text{data}) = \int\int \sum_{l=1}^{L} p_l \left\{ \prod_{i=1}^{N} \text{Bin}(y_{i0}; m_0, \pi(Z_l(x_i))) \prod_{j=1}^{M} \text{Bin}(\tilde{y}_{j0}; m_0, \pi(\tilde{Z}_l(\tilde{x}_j))) \right\} \times \left(\prod_{l=1}^{L} \text{N}_M(\tilde{Z}_l(\tilde{\boldsymbol{x}}); \tilde{\mu}_l, \tilde{\Sigma}) \right) \text{d}\tilde{\boldsymbol{Z}} \text{d}p(\boldsymbol{p}, \boldsymbol{Z}, \alpha, \psi \mid \text{data}).$$

Here, $\tilde{\mu}_l = (\beta_0 \boldsymbol{j}_M + \beta_1 \tilde{\boldsymbol{x}}) + H^{MN}(\phi) H^{-1}(\phi)(Z_l(\boldsymbol{x}) - (\beta_0 \boldsymbol{j}_N + \beta_1 \boldsymbol{x}))$, where $\boldsymbol{j}_K$ denotes a vector of dimension K with all elements equal to 1, and $H^{MN}(\phi)$ is the $M \times N$ matrix with $H_{ij}^{MN}(\phi) = \exp(-\phi|\tilde{x}_i - x_j|^d)$. Moreover, $\tilde{\Sigma} = \sigma^2\{H^{MM}(\phi) - H^{MN}(\phi)H^{-1}(\phi)(H^{MN}(\phi))^T\}$, where $H^{MM}(\phi)$ is the $M \times M$ matrix with $H_{ij}^{MM}(\phi) = \exp(-\phi|\tilde{x}_i - \tilde{x}_j|^d)$.

Using the posterior draws for the parameters at any set of dose levels, we evaluate expression (3.3) to develop inference for the intracluster correlation as a function of dose level. Standard hierarchical extensions of the Binomial model are limited with regard to such inference, e.g., the Beta-binomial model involves the same (positive) correlation for all dose levels. The data of Section 3.1.3 illustrate the practical utility of the DDP mixture model in recovering dose-dependent intracluster correlation patterns with appropriate uncertainty quantification.

Key to quantitative risk assessment is inference for the dose-response relationship. Using the posterior samples for $(\boldsymbol{p}, \boldsymbol{Z}, \tilde{\boldsymbol{Z}})$, we obtain the posterior distribution of

$$\Pr(y^* = 1; G_{x_0}^L) = \sum_{l=1}^{L} p_l \pi(Z_l(x_0))$$

arising from (3.4) for each $x_0 \in (\boldsymbol{x}, \tilde{\boldsymbol{x}})$. These distributions can be summarized with posterior means and two percentiles to provide point and interval estimates for the dose-response curve $\Pr(y^* = 1; G_{\mathcal{X}}^L)$ (as in Figures 3.6 and 3.7). It can be readily shown that

the posterior expectation of $\Pr(y^* = 1; G^L_{x_0})$ is equal to the expectation of y_0/m_0 from the joint posterior predictive distribution $p(m_0, y_0 \mid x_0, \text{data})$, that is, $\mathrm{E}(y_0/m_0 \mid \text{data}) = \mathrm{E}(\Pr(y^* = 1; G^L_{x_0}) \mid \text{data})$. As shown in Figure 3.12, the posterior predictive samples for y_0/m_0 obtained across a range of dose levels also provide useful inference for the dose-response relationship.

Finally, risk assessment can be based on estimation of the dose level x_q that corresponds to a specified probability, q, of a negative outcome, that is, $q = \Pr(y^* = 1; G^L_{x_q})$. For any set of probabilities q, each posterior realization for $\Pr(y^* = 1; G^L_{\mathcal{X}})$ can be (numerically) inverted to obtain the posterior distribution for the corresponding inverted dose levels x_q.

3.1.3 Data illustrations

We illustrate the nonparametric modeling approach with two data sets. We first study a range of inferences under the DDP Binomial mixture model for a data set commonly considered in the literature. The second data example is included to highlight the feature of the DDP modeling framework with regard to recovering non-monotonic dose-response relationships. Chapter 1 presents the two data sets in more detail.

Application to 2,4,5-T data

We first explore posterior inference for the DDP precision parameter, α, and the corresponding number of distinct components, n^*. We also investigate the GP hyperparameters, namely β_0, β_1, σ^2, and ϕ. Under the prior specifications provided in Section 3.1.1, the posterior densities peak significantly from the prior densities. Figure 3.3 includes histograms of the posterior distributions of the hyperparameters with the prior density given in blue. In all cases, there is moderate to substantial learning for the

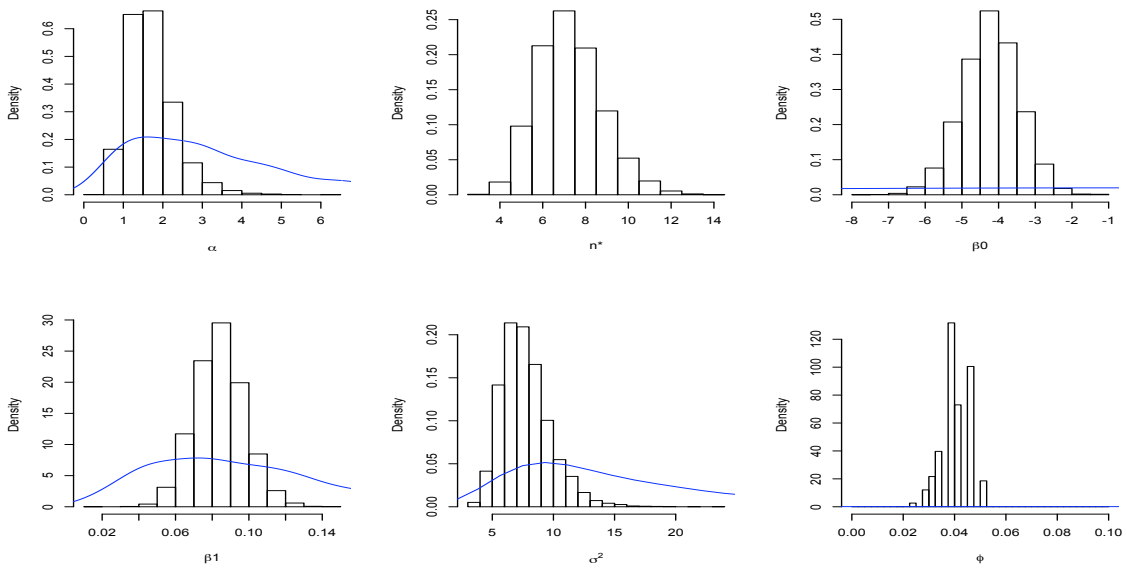

Figure 3.3: 2,4,5-T data. Posterior densities of the precision parameter, α (top left), n^* (top middle), and the four GP hyperparameters: β_0 (top right), β_1 (bottom left), σ^2 (bottom middle), and ϕ (bottom right). Prior densities are provided in blue.

parameters. We observe similar phenomena for the simulated and experimental data sets explored through the rest of the dissertation.

Focusing on inference for conditional response distributions, Figure 3.4 plots the posterior mean and 90% uncertainty bands for $f(y \mid m = 12; G^L_{x_i})$ at all observed dose levels, and for $f(y \mid m = 12; G^L_{\tilde{x}_0})$ at two new doses, one ($\tilde{x}_0 = 50$ mg/kg) within the observed range, and one extrapolated at $\tilde{x}_0 = 100$ mg/kg. The probability mass functions corresponding to low and high dose levels depict shapes that could be captured by traditional parametric models. However, in the mid-range of dose values, the DDP mixture model uncovers non-standard probability mass function shapes, which suggest bimodality. The estimated mass functions at the new dose levels have larger probability bands, and their shape highlights the smooth evolution of the DDP-based response distributions across dose levels.

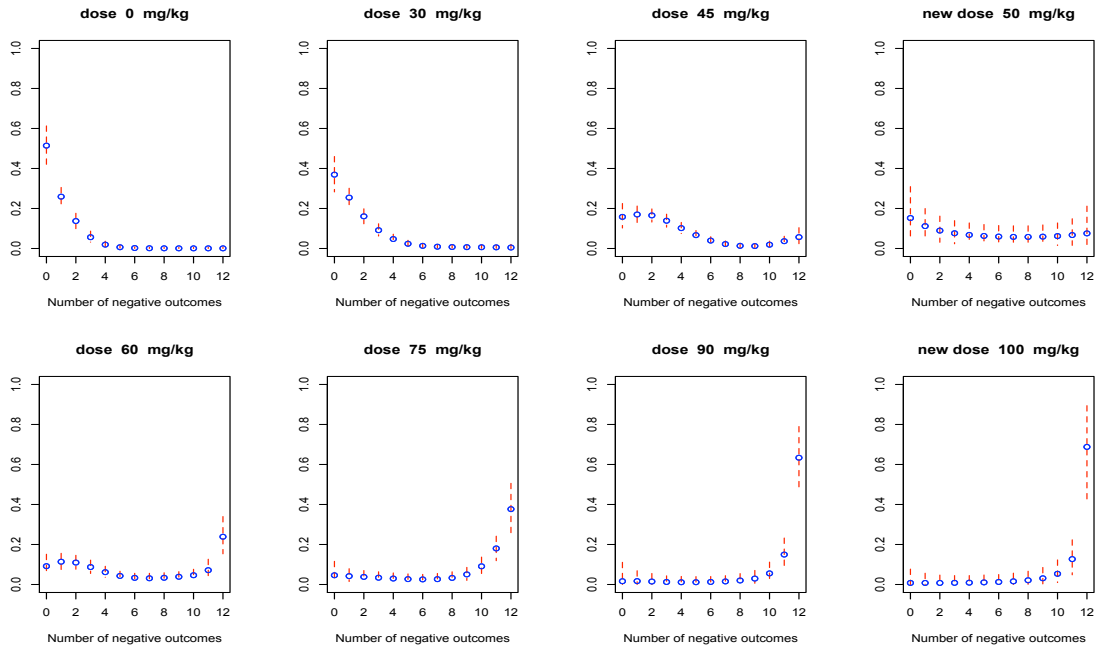

Figure 3.4: 2,4,5-T data. For the 6 observed dose levels and 2 new dose levels, the posterior mean probability mass functions (denoted by "o") and 90% probability bands for the number of negative outcomes conditional on $m = 12$ implants.

The posterior densities for the intracluster correlations at the observed dose levels are given in the left panel of Figure 3.5. The correlations depict an increasing trend up to dose levels 60–75 mg/kg, with increasing uncertainty beyond dose 75 mg/kg consistent with the smaller number of dams at the two higher dose levels. Using for illustration four probabilities, $q = 0.1$, 0.25, 0.40, and 0.50, the right panel of Figure 3.5 shows the posterior densities of the corresponding calibrated dose level x_q, obtained as discussed in Section 3.1.2.

While results for response distributions are not shown here, we also fit the Binomial-logit and Beta-Binomial models to the 2,4,5-T data. The Binomial-logit model is not flexible enough to capture the non-standard distributions and estimates little variation. The Beta-Binomial model also cannot deviate from unimodal probability

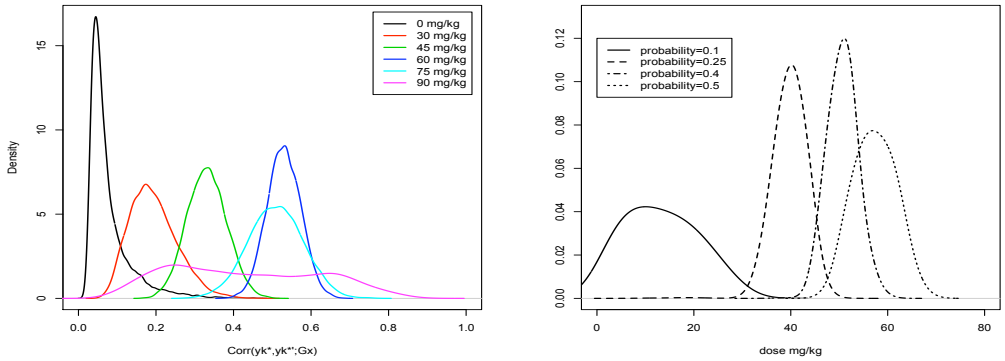

Figure 3.5: 2,4,5-T data. The left panel plots the posterior densities for the intracluster correlations at each of the six observed dose levels. The posterior densities for the calibrated dose level corresponding to four probability thresholds are given in the right panel.

mass functions, but attempts to compensate for the data heterogeneity by increasing the variability in the probability of response, thereby producing large uncertainty bands.

This overcompensation is manifested in the overly wide probability bands for the estimated dose-response curve under the Beta-Binomial model (Figure 3.6, middle panel). On the other extreme, the Binomial-logit model underestimates the uncertainty in the curve, and is also restricted to the logistic function shape (Figure 3.6, left panel). The posterior mean estimate from the DDP mixture model (Figure 3.6, right panel) supports a non-decreasing dose-response relationship with curvature that deviates at smaller doses from the logistic shape, and with larger uncertainty at the interpolated values. The inference results for the dose-response curve provide an interesting illustration of a nonparametric Bayesian model producing more realistic uncertainty quantification for posterior estimates relative to simpler parametric models. In particular, in contrast to continuous mixing that defines the Beta-Binomial model, the discrete nature of the DDP prior enables clustering of the Binomial kernel latent

mixing parameters, thus, controlling more effectively the variability of the estimated conditional response distributions.

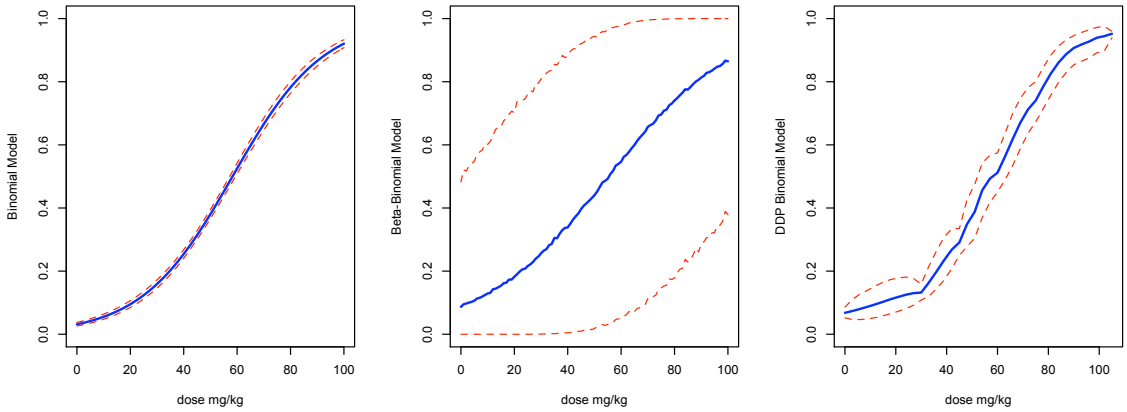

Figure 3.6: 2,4,5-T data. The posterior mean estimate (solid line) and 90% probability bands (dashed lines) for the dose-response curve, using the Binomial-logit model (left panel), the Beta-Binomial model (middle panel), and the DDP Binomial mixture model (right panel).

Application to DEHP data

Here, we present a brief analysis of the DEHP data, mainly to highlight the feature of the DDP modeling framework with regard to recovering non-monotonic dose-response relationships. A more substantial analysis of this data is found in Section 3.2, given a more biologically relevant treatment of the multicategory classification.

First, we note that the data (Figure 3.7, left panel) appear to suggest a drop in the probability of a negative outcome from the control level to level 25 mg/kg $\times$ 1000. In addition to the graphical indication, the drop is suggested by an (admittedly crude) "data-based" analysis, using independent Binomials with common probability for all dams at each dose. The resulting (maximum likelihood) estimates of the probability

of a negative outcome at doses 0 and 25mg/kg × 1000 are equal to 0.200 and 0.116, with respective standard errors 0.0209 and 0.0179. As discussed in Section 1.2, such a dose-response pattern may be associated with hormesis, and thus, it is practically important to be able to quantify how well it is supported by the data.

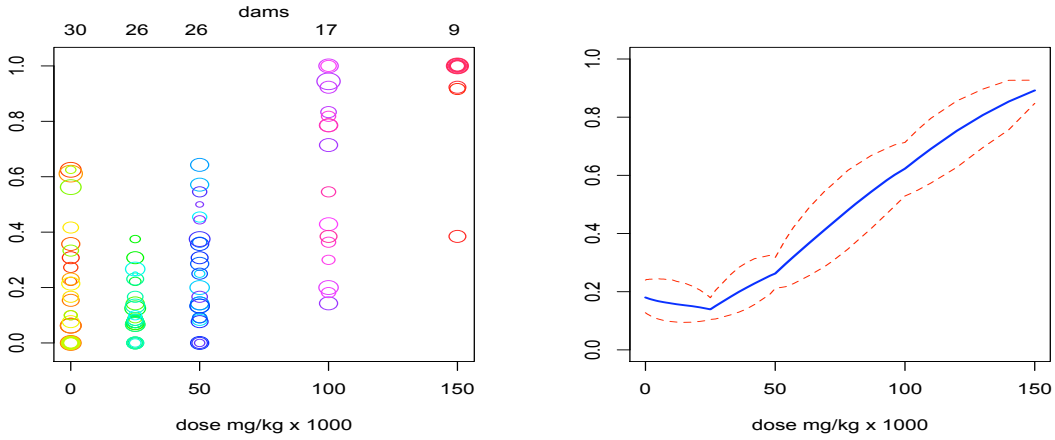

Figure 3.7: DEHP data. The left panel shows the data, where each circle corresponds to a particular dam, the size of the circle is proportional to the number of implants, and the coordinates of the circle are the dose level and the proportion of negative outcomes. The right panel includes the posterior mean estimate (solid line) and 90% probability bands (dashed lines) for the dose-response curve.

Indeed, this particular non-monotonic dose-response shape is apparent in the posterior mean estimate and corresponding uncertainty bands for $\Pr(y^* = 1; G^L_{\mathcal{X}})$ (Figure 3.7, right panel). Under essentially all standard models for developmental toxicology data, this dip in the dose-response curve would not be captured. Moreover, the DDP Binomial mixture model is again able to recover varying shapes for conditional response distributions across dose levels. In particular, point and interval estimates for $f(y \mid m = 12; G^L_x)$ (not shown) support shapes that evolve with increasing dose from right to left skewness, with bimodal probability mass functions uncovered for values of x around observed dose 100 mg/kg × 1000.

3.1.4 Comparison study

Here, we consider two synthetic data sets to check the performance of our model and to compare with simpler semiparametric and nonparametric Bayesian models.

Simulated data

We work with simulated data sets generated under three distinct settings. The first case includes a standard Binomial response distribution and a linear dose response function, $\pi(-4 + 0.12x)$. The second is based on a Binomial response distribution with a non-standard non-linear function for the dose-response curve. In particular, we define the probability of a negative outcome at dose x by $\pi(h(x))$, where $h(x) = -2 + 0.04x - 0.25\sin(2.7x) - 1.1/(1 + x^2)$. The third simulation example is built from a mixture of three Binomial-logit distributions, $\sum_{i=1}^{3} p_i \text{Bin}(y; m, \pi(q_i(x)))$, where $(p_1, p_2, p_3) = (0.1, 0.4, 0.5)$, $q_1(x) = -2 + 0.02x$, $q_2(x) = -10 + 0.20x$, and $q_3(x) = -4 + 0.15x$. For all simulations, we use the values of the dose levels, number of dams, and implant vectors from the 2,4,5-T data (see Section 1.2 and 3.1.3). Figure 3.8 plots the simulated data sets, including the true dose-response curve.

Comparison models

For comparison, consider the semiparametric product of mixtures of Dirichlet process (PMDP) model from Dominici and Parmigiani (2001), the only approach from the Bayesian nonparametrics literature for analysis of developmental toxicology data. The PMDP approach involves a different modeling structure than the DDP Binomial mixture. We also consider comparison with two models that can be viewed as special cases of the model developed in Section 3.1.1, a GP Binomial regression model, and a more structured DDP mixture model which ensures monotonicity for the dose-response

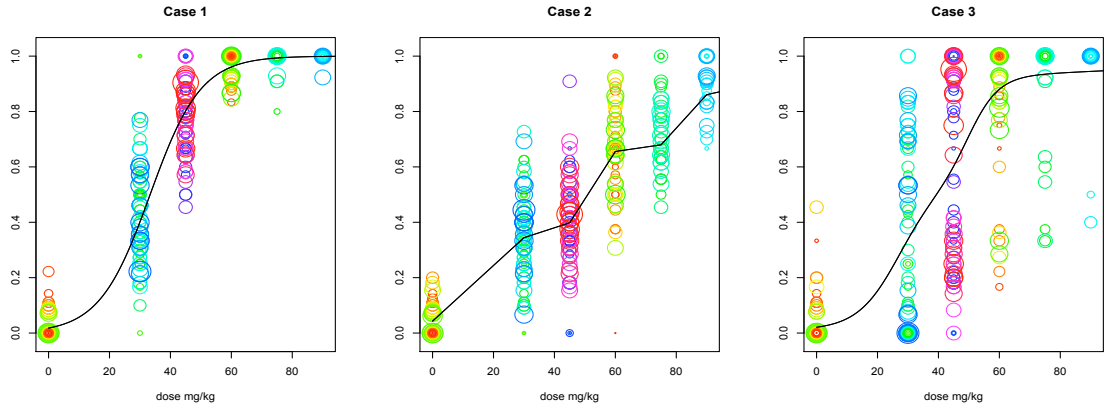

Figure 3.8: Simulation study. Each circle corresponds to a particular dam, the size of the circle is proportional to the number of implants, and the coordinates of the circle are the dose level and the proportion of negative outcomes. The left panel corresponds to the first simulation setting (Binomial with standard logistic dose-response curve), the middle panel to the second setting (Binomial with non-linear dose-response curve), and the right panel to the third setting (three-component mixture of Binomial-logit distributions). In each panel, the solid line denotes the true dose-response curve.

curve.

PMDP model: We implement the PMDP model as in Dominici and Parmigiani (2001) with dose-specific precision parameters and a Binomial-logit centering distribution for the number of negative outcomes given a fixed number of implants. Under the PMDP model,

$$y_{ij} \mid F_{ij} \overset{ind.}{\sim} F_{ij}, \quad j = 1, \ldots, n_i, \ i = 1, \ldots, N \tag{3.6}$$

$$F_{ij} \mid \{A_i\}, (\eta_0, \eta_1) \overset{ind.}{\sim} \mathrm{DP}\left(A_i, \mathrm{Bin}\left(m_{ij}, \pi(\eta_0 + \eta_1 x_i)\right)\right), \quad j = 1, \ldots, n_i, \ i = 1, \ldots, N.$$

By integrating out the infinite dimensional parameters, F_{ij}, MCMC posterior sampling involves an $N + 2$ dimensional Metropolis-Hastings step for the regression coefficients and the N dose-specific precision parameters. Conditional on the A_i and (η_0, η_1), the posterior distribution of F_{ij} is a DP with updated parameters, and inference can

be obtained using its definition. Following Dominici and Parmigiani (2001), we use independent normal priors for η_0 and η_1 with mean 0 and variance 3, and independent priors for the A_i arising from uniform distributions for the $A_i/(10+A_i)$. For all simulated data sets, there is significant learning for the A_i and, especially, for η_0 and η_1, under this prior choice.

The PMDP model appears restrictive for inference outside the observed dose levels as the distributions are dependent in a weak fashion being linked only through the common regression coefficients. In particular, the probability of a negative outcome at a new dose is problematic to define under the version of the PMDP model in (3.6). Taking the precision parameter to be a function of dose, this probability can be shown to follow a Beta$(A(x)\pi(\eta_0+\eta_1 x), A(x)(1-\pi(\eta_0+\eta_1 x)))$ distribution obtaining a result for the PMDP prior on a connection analogous to the one between DDP models (3.1) and (3.2). This connection is established through working with the PMDP model centered on a product of Bernoullis distribution. Through induction, starting with the cases of $m = 1$ and $m = 2$, we find the probability of response for the individual pups, profiting from the properties of the Dirichlet distribution. Evidently, effective interpolation (or extrapolation) at new dose levels requires an appropriate dose-dependent prior model for the DP precision parameter. In general, such a specification does not seem straightforward, for instance, simple choices such as $\log(A(x)) = \gamma_0 + \gamma_1 x$ (Carota and Parmigiani, 2002) may not be sufficiently flexible to capture the degree to which the data deviate from the centering Binomial distribution.

GP Binomial regression model: This model retains the restrictive Binomial response distribution, but is more flexible than the Binomial-logit model in inference for the dose-response curve. The GP model is a limiting case of the DDP mixture model (as $\alpha \to 0^+$), in particular, it is based on a GP prior for the dose-response curve on the

logistic scale,

$$y_{ij} \mid m_{ij}, \{\theta(x_i)\} \overset{ind.}{\sim} \text{Bin}\left(y_{ij}; m_{ij}, \pi(\theta(x_i))\right), \;\; j = 1, \ldots, n_i, \;\; i = 1, \ldots, N$$

where $(\theta(x_1), ..., \theta(x_N))$ has a $\text{N}_N(\xi_0 \mathbf{j}_N + \xi_1 \mathbf{x}, \Lambda)$ prior given hyperparameters $(\xi_0, \xi_1, \tau^2, \rho)$. Here, $\Lambda = \tau^2 H(\rho)$, and $H_{ij}(\rho) = \exp(-\rho|x_i - x_j|^d)$, with fixed $d \in [1, 2]$. We used the exponential correlation function ($d = 1$) for both GP and DDP mixture models. (The inference results discussed below were similar under the Gaussian correlation function, although predictive inference under the GP model was less numerically stable for $d = 2$.) The GP hyperparameters $(\xi_0, \xi_1, \tau^2, \rho)$ are assigned the same priors with the DDP hyperparameters $(\beta_0, \beta_1, \sigma^2, \phi)$ given in Section 3.1.1.

Linear-DDP mixture model: A distinguishing feature of the DDP mixture model is that it supports non-monotonic dose-response relationships. We argue that this is practically relevant in the analysis of developmental toxicology data.

However, if one wishes to enforce monotonicity for the dose-response curve (with prior probability 1 rather than only in prior expectation), this can be accomplished within the DDP mixture framework using a simplified version of the DDP prior. Specifically, setting $\eta_l(x) = \gamma_{0l} + \gamma_{1l} x$ in (2.2) yields the linear DDP prior as discussed in Chapter 2. Here, the $(\gamma_{0l}, \gamma_{1l})$ are i.i.d., given hyperparameters, from a centering distribution, typically, with independent components. Now, with the DP truncation approximation, the linear-DDP Binomial mixture is given by

$$f(y \mid m; G_x^L) = \sum_{l=1}^{L} p_l \text{Bin}(y; m, \pi(\gamma_{0l} + \gamma_{1l} x)).$$

It is straightforward to verify that, if $\gamma_{1l} > 0$ for all l, then the corresponding dose-

response curve is non-decreasing in x. We implement this model assuming $\gamma_{0l} \mid \delta_0, \sigma_0^2$ i.i.d. $\mathrm{N}(\delta_0, \sigma_0^2)$, and, independently, $\gamma_{1l} \mid \varphi$ i.i.d. gamma(c, φ). We fix $c = 1$ and assign hyperpriors to δ_0, σ_0^2 and φ, specifically, we place a zero mean normal prior on δ_0 with variance 20, an inverse-gamma$(2, 25)$ on σ_0^2, and a gamma$(10, 1)$ on φ for the simulated data sets.

Results

The first simulation case is the simplest case. Figure 3.9 gives the posterior mean and 90% probability bands for the probability mass function of the number of negative outcomes given $m = 12$ implants for three dose levels. The GP, linear-DDP, and DDP models approximate the truth relatively well, with only a few areas that have slightly larger probability bands under the general DDP mixture model. The PMDP model roughly follows the true values, however it is greatly influenced by the data at the given dose levels and thereby generates large uncertainty bands.

Under the second simulation case, Figure 3.10 gives the posterior mean and 90% uncertainty bands for the probability mass function of the number of negative outcomes given $m = 12$ implants for three dose levels. The GP and DDP mixture models provide similar inference, with slightly larger uncertainty bands arising from the DDP model. The linear-DDP model produces somewhat less accurate point estimates with narrow interval estimates. While the PMDP model captures the general shape of the mass function, there is substantial uncertainty in its interval estimates.

The mixture of Binomials simulation setting provides more striking differences between the performance of the models, as seen in Figure 3.11. The GP model relies on a Binomial response distribution and therefore can not pick up the bimodality at dose levels 30 and 45, and also misses the larger true probability mass function values

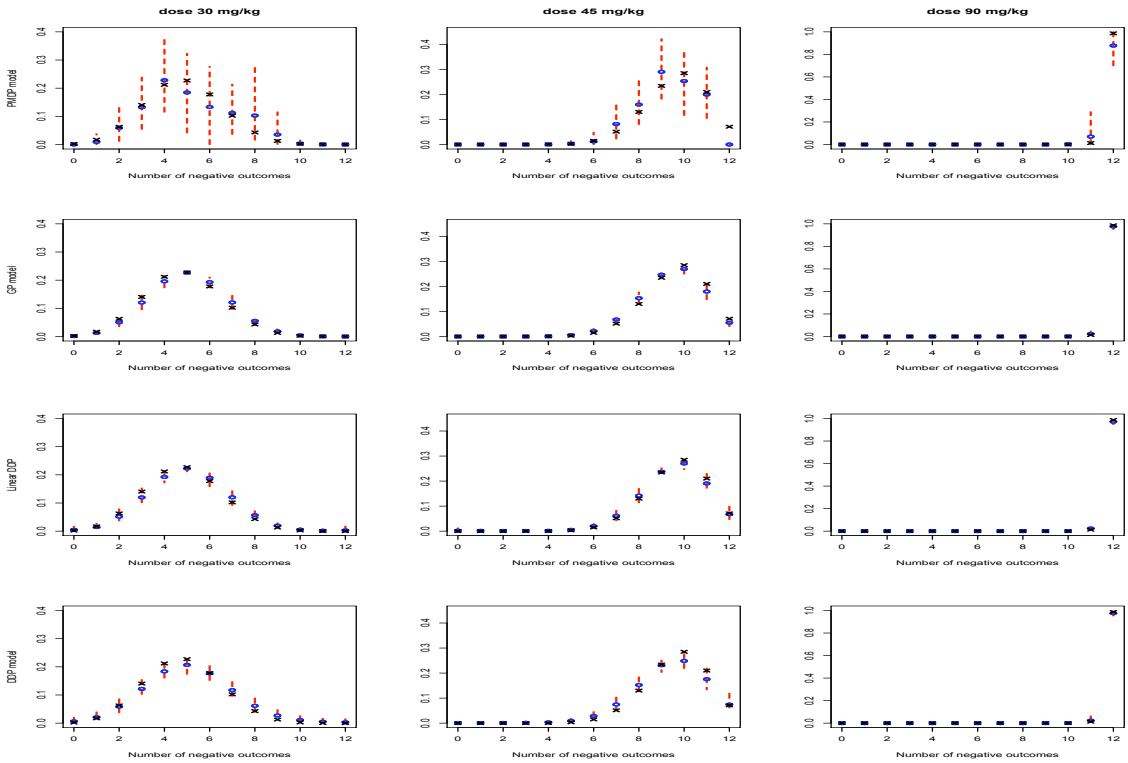

Figure 3.9: Simulation case 1. Posterior mean (denoted by "o") and 90% uncertainty bands for the probability mass function of the number of negative outcomes, given $m = 12$ implants, at three dose levels, using the PMDP, GP, Linear-DDP, and DDP models (first, second, third and fourth row, respectively). In each panel, the values of the true probability mass function are denoted by "x".

at dose 75. The linear-DDP model attempts to capture the essence of the bimodal shapes of the probability mass functions at doses 30 and 45; however its restrictive dependence structure limits the posterior accuracy. Inference at dose level 75 resembles the actual probability mass function, yet fails to include the true values within its narrow uncertainty bands. The PMDP model generally includes the true probabilities within the large uncertainty bands it produces for all three dose levels. However, the changes in the estimates across and within dose levels are quite drastic. In the case of dose level 75, there are 8 data points associated with $m = 12$ implants (compared

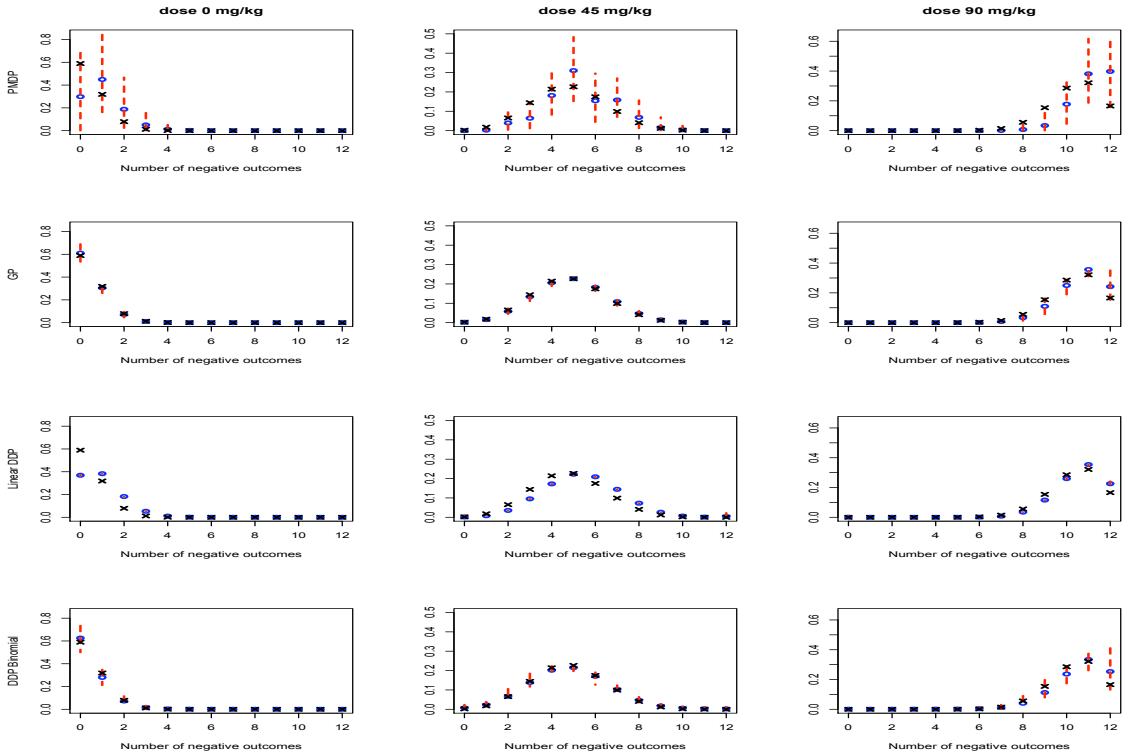

Figure 3.10: Simulation case 2. Posterior mean (denoted by "o") and 90% probability bands for the probability mass function of the number of negative outcomes, given $m = 12$ implants, at three dose levels, using the PMDP, GP, Linear-DDP, and DDP models (first, second, third and fourth row, respectively). In each panel, the values of the true probability mass function are denoted by "x".

to 15-20 observations for doses 30 and 45). Of these 8 observations, two animals had 5 negative outcomes, which is apparent in the PMDP model results. Owing to the smoother evolution of DDP realizations and to its mixture structure, the DDP model has the capacity to avoid such sudden changes in the estimated probabilities. Moreover, the DDP model recovers the true probability mass function shapes with notably tighter uncertainty bands than the PMDP model.

Contrasting the results from Figures 3.9, 3.10, and 3.11 reveals an appealing feature of the proposed modeling framework: the DDP mixture model can uncover non-

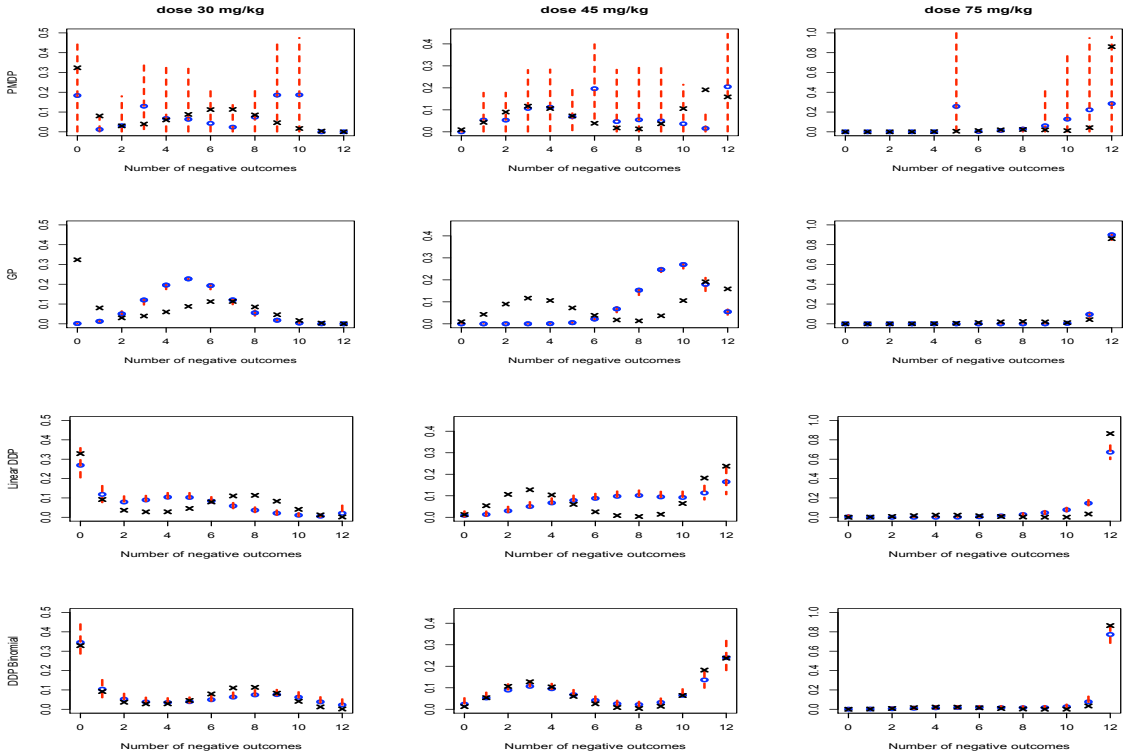

Figure 3.11: Simulation case 3. Posterior mean (denoted by "o") and 90% probability bands for the probability mass function of the number of negative outcomes, given $m = 12$ implants, at three dose levels, using the PMDP, GP, Linear-DDP, and DDP models (first, second, third and fourth row, respectively). In each panel, the values of the true probability mass function are denoted by "x".

standard distributional shapes at different dose levels when such shapes are suggested by the data (Figure 3.11), but at the same time, will recover simpler probability mass functions with a relatively small amount of additional uncertainty relative to parametric models (Figure 3.9).

Because inference for the entire dose-response curve is not readily available under the PMDP model, we focus this aspect of the comparison on the GP, linear-DDP, and DDP mixture models, and looking mainly at the second simulation setting involving the nonstandard dose-response curve. Here, the DDP model represents more accurately

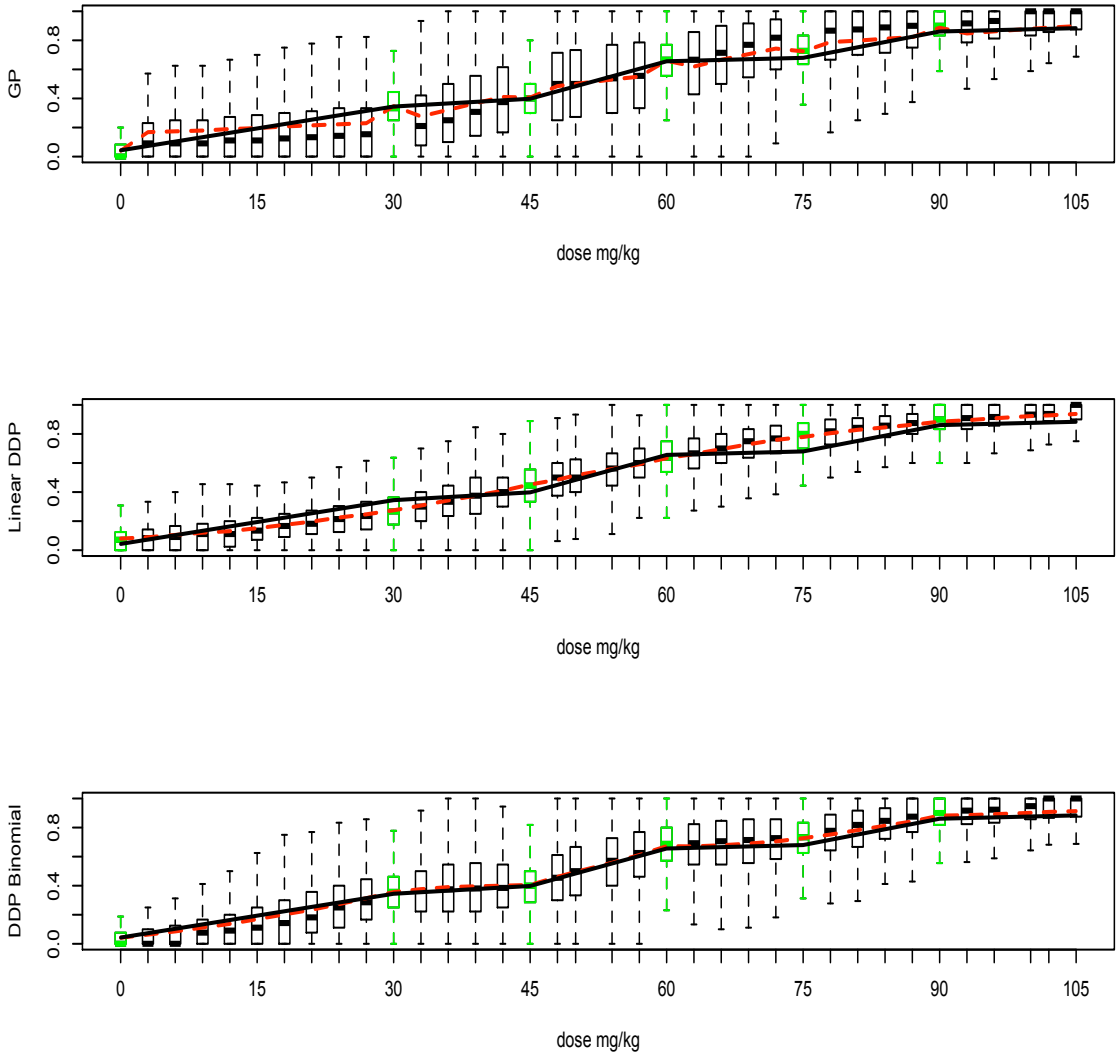

Figure 3.12: Simulation case 2. Boxplots of the posterior predictive samples of y_0/m_0 at the observed and new dose levels, using the GP model (top), Linear-DDP (middle), and DDP mixture model (bottom). In each panel, the posterior mean estimate for the dose-response curve and the true curve are denoted by the dashed and solid line, respectively.

the dose-response curve across dose levels (Figure 3.12). While the GP model performs well at the observed dose levels, it fares worse at the interpolated doses as indicated by

the jumps at the observed doses in the posterior mean dose-response curve. Under the linear-DDP model, dependence across dose x is built solely through the linear functions $\gamma_{0l} + \gamma_{1l}x$, and we thus expect roughly uniform uncertainty in posterior inference results for the dose-response curve (and conditional probability mass functions across dose levels). This is reflected in Figure 3.12, where we also note that the linear-DDP posterior mean estimate tends to smooth out the probability of response. This smoothing results in biased probability mass function estimates. For the third simulation example, the DDP model again produced a more accurate estimate for the dose-response function (results not shown), although in this case with larger posterior predictive uncertainty compared to the GP and linear DDP models.

Finally, we compare the four models with regard to inference for the probability of a negative outcome at the six observed dose levels x_i. Denoting as before by y^* a generic binary outcome at x_i, this probability is given by $\sum_{l=1}^{L} p_l \pi(Z_l(x_i))$ under the DDP model, $\sum_{l=1}^{L} p_l \pi(\gamma_{0l} + \gamma_{1l}x_i)$ under the linear-DDP model, and $\pi(\theta(x_i))$ under the GP model; moreover, under the PMDP model, it arises from a $\text{Beta}(A_i\pi(\eta_0 + \eta_1 x_i), A_i(1 - \pi(\eta_0 + \eta_1 x_i)))$ distribution. Table 1 includes point and 90% interval estimates based on the corresponding posterior distributions for the second and third simulation setting. Noteworthy here are the results under the PMDP model, which produces interval estimates that are too wide to be practical. This level of posterior uncertainty is consistent with the results in Figures 3.10 and 3.11. For the 2,4,5-T data considered in Section 3.1.3 (where results are reported only under the DDP model), we also obtained overly wide 90% interval estimates from the PMDP model (with equivalent interquartile ranges in Table 1 of Dominici and Parmigiani, 2001). Contrarily, for the second simulation case, the linear-DDP model moderates the irregularities of the actual dose-response function into a smooth curve (see Figure 3.12),

resulting in underestimation of the variability in the negative outcome probability. The model produces more realistic interval estimates in the third case, though it generally overestimates the probabilities. As suggested by the posterior predictive results for y_0/m_0, the GP and DDP mixture models result in relatively similar inference for the dose-response curve at the observed dose levels. The DDP model produces wider interval estimates, especially under the setting with the mixture of Binomials true distribution, which are more effective in capturing the true values at the largest dose.

Table 3.1: Simulation study (cases 2 and 3). Posterior mean and 0.05 and 0.95 percentiles (in parentheses) for the probability of a negative outcome at the 6 observed dose levels, using the GP, Linear-DDP, DDP mixture, and PMDP models. The true values of the dose-response curve are given in bold.

Simulation case 2				
Dose	GP	Linear-DDP	DDP	PMDP
0 (**0.04**)	0.04 (0.03,0.05)	0.08 (0.07,0.09)	0.04 (0.03,0.06)	0.09 (0.00,0.06)
30 (**0.34**)	0.35 (0.32,0.38)	0.28 (0.27,0.29)	0.36 (0.33,0.39)	0.30 (0.03,0.72)
45 (**0.39**)	0.41 (0.38,0.43)	0.45 (0.44,0.46)	0.40 (0.37, 0.43)	0.47 (0.08,0.89)
60 (**0.66**)	0.67 (0.63,0.70)	0.63 (0.62,0.64)	0.67 (0.62,0.70)	0.64 (0.24,0.95)
75 (**0.68**)	0.73 (0.69,0.76)	0.78 (0.77,0.79)	0.73 (0.69,0.76)	0.79 (0.34,0.99)
90 (**0.86**)	0.89 (0.86,0.92)	0.88 (0.87,0.89)	0.88 (0.83,0.91)	0.88 (0.45,1.00)
Simulation case 3				
Dose	GP	Linear-DDP	DDP	PMDP
0 (**0.02**)	0.03 (0.02,0.04)	0.04 (0.03,0.06)	0.03 (0.02,0.05)	0.09 (0.00,0.07)
30 (**0.34**)	0.34 (0.31,0.36)	0.27 (0.23,0.32)	0.33 (0.28,0.39)	0.34 (0.01,0.89)
45 (**0.60**)	0.61 (0.59,0.63)	0.65 (0.62,0.69)	0.60 (0.55,0.66)	0.58 (0.06,0.98)
60 (**0.88**)	0.87 (0.86,0.89)	0.87 (0.85,0.90)	0.87 (0.83,0.91)	0.79 (0.09,1.00)
75 (**0.94**)	0.93 (0.91,0.95)	0.94 (0.92,0.95)	0.92 (0.88,0.96)	0.89 (0.28,1.00)
90 (**0.95**)	0.97 (0.96,0.98)	0.97 (0.95,0.98)	0.96 (0.92,0.99)	0.97 (0.85,1.00)

3.1.5 Discussion

We have developed a Bayesian nonparametric mixture framework for modeling and risk assessment in developmental toxicity studies where combined negative outcomes

is the endpoint of interest. The impetus for the proposed modeling approach is that for such studies it is critical to model flexibly both the dam specific distributions and the probability of response to accurately account for the multiple sources of data heterogeneity. The methodology is built from Binomial mixtures with a dependent Dirichlet process prior for the dose-dependent mixing distributions. The resulting nonparametric mixture model provides rich inference for the response distribution as well as for the dose-response curve. Data from two toxicity studies were used to illustrate the variety of inferences that can be obtained from the DDP mixture model, including its practical utility with regard to estimation of non-monotonic dose-response relationships. Finally, using a simulation study, we have shown that, relative to simpler semiparametric Bayesian approaches, the DDP mixture model is the only one that accomplishes both of the inferential goals above.

3.2 Multicategory responses

Although the statistical literature typically considers the form of the data where the response is combined negative outcomes, this is not as informative from a biological point of view. In Section 3.1, we developed the DDP mixture model for this simplified version of the data to set the stage for the more general setting. Here, we consider the first practically important extension which involves modeling developmental toxicology data with responses that include a multicategory classification, for instance, outcomes ("dead", "normal", "malformed"). To our knowledge, the literature does not include any Bayesian nonparametric approaches to modeling developmental toxicology data with a multicategory response classification. Examples of parametric Bayesian hierarchical models for toxicology data, comprising joint discrete-continuous outcomes,

include Dunson et al. (2003) and Faes et al. (2006). Regarding the classical literature, a Dirichlet-trinomial model is presented in Chen et al. (1991); Zhu et al. (1994) develop an extended Dirichlet-multinomial model with Weibull dose-response functions; and Ryan (1992) and Krewski and Zhu (1994) use quasi-likelihood and generalized estimating equations, respectively, to fit multinomial models which incorporate overdispersion.

3.2.1 Modeling approach

For notational purposes, we consider again the Segment II developmental toxicity setting where n_i pregnant dams are exposed to dose level, $x_i,\ i = 1, \ldots, N$. Dam $j = 1, \ldots, n_i$ at dose x_i has m_{ij} implants, of which the number of resorptions (r_{ij}) and prenatal deaths (d_{ij}) are typically recorded as $R_{ij} = r_{ij} + d_{ij}$, and the number of live pups at birth with a certain defect are recorded as y_{ij}. Consequently, the litter size (the number of viable fetuses) for dam j at dose x_i is $m_{ij} - R_{ij}$. The outcomes from the jth dam at dose level x_i are now recorded as $\{(m_{ij}, R_{ij}, y_{ij}) :\ i = 1, \ldots, N,\ j = 1, \ldots, n_i\}$. Note that the experiment provides information on both embryolethality (non-viable fetuses) and fetal malformation, and therefore we seek to jointly model the correlated endpoints.

Given the triplet (m, R, y), we could assume a trinomial kernel for modeling non-viable fetuses, malformation, and normalcy simultaneously, i.e., the trinomial distribution in its original parameterization. While this approach incorporates the correlation between endpoints, it does not highlight nested nature of the data. That is, an implant must be not be either a resorption or a prenatal death before it can be a viable fetus, malformed or not. Therefore, the mixture model for a generic dam with m implants at dose level x can be built more directly from the decomposition of the trinomial kernel into $\text{Bin}(R; m, \pi(\gamma))\text{Bin}(y; m - R, \pi(\theta))$, where $\pi(\gamma)$ is the probability

of a non-viable fetus and $\pi(\theta)$ the conditional probability of a malformation for a viable pup. Hence, the DDP mixture model for the joint distribution of embryolethality and fetal malformation is formulated as follows

$$f(R, y \mid m; G_{\mathcal{X}}) = \int \text{Bin}\left(R; m, \pi(\gamma)\right) \text{Bin}\left(y; m - R, \pi(\theta)\right) \, \mathrm{d}G_{\mathcal{X}}(\gamma, \theta),$$

with $G_{\mathcal{X}} \mid \alpha, \psi \sim \text{DDP}(\alpha, G_{0\mathcal{X}})$. In this case, $G_{0\mathcal{X}}$ is a product of independent GPs, one driving each probability of response, each with linear mean functions ($\xi_0 + \xi_1 x$ and $\beta_0 + \beta_1 x$ for γ and θ, respectively) and exponential correlation functions, with variances τ^2 and σ^2, and correlation functions $\exp\{-\rho|x - x'|\}$ and $\exp\{-\phi|x - x'|\}$. For simplicity, we denote by $\psi = (\xi_0, \xi_1, \tau^2, \rho, \beta_0, \beta_1, \sigma^2, \phi)$ the GP hyperparameters. Here, we have DDP mixing on both γ and θ, where R is the number of non-viable fetuses (including resorptions and prenatal deaths) and y is the number of malformations. A marginal model for m, say a shifted Poisson distribution, can again be added to complete the joint distribution of (m, R, y).

Model properties

Under this modeling approach, of interest will be inference for risk assessment associated with both the probability of "death" and the probability of "malformation" given the fetus is alive. To develop the corresponding dose-response curves, we establish a connection of the mixture model with the clustered Binomial kernels with the model based on products of Bernoullis kernel. Specifically, consider the model for the underlying binary responses. That is, for a generic dam at dose level x with m implants, let $\boldsymbol{R}^* = \{R_k^* : k = 1, \ldots, m\}$ be the individual non-viable fetus indicators and denote by $\boldsymbol{y}^* = \{y_s^* : s = 1, \ldots, m - \sum_{k=1}^m R_k^*\}$ the malformation indicators for the viable

fetuses. Therefore, $R = \sum_{k=1}^{m} R_k^*$ and $y = \sum_{s=1}^{m-\sum_k R_k^*} y_s^*$. Then, the DDP mixture model for the clustered binary responses can be formulated as

$$f^*\left(\boldsymbol{R}^*, \boldsymbol{y}^* \mid m; G_{\mathcal{X}}\right) = \int \prod_{k=1}^{m} \text{Bern}\left(R_k^*; \pi(\gamma)\right) \prod_{s=1}^{m-\sum_k R_k^*} \text{Bern}\left(y_s^*; \pi(\theta)\right) \mathrm{d}G_{\mathcal{X}}(\gamma, \theta),$$

where $G_{\mathcal{X}} \mid \alpha, \psi \sim \text{DDP}(\alpha, G_{0\mathcal{X}})$ as above. Starting with the model for the binary responses given above, we obtain for the moment generating function:

$$\begin{aligned}
\mathrm{E}_{f*}(e^{t_1 \sum R_k^* + t_2 \sum y_s^*} \mid m; G_{\mathcal{X}}) &= \int \prod_{k=1}^{m} \sum_{R_k^*=0,1} e^{t_1 R_k^*} \text{Bern}(R_k^*; \pi(\gamma)) \prod_{s=1}^{m-\sum R_k^*} \sum_{y_s^*=0,1} e^{t_2 y_s^*} \text{Bern}(y_s^*; \pi(\theta)) \mathrm{d}G_{\mathcal{X}}(\gamma, \theta) \\
&= \int \left\{ \prod_{k=1}^{m} \sum_{R_k^*=0,1} e^{t_1 R_k^*} \text{Bern}(R_k^*; \pi(\gamma)) \right\} \{1 + \pi(\theta)(e^{t_2} - 1)\}^{m - \sum R_k^*} \mathrm{d}G_{\mathcal{X}}(\gamma, \theta) \\
&= \int \prod_{k=1}^{m} \sum_{R_k^*=0,1} e^{t_1 R_k^*} \text{Bern}(R_k^*; \pi(\gamma)) \{1 + \pi(\theta)(e^{t_2} - 1)\}^{1 - R_k^*} \mathrm{d}G_{\mathcal{X}}(\gamma, \theta) \\
&= \int \left[(1 - \pi(\gamma)) \{1 + \pi(\theta)(e^{t_2} - 1)\} + e^{t_1} \pi(\gamma) \right]^m \mathrm{d}G_{\mathcal{X}}(\gamma, \theta) \\
&= \int \sum_{R=0}^{m} \binom{m}{R} \left(\pi(\gamma) e^{t_1}\right)^R \left[(1 - \pi(\gamma)) \{1 + \pi(\theta)(e^{t_2} - 1)\} \right]^{m-R} \mathrm{d}G_{\mathcal{X}}(\gamma, \theta) \\
&= \int \sum_{R=0}^{m} e^{t_1 R} \text{Bin}(R; m, \pi(\gamma)) \left[\pi(\theta) e^{t_2} + (1 - \pi(\theta)) \right]^{m-R} \mathrm{d}G_{\mathcal{X}}(\gamma, \theta) \\
&= \int \sum_{R=0}^{m} e^{t_1 R} \text{Bin}(R; m, \pi(\gamma)) \sum_{y=0}^{m-R} e^{t_2 y} \text{Bin}(y; m - R, \pi(\theta)) \mathrm{d}G_{\mathcal{X}}(\gamma, \theta) \\
&\equiv \mathrm{E}_f(e^{t_1 R + t_2 y} \mid m; G_{\mathcal{X}}).
\end{aligned}$$

Hence, as in Section 3.1, we can define dose-response curves working with probabilities of the endpoints ("death" or "malformation") for a generic implant (thus involving implicit conditioning on $m = 1$).

In terms of studying certain properties of the DDP mixture model, we apply finite approximation to $G_{\mathcal{X}}$ (which will also provide a basis for MCMC posterior simulation). Specifically, we work with $G_{\mathcal{X}}^L = \sum_{l=1}^{L} p_l \delta_{(U_{l\mathcal{X}}, Z_{l\mathcal{X}})}$, where the $(U_{l\mathcal{X}}, Z_{l\mathcal{X}})$ are *i.i.d.* realizations from $G_{0\mathcal{X}}$ given ψ, and the weights p_l arise from the truncated

version of the stick-breaking process found in Chapter 2.

Dose-response curves and risk assessment

The first risk assessment quantity of interest is the probability of embryolethality across effective dose levels, which is defined by

$$D(x) \equiv \Pr(R^* = 1; G_x^L) = \int \pi(\gamma) \mathrm{d}G_x^L(\gamma, \theta) = \sum_{l=1}^{L} p_l \pi(U_l(x)), \quad x \in \mathcal{X}.$$

Risk assessment for the malformation endpoint is based on the conditional probability that a generic pup has a malformation given that it is a viable fetus, i.e.,

$$M(x) \equiv \Pr(y^* = 1 \mid R^* = 0; G_x^L) = \frac{\Pr(R^* = 0, y^* = 1; G_x^L)}{\Pr(R^* = 0; G_x^L)}, \quad x \in \mathcal{X}.$$

Here, $\Pr(R^* = 0, y^* = 1; G_x^L) = \sum_{l=1}^{L} p_l(1 - \pi(U_l(x)))\pi(Z_l(x))$. Moreover, a full risk function at any given dose level can be defined through the combination of the probability of a non-viable fetus and the probability of a live, malformed pup; that is, the combined risk at dose level x is given by

$$\begin{aligned} r(x) &\equiv \Pr(R^* = 1 \text{ or } y^* = 1; G_x^L) = \Pr(R^* = 0, y^* = 1; G_x^L) + \Pr(R^* = 1; G_x^L) \\ &= \int (1 - \pi(\gamma))\pi(\theta) \, \mathrm{d}G_x(\gamma, \theta) + \int \pi(\gamma) \, \mathrm{d}G_x(\gamma, \theta) \\ &= 1 - \sum_{l=1}^{L} p_l\{(1 - \pi(U_l(x)))(1 - \pi(Z_l(x))), \quad x \in \mathcal{X}. \end{aligned}$$

With $\xi_1 > 0$ and $\beta_1 > 0$ prior restrictions, we can build to both the probability of a non-viable fetus and to the combined risk function the non-decreasing trend in prior expectation.

Monotonicity of the dose-response curves: For $x < x'$, if $\xi_1 > 0$, then $0 \leq \mathrm{E}(\pi(U_l(x))) \leq \mathrm{E}(\pi(U_l(x'))) \leq 1$, following the result in Section 3.1. Similarly, if $\beta_1 > 0$, then $0 \leq \mathrm{E}(1 - \pi(Z_l(x'))) \leq \mathrm{E}(1 - \pi(Z_l(x))) \leq 1$. Therefore,for any $l = 1, \ldots, L$, $U_l(x)$ and $U_l(x')$ follow $\mathrm{N}(\xi_0 + \xi_1 x, \tau^2)$ and $\mathrm{N}(\xi_0 + \xi_1 x', \tau^2)$ distributions, respectively, and correspondingly, $Z_l(x)$ and $Z_l(x')$ follow $\mathrm{N}(\beta_0 + \beta_1 x, \sigma^2)$ and $\mathrm{N}(\beta_0 + \beta_1 x', \sigma^2)$ distributions. When $\xi_1, \beta_1 > 0$, then $U_l(x)$ is stochastically smaller than $U_l(x')$ and $Z_l(x)$ is stochastically smaller than $Z_l(x')$, independently across l and between U and Z. This implies that $\mathrm{E}(\Pr(R^* = 1; G_x^L)) \leq \mathrm{E}(\Pr(R^* = 1; G_{x'}^L))$, and

$$\sum_{l=1}^{L} \mathrm{E}(p_l)\mathrm{E}(1 - \pi(Z_l(x')))\mathrm{E}(1 - \pi(U_l(x'))) \leq \sum_{l=1}^{L} \mathrm{E}(p_l)\mathrm{E}(1 - \pi(Z_l(x)))\mathrm{E}(1 - \pi(U_l(x))),$$

and therefore, $\mathrm{E}(r(x)) \leq \mathrm{E}(r(x'))$.

Although the same argument does not extend to the conditional probability of malformation, $M(x)$, the restriction $\xi_1 > 0$ and $\beta_1 > 0$ appears sufficient to provide the prior expectation non-decreasing trend for all three dose-response curves. In this respect, it is useful to note that, even though we develop inference about three dose-response relationships, there are only two endpoints and, consequently, the model is driven at any specific dose level by a bivariate random mixing distribution.

Indeed, the linear mean functions are crucial for practicable posterior inference. As suggested by Figure 3.13, if the model is applied using constant mean functions for the DDP prior centering GPs (i.e., setting $\xi_1 = \beta_1 = 0$), we should not expect practically useful results outside the observed dose levels. For illustration, Figure 3.13 plots results from prior simulation for the embryolethality and malformation dose-response curves,

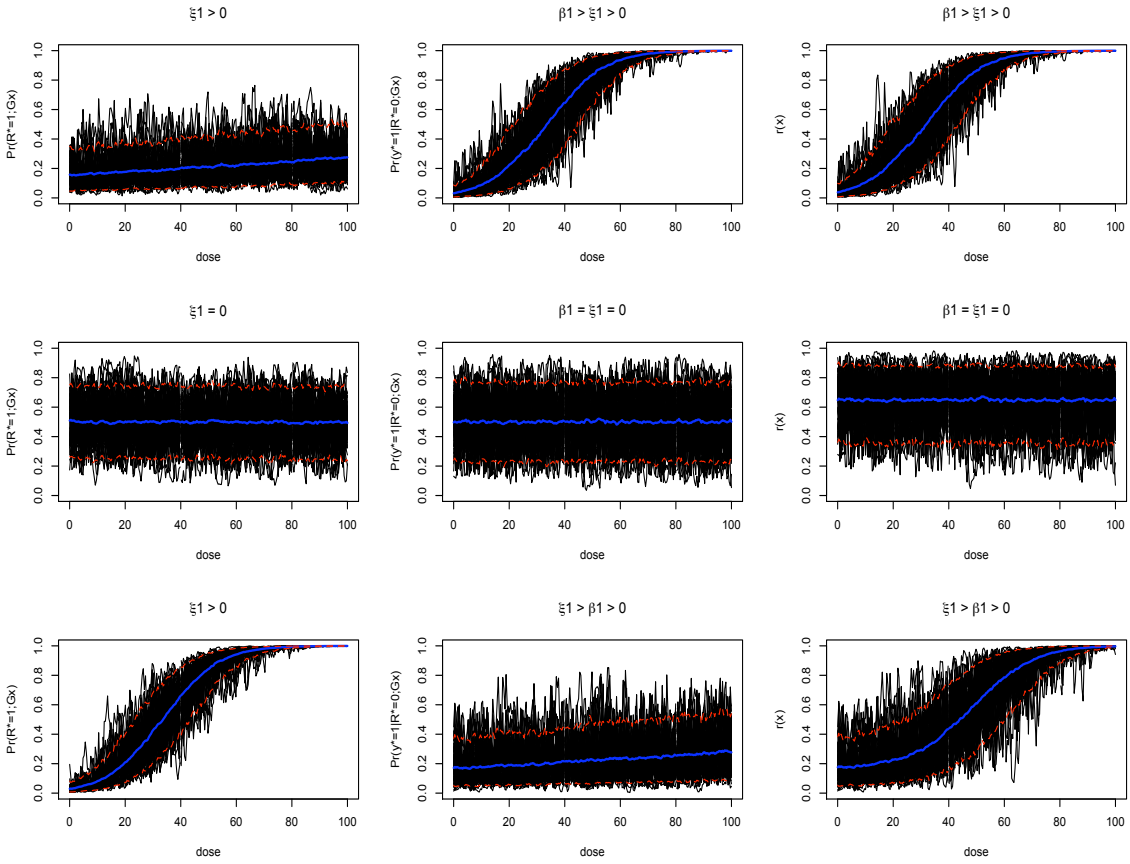

Figure 3.13: Prior mean and 90% interval estimates, along with 5 individual prior realizations, for the three dose-response curves.

and for the combined risk function, using fixed values for α $(= 1)$ and ψ. In particular, $(\xi_1 = 0.0085,\ \beta_1 = 0.12)$ and $(\xi_1 = 0.12,\ \beta_1 = 0.01)$ in the top and middle row, respectively. Although the relative magnitude of ξ_1 and β_1 affects the rate of increase for the different curves, in all cases with $\xi_1 > 0$ and $\beta_1 > 0$, the non-decreasing trend in prior expectation is preserved. As discussed previously, we argue this is an asset of the modeling approach; the analysis of the expanded version of the DEHP data set (see Section 3.2.4) highlights the practical utility of this model feature.

3.2.2 Posterior simulation and inference

Here, we consider the details for implementation and inference. Let y_{ij} be the number of malformed fetuses out of the $m_{ij} - R_{ij}$ viable fetuses. Now, under the truncated version of the mixing distributions, the model includes $G^L_{\boldsymbol{x}} = \sum_{l=1}^{L} p_l \delta_{(U_l(\boldsymbol{x}), Z_l(\boldsymbol{x}))}$, with $\boldsymbol{\gamma}_j = \boldsymbol{U}_l$ and $\boldsymbol{\theta}_j = \boldsymbol{Z}_l$ with probability p_l, given $\boldsymbol{U}_l = U_l(\boldsymbol{x})$ and $\boldsymbol{Z}_l = Z_l(\boldsymbol{x})$ in the same fashion as Chapter 2. Introducing configuration variables, w_j, the first stage of the hierarchical model is

$$\{\boldsymbol{R}_j, \boldsymbol{y}_j\} \mid \{\boldsymbol{m}_j\}, \boldsymbol{w}, \{\boldsymbol{U}_l\}, \{\boldsymbol{Z}_l\} \sim \prod_{j=1}^{n} \prod_{i=1}^{N} \left\{ \text{Bin}(R_{ij}; m_{ij}, \pi(U_{w_j}(x_i))) \text{Bin}(y_{ij}; m_{ij} - R_{ij}, \pi(Z_{w_j}(x_i))) \right\}^{s_{ij}}, \tag{3.7}$$

with $\boldsymbol{p} \mid \alpha \sim f(\boldsymbol{p} \mid \alpha)$ and $(U_l(\boldsymbol{x}), Z_l(\boldsymbol{x}))$ are i.i.d. from $G_{0\boldsymbol{x}}$ for $l = 1, \ldots, L$. Here, we place normal priors on ξ_0 and β_0, exponential priors to promote an increasing trend in the dose-response functions on ξ_1 and β_1, conjugate inverse-gamma priors on the variance terms τ^2 and σ^2, and uniform priors on the range parameters ρ and ϕ.

The change in the MCMC details from the DDP Binomial mixture to the multicategory model is that there are two, independent Gaussian processes to learn about. Therefore, the updates are the same as above, but doubled. The parameters sampled differently are the $\{w_j\}$, which come from a discrete distribution given by $\sum_{l=1}^{L} \tilde{p}_{lj} \delta_l(\cdot)$, where $\tilde{p}_{lj} \propto p_l \prod_{i=1}^{N} \{\text{Bin}(R_{ij}; m_{ij}, \pi(U_l(x_i))) \text{Bin}(y_{ij}; m_{ij} - R_{ij}, \pi(Z_l(x_i)))\}^{s_{ij}}$.

To augment the N known dose levels, we want to interpolate across M new dose levels, $\tilde{\boldsymbol{x}} = (\tilde{x}_1, \ldots, \tilde{x}_M)$. Now, each $U_l(\boldsymbol{x}) = (U_l(\boldsymbol{x}), \tilde{U}_l(\tilde{\boldsymbol{x}}))$ and $Z_l(\boldsymbol{x}) = (Z_l(\boldsymbol{x}), \tilde{Z}_l(\tilde{\boldsymbol{x}}))$, for $l = 1, \ldots, L$, where the $(\tilde{U}_l(\tilde{\boldsymbol{x}}), \tilde{Z}_l(\tilde{\boldsymbol{x}}))$ arise from the standard GP regression conditional normal distributions, discussed in Section 3.1.2.

Inference for the risk assessment quantities are estimated through the

evaluation of the dose-response curves discussed in the previous section. The curves are assessed at the observed doses as well as the unobserved dose levels through the full sample paths.

Prior specification

Prior specification for the centering GP parameters is done under the same process as in the combined negative outcomes. We set the prior means for ξ_0 and β_0 to 0, and the shape parameters of the inverse gamma priors for τ^2 and σ^2 to 2 (implying infinite prior variance). The prior variances for ξ_0 and β_0, and the prior means for ξ_1, β_1, τ^2 and σ^2 are chosen by studying the induced prior distribution for the dose-response curves. For the simulation study and the analysis of the DEHP data, we placed a $\mathrm{N}(0, 10)$ prior on ξ_0 and β_0, an exponential prior with mean $b_\xi^{-1} = b_\beta^{-1} = 0.1$ on ξ_1 and β_1, and an inv-gamma$(2, 10)$ on τ^2 and σ^2. Under this prior choice, the prior means for functions $D(x)$ and $M(x)$ have a relatively weak increasing trend starting around 0.5, with 90% uncertainty bands that cover almost the entire unit interval.

3.2.3 Simulation study

We work with two simulated data sets to study the performance of the model. In both cases, we use the dose levels, animals per dose level, and the number of implants from each dam from the DEHP data. For the first, simple simulation setting, we draw $R \mid m \sim \mathrm{Bin}(R; m, \pi(-3 + 0.001x))$ and $y \mid m, R \sim \mathrm{Bin}(y; m - R, \pi(-3 + 0.004x))$. The second setting involves a mixture of three binomials for the number of malformations. The number of non-viable fetuses are drawn from a binomial given the number of implants with probability of success $\pi(-3+0.08x)$; then, the number of malformations is drawn from a mixture of three binomials $0.1\mathrm{Bin}(y; m-R, \pi(-2+0.02x))+0.4\mathrm{Bin}(y; m-$

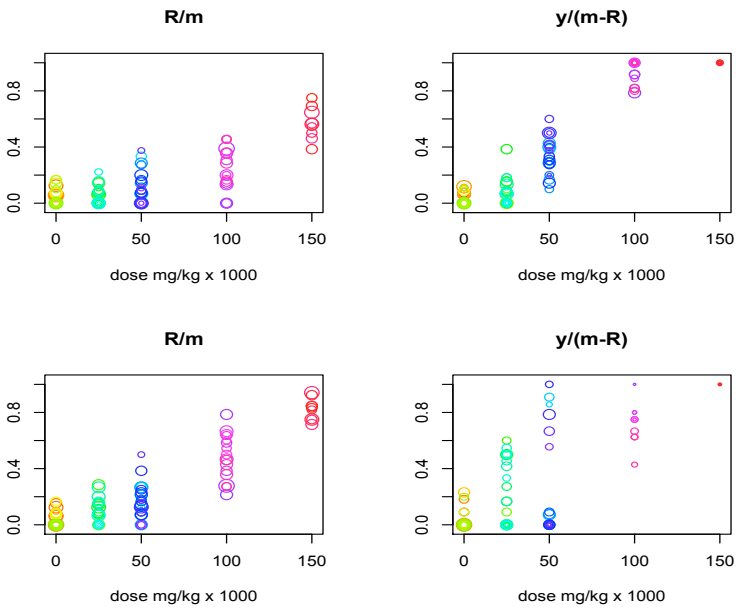

Figure 3.14: Simulation data. For the first simulation setting (top row) and second setting (bottom row), the left column gives the proportions of non-viable fetuses for each dam and the right column gives the conditional proportions of the number of malformations given the pups are viable for each dam. The size of the circle is proportional to the number of implants.

$R, \pi(-10+0.11x))+0.5\text{Bin}(y; m-R, \pi(-4+0.07x))$. The data from the two simulations are found in Figure 3.14.

Using the product of binomials DDP model and the priors found in 3.2.2, we obtain inference for the conditional probability mass functions and the risk assessment quantities discussed above. For the first simulation setting, the top row of Figure 3.15 gives the posterior mean (solid) and 90% intervals for the risk assessment quantities; $\Pr(R^* = 1; G_x^L)$ in the left panel, $\Pr(y^* = 1 \mid R^* = 0; G_x^L)$ in the middle panel, and the combined risk, $r(x)$, in the right panel. The probability of a non-viable fetus has a weak signal and therefore includes larger uncertainty. However, the true curves (given in green) are contained within the 90% intervals for all three quantities. Figure 3.16

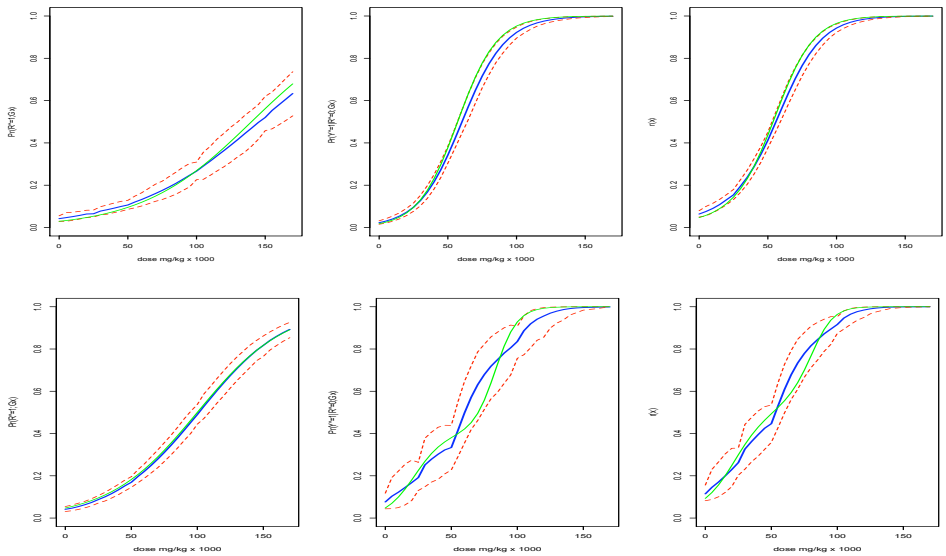

Figure 3.15: Simulation data. For the first simulation setting (top row) and second setting (bottom row), the posterior mean (solid line) and 90% interval bands (dashed lines) for the risk assessment quantities: the probability of a non-viable fetus in the left column, the conditional probability of malformation in the middle column, and the combined risk in the right column. The green lines represent the true curves.

includes the posterior mean ("o") and 90% uncertainty bands (dashed lines) for the probability mass functions of the number of non-viable fetuses given $m = 12$, with the truth shown as black "x". As expected, the shapes are unimodal and roughly follow the underlying distributional shapes. Finally, we plot the conditional probability mass functions for the number of malformations given $m = 12$ and $R = 4$ in Figure 3.17. Again, the DDP model captures the unimodal shapes with relatively small uncertainty.

The second simulation setting is more interesting because of the mixture for the conditional probability of malformation. The risk assessment quantities are picked up by the model, with larger probability bands in the malformation dose response curve to handle the mixture. The true curves, given in green, are basically enclosed by the 90% probability intervals (see bottom row of Figure 3.15). The probability mass functions

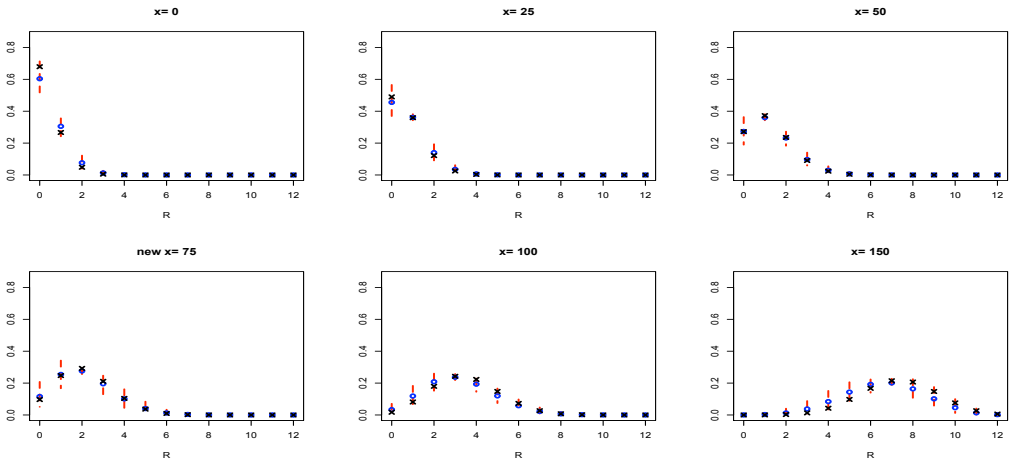

Figure 3.16: Simulation 1. The posterior mean ("o") and 90% probability bands (dashed lines) of the probability mass functions for the number of non-viable fetuses given $m = 12$ implants. The true values are given as "x".

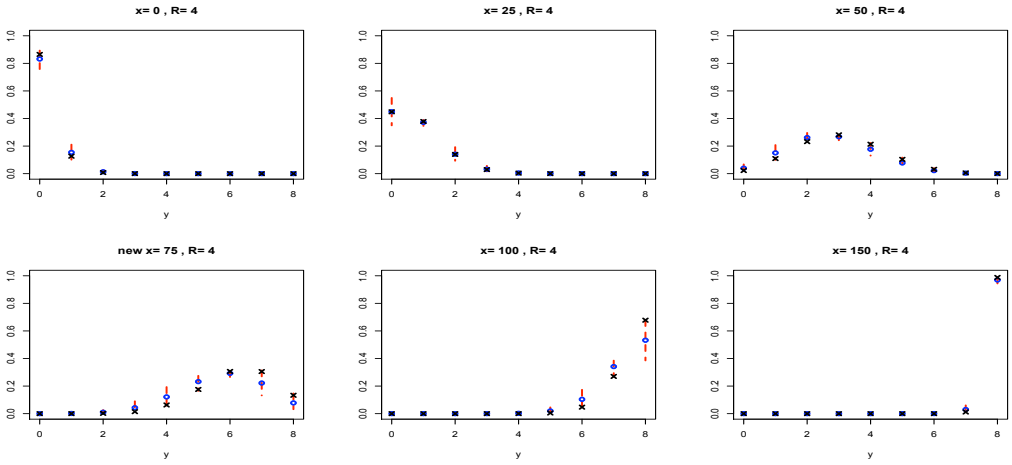

Figure 3.17: Simulation 1. The posterior mean ("o") and 90% probability bands (dashed lines) of the probability mass functions for the number of malformations given $m = 12$ implants and $R = 4$ non-viable fetuses. The true values are given as "x".

for the number of non-viable fetuses are plotted in Figure 3.18 are again unimodal across the range of dose levels, and lie within the 90% uncertainty bands. Finally, the estimated conditional probability mass functions for the number of malformations given $m = 12$ and $R = 3$ are plotted in Figure 3.19. The high and low dose levels have shapes

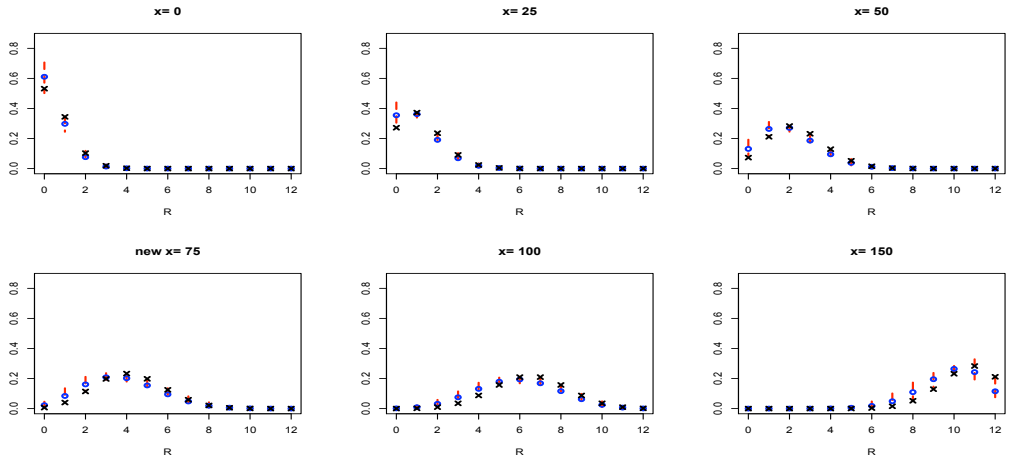

Figure 3.18: Simulation 2. The posterior mean ("o") and 90% probability bands (dashed lines) of the probability mass functions for the number of non-viable fetuses given $m = 12$ implants. The true values are given as "x".

that would be picked up by a simple parametric model, but there are bimodalities in the midrange that would be ignored. The DDP model adapts to appropriately model both the unimodal and multimodal distributions, with increasing variability in the estimated mass functions where needed.

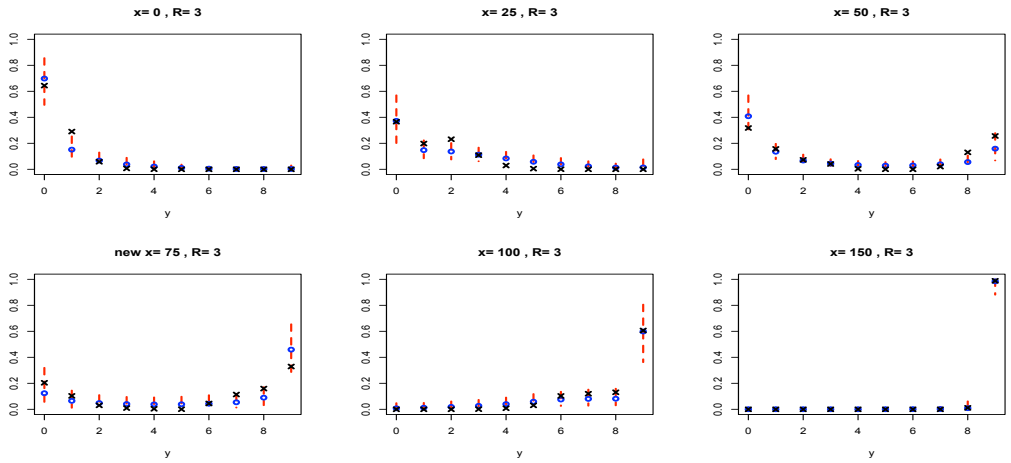

Figure 3.19: Simulation 2:. The posterior mean ("o") and 90% probability bands (dashed lines) of the probability mass functions for the number of malformations given $m = 12$ implants and $R = 3$ non-viable fetuses. The true values are given as "x".

3.2.4 DEHP data revisited

We revisit the analysis of the DEHP data with the expansion of the combined negative outcomes category to the two endpoints "non-viable fetus" (resorption or actual prenatal death) and "malformation" (external, visceral or skeletal malformation of a live fetus). Figure 3.20 gives the posterior mean and 90% interval estimates for the dose-response curves for the risk assessment. In the left panel, the probability of a non-viable fetus across the dose levels is a monotonically increasing function. The conditional probability of malformation, however, retains the non-monotonic behavior at the low dose levels, and this shape carries over to the combined risk which also exhibits the dip in the probability from the control through dose 25 mg/kg $\times$ 1000. The inference for the combined risk function agrees with the estimated dose-response curve for the combined negative outcomes version of the DEHP data, as obtained in Section 3.1.3. The modelling approach is key to uncovering the malformation endpoint as the one that contributes to the non-monotonic, possibly hormetic, combined dose-response relationship.

Inference for response distributions is illustrated with posterior mean and 90% interval estimates for the probability mass function of the number of non-viable fetuses given $m = 12$ implants (Figure 3.21) and the number of malformations given $m = 12$ implants and $R = 3$ non-viable fetuses (Figure 3.22). Results are reported for the control group, the four effective dose levels, and a new dose at $x = 75$ mg/kg $\times$ 1000. As expected, there is more uncertainty in the estimation of the conditional response distributions for malformation. The interpolation at the new dose level appears to be influenced more by the distribution at dose 50, which can be attributed to the larger sample size relative to dose 100. The estimated distributions for the number of non-

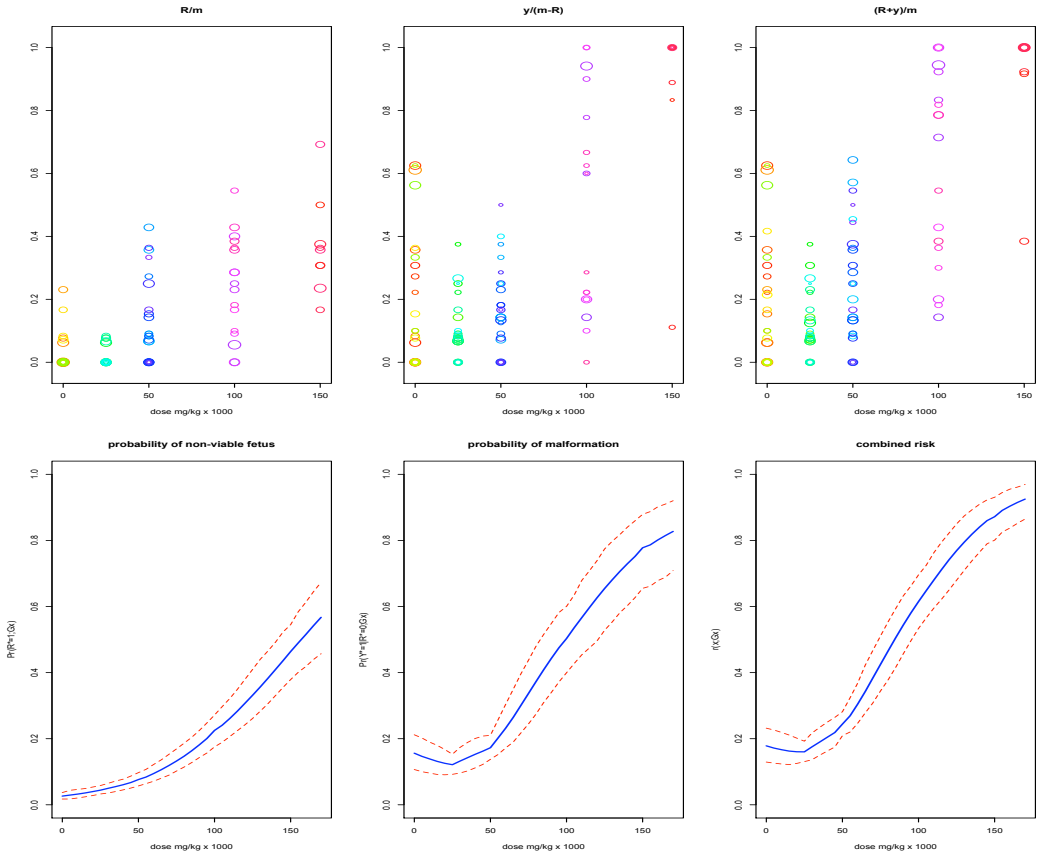

Figure 3.20: Expanded DEHP data. The top row includes the data as in Figure 1.2. The bottom row plots the posterior mean (solid lines) and 90% interval bands (dashed lines) for the risk assessment functions: probability of a non-viable fetus (left panel); conditional probability of malformation (middle panel); combined risk (right panel).

viable fetuses have relatively standard shapes, whereas there is some evidence of a bimodal shape (at dose 100) and skewness in the estimated malformation distributions.

3.2.5 Discussion

We have extended the DDP Binomial mixture model to take into account the expansion of the combined negative outcomes, a scientifically relevant, yet underdeveloped, extension for developmental toxicity studies. The model highlights

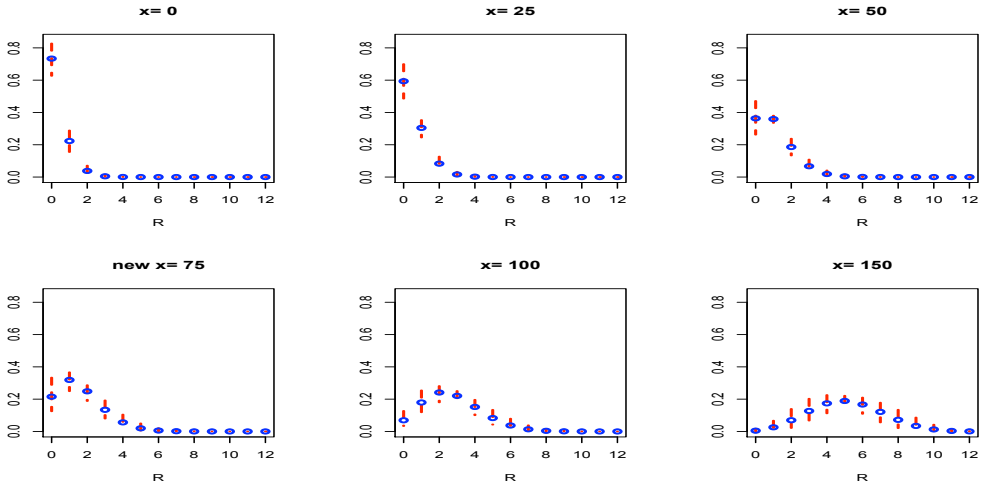

Figure 3.21: Expanded DEHP data. The posterior mean ("o") and 90% probability bands (dashed lines) of the probability mass functions for the number of non-viable fetuses given $m = 12$ implants.

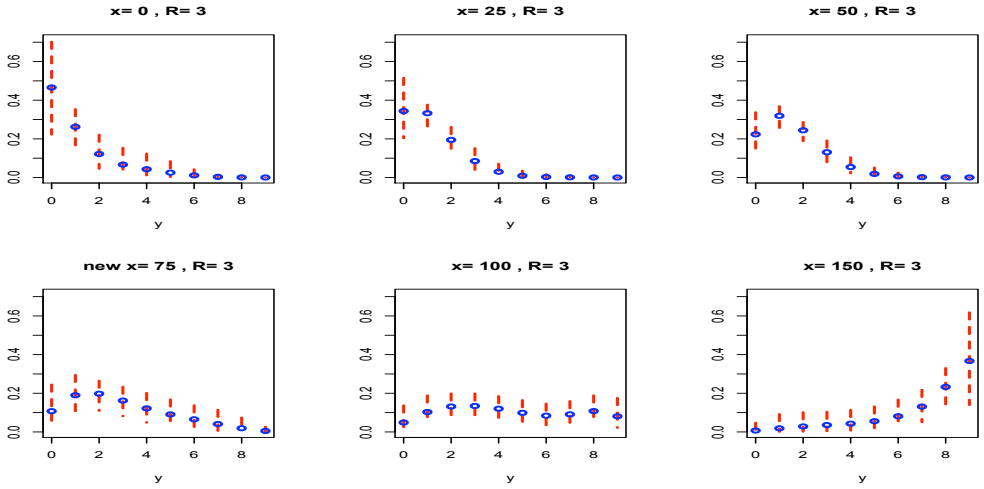

Figure 3.22: Expanded DEHP data. The posterior mean ("o") and 90% probability bands (dashed lines) of the probability mass functions for the number of malformations given $m = 12$ implants and $R = 3$ non-viable fetuses.

the nested structure of the data, and yields rich inference for the response distributions and the different dose-response relationships for different endpoints. A data example previously analyzed under the combined endpoints model expounds the practical utility

of the modeling approach with regard to capturing non-standard, "J"-shaped dose-response curves when applicable.

3.3 Joint continuous-discrete outcomes

Including a continuous outcome, such as fetal weight, for each of the live pups to the DDP mixture model setting for multicategory responses presents a challenge in that there are now clustered outcomes that include discrete and continuous responses. There is a plethora of likelihood based models (e.g., Catalano and Ryan, 1992; Regan and Catalano, 1999; Gueorguieva and Agresti, 2001); however, these models require parametric assumptions that are arguably restrictive and can be limited with regard to uncertainty quantification for risk assessment. Regarding Bayesian work, we are aware of only two parametric approaches. Dunson et al. (2003) propose a joint model for the number of viable fetuses and multiple discrete-continuous outcomes. A continuation-ratio ordinal response model is used for the number of viable fetuses and the multiple outcomes are assigned an underlying normal model with shared latent variables within outcome-specific regression models. In Faes et al. (2006), the proposed model is expressed in two stages; the first to assess the probability that a fetus is non-viable and the second determines the probability that a viable fetus has a malformation and/or suffers from low birth weight as function of the number of viable fetuses while also accounting for intralitter correlation.

3.3.1 The probability model

To incorporate a continuous response, say, fetal weight $\boldsymbol{u}^* = \{u_k^* : k = 1, \ldots, m - R\}$, for each live pup from a generic dam with m implants and R non-viable

fetuses at dose x, we must consider the corresponding vector of binary malformation responses, denoted by $\boldsymbol{y}^* = \{y_k^* : k = 1, \ldots, m - R\}$. Then, a possible DDP mixture model for such clustered binary and continuous outcomes is:

$$f(R, \boldsymbol{y}^*, \boldsymbol{u}^* \mid m; G_{\mathcal{X}}) = \int \text{Bin}\left(R; m, \pi(\gamma)\right) \prod_{k=1}^{m-R} \text{Bern}\left(y_k^*; \pi(\theta)\right) \text{N}\left(u_k^*; \mu, \varphi\right) dG_{\mathcal{X}}(\gamma, \theta, \mu), \tag{3.8}$$

with $G_{\mathcal{X}} \mid \alpha, \psi \sim \text{DDP}(\alpha, G_{0\mathcal{X}})$, where $G_{0\mathcal{X}}$ is a product of three Gaussian processes, one for each mixing parameter. Specifically, the GP prior associated with γ has a linear mean function, $\xi_0 + \xi_1 x$, variance τ^2, and correlation function $\exp\{-\rho|x - x'|\}$; the linear mean function for θ is $\beta_0 + \beta_1 x$, the variance σ^2, and correlation function $\exp\{-\phi|x - x'|\}$; and the GP prior on μ includes a mean function $\eta_0 + \eta_1 x$, variance ν^2, and covariance function $\exp\{-\kappa|x - x'|\}$. Thereby, the GP hyperparameters are denoted by $\psi = (\beta_0, \beta_1, \sigma^2, \phi, \xi_0, \xi_1, \tau^2, \rho, \eta_0, \eta_1, \nu^2, \kappa)$.

Note that the proposed model for $f(R, \boldsymbol{y}^*, \boldsymbol{u}^* \mid m; G_{\mathcal{X}})$ provides the most natural extension of the models from sections 3.2 and 3.1. Without the continuous response and aggregating the binary malformation responses, $y = \sum_{k=1}^{m-R} y_k^*$, the moment generating argument in Section 3.3 essentially yields the multicategory model, $f(R, y \mid m; G_{\mathcal{X}})$. Further collapsing by combining the number of non-viable fetuses and the number of malformations, we arrive at the model for combined negative outcomes in Section 3.1.

We could, in principal, mix on the variance of the fetal weight kernel, φ, as well. This approach sacrifices the ability to promote an increasing trend in the full, combined risk function and involves more complex model fitting as we require a third GP centering process for the DDP prior. Therefore, to strike a balance between flexibility

and feasibility, we adopt the more pragmatic location normal mixture component for the continuous endpoint.

Functionals for risk assessment

Of key importance is study of the various dose-response relationships for risk assessment. We use the truncation approximation found in Chapter 2, such that for any dose level x, $\gamma_j(x) = V_l(x)$, $\theta_j(x) = Z_l(x)$, and $\mu_j(x) = T_l(x)$ with probability p_l. Again, the following functions are in terms of a generic implant, thereby inducing an implicit conditioning on $m = 1$ (though it is excluded from the expressions).

The first dose-response curve is the probability of a non-viable fetus across effective dose levels,

$$D(x) \equiv \Pr(R^* = 1; G_x^L) = \sum_{l=1}^{L} p_l \pi(V_l(x)), \quad x \in \mathcal{X}.$$

As in the multicategory model of Section 3.2, provided $\xi_1 > 0$, $D(x)$ is increasing in prior expectation.

We also have the probability that a generic pup has a malformation given that it is a viable fetus, i.e., $M(x) \equiv \Pr(y^* = 1 \mid R^* = 0; G_x^L) = \Pr(y^* = 1, R^* = 0; G_x^L)/\Pr(R^* = 0; G_x^L)$, where $\Pr(y^* = 1, R^* = 0; G_x^L) = \sum_{l=1}^{L} p_l(1 - \pi(V_l(x)))\pi(Z_l(x))$. Thus, the second dose-response curve for malformation is

$$M(x) = \frac{\sum_{l=1}^{L} p_l(1 - \pi(V_l(x)))\pi(Z_l(x))}{\sum_{l=1}^{L} p_l(1 - \pi(V_l(x)))}, \quad x \in \mathcal{X}.$$

For the continuous outcome, we may look at the conditional expectation of weight, $\mathrm{E}(u^* \mid R^* = 0; G_x^L) = \int u^* f(u^* \mid R^* = 0; G_x^L)\mathrm{d}u^*$. Therefore, evaluating the integral based on the mixture representation for $f(R^* = 0, u^*; G_x^L)$, the third risk

assessment quantity of interest for $x \in \mathcal{X}$ can be expressed as

$$\mathrm{E}(u^* \mid R^* = 0; G_x^L) = \frac{\sum_{l=1}^{L} p_l T_l(x)(1 - \pi(V_l(x)))}{\sum_{l=1}^{L} p_l (1 - \pi(V_l(x)))}.$$

Alternatively, we can look at the probability of the pup being of low birth weight. We follow Regan and Catalano (1999) such that for a pup to be considered of "low fetal weight," we define the cutoff, $\mathcal{U}$, to be two standard deviations below the average birth weight at the control level. This probability is defined as $\Pr(u^* < \mathcal{U} \mid R^* = 0; G_x^L) = \int^{\mathcal{U}} f(u^* \mid R^* = 0; G_x^L) \mathrm{d}u^*$. The curve for all $x \in \mathcal{X}$ is then evaluated through

$$\frac{\sum_{l=1}^{L} p_l (1 - \pi(V_l(x))) \Phi((\mathcal{U} - T_l(x))/\varphi^{1/2})}{\sum_{l=1}^{L} p_l (1 - \pi(V_l(x)))}.$$

We can also study the combined risk of the discrete outcomes; that is, we look at the probability of embryolethality or malformation across the dose levels:

$$r_d(x) = \Pr(R^* = 1 \text{ or } y^* = 1; G_x^L) = 1 - \sum_{l=1}^{L} p_l (1 - \pi(V_l(x)))(1 - \pi(Z_l(x))).$$

This is equivalent to the combined risk function in the multicategory response DDP mixture model (Section 3.2), and therefore the prior expectation $\mathrm{E}(r_d(x))$ is increasing if both $\xi > 0$ and $\beta_1 > 0$.

In terms of the full risk function, $r_f(x)$, we combine the probability of embryolethality and the probability a pup is adversely affected (malformed and/or low birth weight) by the toxin at dose x. We first separate the negative outcomes for the discrete and continuous outcomes, where the first probability is then equivalent to $r_d(x)$ and the second integrates $f(R^* = 0, y^* = 0, u^*; G_x^L)$ with respect to u^* up to $\mathcal{U}$.

Therefore, the expression at any dose level $x \in \mathcal{X}$ becomes

$$\begin{aligned} r_f(x) &= \Pr(R^* = 1 \text{ or } y^* = 1 \text{ or } u^* < \mathcal{U}; G_x^L) \\ &= \Pr(R^* = 1 \text{ or } y^* = 1; G_x^L) + \Pr(R^* = 0, y^* = 0, u_k^* < \mathcal{U}; G_x^L) \\ &= 1 - \sum_{l=1}^{L} p_l \{1 - \pi(V_l(x))\}\{1 - \pi(Z_l(x))\}\{1 - \Phi((\mathcal{U} - T_l(x))/\varphi^{1/2})\}. \end{aligned}$$

Given this expression for the combined risk, we consider the prior expectation

$$\begin{aligned} \mathrm{E}(r_f(x)) = 1 - \sum_{l=1}^{L} \mathrm{E}(p_l) \big\{ & \mathrm{E}_{\mathrm{N}(\xi_0+\xi_1 x, \tau^2)}(1 - \pi(V_l(x))) \mathrm{E}_{\mathrm{N}(\beta_0+\beta_1 x, \sigma^2)}(1 - \pi(Z_l(x))) \\ & \mathrm{E}_{\mathrm{N}(\eta_0+\eta_1 x, \nu^2)}(1 - \Phi((\mathcal{U} - T_l(x))/\varphi^{1/2})) \big\}. \end{aligned}$$

Provided $\xi_1 > 0$, $\beta_1 > 0$, and $\eta_1 < 0$, the three normal distributions are stochastically ordered in x and it follows that $\mathrm{E}(1 - \pi(V_l(x)))$, $\mathrm{E}(1 - \pi(Z_l(x)))$, and $\mathrm{E}(1 - \Phi((\mathcal{U} - T_l(x))/\varphi^{1/2}))$ are decreasing functions of x (with values in the unit interval). Thus, $\mathrm{E}\{r_f(x)\}$ is monotonically increasing in x under the given restrictions for ξ_1, β_1, and η_1.

Thus, we can promote increasing trends in the probability of embryolethality and both combined risk functions. As in the multicategory setting, this argument does not hold in the conditional probabilities and expectation, $M(x)$, $\mathrm{E}(u^* \mid R^* = 0; G_x^L)$, and $\Pr(u^* < \mathcal{U} \mid R^* = 0; G_x^L)$. However, empirical evidence suggests that the $\xi_1 > 0$, $\beta_1 > 0$, and $\eta_1 < 0$ restrictions induce non-decreasing prior expectations for all dose-response curves (see Figure 3.23).

Finally, we investigate the correlation between the discrete malformation endpoints for two viable pups, $\mathrm{Corr}(y_k^*, y_{k'}^* \mid R_k^* = 0, R_{k'}^* = 0; G_x^L)$, between the continuous endpoints of two viable pups, $\mathrm{Corr}(u_k^*, u_{k'}^* \mid R_k^* = 0, R_{k'}^* = 0; G_x^L)$, and

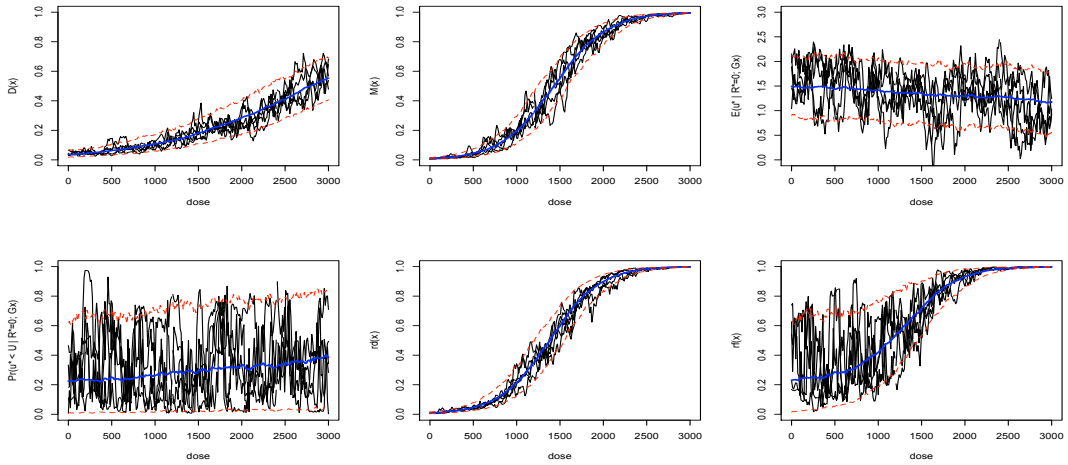

Figure 3.23: Prior mean (blue) and 90% interval bands (red dashed) for the six dose-response relationships with hyperparameters fixed such that $\xi_1 > 0$, $\beta_1 > 0$ and $\eta_1 < 0$: $D(x)$ top left panel, $M(x)$ top middle, $\mathrm{E}(u^* \mid R^* = 0; G_x^L)$ top right, $\Pr(u_k^* < \mathcal{U} \mid R^* = 0; G_x^L)$ bottom left, $r_d(x)$ bottom middle, and $r_f(x)$ bottom right.

between the discrete and continuous endpoints within one viable fetus at a given dose level x, $\mathrm{Corr}(y^*, u^* \mid R^* = 0; G_x^L)$.

We begin with the correlation between the two endpoints within a generic pup, where we have implicit conditioning on $m = 1$. Thus, we have already developed expressions for $\mathrm{E}(y^* \mid R^* = 0; G_x^L)$ (which is equivalent to $\Pr(y^* = 1 \mid R^* = 0; G_x^L)$, or $M(x)$), $\mathrm{E}(y^{*2} \mid R^* = 0; G_x^L)$ (also equal to $M(x)$), and $\mathrm{E}(u^* \mid R^* = 0; G_x^L)$. For $\mathrm{E}(u^{*2} \mid R^* = 0; G_x^L)$, we follow the derivations found above and obtain $\{\sum_{l=1}^{L} p_l(1 - \pi(V_l(x)))[T_l^2(x) + \varphi]\} / \sum_{l=1}^{L} p_l(1 - \pi(V_l(x)))$. The final component is $\mathrm{E}(y^* u^* \mid R^* = 0; G_x^L)$, a double integral over $f(y^*, u^* \mid R^* = 0; G_x^L)$, which is given by $[\sum_{l=1}^{L} p_l(1 - \pi(V_l(x)))\pi(Z_l(x))T_l(x)] / \sum_{l=1}^{L} p_l(1 - \pi(V_l(x)))$. Through the definition of covariance and variance, we construct the correlation with the above expressions.

The correlations of the different endpoints between two pups have an implicit conditioning on $m = 2$. Therefore, we have generic responses $(R_k^*, R_{k'}^*)$, $(y_k^*, y_{k'}^*)$, and

$(u_k^*, u_{k'}^*)$. Expressions have already been derived for $\mathrm{E}(y_k^* \mid R_k^* = 0; G_x^L) = \mathrm{E}(y_{k'}^* \mid R_{k'}^* = 0; G_x^L)$ and $\mathrm{E}(u_k^* \mid R_k^* = 0; G_x^L) = \mathrm{E}(u_{k'}^* \mid R_{k'}^* = 0; G_x^L)$, as well as $\mathrm{E}(u_k^{*2} \mid R_k^* = 0; G_x^L)$ and $\mathrm{E}(y_k^{*2} \mid R_k^* = 0; G_x^L)$. To build the final two correlations, we need two additional expectations: $\mathrm{E}(y_k^* y_{k'}^* \mid R_k^* = 0, R_{k'}^* = 0; G_x^L)$ and $\mathrm{E}(u_k^* u_{k'}^* \mid R_k^* = 0, R_{k'}^* = 0; G_x^L)$. The expectations are $\{\sum_{l=1}^L p_l(1 - \pi(V_l(x)))^2 \pi^2(Z_l(x))\}/[\sum_{l=1}^L p_l(1 - \pi(V_l(x)))^2]$ and $\{\sum_{l=1}^L p_l(1 - \pi(V_l(x)))^2 T_l^2(x)\}/[\sum_{l=1}^L p_l(1 - \pi(V_l(x)))^2]$, respectively. Again, we calculate the dose-dependent correlations through the definition of variance and covariance.

3.3.2 Posterior simulation and inference

Turning to implementation of the model, let $\boldsymbol{m}_j = \{m_{1j}, \ldots, m_{Nj}\}$ denote the number of implants for the jth dam at dose $x_i,\ i = 1, \ldots, N$, and $\boldsymbol{R}_j = (R_{1j}, \ldots, R_{Nj})$ be the number of prenatal deaths in the jth replicate, for $j = 1, \ldots, n$. We also let $\boldsymbol{y}_{ij}^*$ and $\boldsymbol{u}_{ij}^*$ be $(m_{ij} - R_{ij}) \times 1$ vectors of viable pup specific binary and continuous responses, respectively. Following the discussion in Chapter 2 and the previous section, we can write the first stage of the model for the $(\boldsymbol{R}_j, \boldsymbol{y}_j^*, \boldsymbol{u}_j^*)$ as

$$\{\boldsymbol{R}_j, \boldsymbol{y}_j^*, \boldsymbol{u}_j^*\} \mid \boldsymbol{m}, w_j, V(\boldsymbol{x}), Z(\boldsymbol{x}), T_l(\boldsymbol{x}), \varphi \sim \prod_{j=1}^{n} \prod_{i=1}^{N} \left\{\mathrm{Bin}(R_{ij}; m_{ij}, \pi(V_{w_j}(x_i)))\right\}^{s_{ij}} \times \prod_{k=1}^{m_{ij}-R_{ij}} \left(\mathrm{Bern}(y_{ijk}^*; \pi(Z_{w_j}(x_i)))\mathrm{N}(u_{ijk}^*, T_{w_j}(x_i), \varphi)\right)^{s_{ij}}$$

where $G_{0\boldsymbol{x}}(V_l(\boldsymbol{x}), Z_l(\boldsymbol{x}), T_l(\boldsymbol{x}) \mid \psi)$ comprises three independent N-variate normal distributions induced by the corresponding GPs used to define $G_{0\mathcal{X}}$. We complete the model with normal priors on the intercepts of the GP linear mean functions, gamma priors on ξ_1 and β_1, a truncated normal for η_1, inverse-gamma priors on the variance terms, and uniform priors on the parameters associated with the range of dependence.

In terms of MCMC posterior simulation, the updates for the GP corresponding to the non-viable fetus and malformation endpoints are similar to those for the multicategory model. The GP hyperparameters of the fetal weight endpoint are the same as in the other GPs, the points of contrast are in the sample paths, $T_l(\boldsymbol{x})$, φ, and the $\{w_j\}$.

The full conditional for $\boldsymbol{T}_l$ depends on whether l corresponds to one of the distinct components. If $l \notin \{w_k^* : k = 1, \ldots, n^*\}$, then $\boldsymbol{T}_l$ is drawn from its prior distribution, $\mathrm{N}_N(\eta_0 \boldsymbol{j}_N + \eta_1 \boldsymbol{x}, \Omega)$. For $l \in \{w_k^* : k = 1, \ldots, n^*\}$,

$$\boldsymbol{T}_{w_k^*} \mid \boldsymbol{w}, \psi, \text{data} \propto \mathrm{N}_N(\boldsymbol{T}_{w_k^*}; \eta_0 \boldsymbol{j}_N + \eta_1 \boldsymbol{x}, \Omega) \prod_{\{j: w_j = w_k^*\}} \prod_{i=1}^{N} \prod_{k=1}^{m_{ij}-R_{ij}} \left\{ \mathrm{N}\left(u_{ij}^*; T_{w_k^*}(x_i), \varphi\right) \right\}^{s_{ij}}$$

For the w_j, the updates come from a discrete distribution given by $\sum_{l=1}^{L} \tilde{p}_{lj} \delta_l(\cdot)$, where $\tilde{p}_{lj} \propto p_l \prod_{i=1}^{N} \{\mathrm{Bin}(R_{ij}; m_{ij}, \pi(V_l(x_i))) \prod_{k=1}^{m_{ij}-R_{ij}} \mathrm{Bern}(y_{ijk}^*; \pi(Z_l(x_i))) \mathrm{N}(u_{ijk}^*; T_l(x_i), \varphi)\}^{s_{ij}}$. Parameter φ is the variance of the normal kernel of the weights. Therefore, given an inverse-gamma prior, with shape parameter $a_\varphi > 1$ and mean $b_\varphi/(a_\varphi - 1)$, φ has an inverse-gamma full conditional with parameters $a_\varphi + 0.5 \sum_{i=1}^{N} \sum_{j=1}^{n_i} (m_{ij} - R_{ij})$ and $b_\varphi + 0.5 \sum_i \sum_j \sum_k (T_{w_j}(x_i) - u_{ijk}^*)^2$.

Given the MCMC posterior samples, we obtain inference by evaluating the expressions above for risk assessment quantities. Also of interest are the various predictive probability mass functions and density functions at any given dose level, x. For the unobserved dose levels, we use the same procedure as in the previous sections for interpolation of the GP sample paths, including the conditional normal draws for the $(\tilde{\boldsymbol{V}}_l, \tilde{\boldsymbol{Z}}_l, \tilde{\boldsymbol{T}}_l)$.

3.3.3 EG data

In a study found at the National Toxicology Program database, ethylene glycol, a widely used industrial chemical, is evaluated for toxic effects in pregnant mice (Price et al., 1985). See Section 1.2 for a more detailed summary of the data set.

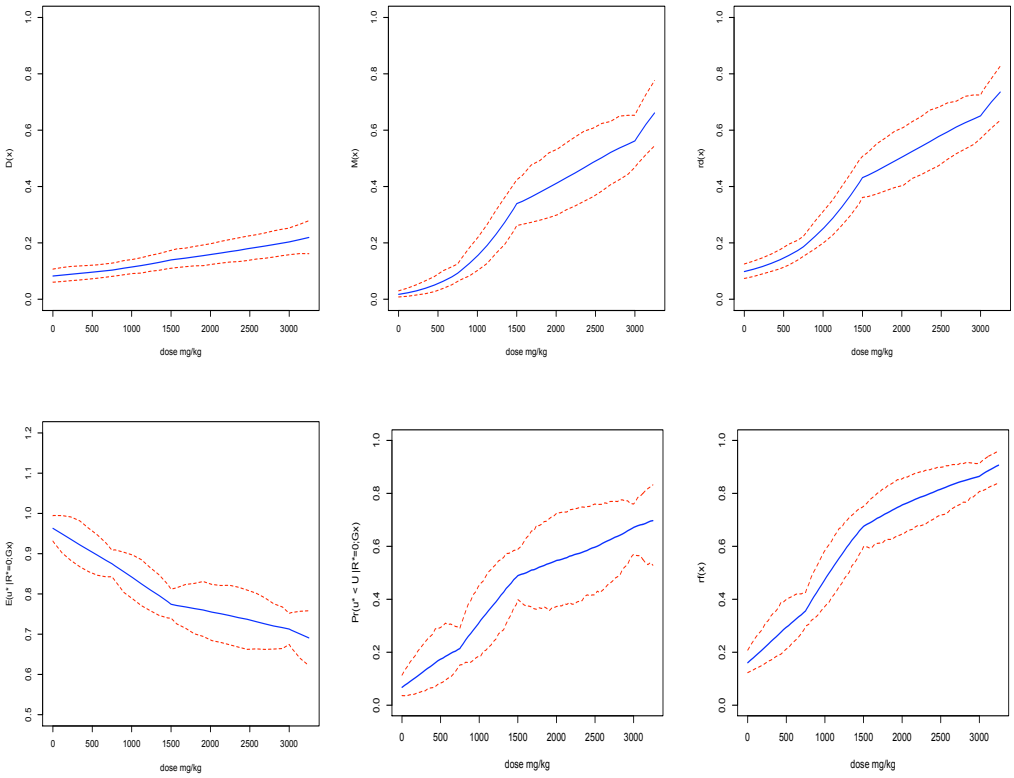

Figure 3.24: EG data. The posterior mean (solid) and 90% uncertainty bands (dashed lines) of the probability of a non-viable fetus (top left), the conditional probability of malformation (top middle), the risk of combined discrete endpoints (top right), expected birth weight (bottom left), the conditional probability of low birth weight (bottom middle), and the full combined risk (bottom right).

Figure 3.24 gives the posterior mean and probability bands of the dose-response relationships. The probability of embryolethality has a moderate increasing trend and is approximately linear. The conditional probability of malformation has a roughly

exponential curve at the lower dose levels, then becomes more linear after dose 1500 mg/kg. The extrapolated values after 3000 mg/kg are linear with a stronger slope. Thus, we have a dose-response relationship with structural changes at multiple points throughout the dose range, a phenomenon not possible under many models. The combined risk function for the discrete outcomes is similar to the $M(x)$ function, though shifted up slightly. The combined risk for the discrete outcomes is mainly controlled by the conditional probability of malformation, which reflects the modest increasing trend in $D(x)$.

The expected weight of a generic pup has a decreasing trend which follows the data, including linear patterns with varying slopes between the observed dose levels. For the EG data set, the cutoff representing low fetal weight is 0.771g. For the probability of low fetal weight, we see roughly increasing, linear forms between the observed dose levels which coincide with the structures found in the expected fetal weight. Moreover, the uncertainty estimates become wider as the dose level increases. The full risk function combines the risk of the discrete outcomes (top right panel) with the probability of low fetal weight, and therefore is always greater than $r_d(x)$. The probability of low fetal weight dominates $r_f(x)$ at the lower dose levels and is tempered by the risk of embryolethality or malformation as dose increases.

By evaluating the expressions involved in the correlations with the posterior samples from the MCMC algorithm, we obtain the posterior densities for the three correlations. The left panel of Figure 3.25 depicts the posterior densities of the correlation between the discrete (malformation) and continuous (fetal weight) outcomes. The densities change across dose levels, with a decreasing mean. At the control level, the correlation is centered about zero with relatively narrow spread. Moving towards the larger doses, the densities shift to the left and the spread gets larger. Hence, at

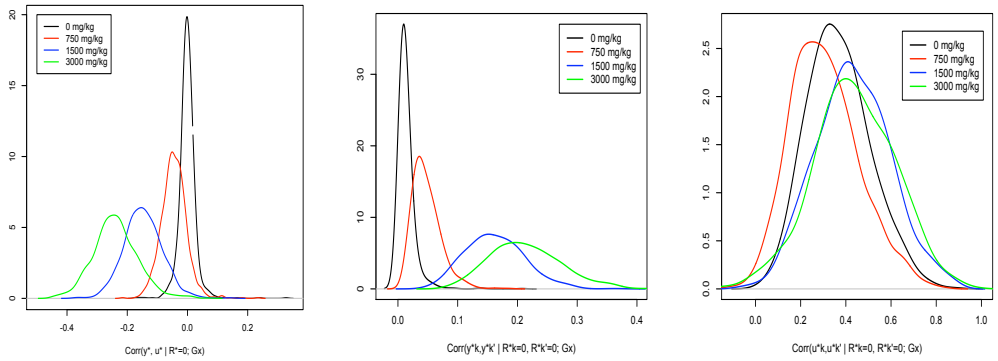

Figure 3.25: EG data. Posterior densities of the correlation between the discrete and continuous outcomes of a generic pup at the observed dose levels (right), the malformation responses of two pups within a litter (middle), and the fetal weight endpoints of two pups within a litter (right).

the high dose levels, a decrease in fetal weight generally corresponds to an increase in the malformation response. The correlation between the malformation responses of two pups within a litter at the four observed dose levels is found in the middle panel of Figure 3.25. While always positive, the center and spread of the correlations increase as the dose level increases. The correlation between the fetal weight endpoints of two pups within a litter do not change drastically at the different dose levels. The breadth of the distributions widens slightly as dose level increases, but the centers are roughly equivalent around 0.4 (see the right panel of 3.25). Traditional parametric models yield a constant correlation for all dose levels, whereas the DDP model allows both evolving and relatively consistent correlations across dose levels.

Also of interest are the predictive response distributions. Similar to the multicategory inferences, there are the posterior probability mass functions for the number of non-viable fetuses given a number of implants (see Figure 3.26) and the number of malformations given a number of implants and the number of non-viable

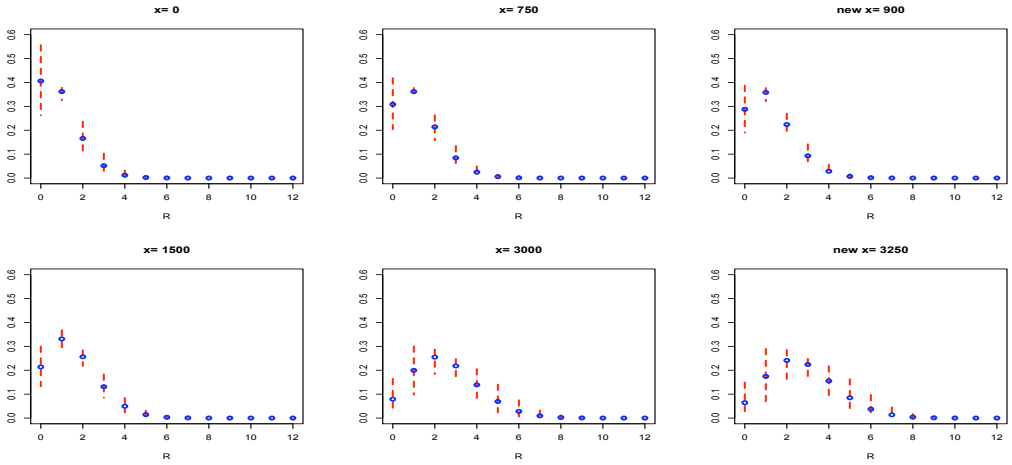

Figure 3.26: EG data. The posterior mean ("o") and 90% uncertainty bands (dashed lines) of the probability mass functions for the number of non-viable fetuses given $m = 12$ implants.

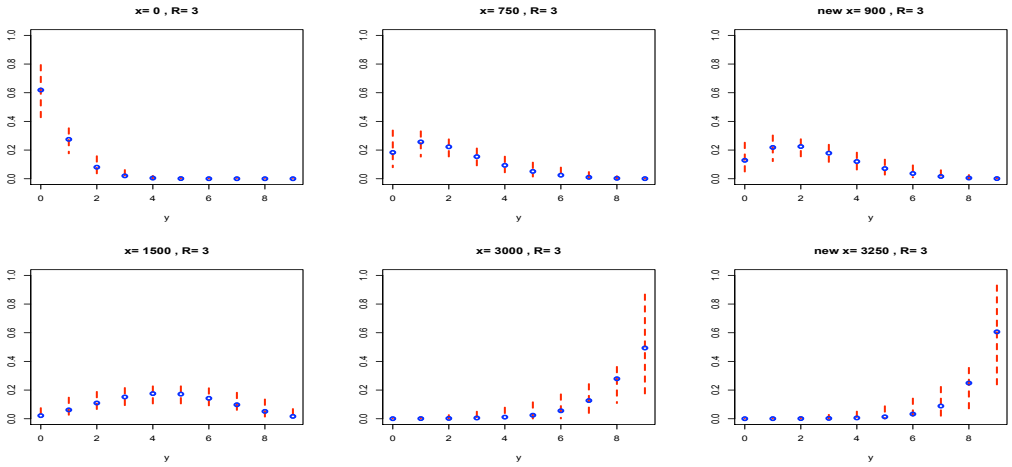

Figure 3.27: EG data. The posterior mean ("o") and 90% probability bands (dashed lines) of the probability density functions for the number of malformations given $m = 12$ implants and $R = 3$ non-viable fetuses.

fetuses (in Figure 3.27). While the data include a lot of variation, the DDP model finds simple, underlying probability mass functions that coincide with the dose-response curves displayed above. With the extra continuous response, we may look at the distribution of fetal weights at each dose level. As seen in Figure 3.28, the spread

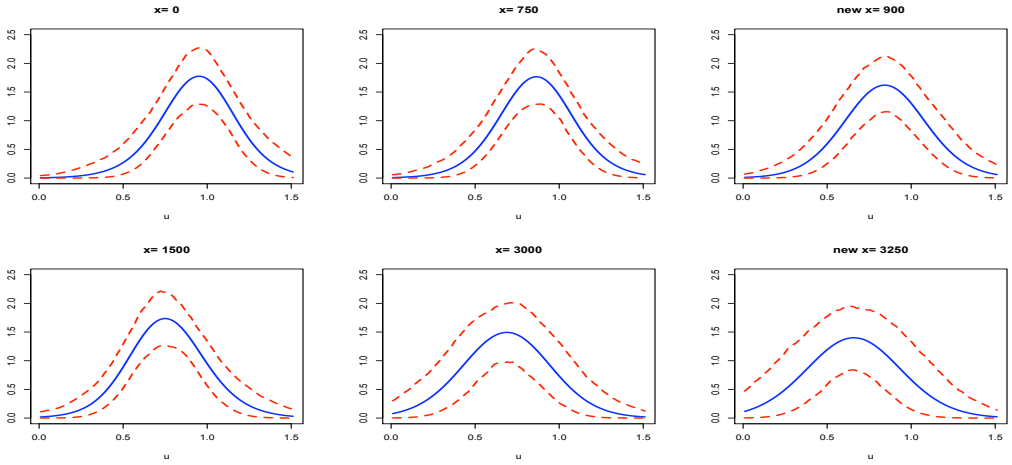

Figure 3.28: EG data: The posterior mean (solid) and 90% uncertainty bands (dashed lines) of the probability mass functions for fetal weight.

of the distributions is the same across the dose levels but the center shifts toward zero.

3.3.4 Discussion

The approach developed within this section is applicable to the extension of the multicategory setting by the inclusion of a continuous response on each of the viable pups. The framework provides flexibility in the multiple response distributions as well as the various risk assessment quantities. Demonstration of the model capabilities is found through data from an experiment investigating the toxic effects of ethelyne glycol.

The model developed in this section involves DDP mixing with respect to three parameters. This results in a relatively complex setting for prior specification as well as posterior simulation. Although the study of model properties along with the data analysis clearly demonstrate it is a feasible model given sufficient amounts of data, it may be useful to entertain simpler versions for some applications. A possible "semiparametric" version of the model leaves the distribution of prenatal deaths outside the DDP mixing; that is, we could build the joint response distribution through $f(y^*, u^* \mid$

$R, m; G_x) f_x(R \mid m) f(m)$, where $f_x(R \mid m)$ is a parametric distribution, such as a Binomial or Beta-Binomial with a logistic form for the probability of embryolethality. Now, the DDP mixture would be reserved for the pup-specific response distribution,

$$f(y^*, u^* \mid R, m; G_{\mathcal{X}}) = \int \prod_{k=1}^{m-R} \text{Bern}(y_k^*; \pi(\theta)) \text{N}(u_k^*; \mu, \varphi) \text{d}G_{\mathcal{X}}(\theta, \mu),$$

with a DDP prior on $G_{\mathcal{X}}$, where the base stochastic process, $G_{0\mathcal{X}}$, would be defined by two independent GPs.

This model has a few practical benefits. There are only two mixing parameters and, consequently, we must only learn about the parameters of two GPs. Also, under this formulation, the conditional probabilities of malformation and being of low birth weight can be shown to be monotonically increasing in prior expectation as the conditional distribution cancels out. The downside is that, as suggested by the EG data, the binomial assumption for the prenatal death distribution may not be sufficiently flexible. However, if the main risk assessment concerns are with respect to the malformation and fetal weight endpoints, this model may be worth exploring. Another advantage to this model lies in the possibility of using latent variables for the binary malformation responses in conjunction with imputing missing values to warrant Gibbs updates for the mixing parameters. We may also consider a more general version of the pup-specific responses, where we introduce a bivariate normal kernel for the latent malformation and fetal weight responses. If the data warrant a model such that the correlation between the two endpoints is expected to be very strong, we can incorporate mixing on this parameter in a structured manner.

3.4 Pre-implantation exposure

3.4.1 Motivation and background

In the previous sections, the experimental setting of the developmental toxicity studies is such that animals are exposed to varying levels of a toxin after implantation. Due to the fact that implantation occurs before exposure, the number of observed implants is assumed to contain no information about the dose-response relationship. However, there is evidence to support that pre-implantation exposure affects the early reproductive process and can cause birth defects or embryolethality. In studies with pre-implantation exposure, the number of observed implants should be dose dependent and, consequently, there is an added risk associated with an unsuccessful implantation.

To handle this type of data, Rai and Ryzin (1985) introduce a parametric dose-response model where the probability of response varies with random, dose-dependent litter size. In particular, a conditional Binomial distribution is used for the number of affected pups with probability that depends on litter size and dose level, along with a Poisson litter size distribution with dose-dependent mean. Dunson (1998) uses a multiple imputation approach to estimate the number of missing fetuses. Allen and Barnhart (2002) model both the number of implants and the outcome of each implant with separate generalized linear models, a multinomial distribution parameterized by the cumulative logit with proportional odds for the number of implantations and a binary response with a logit link for the negative outcomes. Kuk (2003) develops a model that expresses the probability of a successful implant in terms of the parameters of the dose-dependent litter size distribution, including a logit link to model the probability that an implanted pup is adversely affected by the toxin.

While not proposed in terms of pre-implantation studies, there are methods of

analyzing the more common Segment II developmental toxicity studies which treat the number of implants or the litter size as a random variable that depends on the dose level. Ten Have and Chinchilli (1998) propose two-stage negative binomial and overdispersed Poisson models with a log-linear function of dose driving both the probability of success and the mean of the random litter size distribution. Ma et al. (2009) model the number of successes and the number of failures with dual Poisson models including functions of a mixed model on the dose level in the mean.

We propose a Bayesian nonparametric mixture model built from a dependent Dirichlet process prior to provide rich inference for the response distributions and the dose-response relationships. The approach is developed for the distribution of the triplet of the number of potential implants, the number of observed implants, and the number of malformations, with implementation requiring imputing the missing, potential implants. Alternatively, integrating out the unobserved variables generates a nonparametric mixture with structured kernel components for the observed number of implants and the number of malformations. We focus on flexible probabilistic modeling for the set of related distributions of the triplet across the dose levels, which results in general inference for the risk assessment quantities; importantly, such inference captures the inherent complexity of the data while appropriately accounting for uncertainty. To our knowledge, the proposed framework provides the first model-based nonparametric approach to the analysis of pre-implantation studies.

3.4.2 Methods

We begin with a generic dam at a given dose level x. Let m^* denote the number of potential implants for the dam, and y be the response with observed number of implants m. Note that $y \leq m \leq m^*$, where m^* is not observed other than at dose

$x_0 = 0$, that is, we presume $m_0 = m_0^*$. We take a Pois(λ) distribution for the number of potential implants, though this can be extended to a general mixture of negative-binomial distributions. Here, we use the simpler case for more clear exposition of the main component of the modeling approach.

Notwithstanding, both the observed m and y to depend on the dose level. To relax the potentially restrictive assumptions imposed by standard parametric models, we propose a nonparametric mixture modeling framework for the joint distribution of the number of observed implants and negative outcomes. The mixture model for $f(y, m)$ is induced by a model for $f(m^*, m, y)$, which is built from $f(m^*)$ and a nonparametric mixture model for $f(m, y \mid m^*)$ with dose-dependent mixing distributions. The resulting marginal distribution for the number of observed implants, $f(m)$, and conditional distribution for the number of negative outcomes given the number of implants, $f(y \mid m)$, depend on dose level in a fully nonparametric, model-based fashion.

The assumptions regarding m^* induce different models for $x_0 = 0$ and $x > 0$. First focusing on the active dose levels, $x > 0$, we propose the following prior mixture model for the triplet (m^*, m, y),

$$\begin{aligned} f(m^*, m, y; G_{\mathcal{X}}) &= \text{Pois}(m^*; \lambda) f(m, y \mid m^*; G_{\mathcal{X}}) \\ &= \text{Pois}(m^*; \lambda) \int \text{Bin}(m; m^*, \pi(\gamma)) \text{Bin}(y; m, \pi(\theta)) \, \text{d}G_{\mathcal{X}}(\gamma, \theta), \end{aligned} \tag{3.9}$$

where $G_{\mathcal{X}} \mid \alpha, \psi \sim \text{DDP}(\alpha, G_{0\mathcal{X}})$. In the mixture model for $f(m, y \mid m^*; G_{\mathcal{X}})$, the parameter of the first Binomial in the kernel, $\pi(\gamma)$, is the probability of successful implantation, whereas the parameter of the second Binomial kernel, $\pi(\theta)$, is the probability of a negative outcome after implantation. Here, $G_{0\mathcal{X}}$ is defined through two independent GPs each with a linear mean function and exponential correlation

function. Specifically, we use a GP prior with mean function $\beta_0 + \beta_1 x$, variance σ^2, and covariance function $\exp(-\phi|x - x'|)$ for the component θ, and an independent GP with mean $\xi_0 + \xi_1 x$, variance τ^2, and covariance function $\exp(-\rho|x - x'|)$ for the γ component of $G_{0\mathcal{X}}$. The model is completed with priors on the GP hyperparameters and for α. In particular, we place normal priors on the intercept terms, exponential priors on the slope coefficients, inverse gamma priors on the GP variance terms, and uniform priors on the range parameters.

The marginal model for the number of potential implants has support over $m^* = 0, 1, \ldots$. The event $m^* = 0$ is trivial, however, it is necessary such that if we marginalize over the unobserved m^*, the model reduces to a proper probability model. Specifically, using the result $\sum_{m^*=m}^{\infty} \text{Pois}(m^*; \lambda)\text{Bin}(m; m^*, \pi(\gamma)) = \text{Pois}(m; \lambda\pi(\gamma))$, and marginalizing over m^* in 3.9, the model for the effective dose levels becomes

$$f(y, m; G_{\mathcal{X}}) = \int \text{Pois}(m; \lambda\pi(\gamma))\text{Bin}(y; m, \pi(\theta))\, \mathrm{d}G_{\mathcal{X}}(\gamma, \theta), \quad m = 0, 1, ...; \; y \leq m.$$

Therefore, the marginal mixture model for the number of observed implants becomes $f(m; G_{\mathcal{X}}) = \int \text{Pois}(m; \lambda\pi(\gamma))\, \mathrm{d}G_{\mathcal{X}}(\gamma)$. The joint mixture model provides directly the conditional model for the distribution of the number of malformations given the number of observed implants, that is, $f(y \mid m; G_{\mathcal{X}}) = f(y, m; G_{\mathcal{X}})/f(m; G_{\mathcal{X}})$.

For the control group (dose $x_0 = 0$), we have assumed $m_0^* = m_0$. Thus, the model at the control level is given by the marginal $\text{Pois}(m_0; \lambda)$ distribution for m_0, not included in the nonparametric mixture, and the conditional for y_0 given m_0, defined by the DDP mixture of Binomial distributions, $f(y_0 \mid m_0; G_{\mathcal{X}}) = \int \text{Bin}(y_0; m_0, \pi(\theta))\, \mathrm{d}G_{\mathcal{X}}(\theta)$.

Model properties

The DDP mixture structure provides the joint hierarchical model for the data at the effective doses and at the control group, which is used for posterior simulation. For predictive inference for risk assessment, it is useful to establish an equivalent mixture model formulation in terms of the underlying binary implantation responses for a generic dam at dose level x, m'_k, $k = 1, 2, \ldots, m^*$, and the product of the underlying binary malformation responses, y^*_k, $k = 1, \ldots, m$. We show that the DDP mixture model in 3.9 is equivalent to

$$f^*(m^*, \{m'_k\}, \{y^*_k\}; G_{\mathcal{X}}) = \text{Pois}(m^*; \lambda) \int \prod_{k=1}^{m^*} \text{Bern}(m'_k; \pi(\gamma)) \prod_{k=1}^{\sum_{k=1}^{m^*} m'_k} \text{Bern}(y^*_k; \pi(\theta)) \text{d}G_{\mathcal{X}}(\gamma, \theta),$$

where the same mixing parameters are used for all binary responses corresponding to the same dam.

Equivalence of the model with the underlying binary responses: To show the equivalence of the DDP model with the Poisson-Binomial-Binomial kernel ($M_1 = \text{Pois}(m^*; \lambda) \int \text{Bin}(m; m^*, \pi(\gamma)) \text{Bin}(y; m, \pi(\theta)) \text{d}G_{\mathcal{X}}(\gamma, \theta)$) and the Poisson-product of Bernoullis-product of Bernoullis kernel
($M_2 = \text{Pois}(m^*; \lambda) \int \prod_{k=1}^{m^*} \text{Bern}(m'_k; \pi(\gamma)) \prod_{k=1}^{\sum_{k=1}^{m^*} m'_k} \text{Bern}(y^*_k; \pi(\theta)) \text{d}G_{\mathcal{X}}(\gamma, \theta)$), we first

find the joint MGF of (m, y).

$$
\begin{aligned}
E_{M_1}(e^{t_1 m + t_2 y}; G_{\mathcal{X}}) &= \sum_{m^*=0}^{\infty} \sum_{m=0}^{m^*} \sum_{y=0}^{m} e^{t_1 m + t_2 y} f(m^*, m, y; G_{\mathcal{X}}) \\
&= \int \sum_{m=0}^{\infty} \sum_{y=0}^{m} e^{t_1 m + t_2 y} \text{Pois}(m; \lambda\pi(\gamma)) \text{Bin}(y; m, \pi(\theta)) \text{d}G_{\mathcal{X}}(\gamma, \theta) \\
&= \int \sum_{m=0}^{\infty} e^{t_1 m} \text{Pois}(m; \lambda\pi(\gamma))(1 + \pi(\theta)(e^{t_2} - 1))^m \text{d}G_{\mathcal{X}}(\gamma, \theta) \\
&= \int \exp\{\lambda\pi(\gamma)(e^{t_1}(1 + \pi(\theta)(e^{t_2} - 1)) - 1)\} \text{d}G_{\mathcal{X}}(\gamma, \theta)
\end{aligned}
$$

Now, the joint MGF of the $(\sum_{k=1}^{m^*} m'_k, \sum_{k=1}^{\sum_{k=1}^{m^*} m'_k} y^*_k)$ is as follows.

$$
\begin{aligned}
E_{M_2}(e^{t_1 \sum_{k=1}^{m^*} m'_k + t_2 \sum_{k=1}^{\sum_{k=1}^{m^*} m'_k} y^*_k}; G_{\mathcal{X}}) &= \sum_{m^*=0}^{\infty} \sum_{\sum m'_k = 0} \sum_{y=0}^{\sum m'_k} e^{t_1 \sum_{k=1}^{m^*} m'_k + t_2 \sum_{k=1}^{\sum_{k=1}^{m^*} m'_k} y^*_k} f^*(m^*, \{m'_k\}, \{y^*_k\}; G_{\mathcal{X}}) \\
&= \int \sum_{m^*=0}^{\infty} \text{Pois}(m^*; \lambda) \sum_{\sum m'_k = 0}^{m^*} \sum_{y=0}^{\sum m'_k} e^{t_1 \sum_{k=1}^{m^*} m'_k + t_2 \sum_{k=1}^{\sum_{k=1}^{m^*} m'_k} y^*_k} \times \\
&\quad \prod_{k=1}^{m^*} \text{Bern}(m'_k; \pi(\gamma)) \prod_{k=1}^{\sum_{k=1}^{m^*} m'_k} \text{Bern}(y^*_k; \pi(\theta)) \text{d}G_{\mathcal{X}}(\gamma, \theta) \\
&= \int \sum_{m^*=0}^{\infty} \text{Pois}(m^*; \lambda) \prod_{k=1}^{m^*} \sum_{m'_k=0}^{1} \text{Bern}(m'_k; \pi(\gamma))(1 + \pi(\theta)(e^{t_2} - 1))^{m'_k} \text{d}G_{\mathcal{X}}(\gamma, \theta) \\
&= \int \sum_{m^*=0}^{\infty} \text{Pois}(m^*; \lambda)\{1 + \pi(\gamma)(e^{t_1} + e^{t_1}\pi(\theta)(e^{t_2} - 1) - 1)\}^{m^*} \text{d}G_{\mathcal{X}}(\gamma, \theta) \\
&= \int \exp\{\lambda\pi(\gamma)(e^{t_1}(1 + \pi(\theta)(e^{t_2} - 1)) - 1)\} \text{d}G_{\mathcal{X}}(\gamma, \theta) \\
&\equiv E_{M_1}(e^{t_1 m + t_2 y}; G_{\mathcal{X}})
\end{aligned}
$$

This alternative model formulation gives a basis for risk assessment based on the probability a generic pup at dose level x will not be successfully implanted, the conditional probability that the pup will have a malformation given it is successfully

implanted, and a combined risk of unsuccessful implantation and malformation.

Risk assessment

For risk assessment, we have an implicit conditioning on $m^* = 1$. The first dose-response curve of interest is the probability of an unsuccessful implant across effective dose levels,

$$\Pr(m' = 0 \mid m^* = 1; G_x^L) = \sum_{l=1}^{L} p_l[1 - \pi(U_l(x))], \quad x \in \mathcal{X}.$$

Following the procedures in the previous sections, this curve is increasing in prior expectation given a restriction on the slope coefficient in the underlying GP linear mean, namely $\xi_1 < 0$.

We are also interested in the probability that a generic pup has a malformation (in general, negative outcome) given that it is successfully implanted. This probability is naturally defined through conditional probability given a successful implant for all $x \in \mathcal{X}$:

$$\Pr(y^* = 1 \mid m' = 1, m^* = 1; G_x^L) = \Pr(m' = 1, y^* = 1 \mid m^*; G_x^L)/\Pr(m' = 1 \mid m^* = 1; G_x^L),$$

which is written as $\sum_{l=1}^{L} q_l(x)\pi(Z_l(x))$ where $q_l(x) = \{p_l\pi(U_l(x))\}/(\sum_{m=1}^{L} p_m\pi(U_l(x)))$.

For full risk assessment at any active dose level, $x > 0$, we want to combine the probability of a live, malformed pup and the probability of an unsuccessful implant.

Hence, the combined risk, $r(x)$, at dose level $x \in \mathcal{X}$ is given by

$$\begin{aligned} r(x) &= \Pr(y^* = 1 \text{ or } m' = 0 \mid m^* = 1; G_x^L) \\ &= \Pr(m' = 1, y^* = 1 \mid m^* = 1; G_x^L) + \Pr(m' = 0 \mid m^* = 1; G_x^L) \\ &= \sum_{l=1}^{L} p_l \pi(Z_l(x)) \pi(U_l(x)) + \sum_{l=1}^{L} p_l (1 - \pi(U_l(x))) \\ &= 1 - \sum_{l=1}^{L} p_l \pi(U_l(x)) \{1 - \pi(Z_l(x))\}. \end{aligned}$$

Therefore, through investigation of the prior expectation, the full risk function is monotonically increasing provided both $\xi_1 < 0$ and $\beta_1 > 0$. Prior simulation suggests that these restrictions also promote an increasing trend in prior expectation for the conditional risk of malformation.

Prior specification

We set the means of the normal prior for the centering Gaussian process linear mean parameters, ξ_0 and β_0 to 0, and the shape parameter of the inverse gamma priors for the Gaussian process variance, τ^2 and σ^2, to 2 (implying infinite prior variance). The prior variances for ξ_0 and β_0, the prior mean for the exponential priors on ξ_0 and β_0, and the means for the inverse-gamma priors on τ^2 and σ^2 are chosen by studying the induced prior distribution for the dose-response curve for malformations and successful implants, for which prior realizations can be readily sampled. Specifically, under the prior choice, the prior mean for both $\Pr(y^* = 1; G_{\mathcal{X}}^L)$ and $\Pr(m' = 1; G_{\mathcal{X}}^L)$ have a slight increasing trend beginning around 0.5 and the corresponding 95% interval bands are essentially spanning the $(0, 1)$ interval.

3.4.3 Posterior inference

Let $\boldsymbol{m}^*{}_j = \{m^*_{0j}, m^*_{1j}, \ldots, m^*_{Nj}\}$ denote the number of potential implants for the jth dam at dose x_i, $i = 0, 1, \ldots, N$, and $\boldsymbol{y}_j = (y_{0j}, y_{1j}, \ldots, y_{Nj})$ be the jth response replicate with observed number of implants vector $\boldsymbol{m}_j = (m_{0j}, m_{1j}, \ldots, m_{Nj})$, for $j = 1, \ldots, n$. Now, let $\boldsymbol{\theta}_j \equiv \boldsymbol{\theta}_j(\boldsymbol{x}_{N+1}) = (\theta_j(x_0), \theta_j(x_1), \ldots, \theta_j(x_N))$ be the $(N+1)$-dimensional latent mixing vector for $\boldsymbol{y}_j$ and $\boldsymbol{x}_{N+1} = (x_0, x_1, \ldots, x_N)$, and $\boldsymbol{\gamma}_j \equiv \boldsymbol{\gamma}_j(\boldsymbol{x}_N) = (\gamma_j(x_1), \ldots, \gamma_j(x_N))$ be the N-dimensional latent mixing vector for m_{ij}, $i > 0$ and $\boldsymbol{x}_N = (x_1, \ldots, x_N)$, for $j = 1, \ldots, n$.

Given the truncation approximation as in Chapter 2, $\gamma_j = \boldsymbol{U}_l$ and $\theta_j = \boldsymbol{Z}_l$ with probability p_l, the first stage of the hierarchical model for the (m_{ij}, y_{ij}) is expressed as

$$\{\boldsymbol{m}_j, \boldsymbol{y}_j\} \mid \lambda, w_j, U(\boldsymbol{x}), Z(\boldsymbol{x}) \overset{ind.}{\sim} \prod_{j=1}^{n} \left\{\text{Pois}(m_{0j}; \lambda)\text{Bin}(y_{0j}; m_{0j}, \pi(Z_{w_j}(x_0)))\right\}^{s_{0j}} \times \prod_{i=1}^{N} \left\{\text{Pois}(m_{ij}; \lambda\pi(U_{w_j}(x_i)))\text{Bin}(y_{ij}; m_{ij}, \pi(Z_{w_j}(x_i)))\right\}^{s_{ij}}, \tag{3.10}$$

where $f(\boldsymbol{p} \mid \alpha)$ is defined as in Chapter 2 and $(U_l(\boldsymbol{x}_N), Z_l(\boldsymbol{x}_{N+1}))$ arise i.i.d. from $G_{0\boldsymbol{x}}(U_l(\boldsymbol{x}_N), Z_l(\boldsymbol{x}_{N+1}) \mid \psi)$.

MCMC details

For the studies with pre-implantation exposure, the two centering GPs driving the parameters of the kernel for the observed implants and the kernel for the number of negative outcomes. The updates for the parameters of the negative outcomes are equivalent to those from the DDP Binomial model, with an $N+1$-variate proposal (N active dose levels and one control). The observed implants include similar updates for

the GP hyperparameters, plus the updates for the N-variate sample paths, $U_l(\boldsymbol{x})$, $\{w_j\}$, and the mean of the potential implant distribution, λ.

The full conditional for the vector $(U_l(\boldsymbol{x}))$ depends on whether l corresponds to one of the distinct components. Therefore, if $l \notin \{w_k^* : k = 1, \ldots, n^*\}$, then $U_l(\boldsymbol{x}) \sim N_N((\xi_0 + \xi_1 x_1, \ldots, \xi_0 + \xi_1 x_N)^T, \Lambda)$ For $U_l : l \in \{w_k^* : k = 1, \ldots, n^*\}$,

$$p(U_{w_k^*}(\boldsymbol{x}) \mid \boldsymbol{w}, \psi, \text{data}) = g_0(U_{w_k^*}(\boldsymbol{x})|\psi) \prod_{\{j:w_j=w_k^*\}} \prod_{i=1}^{N} \text{Pois}(m_{ij}; \lambda\pi(U_{w_k^*}(x_i)))^{s_{ij}}.$$

A Metropolis-Hastings step is used with an N-variate normal distribution as the proposal, with covariance matrix of similar form to the base GP covariance function. Specifically, for the real data, we use $D_{ij} = 1.3\exp(-0.1|x_i - x_j|)$ which yields acceptance rates of approximately 0.15. For the w_j, the updates come from a discrete distribution with weights proportional to $p_l \text{Bin}(y_{0j}; m_{0j}, \pi(Z_l(x_0)))^{s_{0j}} \prod_{i=1}^{N} \{\text{Pois}(m_{ij}; \lambda\pi(U_l(x_i)))\text{Bin}(y_{ij}; m_{ij}, \pi(Z_l(x_i)))\}^{s_{ij}}$.

The parameter λ is the mean of the Poisson distribution for the control group and coupled with a function of the $\boldsymbol{U}_l(x)$ in the mean of the Poisson distribution for the active dose levels. Therefore, with a gamma(a_λ, b_λ) prior, the conditional posterior is also a gamma distribution with parameters $a_\lambda + \sum_{j=1}^{n}\sum_{i=0}^{N} m_{ij}^{s_{ij}}$ and $b_\lambda + n_0 + \sum_{j=1}^{n}\sum_{i=1}^{N} \pi(U_{w_j}(x_i))^{s_{ij}}$, where $n_0 = \sum_{j=1}^{n} s_{0j}$.

Posterior inference

Under this experimental design, we briefly discuss predictive inference for the response distributions, where we want to predict a new vector of responses at the

observed dose levels, $\boldsymbol{y}_0 = (y_{00}, \ldots, y_{N0})$ and implants $\boldsymbol{m}_0 = (m_{00}, \ldots, m_{N0})$ from

$$p(\boldsymbol{y}_0, \boldsymbol{m}_0 \mid \text{data}) = \int \left[\sum_{l=1}^{L} p_l \text{Pois}(m_{00}; \lambda) \text{Bin}(y_{00}; m_{00}, \pi(Z_l(x_0))) \times \right.$$
$$\left. \prod_{i=1}^{N} \text{Pois}(m_{i0}; \lambda\pi(U_l(x_i))) \text{Bin}(y_{i0}; m_{i0}, \pi(Z_l(x_i))) \right] \text{d}p(\Theta \mid \text{data}),$$

which accentuates the distinction of the DDP model at the control level from the active doses. Thus, $p(y_0 \mid m_0, \text{data}) = p(y_0, m_0 \mid \text{data})/p(m_0 \mid \text{data})$, where

$$p(m_0 \mid \text{data}) = \int \text{Pois}(m_{00}; \lambda) \sum_{l=1}^{L} p_l \prod_{i=1}^{N} \text{Pois}(m_{i0}; \lambda\pi(U_l(x_i))) p(\Theta \mid \text{data}) \text{d}\Theta.$$

To augment the N known dose levels, we want to interpolate across M new dose levels, $\tilde{\boldsymbol{x}} = (\tilde{x}_1, \ldots, \tilde{x}_M)$. Now, each $U_l(\boldsymbol{x}) = (U_l(\boldsymbol{x}_N), \tilde{U}_l(\tilde{\boldsymbol{x}}))$ and $Z_l(\boldsymbol{x}) = (Z_l(\boldsymbol{x}_{N+1}), \tilde{Z}_l(\tilde{\boldsymbol{x}}))$, for $l = 1, \ldots, L$. These are obtained through the standard conditional normal regression distributions.. The full vectors of the GP sample paths are utilized to obtain inference for the dose-response relationships by evaluating the expressions found in the previous sections.

3.4.4 Data illustrations

We use a simulation study to check the performance of our model, the details and results of which are given in the next section. To further illustrate the interpolation and extrapolation capabilities of the model, we also analyze data from a dominant lethal assay experiment.

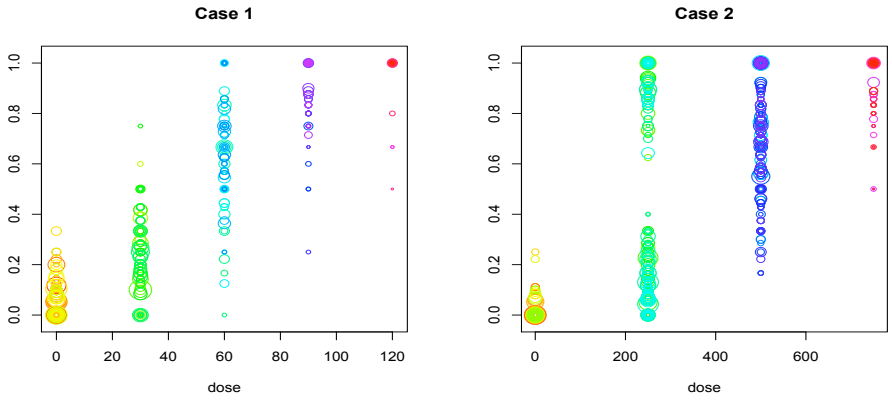

Figure 3.29: Simulation data. Each circle corresponds to a particular dam, the size of the circle is proportional to the number of implants, and the coordinates of the circle are the dose level and the proportion of malformations for the three simulation settings.

Simulation study

To determine the effectiveness of the proposed model in capturing inferential objectives, we use two simulation settings. The first involves observed implants from a single Poisson distribution with a mean of $\lambda\pi(h_1(x))$ with $h_1(x) = 2 - 0.025x$ and $\lambda = 11$, and the number of malformed pups arising from a single Binomial model with dose-dependent probability of $\pi(f_1(x))$ where $f_1(x) = -3 + 0.06x$. Here, there are 4 equally spaced active dose levels ranging from 30 to 120 and $n = 100$ animals at each dose level. The second uses a three component mixture for both the observed implants and the number of malformed pups, with the three Poisson distributions being very close and the three binomial distributions well separated; that is, m is drawn from $\sum_{i=1}^{3} w_i \text{Pois}(\lambda\pi(h_i(x)))$ with $\boldsymbol{w} = (0.35, 0.40, 0.25)$, $\lambda = 11$, $h_1(x) = 5.5 - 0.008x$, $h_2(x) = 6 - 0.009x$, and $h_3(x) = 5 - 0.006x$. Conditioned on m, y arises from $\sum_{i=1}^{3} w_i \text{Bin}(m, \pi(m_i(x)))$, where $\boldsymbol{w}$ is the same, $m_1(x) = -10 + 0.05x$, $m_2(x) = -5 + 0.012x$, and $m_1(x) = -4 + 0.008x$. In this setting, we have 3 active dose

levels (250, 500, and 750) and $n = 300$ animals at each dose level. The data from each of the simulation settings is given in Figure 3.29.

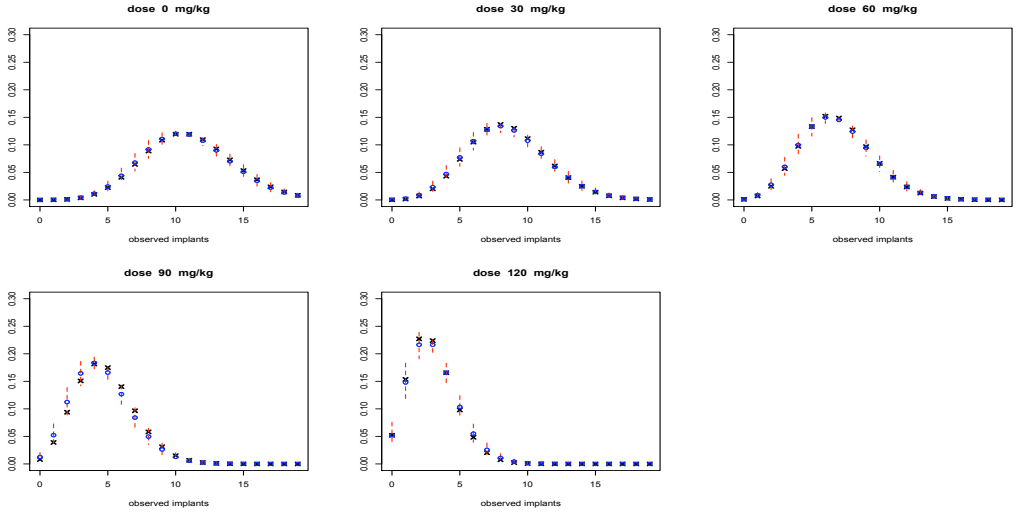

Figure 3.30: For the 5 observed dose levels in simulation 1, the posterior mean probability mass functions (denoted by "o") and 90% probability bands for the number observed implants with the true pmf denoted by "x".

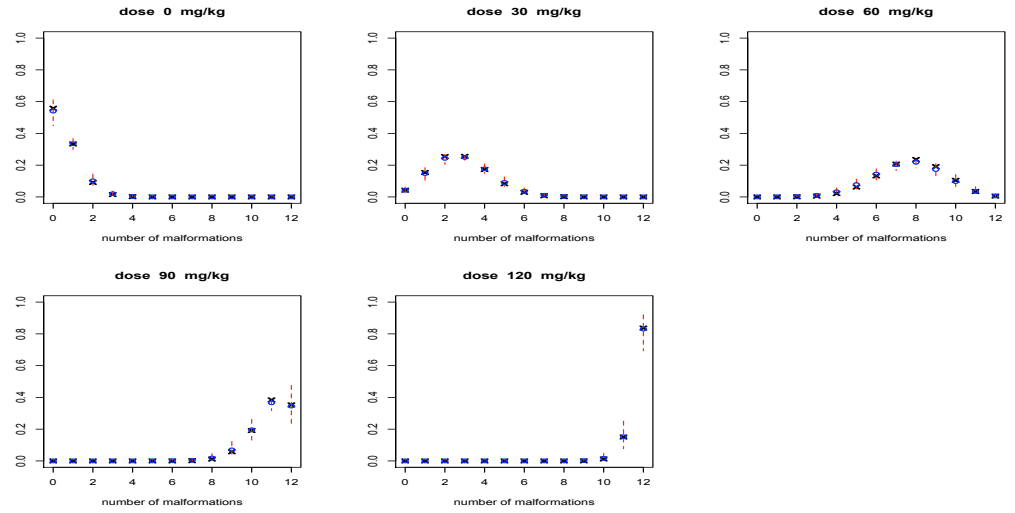

Figure 3.31: For the 5 observed dose levels in simulation 1, the posterior mean probability mass functions (denoted by "o") and 90% probability bands for the number of malformations conditional on $m = 12$ implants with the true pmf denoted by "x".

In the first simulation setting, the DDP model recovers the true marginal probability mass functions for the number of observed implants with narrow probability bands (Figure 3.30). Inference for the conditional probability mass functions of the number of malformations given a specified number of observed implants is also comparable to the truth with relatively small uncertainty bands, as shown in Figure 3.31. Even with the simple structure of a single Binomial and Poisson component, the DDP model has the capacity to yield appropriate inference, as seen in Figure 3.32, with satisfactory uncertainty quantifications.

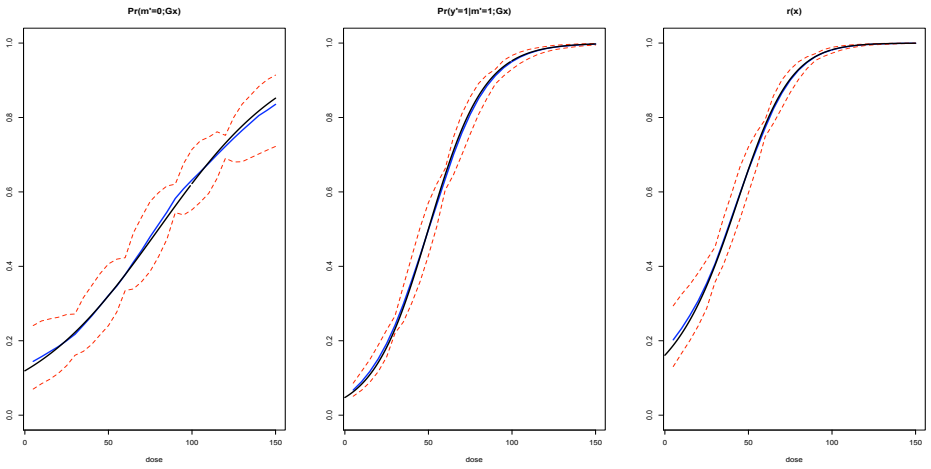

Figure 3.32: The posterior mean estimate (solid line) and 90% probability bands (dashed lines) for $\Pr(m' = 0; G_{\mathcal{X}}^{L})$ (left panel), $\Pr(y^* = 1; G_{\mathcal{X}}^{L})$ (middle panel), and the risk function $r(x)$ (right panel). The true curves given the first simulation setting are given in blue.

The second simulation setting results in probability mass functions for the number of observed implants which are comparable to the true values with little uncertainty, and are omitted here. The conditional pmfs for the number of malformations given a specific number of successful implants ($m = 12$ in Figure 3.33) are multimodal in parts of the dose range at both observed and unobserved dose levels.

The DDP model follows the true values with sensible probability bands.

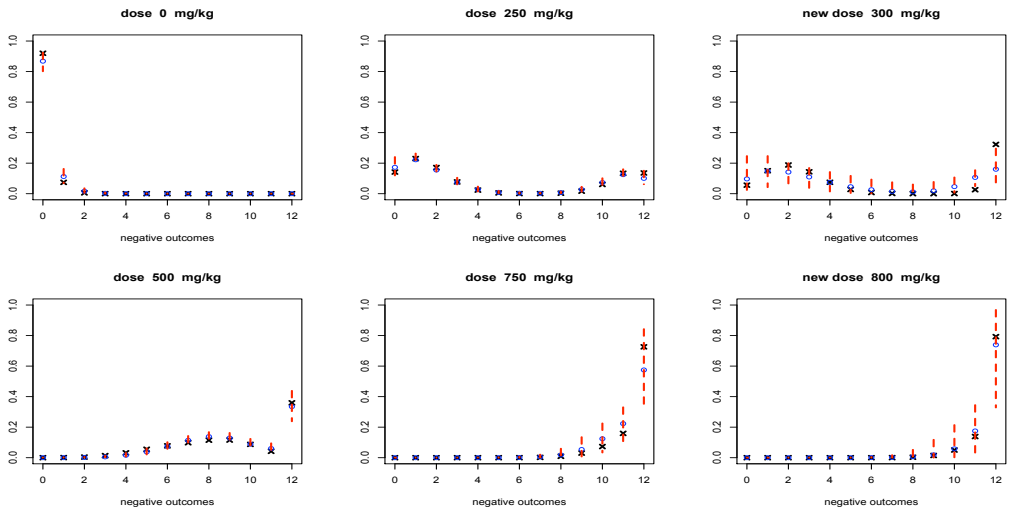

Figure 3.33: For the 4 observed dose levels in simulation 2, the posterior mean probability mass functions (denoted by "o") and 90% uncertainty bands for the number of malformations conditional on $m = 12$ implants with the true pmf denoted by "x".

The risk assessment quantities are found in Figure 3.34. The DDP model captures both the simple and non-standard dose-response curves, as the true curves, given in black, lie within the relatively tight probability intervals. Because there is no data between 0 mg/kg and 250 mg/kg, the model almost misses the true conditional probability of malformation and, subsequently, the combined risk between these points. However, the uncertainty bands still cover the true curves.

Dominant lethal assay experiment

We further illustrate the utility of the DDP model through a data set from a dominant lethal assay experiment discussed in Lüning et al. (1966). In this experiment, a number of male mice were given 0, 300, and 600 R of radiation and, within 7 days, were mated with females. For the 1773 females, the number of implants and the number

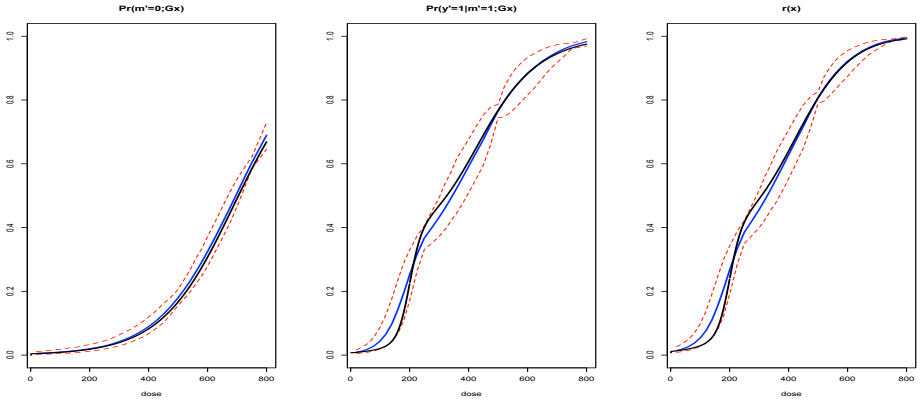

Figure 3.34: The posterior mean estimate (solid line) and 90% probability bands (dashed lines) for $\Pr(m' = 0; G_{\mathcal{X}}^L)$ (left column), $\Pr(y^* = 1; G_{\mathcal{X}}^L)$ (middle column), and the risk function $r(x)$ (right column). The true curves given the second simulation setting are given in black.

of non-viable fetuses were recorded; the data are shown in the left panel of Figure 3.35.

Figure 3.36 gives the posterior mean and 90% probability intervals for the marginal probability mass function for the number of observed implants. While the data do not suggest bimodal or vast deviations from standard shapes, the DDP model has the ability to interpolate and extrapolate and produces probability mass functions with a smooth evolution. This progression is also apparent in the conditional probability mass functions for the number of non-viable fetuses given $m = 8$ at the three observed dose levels and six new dose levels within and outside the range of observed levels (see Figure 3.37). Finally, the risk assessment quantities are found in Figure 3.35. As the data has only three dose levels and the variability is high, the risk assessment quantities have a fairly weak signal and large uncertainty bands.

3.4.5 Discussion

We have presented a Bayesian nonparametric mixture model for developmental toxicity studies with pre-implantation exposure. The model characterizes the data

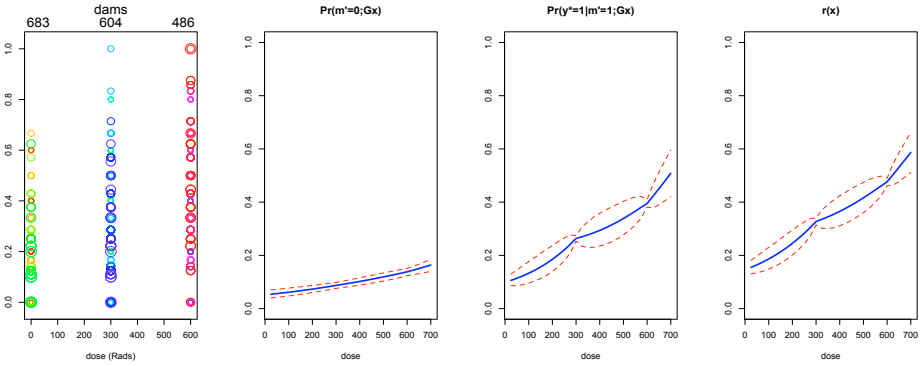

Figure 3.35: In the left panel, the radiation data. The posterior mean estimate (solid line) and 90% probability bands (dashed lines) for $\Pr(m' = 0; G_{\mathcal{X}}^L)$ (left middle panel), $\Pr(y^* = 1; G_{\mathcal{X}}^L)$ (right middle panel), and the risk function $r(x)$ (right panel).

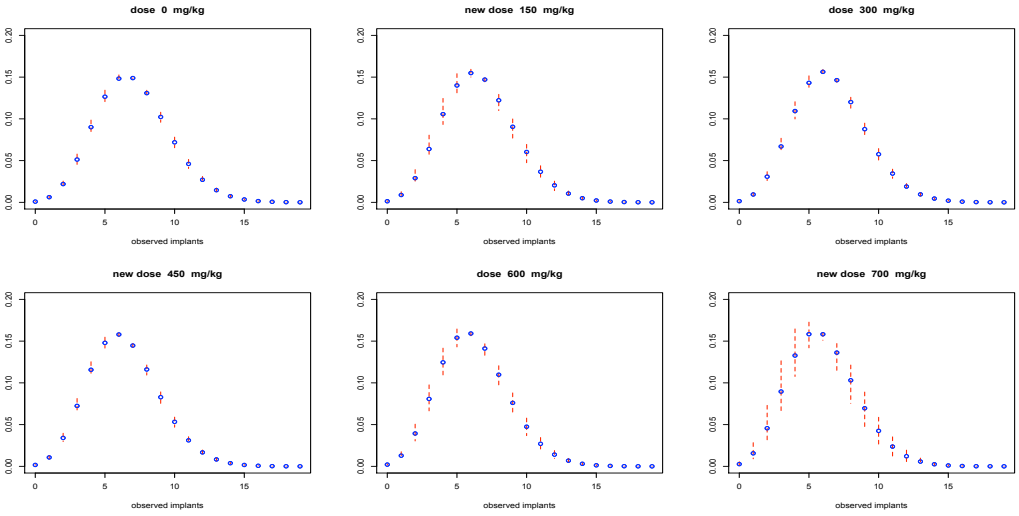

Figure 3.36: For the 3 observed dose levels and 3 new dose levels, the posterior mean probability mass functions (denoted by "o") and 90% uncertainty bands for the number observed implants.

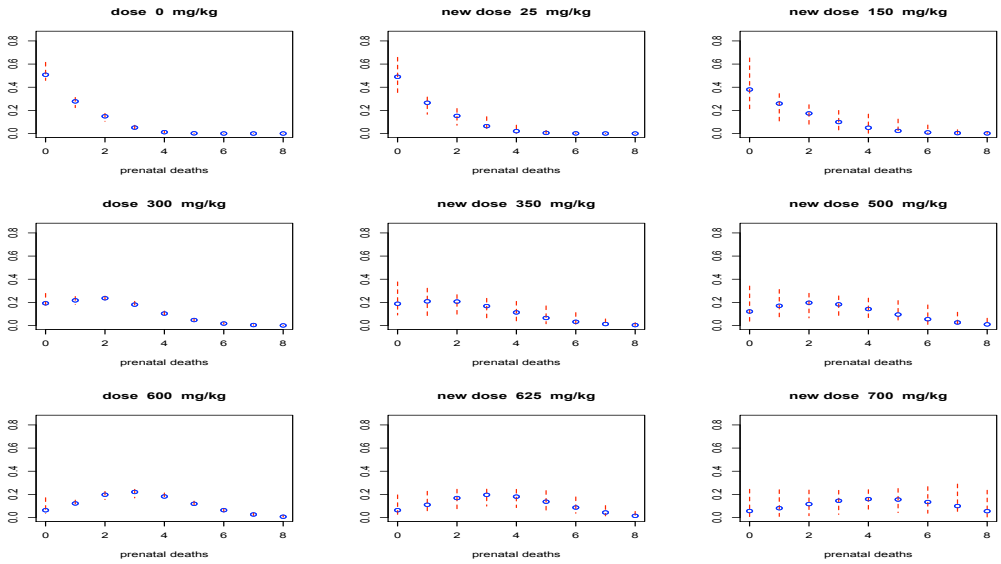

Figure 3.37: For the 3 observed dose levels and 6 new dose levels, the posterior mean probability mass functions (denoted by "o") and 90% probability bands for the number of non-viable fetuses conditional on $m = 8$ implants.

structure, and avoids the imputation of the unobserved variables. Simulated data as well as data from a dominant lethal assay experiment were used to show the ability to interpolate and extrapolate inferences at unobserved dose levels.

3.5 Conclusions

We have presented a comprehensive framework for the analysis of the many types of developmental toxicity studies. The framework is based on the single-p dependent Dirichlet process prior to provide flexibility in the response distributions and the various probabilities of response. The replication required for the DDP prior arises in the multiple animals at each dose level. The dose-response curves are imbued with an increasing trend to aid in interpolation and extrapolation, nevertheless they are capable of producing non-monotonic shapes if suggested by the data.

Chapter 4

Nonparametric mixture modeling for bioassay settings

4.1 Motivation and background

The quantal bioassay dose-response setting has been found valuable in many different fields, from ecology to medicine. Though quantitative measurement of a response is preferred, there are certain responses which can only be expressed as a binary outcome, either occurring or not occurring. The dose-response relationship is based on observed data from experimental animal, human clinical, or cell studies. In these types of settings, as dose increases, the responses generally become more severe. Within a population, the majority of the subjects' responses are similar; however, a wide variance of responses may be encountered as some individuals are more susceptible and others resistant.

At each of N dose levels $x_i,\ i = 1, \ldots, N,\ n_i$ subjects are exposed to the substance, and the number of subjects that have the response of interest is recorded as y_i. The dose-response curve, denoted by $D(x)$, is defined as the probability of positive

response as a function of dose x. One of the standard assumptions in the analysis of quantal bioassay problems is that $D(x)$ is a non-decreasing function and can therefore be modeled as a cumulative distribution function (CDF). Under this assumption, the distribution corresponding to CDF $D(x)$ is referred to as the tolerance distribution. In parametric modeling, $D(x)$ is assumed to be a member of a parametric family of CDFs. Standard parametric models are intuitive and simple to implement, but do not have the ability to capture complex curves, including skewness and multimodality.

To extend the inferential scope of parametric models, authors have proposed finite mixtures for the tolerance distribution. Lwin and Martin (1989) presents a location-scale mixture model, but develops the model based on the idea of a finite number of unobserved, underlying subpopulations. Geweke and Keane (1999) also propose full mixtures of probit models to more accurately describe the shape of the dose-response curve, while Basu and Mukhopadyay (2000) use a finite scale mixture of normal CDFs. These finite mixtures are, in general, more flexible than parametric models, however one must choose the number of components. One may take the route of increasing the number of components in the mixture model, to move in the direction of semiparametric models. This can be inefficient and entails multiple runs of an MCMC algorithm to ensure a pertinent cutoff point. In addition to theoretical advantages, Bayesian nonparametric priors offer the practical benefit of removing the requirement of setting the cutoff.

The quantal bioassay setting has been studied extensively in the Bayesian nonparametric arena. One of the more common approaches involves assign a nonparametric prior to the distribution function $D(x)$. Variants of the DP prior have been investigated previously for the dose-response curve (e.g., Antoniak, 1974; Bhattacharya, 1981; Disch, 1981; Kuo, 1983, 1988; Gelfand and Kuo, 1991;

Mukhopadhyay, 2000). Also along these lines is Muliere and Walker (1997), where Pólya tree priors are used to determine the maximum tolerated dose. A monotone nonparametric regression framework is presented in Bornkamp and Ickstadt (2009), where the monotone function is modeled as a mixture of shifted and scaled two-sided power distribution functions. An alternative approach is to model the binary responses with latent continuous variables, and assume flexible mixtures of probit or logit functions for the dose-response curve. MacEachern (1998) assumes a DP prior directly on the distribution of the latent variables. Presented in terms of multivariate probit regression, Jara et al. (2007) mix on the intercept of the linear mean and the covariance matrix of the distribution of the subject specific latent variable distribution. This model requires constraints to ensure identifiability. Casanova et al. (2010) place a scale DP mixture of normal distributions prior on the distribution of the latent variables.

We aim to extend these versions of the product of bernoullis or binomial models by assuming the data at each dose level arises from a DP mixture model, where the collection of mixing distributions are related but not identical. That is, we employ the brilliant services of the dependent Dirichlet process prior.

4.2 Methods

Here, we discuss possible models for the analysis of quantal response data, including a general DDP approach and modifications thereof. For all subsequent approaches, we let y_{ij} be the binary responses corresponding to the n_i subjects at dose level, x_i, for $i = 1, \ldots, N$; thus, $y_i = \sum_{j=1}^{n_i} y_{ij}$.

4.2.1 General DDP prior

In the same spirit with the mixture models developed in Chapter 3 for developmental toxicology data, we first consider a general DDP mixture model for the binary responses. In particular, we use a Bernoulli kernel with a logistic transform for the probability of positive response. A single-p DDP prior is placed on the collection of mixing distributions of the logit parameter across the dose levels, giving the prior model

$$f(y; G_{\mathcal{X}}) = \int \text{Bern}\left(y; \frac{\exp(\theta)}{1+\exp(\theta)}\right) \mathrm{d}G_{\mathcal{X}}(\theta),$$

where $G_{\mathcal{X}} \sim \text{DDP}(\alpha, G_{0\mathcal{X}})$. Alternatively, we may take the probit link and place the DDP prior on the dose-specific, mixing distributions of the probit parameter, i.e., $f(y; G_{\mathcal{X}}) = \int \text{Bern}\left(y; \Phi(\theta)\right) \mathrm{d}G_{\mathcal{X}}(\theta)$. In either formulation, we may take the base stochastic process to be a Gaussian process with linear mean function, $\mu(x) = \beta_0 + \beta_1 x$, and isotropic correlation function, $\Sigma = \sigma^2 H(\phi)$, where $H_{ij}(\phi) = \exp(-\phi|x_i - x_j|)$.

By definition, the dose-response curve is the probability of a response across the dose levels. Therefore, under the general DDP model, the dose-response curve is given by

$$\Pr(y = 1; G_{\mathcal{X}}) = \int \pi(\theta) \mathrm{d}G_{\mathcal{X}}(\theta) \equiv \sum_{l=1}^{L} p_l \pi(\theta_l(x)), \tag{4.1}$$

where $\pi(\cdot)$ may take either the logit or probit transform. This model does not strictly enforce a monotonicity restriction, although it may be shown to be monotonically increasing in prior expectation provided the GP base stochastic process has a linear mean and the coefficient associated with the dose level, β_1, is strictly positive. Section 3.1.1 includes prior realizations of the dose-response curve under this model, and emphasizes the monotonicity of the curve in expectation.

The relationship between the dose level and the observed response can often be extremely complex. In general, however, at relatively low doses, the response to a drug generally increases in direct proportion to increases in the dose. At higher doses of the drug, the amount of change in response to an increase in the dose gradually decreases until a dose is reached that produces no further increase in the observed response. The relationship between the concentration of the drug and the observed effect is consequently assumed to be monotonically increasing.

In the general DDP model, we take each replicate j as a vector of N data points, $\boldsymbol{y}_j = (y_{1j}, y_{2j}, \ldots, y_{Nj})$, thereby linking the neighboring data. In the developmental toxicology case, this is warranted as the animals are most often labeled and recorded in increasing order across the dose level. However, in the traditional dose-response setting, the subjects are arbitrarily assigned to the elements of the response vector at any given dose level. There is no reason to believe the subjects within the jth vector of length N are associated in any fashion. Accordingly, the subjects are assumed to be fully exchangeable across both i and j.

These restrictions and assumptions support the use of the linear-DDP as a special case of the general DDP prior. Consequently, we develop a more structured version of the general DDP mixture model, using the linear-DDP prior for the collection of dose-dependent mixing distributions.

4.2.2 Linear-DDP mixture model

The linear-DDP (e.g., DeIorio et al. (2009)) is a simplified version of the DDP prior. That is, we consider $f(y; G_x) = \int \text{Bern}(y; \pi(\theta)) dG_x(\theta)$, where $G_x = \sum_{l=1}^{L} p_l \delta_{\theta_l(x)}$ and $\theta_l(x) = \gamma_{0l} + \gamma_{1l} x$. The $\{\gamma_{0l}\}$ and $\{\gamma_{1l}\}$ are assumed to be mutually independent, which begets a weaker dependence on the collection of distributions, and enforces the

functional dependence of the distributions on x. With the constraint $\gamma_{1l} > 0$ for all $l = 1, \ldots, L$, we force a monotonic increasing trend for the dose-response curve. That is, with the restriction $\gamma_{1l} > 0$, each $\pi(\gamma_{0l} + \gamma_{1l}x)$ is increasing in x, and therefore $\sum_{l=1}^{L} p_l \pi(\gamma_{0l} + \gamma_{1l}x)$ is monotonically increasing with probability one.

Modeling details

For most dose-response experiments, one can argue that the observed binary response is a proxy for a continuous variable. As a result, an unobserved continuous variable underlies the binary response, say, arising from a normal distribution. This procedure offers a tractable way to estimate the parameters of a model. We assume a probit mixture approach for modeling the intrinsic degree of response variables via the linear-DDP prior.

The essential observation from the construction of the linear-DDP prior is the correspondence of $\{G_x : x \in \mathcal{X}\}$ and $G = \sum_{l=1}^{L} p_l \delta_{(\gamma_{0l}, \gamma_{1l})}$, where
$G \sim \mathrm{DP}(\alpha, G_0 = G^{(0)} \times G^{(1)})$, with $G^{(0)}$ and $G^{(1)}$ denoting the component of the DP base distribution for the γ_{0l} and γ_{1l}, respectively. That is, $(\gamma_{0l}, \gamma_{1l}) \sim G_0$ with probability p_l (definition of G under the DP prior) if and only if $\theta_l(x) = \gamma_{0l} + \gamma_{1l}x$ with probability p_l (definition of G_x under the linear-DDP prior). Hence, using the probit transform for the Bernoulli kernel probability, the linear-DDP mixture model, $f(y; G_x) = \int \mathrm{Bern}(y; \Phi(\theta))\mathrm{d}G_x(\theta)$, can be equivalently formulated as a DP mixture model

$$f(y; G, x) = \int \mathrm{Bern}(y; \Phi(\gamma_0 + \gamma_1 x))\mathrm{d}G(\gamma_0, \gamma_1); \quad G \sim \mathrm{DP}(\alpha, G_0).$$

This model can now be represented through latent continuous responses, say, u_{ij} associated with y_{ij}, with a location normal DP mixture defining the latent response

distribution. Specifically,

$$
\begin{aligned}
y_{ij} \mid u_{ij} &\overset{ind.}{\sim} \mathbf{1}\{y_{ij} = 1 \text{ iff } u_{ij} > 0\};\ i = 1, \ldots, N;\ j = 1, \ldots, n_i \\
u_{ij} \mid G, x_i &\overset{ind.}{\sim} \int \mathrm{N}(u_{ij}; \gamma_0 + \gamma_1 x_i, 1) \mathrm{d}G(\gamma_0, \gamma_1);\ i = 1, \ldots, N;\ j = 1, \ldots, n_i \\
G \mid \alpha, \psi &\sim \mathrm{DP}(\alpha, G_0(\psi)),
\end{aligned} \tag{4.2}
$$

where G_0 is the bivariate centering distribution for (γ_0, γ_1) that includes the $\gamma_1 > 0$ restriction. Comparison to more standard binomial models for the aggregate responses is ideal; some of the literature has presented models with a bernoulli or binomial kernel and a finite, somewhat structured mixture of probit curves for the probability of response. Despite basing the model on the underlying binary outcomes, model 4.2 is equivalently developed in terms of a binomial distribution for the set of $y_i = \sum_{j=1}^{n_i} y_{ij},\ i = 1, \ldots, N$, with a probit mixture for the probability of response.

Connection with the probit mixture: Taking the MGF approach, we show that M_1, the model for the sum of the responses starting from the underlying binary outcomes, is equivalent to M_2, the model for the aggregated responses $\{y_i'\}$ with a probit mixture probability of response and a binomial kernel.

$$
\begin{aligned}
E_{M_1}(e^{t\sum_{j=1}^{n_i} y_{ij}}; G) &= \int \sum_{\sum_{j=1}^{n_i} y_{ij}=0}^{n_i} e^{t\sum_{j=1}^{n_i} y_{ij}} \prod_{j=1}^{n_i} \text{Bern}(y_{ij}; \Phi(\gamma_0 + \gamma_1 x_i)) \text{d}G(\gamma_0, \gamma_1) \\
&= \int \prod_{j=1}^{n_i} \sum_{y_{ij}=0}^{1} e^{t y_{ij}} \text{Bern}(y_{ij}; \Phi(\gamma_0 + \gamma_1 x_i)) \text{d}G(\gamma_0, \gamma_1) \\
&= \int \prod_{j=1}^{n_i} \left(1 - \Phi(\gamma_0 + \gamma_1 x_i) + e^t \Phi(\gamma_0 + \gamma_1 x_i)\right) \text{d}G(\gamma_0, \gamma_1) \\
&= E_{M_2}(e^{t y_i'}; G)
\end{aligned}
$$

With the restriction on $\gamma_1 \in \mathbb{R}^+$, we may reparameterize the bernoulli kernel, and, consequently, the underlying latent response distribution, to a location-scale mixture of normal distributions. Specifically, let $\mu = -\gamma_0/\gamma_1$ and $\tau = 1/\gamma_1$, $\mu \in \mathbb{R}$ and $\tau > 0$. Given the above formulation and letting $G^* \sim \text{DP}(\alpha, G_0^*(\mu, \tau^2))$, we obtain

$$
\begin{aligned}
\Pr(y = 1; G, x) &= \int \Phi(\gamma_0 + \gamma_1 x) \text{d}G(\gamma_0, \gamma_1) \\
&= \int \Phi\left((-\mu/\tau) + (1/\tau)x\right) \text{d}G^*(\mu, \tau^2) = \int \Phi\left(\frac{x - \mu}{\tau}\right) \text{d}G^*(\mu, \tau^2) \\
&= \int \int_{-\infty}^{x} \text{N}(z; \mu, \tau^2) \text{d}z \text{d}G^*(\mu, \tau^2) = \int_{-\infty}^{x} \int \text{N}(z; \mu, \tau^2) \text{d}G^*(\mu, \tau^2) \text{d}z \\
&= \Pr(z \leq x; G^*),
\end{aligned}
$$

where the final probability arises under the location-scale normal DP mixture, $\int \text{N}(z; \mu, \tau^2) \text{d}G^*(\mu, \tau^2)$, for the latent response distribution. Because the (μ, τ^2) normal parameterization expedites prior specification, we work with the location-scale normal DP formulation (suppressing the G^* notation in the following).

As a computational tool, we choose to truncate the collection of mixing distributions at a suitable level. To represent the component to which each data point is

associated, configuration variables are introduced. Contrary to the methods discussed in Chapter 3, each binary response, y_{ij}, is assigned to any component and is thusly given an individual configuration variable, w_{ij}. Therefore, including the latent continuous responses z_{ij} associated with the y_{ij}, the hierarchical model for the data becomes:

$$\begin{aligned}
y_{ij} \mid z_{ij} &\overset{ind.}{\sim} \mathbf{1}\{y_{ij} = 1 \text{ iff } z_{ij} \leq x_i\};\ i = 1, \ldots, N;\ j = 1, \ldots, n_i \\
z_{ij} \mid \boldsymbol{\mu}, \boldsymbol{\tau}^2, \boldsymbol{w} &\overset{ind.}{\sim} \mathrm{N}(z_{ij}; \mu_{w_{ij}}, \tau^2_{w_{ij}}),\quad i = 1, \ldots, N;\ j = 1, \ldots, n_i \\
w_{ij} \mid \boldsymbol{p} &\overset{ind.}{\sim} \sum_{l=1}^{L} p_l \delta_l(w_{ij}),\quad i = 1, \ldots, N;\ j = 1, \ldots, n_i \\
(\mu_l, \tau_l^2) \mid \psi &\overset{ind.}{\sim} G_0(\mu_l, \tau_l^2 \mid \psi),\ l = 1, \ldots, L,
\end{aligned} \tag{4.3}$$

with $\boldsymbol{p} \mid \alpha \sim f(\boldsymbol{p} \mid \alpha)$, and $G_0 = \mathrm{N}(\mu_l; \beta_0, \sigma_0^2) \times$ inverse-gamma$(\tau^2; c, \delta)$. Here, $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_L)$, $\boldsymbol{\tau}^2 = (\tau_1^2, \ldots, \tau_L^2)$, and $\boldsymbol{w} = \{w_{ij} : i = 1, \ldots, N;\ j = 1, \ldots, n_i\}$. We place conjugate normal, inverse-gamma priors on the parameters of the prior for the μ_l, and a gamma prior on δ (with c fixed). As the model requires learning about the latent variables, the mixture parameters, and the hyperparameters, prior specification is key and expected to influence posterior inference.

Prior specification

As the number of parameters in the model is rather large, careful prior specification is seemingly important. However, we are able to define the hyperpriors on the parameters of G_0 only through the range of the dose levels. The mean of the normal prior on the μ_l, β, is centered at the mean of the dose levels and the variance is determined by the taking range of the dose levels, divide by four, and squaring. The variance parameter of the normal prior, σ^2, is given an inverse-gamma prior with shape

parameter defined by half the number of dose levels and scale parameter defined by the range of the dose levels. The τ_l^2 are given an inverse-gamma prior with fixed shape parameter (the number of dose levels minus 2) and scale parameter, δ. The exponential prior on δ is provided such that the mean is roughly a fourth of the dose range.

Specifically, for the trypanosome data where the dose range is from 4.7 to 5.4, the normal prior on β has mean 5.1 and variance 0.04, and σ^2 is given an inverse-gamma$(4, 0.7)$. The prior on the τ_l^2 has a fixed shape parameter at 6, and the scale parameter, δ, is given an exponential(6).

MCMC posterior simulation

The blocked Gibbs sampling approach (Ishwaran and Zarepour, 2000; Ishwaran and James, 2001) yields straight forward updates for all unknown parameters. For the quantal bioassay settings, the model simplifies to a normal mean-scale DP mixture. Also, with the assumption of full exchangeability across dose levels and within, some notation needs to be updated. Here, we denote the n^* distinct values of matrix $\boldsymbol{w}$ by $w_1^*, \ldots, w_{n^*}^*$, $M_k^* = |\{i, j : w_{ij} = w_k^*\}|$, $k = 1, \ldots, n^*$, and $M_l = |\{w_{ij} : w_{ij} = l\}|$, $l = 1, \ldots, L$. The weights of the DP and the mass parameter α are updated in the same fashion as in Chapter 2, the other parameter full conditionals are given below.

The latent variables are updated individually, depending on the value of y_{ij}. If $y_{ij} = 0$, the z_{ij} are drawn from $\mathrm{N}(\mu_{w_{ij}}, \tau_{w_{ij}}^2)$ truncated above at 0. When $y_{ij} = 1$, the z_{ij} are drawn from the normal distribution truncated below at 0. The full conditionals for the distinct components of the (μ_l, τ_l^2) are given in closed form. That is, for the $l \in \{w_k^* : k = 1, \ldots, n^*\}$, the $\mu_{w_k^*}$ arise from a normal with mean $\left(\{\sum\sum_{i,j:w_{ij}=w_k^*} z_{ij}\}/\tau_{w_k^*}^2 + \beta_0/\sigma_0^2\right) / \left(M_k^*/\tau_{w_k^*}^2 + 1/\sigma_0^2\right)$ and variance $\left(M_k^*/\tau_{w_k^*}^2 + 1/\sigma_0^2\right)^{-1}$. Similarly, the $\tau_{w_k^*}^2$ are drawn from an inverse-gamma distribution

with shape $c + 0.5M_k^*$ and scale $\delta + 0.5 \sum \sum_{i,j:w_{ij}=w_k^*} (\mu_k^* - z_{ij})^2$. Finally, the subject specific configuration variables are individually from a discrete distribution with weights proportional to $p_l \mathrm{N}(z_{ij}; \mu_l, \tau_l^2)$, for $l = 1, \ldots, L$

The hyperparameters of G_0 are given conjugate priors. The mean of the prior on the μ_l, β_0, is given a normal prior and is drawn from a normal distribution with mean $\left(\sum_{l=1}^{L} \mu_l/\sigma_0^2 + m_{\beta_0}/s_{\beta_0}^2\right) / \left(L/\sigma_0^2 + 1/s_{\beta_0}^2\right)$ and variance $\left(L/\sigma_0^2 + 1/s_{\beta_0}^2\right)^{-1}$. The variance of the prior on the μ_l, σ_0^2, is given an inverse-gamma prior and is drawn from an inverse-gamma with updated parameters $0.5L + a_{\sigma_0^2}$ and $b_{\sigma_0^2} + 0.5 \sum_{l=1}^{L} (\mu_l - \beta_0)^2$. Finally, δ, the scale parameter of the inverse-gamma prior for the τ_l^2, is given a gamma prior and the updated shape parameter is $Lc + a_\delta$ and scale $b_\delta + \sum_{l=1}^{L} \tau_l^2$.

Samples from the posteriors of these parameters are used in posterior inference for risk assessment.

Posterior inference

The most important inference in terms of the quantal bioassay setting is for the dose-response curve. This curve, defined as the probability of response as a function of dose (x), is described by a finite mixture of probit functions.

$$D(x) = \Pr(y = 1; G, x) = \Pr(z \leq x; G) = \sum_{l=1}^{L} p_l \Phi\left(\frac{x - \mu_l}{\tau_l}\right)$$

Section 4.2.1 elucidates the monotonicity restriction in terms of the (γ_0, γ_1) origination. This restriction carries over to the tolerance formulation by assumption, as τ^2 is a variance parameter and thus positive. Hence, the dose-response curve is necessarily monotonic by virtue of the inherent properties of the linear-DDP prior model.

Another inferential objective is inversion, where interest lies in estimation of

the dose-level, x_q, corresponding to a specific probability, q. Under the linear DDP model, this is obtained by numerically inverting the posterior realizations of the dose-response curve at each iteration of the MCMC; that is, given the posterior sample path of the dose-response curve at each step of the MCMC algorithm, invert $D(x_q) = q$.

In terms of calibration, we consider a specified vector of responses, $\boldsymbol{y}_0 = \{y_{0j} : j = 1, ..., n_0\}$ and seek to estimate the dose level, x_0, which is associated with this new vector of responses. This inference can be obtained by augmenting the hierarchical model with the components associated with $\boldsymbol{y}_0$ and expanding the parameter vector to include the unknown dose x_0. Key is the fact that the DDP prior is defined over the uncountable space $\mathcal{X}$, and it thus induces a proper hierarchical model for any x_0. The calibration problem is challenging for many models, however, our approach lends itself for a seemingly straightforward update. The full hierarchical model for calibrating dose x_0 is given by:

$$\begin{aligned}
\{y_{ij}\} \mid \{z_{ij}\}, x_0 &\sim \prod_{i=0}^{N} \prod_{j=1}^{n_i} \mathbf{1}\{y_{ij} = 1 \text{ iff } z_{ij} \leq x_i\} \\
\{z_{ij}\} \mid \{w_{ij}\}, \boldsymbol{\mu}, \boldsymbol{\tau}^2 &\sim \prod_{i=0}^{N} \prod_{j=1}^{n_i} \mathrm{N}(z_{ij}; \mu_{w_{ij}}, \tau^2_{w_{ij}}) \\
\{w_{ij}\} \mid \boldsymbol{p} &\sim \prod_{i=0}^{N} \prod_{j=1}^{n_i} \sum_{l=1}^{L} p_l \delta_l(w_{ij}) \\
(\mu_l, \tau_l^2) \mid \psi &\overset{i.i.d.}{\sim} G_0(\mu_l, \tau_l^2 \mid \psi),\ l = 1, \ldots, L \\
x_0 &\sim \mathrm{N}(x_0; h, v^2)
\end{aligned}$$

where $\boldsymbol{p} \mid \alpha \sim f(\boldsymbol{p} \mid \alpha)$, and G_0 is defined as before. Conditioned on x_0, the main MCMC

updates are the same. The dose level, x_0, has posterior full conditional distribution,

$$x_0 \mid \boldsymbol{z}, \text{data} \propto \mathrm{N}(x_0; h, v^2) \prod_{j=1}^{n_0} \mathbf{1}\{y_{0j} = 1 \text{ iff } z_{0j} \leq x_0\}.$$

Therefore, x_0 must satisfy $x_0 \geq a_{x_0} = max_{j:y_{0j}=1}\{z_{0j}\}$, and $x_0 < b_{x_0} = min_{j:y_{0j}=0}\{z_{0j}\}$. Combining these conditions, the update for x_0 follows a normal distribution with mean h, variance v^2, truncated below at a_{x_0}, and truncated above at b_{x_0}. Clearly, a conscientious prior is essential for sufficient learning. To this end, we find the quartile of the dose range that roughly corresponds to the proportion of positive responses to center the prior and divide the variance of the prior placed on β_0 by a factor of four to estimate the spread of the prior on x_0.

4.2.3 Data illustrations

We illustrate the linear-DDP model on two data sets, one studied throughout the literature and the other an application to cytogenetic dosimetry.

Trypanosome data application

We highlight the effectiveness of our model with a data set commonly studied in the literature. Ashford and Walker (1972) first analyzed the data related with death rate of protozoan trypanosome found in Table 4.1. Trypanosomes are parasites which do not grow readily in artificial culture (when they do grow they change from the blood form to the insect form) and no reliable quantitative genetic characters have been found nor have accurate methods been available to assess characters such as drug resistance. Here, we look at results from the trypanosome sensitivity to the log concentrations of pure neutral acriflavine (Walker, 1966).

Table 4.1: Trypanosome data reported in Ashford and Walker (1972).

Log concentration	4.7	4.8	4.9	5.0	5.1	5.2	5.3	5.4
Exposed	55	49	60	55	53	53	51	50
Dead	0	8	18	18	22	37	47	50

Using the linear-DDP model, we obtain the posterior mean (solid) and 90% uncertainty bands (dashed) for the dose-response curve; see the left panel of Figure 4.1. The smooth posterior mean curve follows closely the path of the observed data, and the interval bands include the majority of the points. The extremes of the observed proportions are at zero and one, which are known to be burdensome for most parametric models. In general, the proposed nonparametric mixture model captures well the non-standard, bimodal shape for the tolerance distribution suggested by the data.

The middle panel of Figure 4.1 gives the posterior estimates for dose levels, x_q, for inversion probabilities $q = 0.05$, 0.15, 0.25, and 0.5. The posterior range for $x_{0.25}$ is comparable to that found in Mukhopadhyay (2000), and the spread of the densities depends on the width of the probability intervals around the dose-response curve at the given response probability. The final inference reported in Figure 4.1 (right panel) is for the calibrated dose level, x_0, corresponding to a new response vector of $y_0 = 26$ and $n_0 = 52$. This new response vector resembles the observed counts at dose 5.1, a value captured by the posterior density for x_0.

In terms of the calibrated dose level, we look into sensitivity of the prior specification. We begin with a normal prior which includes the majority of the dose range in its support, see the right panel of Figure 4.1. The mean is roughly 5.18 with a standard deviation of about 0.04. Figure 4.2 gives the calibrated dose estimates for the normal prior (left), a uniform prior from [4.9, 5.25] in the middle panel, and a uniform

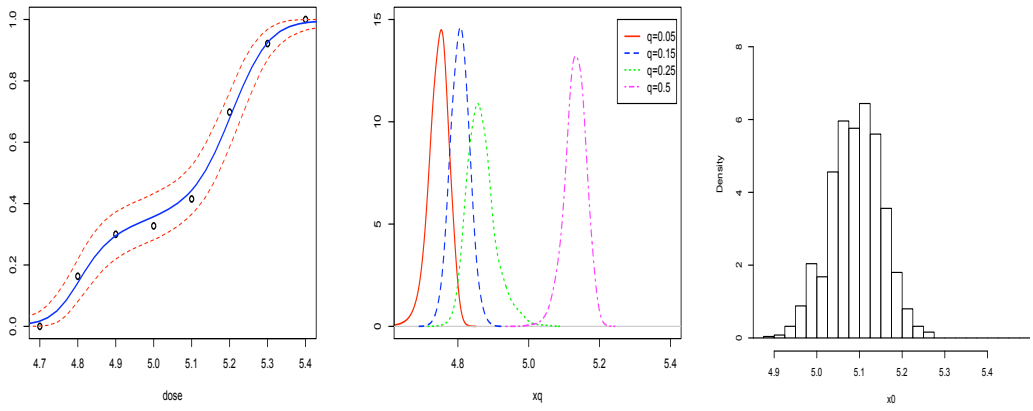

Figure 4.1: Trypanosome data. The posterior mean estimate (solid line) and 90% uncertainty bands (dashed lines) for $\Pr(y = 1; G_{\mathcal{X}})$ (left panel); inversion estimates, x_q for $q = 0.05, 0.15, 0.25$ and 0.5 (middle panel); and the estimated calibrated dose level given $y_0 = 26$ and $n_0 = 52$ (right panel).

prior over the entire dose range (right panel). In all three cases, the mean is close to 5.18, but the spread differs. The uniform prior across the full dose range has the largest spread and the spread decreases as the uniform support shrinks or switches to normal distribution. The calibrated dose level is relatively robust to the prior assumed, though the uniform priors result in better mixing.

Cytogenetic dosimetry application

Cytogenetic dosimetry is a biological tool for dose assessment in cases of radiological accidents and suspected overexposures. The main focus of these studies is to determine the relationship between the dosage of exposure to radiation and some measure of genetic aberration. The typical experiment considers samples of cell cultures exposed to a range of levels of an agent with some measure of cell disability recorded as the response. The two main inferential objectives include the prediction of response

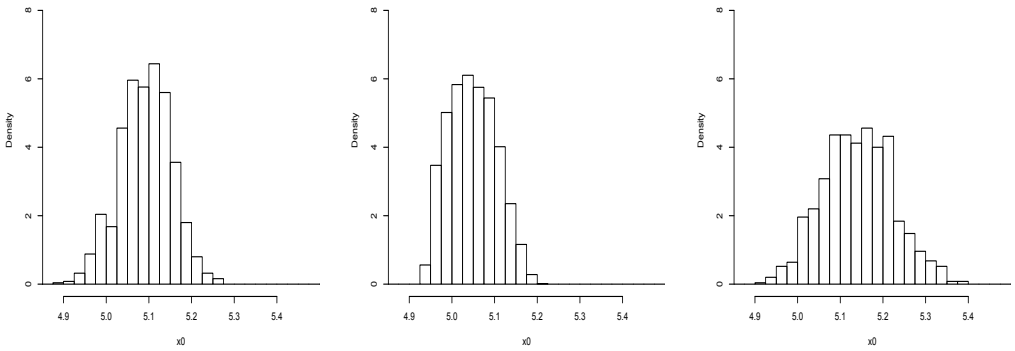

Figure 4.2: Trypanosome data. The posterior estimated calibrated dose level given $y_0 = 26$ and $n_0 = 52$ under a normal prior (left panel), uniform prior from [4.9, 5.25] (middle panel), and a uniform prior across the full dose range (right panel).

at unobserved exposure levels and inference for unknown dose levels given observed responses.

For an illustration, we consider part of a larger data set where the blood samples from individuals were exposed to ^{60}Co radiation at doses of 0, 20, 50, 100, 200, 300, 400, and 500 cGy (centograms). The resulting cultures were tested for binucleated cells with a recorded number of micronuclei. The full data set, found in Madruga et al. (1996), groups the individuals into none, one, or two or more micronuclei, see Table 4.2. The data is grouped into these categories as, when many micronuclei are present, they can be difficult to count exactly. We collapse to two classification groups, no micronuclei and one or more micronuclei (the two columns on the left of Table 4.2), to apply the nonparametric mixture model developed in this chapter for quantal response data. Further exploration of the ordinal response vector of the number of cells with 0 MN, exactly 1 MN, and 2 or more MN is found in Section 4.3.

Given in Figure 4.3 are the posterior distributions of the dose-response curve

Table 4.2: Cytogenetic dosimetry data. The cell counts for the binucleated cells in the healthy, older patients as reported in Madruga et al. (1996).

^{60}Co radiation	≥ 2 MN	exactly 1 MN	0 MN
0 cGy	2	31	920
20 cGy	8	41	989
50 cGy	14	56	933
100 cGy	32	114	939
200 cGy	67	176	794
300 cGy	59	209	683
400 cGy	107	256	742
500 cGy	143	327	771

(left panel), the inversion dose levels given $q = 0.05, 0.15, 0.25$ and 0.3 (middle panel), and the estimated calibrated dose level given $y_0 = 146$ and $n_0 = 1085$. The 90% probability bands of the dose-response curve are much tighter than those found in the Trypanosome data, as the number of subjects per dose level is roughly 20 times larger in the cytogenetic dosimetry data. The data is only observed to about forty percent of the full curve, and therefore the intervals get much wider near the end of the observed data. The data points all lie within the uncertainty bands. The inversion inference is interesting in that at the smaller levels of q (0.05 and 0.15), the estimates of the corresponding dose level are unimodal and have relatively narrow distributions. The larger values (0.25 and 0.3) result in distributions with heavier tails and substantially larger spread. This coincides with the uncertainty bands around the dose-response curve in these areas. Finally, the calibrated dose level for $y_0 = 146$ and $n_0 = 1085$, the observed counts at 100 cGy, is estimated to be lower. The mean is approximately 77.5 cGy, with 90% probability interval (59.2, 94.3).

As discussed above, we only observe part of the dose response curve in cytogenetic dosimetry experiments. We look at calibration for $y_0 = 146$ and $n_0 = 1085$,

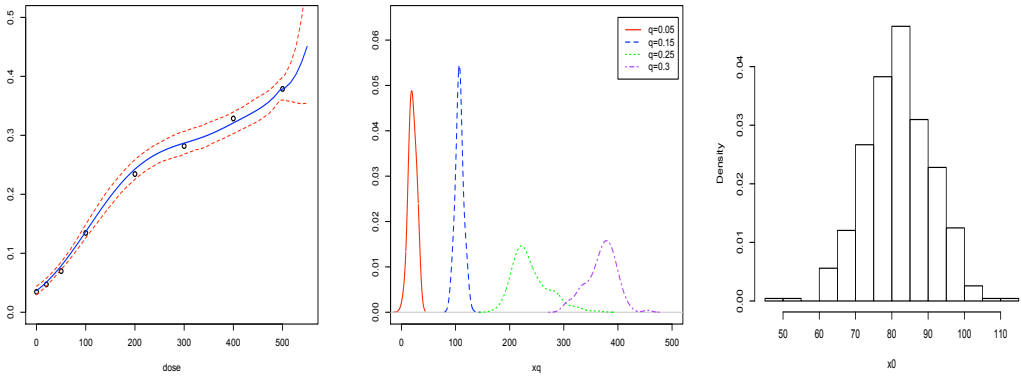

Figure 4.3: Cytogenetic dosimetry. The posterior mean estimate (solid line) and 90% uncertainty bands (dashed lines) for $\Pr(y = 1; G_{\mathcal{X}})$ (left panel); inversion estimates, x_q for $q = 0.05, 0.15, 0.25$ and 0.3 (middle panel); the estimated calibrated dose level given $y_0 = 146$ and $n_0 = 1085$ (right panel).

an observed dose level, and also for two configurations that lie beyond the observed dose-response curve (i.e. $y_0 = 500,\ n_0 = 1000$ and $y_0 = n_0 = 975$). As seen in Figure 4.4, we expect the spread to widen as the proportions become farther from those observed.

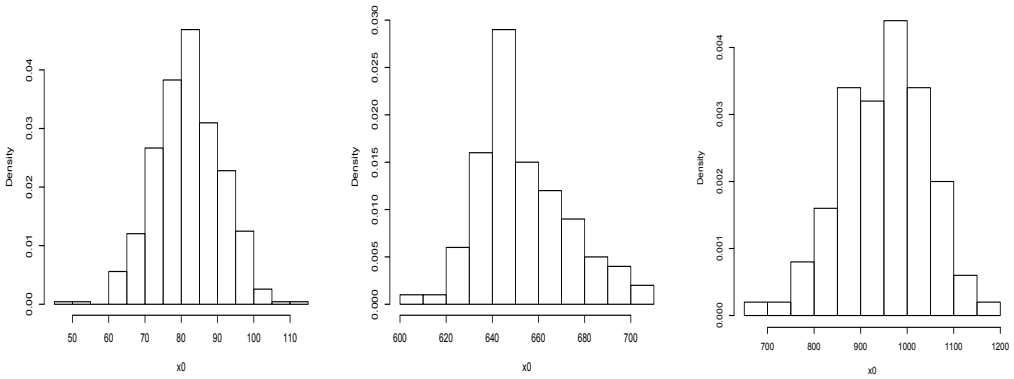

Figure 4.4: Cytogenetic dosimetry. The estimated calibrated dose level given $y_0 = 146$ and $n_0 = 1085$ (left panel), given $y_0 = 500$ and $n_0 = 1000$ (middle panel), and given $y_0 = n_0 = 975$ (right panel).

4.3 Modeling extensions for bioassay experiments with ordinal responses

Ordinal response data are observed frequently in laboratory experiments. For example, the cytogenetic dosimetry application from the previous section resulted in three ordered classifications. The approaches presented here aim to model the general case of bioassay experiments with R ordinal categories. However, we focus on $R = 3$, as in the cytogenetic dosimetry application, and refer to the example from Section 4.2.3 for context. Let $Y = r$, with $r = 1$ if the cell has 2 or more MN, 2 if the cell has exactly 1 MN, and 3 if the cell has no MN, such that 0 MN can be viewed as the reference category. We also denote $\pi_r(x) = \Pr(Y = r; x)$, for $r = 1, 2, 3$, and $F_r(x) = \Pr(Y \leq r; x)$, for $r = 1, 2$ (where $F_3(x) = 1$).

4.3.1 Semiparametric and nonparametric strategies

The modeling approach of 4.2.2 involved latent continuous responses with a flexible DP mixture supporting the rich inference for the quantal dose-response relationship. A logical starting point for the generalization to ordinal responses is to explore the cumulative link regression setting. Parametric cumulative link regression models are built through the F_r using latent continuous responses taking values on the real line (e.g., Albert and Chib (1993), Erkanli et al. (1993), Albert and Chib (1995), Agresti (2002)). That is, $Y = r$ if and only if $a_{r-1} < U \leq a_r$ for $r = 1, 2, 3$, with $-\infty = a_0 < a_1 < a_2 < a_3 = \infty$. Given the continuous latent response U has cumulative distribution function $G(u) \equiv G(u - \theta)$ with location parameter $\theta = x\beta$ and a fixed scale parameter, $F_r(x) = \Pr(Y \leq r; x) = \Pr(U \leq a_r; x) = G(a_r - x\beta)$, for $r = 1, 2$. Accordingly, for the cumulative probit model, we have $F_r(x) = \Phi(a_r - x\beta)$,

$r = 1, 2$, and $\Pr(Y = 1; x) = \Phi(a_1 - x\beta)$, $\Pr(Y = 2; x) = \Phi(a_2 - x\beta) - \Phi(a_1 - x\beta)$, and $\Pr(Y = 3; x) = 1 - \Phi(a_2 - x\beta)$. Note the cutoffs, a_r, $r = 1, 2$, are assumed to be random, and consequently there is no need for an intercept in the location parameter; alternatively, the first cutoff point can be fixed. These cutoffs can be particularly difficult to ascertain during posterior simulation as they are highly correlated with the other parameters (see Lawrence et al. (2008) and further references therein for computational techniques to efficiently update the cutoff values). The parametric latent response distribution may also be too restrictive to appropriately capture skewness or multimodality that should be reflected in the dose-response relationships. Hence, we move to the arena of Bayesian nonparametrics.

We replace the parametric latent response distribution with a mixture of normals, a common distribution driving the cumulative probabilities. Similar to the argument found in Kottas et al. (2005), we fix the cutoff points and let the mixture distribution adapt to give the appropriate mass under the curve between each cutoff. Therefore, we may assume $f(u; G_x) = \int \mathrm{N}(u; \theta, 1)\mathrm{d}G_x(\theta)$, where we give G_x a DDP prior with a GP base stochastic process. Using the special case of the linear-DDP, we simplify the truncated mixing distribution to $G_x^L = \sum_{l=1}^{L} p_l \delta_{(\gamma_{0l} - \gamma_{1l} x)}$, which is equivalent to $G^L = \sum_{l=1}^{L} p_l \delta_{(\gamma_{0l}, \gamma_{1l})}$, and subsequently

$$f(u; G, x) = \int \mathrm{N}(u; \gamma_0 - \gamma_1 x, 1)\mathrm{d}G(\gamma_0, \gamma_1), \quad G \sim \mathrm{DP}(\alpha, G_0).$$

Now, $\pi_1(x) = \Pr(u \leq a_1; G, x) = \sum_{l=1}^{L} p_l \Phi(a_1 - \gamma_{0l} + \gamma_{1l} x)$, which is increasing in x provided $\gamma_{1l} > 0$, for all $l = 1, \ldots, L$. Also, $\pi_2(x) = \Pr(a_1 < u \leq a_2; G, x) = \sum_{l=1}^{L} = p_l\{\Phi(a_2 - \gamma_{0l} + \gamma_{1l} x) - \Phi(a_1 - \gamma_{0l} + \gamma_{1l} x)\}$, and $\pi_3(x) = \Pr(u > a_2; G, x) = 1 - \sum_{l=1}^{L} p_l \Phi(a_2 - \gamma_{0l} + \gamma_{1l} x)$. Here, $\pi_3(x)$ is decreasing in x given the $\gamma_{1l} > 0$ restriction,

and $\pi_2(x)$ is not restricted to a monotonic shape; this agrees with the hypothesis that the number of cells with exactly 1 MN increases to a point then decreases as we expect exposure to the higher dose levels to produce cells with 2 or more MN. Finally, the cumulative probabilities, $F_1(x)$ and $F_2(x)$ are stochastically ordered by definition.

By replacing the parametric response distribution, we provide flexibility for the dose-response curves through the cumulative probabilities. As it is an extension of the model found in the preceding section, implementation of the model is only marginally more complicated. Also, the calibration objective is obtained in a straightforward fashion. However, this formulation for the latent response distribution does not lend itself to reparameterization as a tolerance distribution as in the binary case. And, developing the model for the ordered responses in this manner may carry over the potential restriction of parametric cumulative link models; that is, the cumulative probabilities, $F_r(x)$, have the same shape and change only with a shift. In fact, we should expect some inflexibility in inference for two dose-response curves driven by a single random distribution.

To rectify the conceivable limitations of the cumulative link regression setting, we turn to the approach of modeling the dose-response curves through flexible response distributions as highlighted throughout the dissertation. That is, given generic trinary response vector $\boldsymbol{y} = (y_1, y_2, y_3)$, such that $y_i = 0, 1$ with $\sum_{r=1}^{3} y_r = 1$, we assume the prior model

$f_0(\boldsymbol{y}; G_x) = \int k(\boldsymbol{y}; \pi(\theta_1), \pi(\theta_2)) \mathrm{d}G_x(\theta_1, \theta_2)$, where

$$k(\boldsymbol{y}; \pi(\theta_1), \pi(\theta_2)) = \pi(\theta_1)^{y_1} \pi(\theta_2)^{y_2} (1 - \pi(\theta_1) - \pi(\theta_2))^{y_3}.$$

Here, $\pi(\theta_r)$, $r = 1, 2$, is an appropriate scale for the probabilities (e.g., logit or probit)

such that θ_r are real-valued parameters. In the same spirit as the models of Chapter 3 for developmental toxicity studies and the discussion at the beginning of Chapter 4, we place a general DDP prior on $G_{\mathcal{X}} = \{G_x : x \in \mathcal{X}\}$ with a product of GPs each with a linear mean function as the base stochastic process such that $G^L_{\mathcal{X}} = \sum_{l=1}^{L} p_l \delta_{(\eta_{1l\mathcal{X}}, \eta_{2l\mathcal{X}})}$, under the standard truncation approximation. Here, $(\eta_{1l\mathcal{X}}, \eta_{2l\mathcal{X}})$ are realizations from the independent GPs driving each of the probabilities of response.

Bearing in mind the dose-response curve monotonicity assumptions, we once more simplify the DDP prior to the linear-DDP, whereby $\eta_{1l\mathcal{X}} \equiv \eta_{1l}(x) = \beta_{0l} + \beta_{1l}x$, $\eta_{2l\mathcal{X}} \equiv \eta_{2l}(x) = \gamma_{0l} + \gamma_{1l}x$, and $G^L_x \equiv G^L = \sum_{l=1}^{L} p_l \delta_{((\beta_{0l}, \beta_{1l}), (\gamma_{0l}, \gamma_{1l}))}$. Then, the mixture model for the vector of trinary responses becomes

$$f_0(\boldsymbol{y}; G, x) = \int k(\boldsymbol{y}; \pi(\beta_0 + \beta_1 x), \pi(\gamma_0 + \gamma_1 x)) \mathrm{d}G((\beta_0, \beta_1), (\gamma_0, \gamma_1)),$$

with $G \sim \mathrm{DP}(\alpha, G_0)$. Given this construction, $\pi_1(x) = \sum_{l=1}^{L} p_l \pi(\beta_{0l} + \beta_{1l}x)$ and $\pi_2(x) = \sum_{l=1}^{L} p_l \pi(\gamma_{0l} + \gamma_{1l}x)$, which are both increasing in x provided $\beta_{1l} > 0$ and $\gamma_{1l} > 0$, for $l = 1, \ldots, L$. That is, the first two dose-response curves are monotonically increasing in x, and the curve for the third classification is decreasing. In some experiments, allowing the second classification to increase in probability as dose level increases may be a desirable assumption. Also, the cumulative probabilities, $F_1(x) = \sum_{l=1}^{L} p_l \pi(\beta_{0l} + \beta_{1l}x)$ and $F_2(x) = \sum_{l=1}^{L} p_l \{\pi(\beta_{0l} + \beta_{1l}x) + \pi(\gamma_{0l} + \gamma_{1l}x)\}$, are both increasing functions of x under the positive slope coefficient restriction.

By flexibly modeling the response distribution, we induce flexibility in the probabilities of response. Indeed, with these assumptions we may obtain entirely different shapes for the dose-response curves, a feature that is not possible under the semiparameteric cumulative link regression model. There are two random distributions

driving the two dose-response curves, providing more robust inference. On the other hand, the MCMC details become more complex. As it is impossible to separate the full conditionals for the distinct components of the mixture, a four-dimensional Metropolis-Hastings update is required. Moreover, the calibration objective is no longer routine and will likely be difficult to learn due to highly correlated parameters. The proposed model, developed in the next section, is based on a different construction, which retains and potentially expands on the flexibility of the model under $f_0(\boldsymbol{y}; G, x)$ while avoiding some of the implementation issues.

4.3.2 The nonparametric mixture modeling approach

Each of the two approaches we have presented in Section 4.3.1 have strong and weak points. The drawbacks are not detrimental if considered before the analysis and are outweighed by the advantages. However, an alternate method is which incorporates the continuation-ratio logits structure into the mixture kernel of model $f_0(\boldsymbol{y}; G, x)$ may be optimal. That is, we separate the kernel, $k(\boldsymbol{y}; \pi(\theta_1), \pi(\theta_2))$, into a marginal bernoulli kernel for the probability of a cell having 2 or more MN, $\rho(\varphi_1)$, and a conditional bernoulli kernel with probability of having either exactly 1 MN given that there are 1 or 0 MN, $\rho(\varphi_2)$. Here, $\rho(\cdot)$ is a transform from the real line to the unit interval; the logistic transformation is used for the application in 4.3.3. Note that $\rho(\varphi_1) = \pi(\theta_1)$ and $\rho(\varphi_2) = \pi(\theta_2) \mid (1 - \pi(\theta_1))$. Therefore, while mixing is with respect to $\rho(\varphi_1)$ and $\rho(\varphi_2)$, the kernel can be written in either the product Bernoulli or trinomial form.

We again employ the linear-DDP prior for the mixing distributions, thereby maintaining the possibility for monotonic dose-response curves. Under this prior, we

arrive at the mixture model

$$f(\boldsymbol{y}; G, x) = \int \text{Bern}(y_1; \rho(\beta_0 + \beta_1 x))\text{Bern}(y_2 \mid y_1 = 0; \rho(\gamma_0 + \gamma_1 x))\text{d}G((\beta_0, \beta_1), (\gamma_0, \gamma_1))$$
$$= k\left(\boldsymbol{y}; \rho(\beta_0 + \beta_1 x), \rho(\gamma_0 + \gamma_1 x)(1 - \rho(\beta_0 + \beta_1 x))\right) \text{d}G((\beta_0, \beta_1), (\gamma_0, \gamma_1)),$$

with $G \sim \text{DP}(\alpha, G_0)$ and G_0 is a product of normal distributions for the intercept terms, (β_0, γ_0), and a product of exponential distributions for the slope parameters, (β_1, γ_1). Consequently, the hyperparameters of G_0 are the means, μ and θ, and variances, τ^2 and σ^2, for the normal distributions, and the means, δ and φ, of the exponential distributions (thereby, $\psi = (\mu, \tau^2, \delta, \theta, \sigma^2, \varphi)$). The hyperparameters are given conjugate priors.

Dose-response relationships

Regarding the dose-response curves implied by the mixture model $f(\boldsymbol{y}; G, x)$, the probability of a generic cell having 2 or more MN across all $x \in \mathcal{X}$,

$$\Pr(y_1 = 1; G, x) = \sum_{l=1}^{L} p_l \rho(\beta_{0l} + \beta_{1l} x),$$

is monotonically increasing given the restriction $\beta_{1l} > 0$, $l = 1, \ldots, L$. Contrarily, the probability of the cell having exactly 1 MN,

$$\Pr(y_2 = 1; G, x) = \sum_{l=1}^{l} p_l \rho(\gamma_{0l} + \gamma_{1l} x)[1 - \rho(\beta_{0l} + \beta_{1l} x)], \quad x \in \mathcal{X},$$

which, in common with the cumulative link probabilities model, does not have the monotonic property. The dose-response curve for 0 MN, can be expressed as

$$\Pr(y_3 = 1; G, x) = \sum_{l=1}^{L} p_l(1 - \rho(\beta_{0l} + \beta_{1l}x))(1 - \rho(\gamma_{0l} + \gamma_{1l}x)), \quad x \in \mathcal{X},$$

and is therefore decreasing in x under the additional restriction $\gamma_{1l} > 0$ for $l = 1, \ldots, L$. Despite the fact that $\Pr(Y_2 = 1; G, x)$ is not forced to be non-decreasing, restricting both $\beta_{1l} > 0$ and $\gamma_{1l} > 0$, for $l = 1, \ldots, L$, results in monotonically increasing cumulative probabilities, $F_1(x) = \Pr(y_1 = 1; G, x)$ and $F_2(x) = \Pr(y_1 = 1; G, x) + \Pr(y_2 = 1; G, x)$.

MCMC posterior simulation

Let $\boldsymbol{y}_{ij} = (y_{ij1}, y_{ij2}, y_{ij3})$ be the $j = 1, \ldots, n_i$ response vectors at dose level x_i, $i = 1, \ldots, N$. Moreover, let (β_{0j}, β_{1j}) and $(\gamma_{0j}, \gamma_{1j})$ be the latent mixing parameters driving the probabilities of the product Bernoulli mixture kernel.

For posterior inference, we implement a blocked Gibbs sampler under the truncation approximation as in the quantal response case. We represent the hierarchical model for the data, with subject specific configuration variables, w_{ij}, as

$$\begin{aligned}
\{\boldsymbol{y}_{ij}\} \mid \boldsymbol{\beta}_0, \boldsymbol{\beta}_1, \boldsymbol{\gamma}_0, \boldsymbol{\gamma}_1, \boldsymbol{w} &\sim \prod_{i=1}^{N}\prod_{j=1}^{n_i} \text{Bern}(y_{ij1}; \rho(\beta_{0w_{ij}} + \beta_{1w_{ij}}x_i))\text{Bern}(y_{ij2}; \rho(\gamma_{0w_{ij}} + \gamma_{1w_{ij}}x_i)) \\
w_{ij} \mid \boldsymbol{p} &\sim \prod_{i=1}^{N}\prod_{j=1}^{n_i}\sum_{l=1}^{L} p_l\delta_l(w_{ij}) \\
\boldsymbol{p}, \boldsymbol{\beta}_0, \boldsymbol{\beta}_1, \boldsymbol{\gamma}_0, \boldsymbol{\gamma}_1 \mid \alpha, \psi &\sim f(\boldsymbol{p} \mid \alpha) \times \prod_{l=1}^{L} G_0(\beta_{0l}, \beta_{1l}, \gamma_{0l}, \gamma_{1l} \mid \psi), \qquad (4.4)
\end{aligned}$$

where $f(\boldsymbol{p} \mid \alpha)$ follows the generalized Dirichlet in Chapter 2 and G_0 is defined as before.

The model is completed with priors on ψ, the parameters of G_0, and α. The MCMC details for the hyperparameters of the base distribution are similar to the quantal response linear-DDP mixture model (4.2.2). The differences lie in the updates for the distinct components of the mixing parameters and for the w_{ij}.

The n^* distinct components for the mixing parameters $(\beta_{0w_k^*}, \beta_{1w_k^*})$ and $(\gamma_{0w_k^*}, \gamma_{1w_k^*})$, for $k = 1, \ldots, n^*$, are updated in two blocks in similar fashions. The first full conditional is proportional to

$g_0((\beta_{0w_k^*}, \beta_{1w_k^*}) \mid \psi) \prod_{\{i,j:w_{ij}=w_k^*\}} \prod \text{Bern}(y_{ij1}; \rho(\beta_{0w_k^*} + \beta_{1w_k^*} x_i))\text{Bern}(y_{ij2}; \rho(\gamma_{0w_k^*} + \gamma_{1w_k^*} x_i))$, and analogously for the $(\gamma_{0w_k^*}, \gamma_{1w_k^*})$. For each block, we implement a random-walk Metropolis-Hastings update, with a bivariate normal proposal (with $\beta_{1w_k^*}$ and $\gamma_{1w_k^*}$ on the log-scale) and a scaled identity covariance matrix. The configuration variables are updated on a subject specific basis from a discrete distribution on $\{1, \ldots, L\}$ with weights proportional to $p_l \text{Bern}(y_{ij1}; \rho(\beta_{0l} + \beta_{1l} x_i))\text{Bern}(y_{ij2}; \rho(\gamma_{0l} + \gamma_{1l} x_i))$ for $l = 1, \ldots, L$.

Posterior inference

Given the samples from the posterior distribution of model (4.2), we obtain inference for the dose-response curves by evaluating their expressions developed above. For instance, for the probability of having 2 or more MN, posterior samples are obtained through $\sum_{l=1}^{L} p_l \rho(\beta_{0l} + \beta_{1l} x)$, for all x whether observed or unobserved. The dose-response curves and the cumulative probability functions are used for risk assessment, along with the scientifically important inferential objective of calibration.

The approach to calibration requires augmenting the data with a new vector of responses, and the goal is to estimate the dose level, x_0, corresponding to that vector. This problem is challenging under many modeling techniques, including classical

nonparametric estimation. Parametric Bayesian models (e.g., Madruga et al., 1994, 1996) are limited in that the dose-response relationship is introduced with a functional form. The corresponding calibration inference relies heavily on the assumption of the restriction on the shape of the curve. Even less attention has been paid to this type of inference in the Bayesian nonparametric literature; the only available reference being Kottas et al. (2002), where the dose-response curves are modeled directly through a combination of DP priors. The proposed linear-DDP mixture model improves on inference in that it avoids the discreteness of the DP prior, is simpler to extend to $R > 3$ ordinal classifications, and has a more tenable computational approach.

Given new responses $\boldsymbol{y}_0 = (\boldsymbol{y}_{01}, \ldots, \boldsymbol{y}_{0n_0})$, the hierarchical model becomes

$$\{\boldsymbol{y}_{ij}\} \mid \boldsymbol{\beta}_0, \boldsymbol{\beta}_1, \boldsymbol{\gamma}_0, \boldsymbol{\gamma}_1, \boldsymbol{w} \sim \prod_{i=0}^{N} \prod_{j=1}^{n_i} \text{Bern}(y_{ij1}; \rho(\beta_{0w_{ij}} + \beta_{1w_{ij}} x_i)) \text{Bern}(y_{ij2}; \rho(\gamma_{0w_{ij}} + \gamma_{1w_{ij}} x_i))$$

$$w_{ij} \mid \boldsymbol{p} \sim \prod_{i=0}^{N} \prod_{j=1}^{n_i} \sum_{l=1}^{L} p_l \delta_l(w_{ij}),$$

where the weights and mixing parameters arise as in (4.2), G_0 is defined as before, and a prior $\pi(x_0)$ is placed on x_0. The main MCMC updates remain the same conditioned on x_0, with the added vector of responses. Conditioned on the other parameters, the update for x_0 is proportional to $\pi(x_0) \prod_{l=1}^{L} \text{Bern}(y_{0j1}; \rho(\beta_{0w_{0j}} + \beta_{1w_{0j}} x_0)) \text{Bern}(y_{0j2}; \rho(\gamma_{0w_{0j}} + \gamma_{1w_{0j}} x_0))$, requiring a Metropolis-Hastings step. Here, we use a random walk with a normal proposal, under either dispersed normal or uniform priors; however, mixing proves to be marginally preferable under the uniform prior. Careful tuning bequeaths acceptance rates around 0.20.

4.3.3 Cytogenetic dosimetry application amendment

As in the previous analysis, we consider the same part of a larger data set where the blood samples from individuals were exposed to ^{60}Co radiation at doses of 0, 20, 50, 100, 200, 300, 400, and 500 cGy (see Table 4.2).

Figure 4.5 gives the mean and 90% interval bands of the smooth dose-response curves (left panel). The probability of a cell culture having 2 or more MN is monotonically increasing across the dose levels and follows the data (given by "o"). The probability of the second group (exactly 1 MN, data shown as "x") is increasing through dose 500 cGy. The data fall within the probability bands. The last group, 0 MN (data "$\diamond$"), is defined through the other two dose-response curves and is decreasing monotonically. Again, the data all lie close within the interval bands. The curves capture the non-standard forms suggested by the data with reasonable uncertainty, and, as the three curves are driven by two distributions, they have different shapes. The extrapolated dose response curves are given in the middle panel of Figure 4.5. The curves are equivalent to the left panel up to dose 500 cGy, then the curves level off and the interval bands widen significantly due to the lack of data.

The right panel of Figure 4.5 plots estimates for the cumulative probabilities, $\Pr(Y = 1; G, x)$ (lower curve, data given by "o") and $\Pr(Y = 2; G, x)$ (upper curve, data shown as "x"). Note that the first cumulative probability is equivalent to the first dose-response curve. Both probabilities are nondecreasing, even with the non-monotonic nature of the second dose-response curve, and the corresponding data points follow the curve.

Finally, we have the posterior distribution of the calibrated dose level given a new vector of responses equivalent to the observed vector at 100 cGy (top left panel of

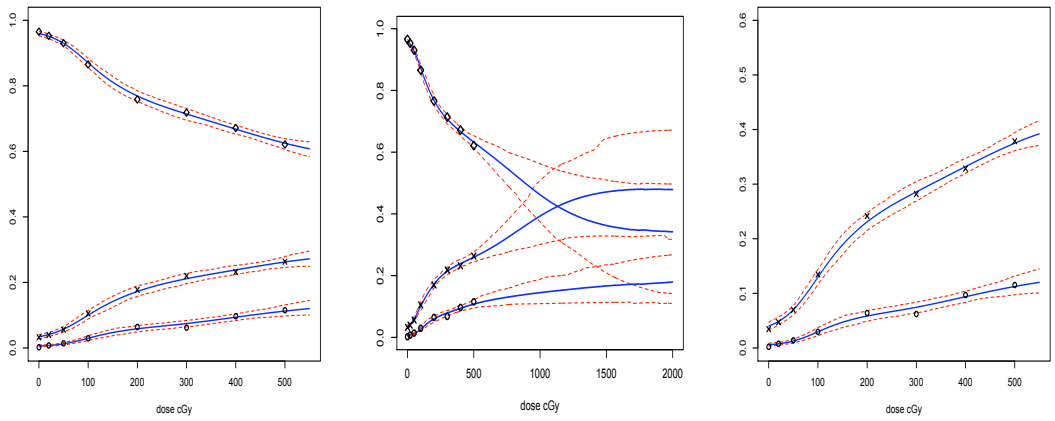

Figure 4.5: Cytogenetic dosimetry application. The posterior mean estimate (solid line) and 90% uncertainty bands (dashed lines) for $\Pr(Y_1 = 1; G, x)$ (bottom curve, left panel, data "o"), $\Pr(Y_2 = 1; G, x)$ (middle curve, left panel, data "x"), and $\Pr(Y_3 = 1; G, x)$ (top curve, left panel, data "$\diamond$"); the extrapolated dose response curves (middle panel), $\Pr(Y = 1; G, x)$ (lower curve, right panel, data "o") and $\Pr(Y = 2; G, x)$ (upper curve, right panel, data "x").

Figure 4.6). The estimate of the dose level is approximately 85.1 cGy with a 90% credible interval of $(61.5, 109.7)$, which includes the observed dose 100 cGy, but is generally underestimated. This is comparable to the results found in Kottas et al. (2002). The observed vector at 500 cGy is estimated in the bottom left panel of Figure 4.6. Here, 500 cGy is closer to the middle of the distribution; the posterior mean is 507.3 cGy with 90% interval of $(476.8, 538.9)$. The right column of Figure 4.6 provides the posterior distribution of the same calibration response vectors, leaving the observed vector out of the fitted data. The centers are roughly comparable, but the spread is larger as the data has been removed.

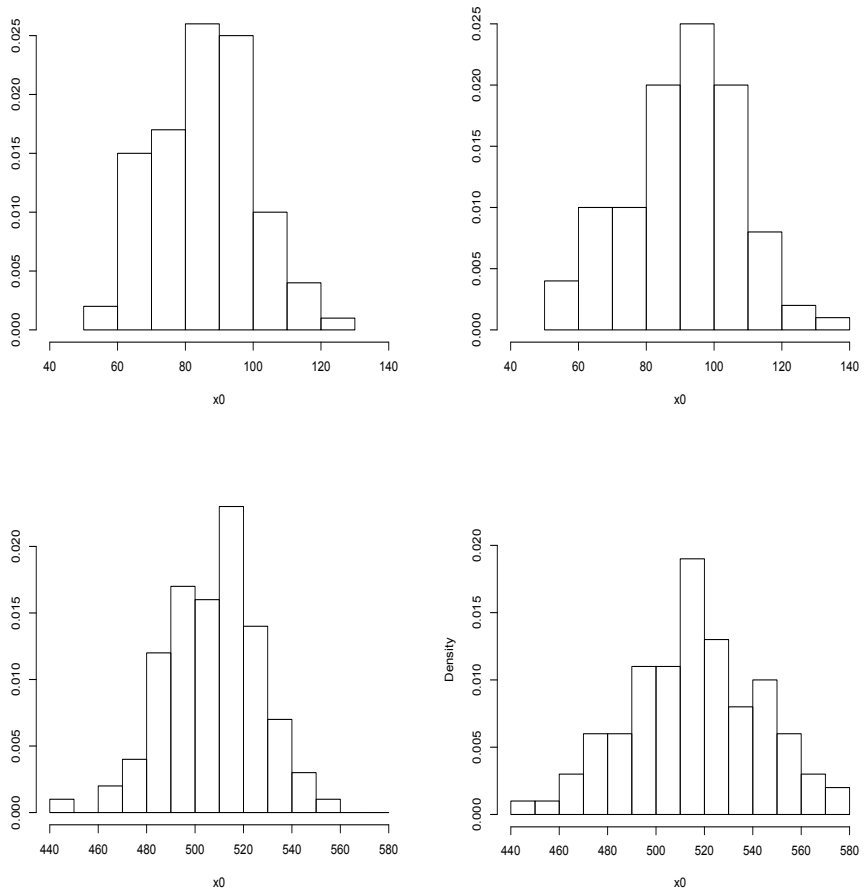

Figure 4.6: Cytogenetic dosimetry application. The posterior distributions for the estimated calibrated dose level given the observed vector at dose 100 cGy including the data vector (top left panel) and leaving the vector out (top right panel), and given the observed vector at dose 500 cGy including the data vector (bottom left panel) and leaving the vector out (bottom right panel).

4.3.4 Discussion

The linear-DDP model for ordinal responses allows for smooth estimates of the dose-response curves which allow for different shapes. The cumulative probabilities are non-decreasing functions of dose level. The details for implementation are not exceedingly difficult, as the updates for the distinct components in the mixture driving

the two probabilities can be separated. The calibration objective requires a Metropolis-Hastings update; nevertheless, under relatively dispersed priors, this inference can be obtained with suitable mixing and acceptance rates. The smooth, flexible dose-response relationships and the aptitude for calibration and extrapolation are the main contributions to the application area.

4.4 Conclusions

We have presented a Bayesian nonparametric approach to the analysis of quantal data from bioassay studies. The framework is driven by a simplified version of a dependent Dirichlet process prior, the linear-DDP prior. The consequent dose-response curve is monotonically increasing with probability one in the prior model, the inversion inferential objective becomes straightforward to address, and the calibration objective can be implemented without immense computational complexity. We provide illustration with a commonly used data example which bears evidence of multiple modes and skewness in the dose-response curve. Further demonstration through portions of a cytogenetic dosimetry data set reveals the practicality of the model.

Prompted by the cytogenetic dosimetry application and the scientific importance of calibration in this setting, we have considered dose-response experiments with ordinal responses. Specifically, we have developed one such linear-DDP mixture model built from a structured product of Bernoulli kernel. This construction results in desirable properties for the implied dose-response relationships and a relatively straightforward approach to implementation for the calibration objective.

Note the linear-DDP prior for the quantal responses setting is ideal: it connects with the approach in Chapter 3 and Binomial models with mixtures of non-decreasing

functions for the probability of response; it is easily implemented as a DP mixture model; and it provides a tractable method for obtaining the calibration inference. The model supports a monotonic dose-response curve, $D(x) = \int \pi(\theta) \mathrm{d}G_x(\theta) = \mathrm{E}^{G_x}(\pi(\theta))$, by stochastically ordering the mixing distributions. That is, if $G_x \leq_{st} G_{x'}$ for $x < x'$, then $D(x)$ is a non-decreasing function. A more general approach to ordering the mixing distributions can be accomplished through replacing the linear function of x with a generic function, $\mu(x)$, with a monotonicity restriction. Dunson and Peddada (2008) proposed a structured function which borrows information across groups, Wang and Dunson (2009) suggest an I-spline basis expansion to promote monotone functions, and Bornkamp et al. (2010) use a linear combination of ridge functions given multiple covariates. While this extension may not add much to the quantal case, in the context of the cytogenetic dosimetry application, the model may be extended with more general stochastically ordered distributions to improve extrapolation. This is an area of future research.

Chapter 5

Conclusions

Bayesian nonparametric mixture models are ideal for the analysis of developmental toxicity studies as well as quantal bioassay studies. The current state of the literature is based on classical estimation techniques, which do not include any distributional assumptions or the distributional assumptions are too restrictive to appropriately capture the complexity of the data. The framework presented builds on the single-p dependent Dirichlet process (DDP) prior to eliminate the limiting assumptions of parametric models while retaining a proper probabilistic setting for inference.

Note that in the developmental toxicity studies, it is very important to model both the subject specific distributions and the probability of response in an appropriate manner to more accurately account for the multiple sources of heterogeneity. The nonparametric mixture models developed in this dissertation provide flexible inference for the response distribution as well as for the various dose-response curves. Using simulation studies, we have shown that the DDP mixture models accomplish both of these inferential goals. Finally, data from assorted toxicity studies were used to

illustrate the variety of inferences that can be obtained from the DDP mixture modeling framework, including its practical utility with regard to estimation of non-monotonic dose-response relationships.

The models are in essence simple, yet effective in providing flexible inference with sensible uncertainty quantification. Implementation can be daunting with multiple independent Gaussian processes about which to learn and thoughtful prior specification is undoubtably imperative, however posterior simulation is relatively routine and posterior distributions peak despite the generally small sample sizes. As the data examples have demonstrated, the single-p DDP prior mixture model is sufficiently flexible to capture the dependence structure of the distributions across dose levels, while remaining interpretable and manageable to implement. While some nonparametric priors are more heavily influenced by individual data points, the single-p DDP prior setting induces a smooth evolution across the dose levels.

The bioassay setting has been studied profusely in the literature, both classical and Bayesian. The main attraction of our approach is that the framework is based upon modeling the response distributions. However, as shown in Chapter 4, this is equivalent to modeling the dose-response curve with a nonparametric mixture model. We have considered a more structured version of the DDP prior to induce monotonicity in the dose-response curve as well as to facilitate full exchangeability of the subjects. Through data previously studied in the literature and an application to the area of cytogenetic dosimetry, we have shown the utility of our method with regard to flexible inference for the dose-response curve as well as for inversion and calibration.

Ordinal responses are routinely found in bioassay studies, and motivated by the full cytogenetic dosimetry application, we have developed a model to accommodate data resulting in multiple, ordered classifications. We assume a continuation-ratio type

structure to promote flexibility in both dose-response curves and maintaining relative ease of implementation. The expanded cytogenetic dosimetry application is used to illustrate the method with respect to the prime inferential objectives, the various dose-response curves and calibration.

Bibliography

Agresti, A. (2002), *Categorical Data Analysis*, Hoboken, NJ: Wiley, 2nd ed.

Albert, J. and Chib, S. (1993), "Bayesian analysis of binary and polychotomous response data," *Journal of the American Statistical Association*, 88, 669–679.

— (1995), "Bayesian residual analysis for binary response regression models," *Biometrika*, 82, 747–759.

Allen, A. and Barnhart, H. (2002), "Joint models for toxicology studies with dose-dependent number of implantations," *Risk Analysis*, 22, 1165–1173.

Antoniak, C. (1974), "Mixtures of Dirichlet processes with applications to Bayesian nonparametric problems," *The Annals of Statistics*, 2, 1152–1174.

Ashford, J. and Walker, P. (1972), "Quantal response anaylsis for mixture populations," *Biometrics*, 28, 981–988.

Basu, S. and Mukhopadyay, S. (2000), "Binary response regression with normal scale mixture links," in *Generalized Linear Models: A Bayesian Perspective*, eds. Dey, D., Ghosh, S., and Mallick, B., New York: Marcel Dekker, pp. 231–241.

Bhattacharya, P. (1981), "Posterior distribution of a Dirichlet process from quantal response data," *The Annals of Statistics*, 9, 803–811.

Bornkamp, B. and Ickstadt, K. (2009), "Bayesian nonparametric estimation of continuous monotone functions with applications to dose-response analysis," *Biometrics*, 65, 198–205.

Bornkamp, B., Ickstadt, K., and Dunson, D. (2010), "Stochastically ordered multiple regression," *Biostatistics*, 11, 419–431.

Bowman, D. and George, E. (1995), "A saturated model for analyzing exchangeable binary data: applications to clinical and developmental toxicity studies," *Journal of the American Statistical Association*, 90, 871–879.

Bush, C. and MacEachern, S. (1996), "A semiparametric Bayesian model for randomised block designs," *Biometrika*, 83, 275–285.

Calabrese, E. J. (2005), "Paradigm lost, paradigm found: The re-emergence of hormesis as a fundamental dose response model in the toxicological sciences," *Environmental Pollution*, 138, 378–411.

Carota, C. and Parmigiani, G. (2002), "Semiparametric regression for count data," *Biometrika*, 89, 265–281.

Casanova, M. P., Iglesias, P., and Bolfarine, H. (2010), "A Bayesian semiparametric approach for solving the discrete calibration problem," *Communications in Statistics - Simulation and Computation*, 39, 347–360.

Catalano, P. and Ryan, L. (1992), "Bivariate latent variable models for clustered discrete and continuous outcomes," *Journal of the American Statistical Association*, 87, 651–658.

Chen, J., Kodell, R., Howe, R., and Gaylor, D. (1991), "Analysis of trinomial responses

from reproductive and developmental toxicity experiments," *Biometrics*, 47, 1049–1058.

Cifarelli, D. and Regazzini, E. (1978), "Nonparametric statistical problems under partial exchangeability. The use of associative means," *Annali dell' Istituto di Matematica Finanaziaria dell'Universita di Torino, Serie III*, 12, 1–36.

Connor, R. and Mosimann, J. (1969), "Concepts of independence for proportions with a generalization of the Dirichlet distribution," *Journal of the American Statistical Association*, 64, 194–206.

DeIorio, M., Johnson, W. O., Müller, P., and Rosner, G. L. (2009), "Bayesian nonparametric non-proportional hazards survival modelling," *Biometrics*, 63, 762–771.

DeIorio, M., Müller, P., Rosner, G., and MacEachern, S. (2004), "An ANOVA model for dependent random measures," *Journal of the American Statistical Association*, 99, 205–215.

Disch, D. (1981), "Bayesian nonparametric inference for effective doses in a quantal-response experiment," *Biometrics*, 37, 713–722.

Dominici, F. and Parmigiani, G. (2001), "Bayesian semiparametric analysis of developmental toxicology data," *Biometrics*, 57, 150–157.

Dunson, D. (1998), "Dose-dependent number of implants and implications in developmental toxicity," *Biometrics*, 54, 558–569.

Dunson, D., Chen, Z., and Harry, J. (2003), "A Bayesian approach for joint modeling of cluster size and subunit-specific outcomes," *Biometrics*, 59, 521–530.

Dunson, D. and Park, J. (2008), "Kernel stick-breaking processes," *Biometrika*, 95, 307–323.

Dunson, D. and Peddada, S. (2008), "Bayesian nonparametric inference on stochastic ordering," *Biometrika*, 95, 859–874.

Erkanli, A., Stangl, D., and Müller, P. (1993), "A Bayesian analysis of ordinal data using mixtures," in *ASA Proceedings of the Section on Bayesian Statistical Science*, Alexandria, VA: American Statistical Association.

Escobar, M. and West, M. (1995), "Bayesian density estimation and inference using mixtures," *Journal of the American Statistical Association*, 90, 577–588.

Faes, C., Geys, H., Aerts, M., and Molenberghs, G. (2006), "A hierarchical modeling approach for risk assessment in developmental toxicity studies," *Computational Statistics & Data Analysis*, 51, 1848–1861.

Ferguson, T. S. (1973), "A Bayesian analysis of some nonparametric problems," *The Annals of Statistics*, 1, 209–230.

Fronczyk, K., Kottas, A., and Munch, S. (2011), "Flexible modeling for stock-recruitment relationships using Bayesian nonparametric mixtures," *Environmental and Ecological Statistics*, to appear.

Fuentes-García, R., Mena, R. H., and Walker, S. G. (2009), "A nonparametric dependent process for Bayesian regression," *Statistics and Probability Letters*, 79, 1112–1119.

Gelfand, A., Kottas, A., and MacEachern, S. (2005), "Bayesian Nonparametric Spatial Modeling With Dirichlet Process Mixing," *Journal of the American Statistical Association*, 100, 1021–1035.

Gelfand, A. and Kuo, L. (1991), "Nonparametric Bayesian bioassay including ordered polytomous response," *Biometrika*, 78, 657–666.

George, E. and Bowman, D. (1995), "A full likelihood procedure for analysing exchangeable binary data," *Biometrics*, 51, 512–523.

Geweke, J. and Keane, M. (1999), "Mixture of normals probit models," in *Analysis of Panels and Limited Dependent Variable Models*, eds. Hsiao, C., Pesaran, M. H., Lahiri, K., and Lee, L. F., Cambridge University Press, pp. 49–78.

Gueorguieva, R. and Agresti, A. (2001), "A correlated probit model for joint modeling of clustered binary and continuous responses," *Journal of the American Statistical Association*, 96, 1102–1112.

Holson, J., Gaines, T., Nelson, C., LaBorde, J., Gaylor, D., Sheehan, D., and Young, J. (1991), "Developmental toxicity of 2,4,5-trichlorophenoxiacetic acid I: Dose response studies in four inbred strains and one outbred stock of mice," *Fundamentals of Applied Toxicology*, 19, 286–297.

Ishwaran, H. and James, L. (2001), "Gibbs sampling methods for stick-breaking priors," *Journal of the American Statistical Association*, 96, 161–173.

Ishwaran, H. and Zarepour, M. (2000), "Markov chain Monte Carlo in approximate Dirihlet and beta two-parameter process hierarchical models," *Biometrika*, 87, 371–390.

Jara, A., García-Zattera, M., and Lesaffre, E. (2007), "A Dirichlet process mixture model for the analysis of correlated binary responses," *Computational Statistics & Data Analysis*, 51, 5402–5415.

Kottas, A., Branco, M., and Gelfand, A. (2002), "A nonparametric Bayesian modeling approach for cytogenetic dosimetry," *Biometrics*, 58, 593–600.

Kottas, A. and Krnjajić, M. (2009), "Bayesian Semiparametric Modelling in Quantile Regression," *Scandinavian Journal of Statistics*, 36, 297–319.

Kottas, A., Müller, P., and Quintana, F. (2005), "Nonparametric Bayesian modeling for multivariate ordinal data," *Journal of Computational and Graphical Statistics*, 14, 610–625.

Krewski, D. and Zhu, Y. (1994), "Applications of multinomial dose-response models in developmental toxicity risk assessment," *Risk Analysis*, 14, 613–627.

Kuk, A. (2003), "A generalized estimating equation approach to modelling foetal response in developmental toxicity studies when the number of implants is dose dependent," *Applied Statistics*, 52, 51–61.

— (2004), "A litter-based approach to risk assessment in developmental toxicity studies via a power family of completely monotone functions," *Applied Statistics*, 53, 369–386.

Kuo, L. (1983), "Bayesian bioassay design," *The Annals of Statistics*, 11, 886–895.

— (1988), "Linear Bayes estimators of the potency curve in bioassay," *Biometrika*, 75, 91–96.

Lawrence, E., Bingham, D., Liu, C., and Nair, V. (2008), "Bayesian inference for multivariate ordinal data using parameter expansion," *Technometrics*, 50, 182–191.

Lüning, K., Sheridan, W., Ytterborn, K., and Gullberg, U. (1966), "The relationship between the number of implantations and the rate of intrauterine death in mice," *Mutation Research*, 3, 444–451.

Lwin, T. and Martin, P. (1989), "Probits of mixtures," *Biometrics*, 45, 721–732.

Ma, R., Jørgenson, B., and Willms, J. (2009), "Clustered binary data with random cluster sizes: a dual poisson modelling approach," *Statistical Modelling*, 9, 137–150.

MacEachern, S. (1998), "Computational Methods for Mixture of Dirichlet Process Models," in *Practical Nonparametric and Semiparametric Bayesian Statistics*, eds. Dey, D., Müller, P., and Sinha, D., Springer-Verlag, pp. 23–44.

— (1999), "Dependent nonparametric processes," in *ASA Proceedings of the Section on Bayesian Statistical Science*, Alexandria, VA: American Statistical Association.

— (2000), "Dependent Dirichlet processes," Tech. rep., Ohio State University, Deptartment of Statistics.

MacEachern, S. and Müller, P. (1998), "Estimating mixture of Dirichlet process models," *Journal of Computational and Graphical Statistics*, 7, 223–238.

Madruga, M., de B. Pereira, C. A., and Rabello-Gay, M. N. (1994), "Bayesian dosimetry: radiation dose versus frequencies of cells with aberrations," *Environmetrics*, 5, 47–56.

Madruga, M. R., Ochi-Lohmann, T. H., Okazaki, K., de B. Pereira, C. A., and Rabello-Gay, M. N. (1996), "Bayesian dosimetry II: credibility intervals for radiation dose," *Environmetrics*, 7, 325–331.

Molenberghs, G. and Ryan, L. (1999), "An exponential family model for clustered multivariate binary data," *Environmetrics*, 10, 279–300.

Mukhopadhyay, S. (2000), "Bayesian nonparametric inference on the dose level with specified response rate," *Biometrics*, 56, 220–226.

Muliere, P. and Walker, S. (1997), "A Bayesian nonparametric approach to determining a maximum tolerated dose," *Journal of Statistical Planning and Inference*, 61, 339–353.

Müller, P., Erkanli, A., and West, M. (1996), "Bayesian curve fitting using multivariate Normal mixtures," *Biometrika*, 83, 67–79.

Neal, R. (2000), "Markov chain sampling methods for Dirichlet process mixture models," *Journal of Computational and Graphical Statistics*, 9, 249–265.

Pang, Z. and Kuk, A. (2005), "A shared response model for clustered binary data in developmental toxicity studies," *Biometrics*, 61, 1076–1084.

Price, C., Kimmel, C., Tyl, R., and Marr, M. (1985), "The developmental toxicity of ethylene glycol in rats and mice," *Toxicology and Applied Pharmacology*, 81, 113–127.

Rai, K. and Ryzin, J. V. (1985), "A dose-response model for teratological experiments involving quantal responses," *Biometrics*, 41, 1–9.

Regan, M. and Catalano, P. (1999), "Likelihood models for clustered binary and continuous outcomes: application to developmental toxicology," *Biometrics*, 55, 760–768.

Reich, B. and Fuentes, M. (2007), "A multivariate semiparametric Bayesian spatial modeling framework for hurricane surface wind fields," *The Annals of Applied Statistics*, 1, 249–264.

Rodriguez, A., Dunson, D., and Gelfand, A. (2009), "Bayesian nonparametric functional data analysis through density estimation," *Biometrika*, 96, 149–162.

Rodriguez, A. and Dunson, D. B. (2011), "Nonparametric Bayesian models through probit stick-breaking processes," *Bayesian Analysis*, 6, 145–178.

Rodriguez, A. and ter Horst, E. (2008), "Bayesian dynamic density estimation," *Bayesian Analysis*, 3, 339–366.

Ryan, L. (1992), "Quantitative risk assessment for developmental toxicity," *Biometrics*, 48, 163–174.

Sethuraman, J. (1994), "A constructive definition of Dirichlet priors," *Statistica Sinica*, 4, 639–650.

Shaked, M. (1980), "On mixtures of exponential families," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 42, 192–198.

Taddy, M. (2010), "An auto-regressive mixture model for dynamic spatial Poisson processes: Application to tracking the intensity of violent crime." *Journal of the American Statistical Association*, 105, 1403–1417.

Taddy, M. and Kottas, A. (2010), "A Bayesian nonparametric approach to inference for quantile regression," *Journal of Business and Economic Statistics*, 28, 357–369.

Ten Have, T. and Chinchilli, V. (1998), "Two-stage negative binomial and overdispersed poisson models for clustered developmental toxicity data with random cluster size," *Journal of Agricultural, Biological, and Environmental Statistics*, 3, 75–98.

Tyl, R., Jones-Price, C., Marr, M., and Kimmel, C. (1983), "Teratologic evaluation of diethylhexyl phthalate (CAS No. 111-81-7)," Final Study Report for NCTR/NTP contract 222-80-2031(c), NITS PB85105674, National Technical Information Service, Springfield, Virginia.

Walker, P. (1966), "A method of measuring the sensitivity of trypanosomes to acriflavine and trivalent tryparsamide," *Journal of General Microbiology*, 43, 45–58.

Wang, L. and Dunson, D. (2009), "Bayesian isotonic density regression," Tech. rep., Biometrika (under revision).

Zhu, Y., Krewski, D., and Ross, W. (1994), "Dose-response models for correlated multinomial data from developmental toxicity studies," *Applied Statistics*, 43, 583–598.