Chilean Journal of Statistics
Vol. 5, No. 2, September 2014, 3–14

Generalized Linear Models
Research Paper

Modeling Bounded Outcome Scores Using The Binomial-Logit-Normal Distribution

Ye Liang[1,*], Dongchu Sun[2], Chong He[2] and Mario Schootman[3]

[1]Department of Statistics, Oklahoma State University, Stillwater, USA,
[2]Department of Statistics, University of Missouri, Columbia, USA,
[3]Washington University School of Medicine, St. Louis, USA

(Received: 23 September 2013 · Accepted in final form: 30 September 2014)

Abstract

Bounded outcome scores are often encountered in health-related survey studies. Such scores are usually bounded and discrete and are often treated as categorical or ordinal data, which is not satisfactory in some scenarios. The binomial-logit-normal distribution, as a parametric model, is a useful alternative for the bounded outcome scores. The proposed model converges to the continuous logit-normal model and hence bridges the gap between discrete modeling and continuous modeling. This result is useful when the score is dense and smooth within its bounded support. A quality of life data is shown as an application.

Keywords: health outcome · binomial-logit-normal · ordinal data · logistic regression · quality of life.

Mathematics Subject Classification: Primary 60F05 · Secondary 62J12.

1. Introduction

Bounded outcome scores are encountered in many survey studies, especially in survey studies involving health assessment. Examples are the Barthel index, drug compliance studies, pain relief studies and quality of life (QOL) studies (Lesaffre et al., 2007). In this paper, we present a QOL study as our motivation example. This is a survey study on breast cancer survivors, investigating their health-related QOL and comparing their QOL with that of community-dwelling women without any type of reportable cancer. With declining breast cancer mortality rates and improved survival following breast cancer, there is increasing focus on the QOL. The study was conducted in the State of Missouri. QOL is defined as a multi-dimensional construct that includes the subjective evaluation of several important aspects of a person's situation, including physical well-being, disease symptoms and treatment side effects, emotional well-being, physical functioning, and social functioning. A common approach to assess QOL is to utilize the RAND 36-Item Health Survey (Ware and Sherbourne, 1992). The QOL scores are summarized from item responses

*Corresponding author. Email: ye.liang@okstate.edu

ISSN: 0718-7912 (print)/ISSN: 0718-7920 (online)

http://www.soche.cl/chjs

based on certain scoring processes.

Figure 1 shows four selected bounded outcome scores in the QOL study. The left panel of the graph is for the case group and the right panel is for the control group. From top to bottom, the graph shows "role emotional", "role physical", "emotional" and "physical". It is worth noting some characteristics of the data. First, they are all bounded between 0 and 100. In literature, they are usually referred as $[0, 1]$ bounded outcome scores for any bounded score can be standardized to be within $[0, 1]$. Second, they are, in nature, all discrete. However, some scores are dense within the interval $[0, 1]$, such as "emotional" and "physical". Third, most scores have smooth shapes and are equally distanced. Lesaffre et al. (2007) described them "J-shaped" or "U-shaped".

There are several possible choices of modeling bounded outcome scores. First, choose a continuous model regardlessly. Two appropriate choices for $[0, 1]$ data are the logit-normal (LN) model and the beta model. Note that the LN model, initiated by Johnson (1949), is different from the lognormal model, and its density function is given by (1). The beta model can be found in Johnson et al. (1995). According to Aitchison and Begg (1976), the beta model can be approximated by the LN model. In other words, the LN model is more flexible for $[0, 1]$ data due to a richer class of shapes. However, neither the beta model nor the LN model is a good choice here because all scores are discrete in nature, though some might look like continuous. Second, treat all scores categorical and choose a discrete model, for instance, the proportional odds (PO) model introduced by McCullagh (1980). The PO model may be appropriate for some sparse scores, such as "role emotional" and "role physical". However, it might be awkward to apply the PO model to scores that are dense within the bounded support, such as "emotional" and "physical". The PO model assumes an underlying continuous model that can be chopped into categories, and introduces cut points, ignoring the parametric shape of the underlying distribution. It may be more natural to use a parametric continuous model for "emotional" and "physical" rather than to determine cut points and group information between cut points. Moreover, the PO model is usually for ordinal data where the distance between categories is meaningless. However, it would be questionable to claim that distance does not matter in our case. Third, propose a parametric discrete model on $[0, 1]$ that converges to a parametric continuous model as the variable becomes dense. Through fitting such a model, the gap between continuous modeling and discrete modeling is eliminated. We propose the third approach in this paper as an alternate choice to the second approach. We do not claim that the third approach is superior to other approaches, but we argue that it could be more effective sometimes.

The model that we propose in this paper is the binomial-logit-normal (BLN) model. This terminology is not completely new. The BLN model was used by Coull and Agresti (2000) and Lesaffre et al. (2007) for modeling binomial counts, because the lowest level in this model is a binomial distribution. Though QOL scores are not binomial counts that are induced from independent Bernoulli trials, the BLN distribution, as a bounded discrete distribution, can still be a reasonable choice for QOL scores. We show that the application of the BLN model can be more general than those in past literature. Indeed, one can always assume an underlying Bernoulli process on a number of dichotomous items and a bounded outcome score is an aggregation of those Bernoulli outcomes.

The paper is organized as follows. We introduce the BLN distribution and its properties, followed by a logistic linear mixed model for QOL data, in Section 2. In Section 3, we show two simulation studies with model comparisons. Data analysis for the QOL data is in Section 4. A discussion is in Section 5.

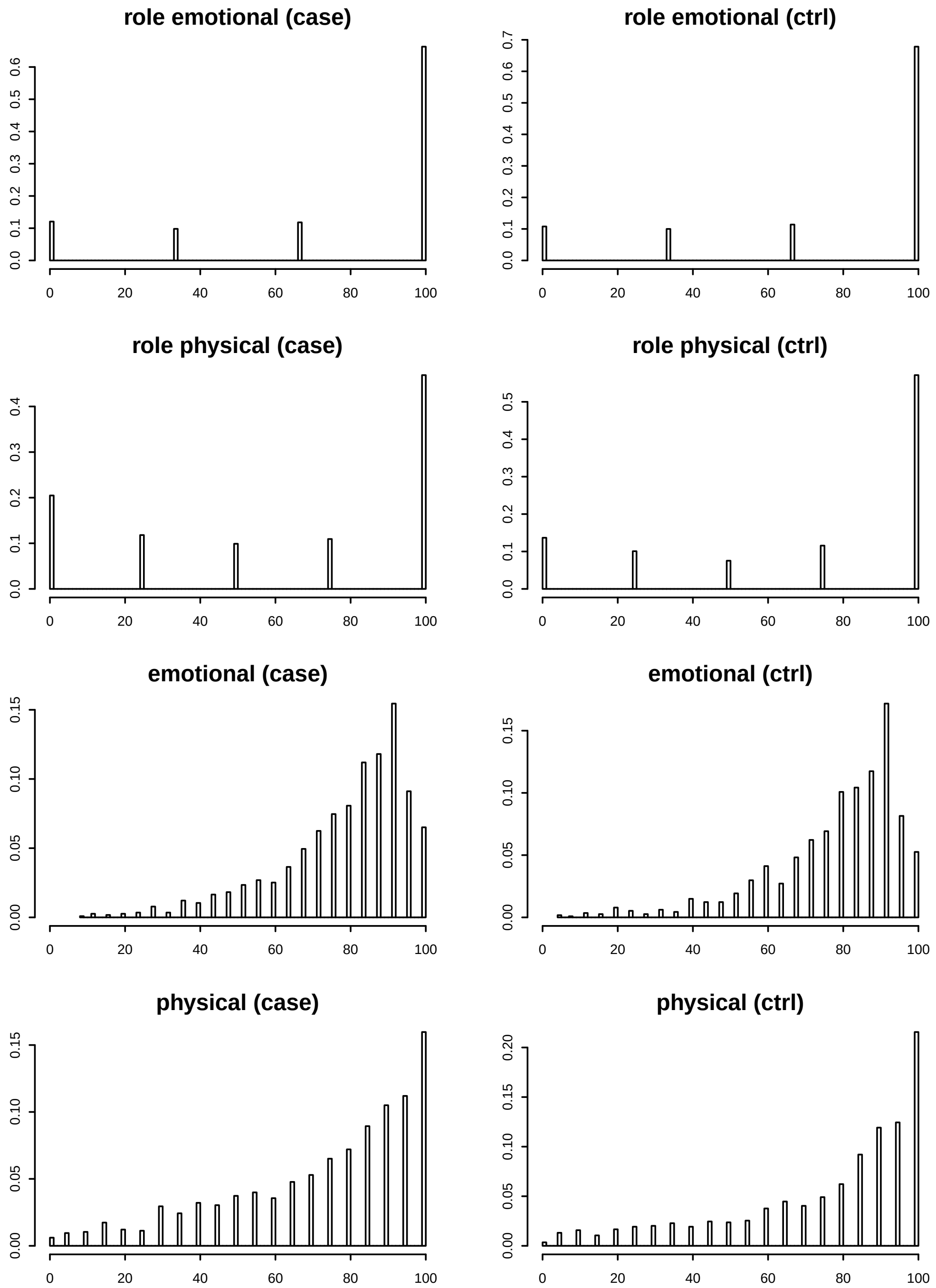

Figure 1. Distributions of Bounded Outcome Scores

2. Model Description

2.1 A Mixture Distribution

We start from reviewing the LN distribution which was initiated by Johnson (1949). Let Z be a normal random variable, $N(\mu, \sigma^2)$. Let $Y = 1/(1 + e^{-Z})$. Then Y is an LN random variable, denoted by $Y \sim LN(\mu, \sigma^2)$. Its density function is given by

$$f(y \mid \mu, \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma y(1-y)} \exp\left\{-\frac{[logit(y)-\mu]^2}{2\sigma^2}\right\}, \tag{1}$$

where $logit(y) = log[y/(1-y)]$. The LN distribution can take a variety of shapes on $[0,1]$ by changing μ and σ^2, including the "J-shape" and "U-shape". It is widely used to model continuous $[0,1]$ variables.

The BLN distribution is a discretized version of the LN distribution. It is defined hierarchically,

$$\begin{cases} (Y_n \mid p) \sim binomial(n, p); \\ (p \mid \mu, \sigma^2) \sim LN(\mu, \sigma^2). \end{cases} \tag{2}$$

The distribution of Y_n, clearly, is a mixture distribution whose probability mass function can be obtained by integrating out the parameter p,

$$P(Y_n = y \mid \mu, \sigma^2) = \int_0^1 P(Y_n = y \mid p) f(p \mid \mu, \sigma^2) dp. \tag{3}$$

Obviously, Y_n is discrete on $\{0, 1, \ldots, n\}$ and Y_n/n is discrete on $[0,1]$. We denote this distribution $Y_n \sim BLN_n(\mu, \sigma^2)$. For any fixed n, the BLN distribution only depends on μ and σ^2. Figure 2 shows different shapes of the BLN distribution with different values of μ and σ^2. The parameter μ controls the location and skewness of the distribution. The parameter σ^2 controls the curvature of the distribution.

Theorem 2.1 Let $\{Y_1, Y_2, \ldots, Y_n\}$ be a sequence of binomial-logit-normal random variables, i.e. $Y_i \sim BLN_i(\mu, \sigma^2)$, $i = 1, \ldots, n$. Then as $n \to \infty$, Y_n/n converges in law to a logit-normal random variable $LN(\mu, \sigma^2)$.

$$\frac{Y_n}{n} \xrightarrow{\mathcal{L}} LN(\mu, \sigma^2). \tag{4}$$

The proof of Theorem 2.1 is in Appendix A. We see that the BLN distribution is a discretized version of the LN distribution while the LN is the limiting distribution of the BLN. It is not difficult to obtain the following facts,

$$E\left(\frac{Y_n}{n}\right) = E(Y); \tag{5}$$

$$Var\left(\frac{Y_n}{n}\right) = Var(Y) + \frac{1}{n}[E(Y) - E(Y^2)], \tag{6}$$

where $Y_n \sim BLN_n(\mu, \sigma^2)$ and $Y \sim LN(\mu, \sigma^2)$.

We want to further give some properties of the BLN distribution on its convergence rate. This is done by investigating the Kolmogorov-Smirnov distance

$$d_n = \sup_q |G_n(q) - G(q)|, \tag{7}$$

where $G_n(\cdot)$ is the cumulative distribution function (CDF) of the BLN random variable divided by n and $G(\cdot)$ is the CDF of the LN random variable.

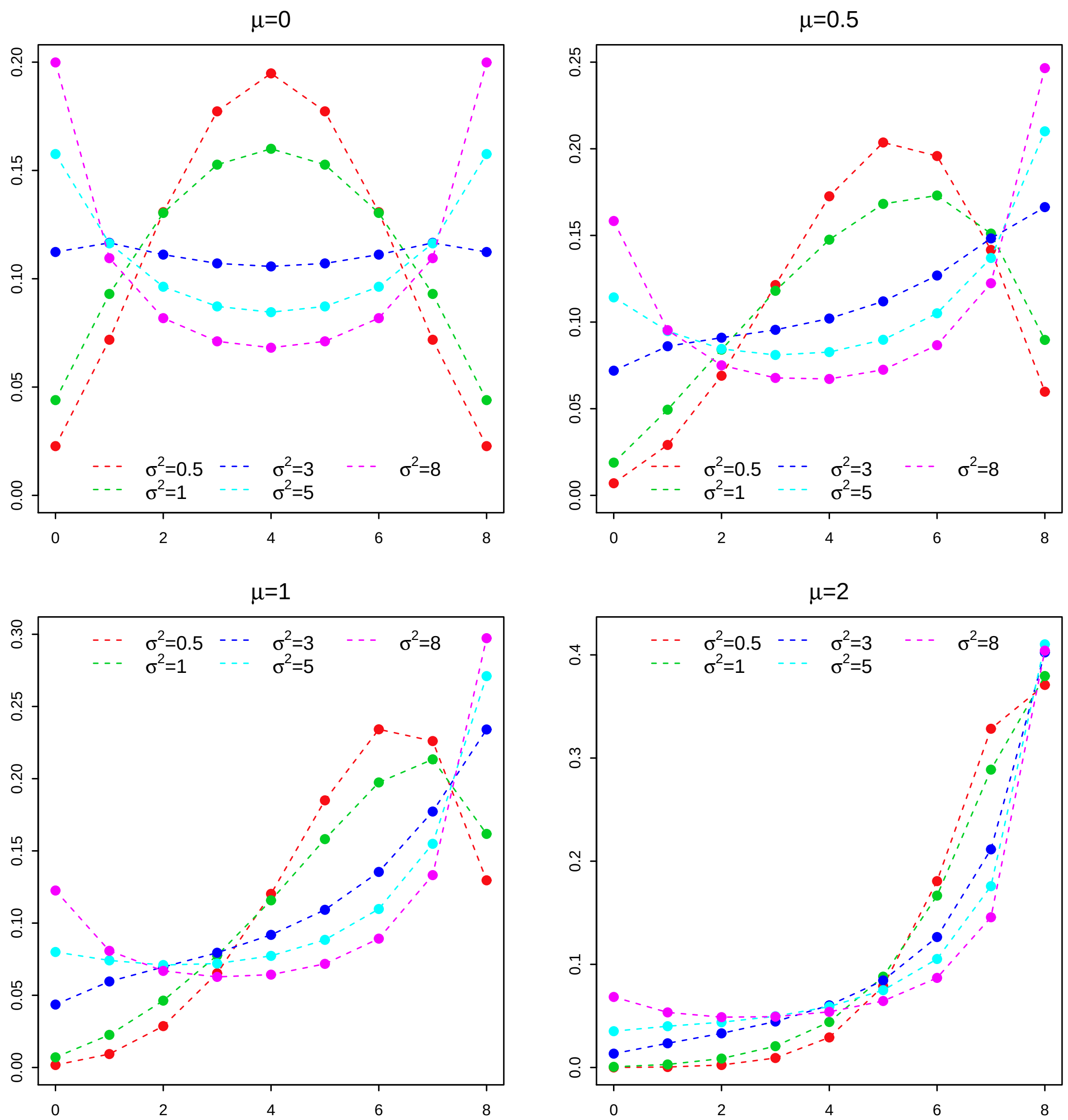

Figure 2. Probability Mass Functions of the Binomial-Logit-Normal Distribution

THEOREM 2.2 Let d_n be the distance defined by (7). Then

$$d_n \leq b_n = M(\mu, \sigma^2)n^{-\alpha} + Cn^{-\frac{1}{2}} + \Phi(-2n^{\frac{1}{2}-\alpha}) \text{ for any } \alpha \in (0, 1/2), \tag{8}$$

where $M(\mu, \sigma^2)$ is the mode of $LN(\mu, \sigma^2)$, C is the constant of Berry-Esseen theorem and $\Phi(\cdot)$ is the CDF of $N(0, 1)$.

The proof of Theorem 2.2 is in Appendix B. The density function of $LN(\mu, \sigma^2)$ shown in (1) is proven to be bounded when $y \in (0, 1)$, which guarantees the existence of its mode $M(\mu, \sigma^2)$. The constant C is from Berry-Esseen theorem, which is in the range of $(0.4, 0.5)$ according to Korolev and Shevtsova (2010). The value of α can be arbitrarily chosen from 0 to 1/2. However, among the three components of b_n, $\Phi(-2n^{\frac{1}{2}-\alpha})$ is typically small while $M(\mu, \sigma^2)n^{-\alpha}$ is relatively large. One could choose a value close to 1/2 for α to lower this bound.

The BLN distribution is a discrete distribution on a bounded support. It has two free parameters and has a rich class of shapes. When the standardized (divided by n) BLN random variable becomes dense within its support, it converges in law to the continuous LN random variable. We shall mention that, later we will always refer to the standardized BLN random variable which is on $[0,1]$.

2.2 A Logistic Linear Mixed Model

The QOL scores, described in Section 1, are discrete random variables distributed on the same bounded support, either sparsely or densely. Their distributions are of smooth shapes. The two-parameter BLN distribution, with a variety of shapes, is appropriate to describe the QOL scores. More importantly, the BLN distribution eliminates the gap between discrete modeling and continuous modeling because of the fact of convergence, allowing modeling both sparse and dense scores in the same framework. Though the QOL scores are not induced from independent Bernoulli trials, they can be explained by underlying Bernoulli processes. For example, the "role emotional" score has 4 levels. Imagine 3 independent underlying Bernoulli items $X_l \mid p \sim Ber(p),\ l = 1,2,3$, measuring 3 latent dichotomous traits, respectively. A subject can "succeed" in the l-th item if $X_l = 1$. The eventual score of "role emotional" is the sum of all 3 items, $Y = \sum_l X_l$, measuring how "successful" the subject is regarding the 3 underlying traits. The variation of the latent Bernoulli probability p is described by an LN distribution.

We now specify the model for QOL scores. Let the index $i = 1$ for the case group and $i = 0$ for the control group. Consider a QOL score $y_{ij},\ j = 1,\ldots,m_i$, where m_i is the number of subjects in group i. Specify the following model for $i = 0$ and 1, separately,

$$y_{ij} \sim BLN_n(\mu_i, \sigma_i^2), \qquad j = 1,\ldots,m_i, \tag{9}$$

which is equivalent to a hierarchical specification,

$$\begin{cases} y_{ij} \sim binomial(n, p_{ij}), \\ logit(p_{ij}) = \mu_i + \epsilon_{ij}, \\ \epsilon_{ij} \sim N(0, \sigma_i^2). \end{cases} \qquad j = 1,\ldots,m_i. \tag{10}$$

It is worth noting that the specification (10) is nothing but a special logistic linear mixed model (LLMM). However, in a usual LLMM, n is n_{ij} instead of a common number, meaning the number of Bernoulli trials. In our case, the common n for all subjects can be explained as the number of underlying Bernoulli trials. In the hierarchical setting, one may further consider modeling μ and σ^2 with explanatory variables and other scientific models. We only illustrate the simple model (9) in this paper.

Without further modeling for μ and σ^2, statistical inference is rather easy for the BLN model. Maximum likelihood estimate (MLE) can be numerically computed through routine algorithms. The only computational issue might be evaluating the integral (3). This one-dimensional integral can be evaluated by numerical integration or Monte Carlo integration (it is easy to sample from the LN density). We did not experience any difficulty in this simple scenario. It would be more complicated if μ and σ^2 are modeled by higher hierarchies. However, if a linear model is used for μ, the model still falls into the framework of LLMM and hence inference can be done through standard packages in statistical software such as SAS and R. Bayesian methods are also readily implementable by assigning prior distributions to μ and σ^2. Bayesian inference can be done by using Markov chain Monte Carlo (MCMC). In particular, one can specify this model easily in the popular Bayesian software WinBUGS, which is convenient for practitioners.

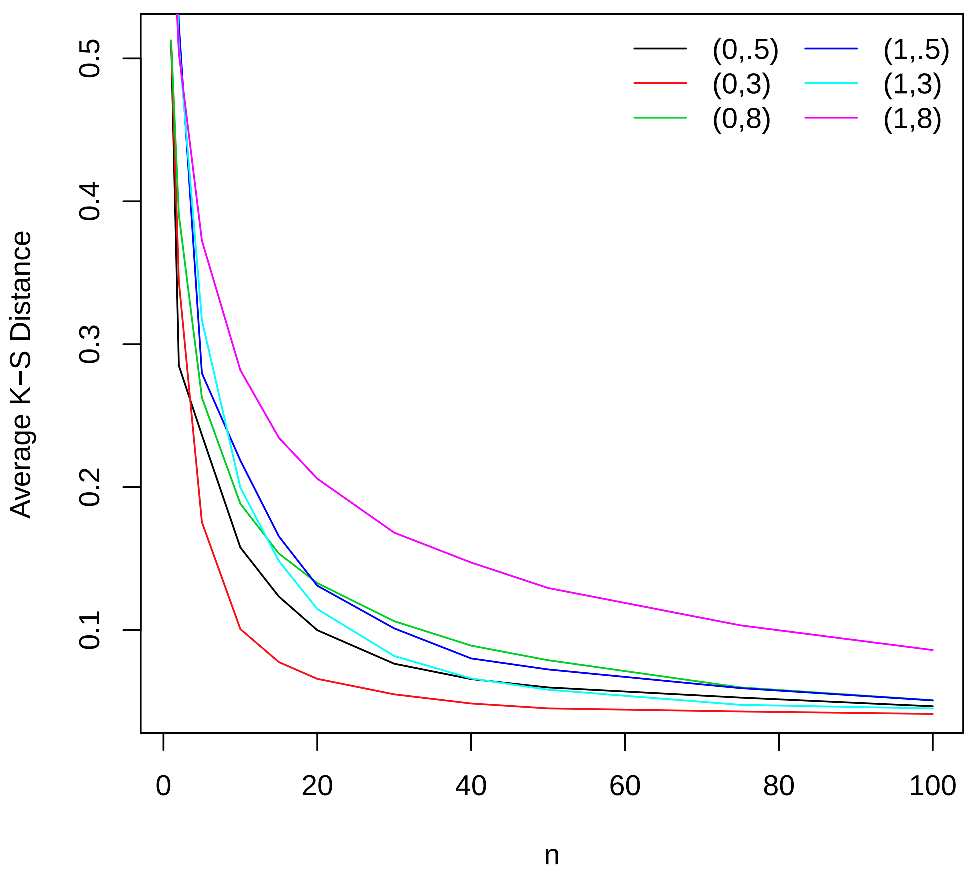

Figure 3. Decay of The Average K-S Distance between BLN and LN Samples for Six Combinations of (μ, σ^2).

3. Simulation Studies

We performed two simulation studies. In the first simulation study, we examined the Kolmogorov-Smirnov distance between empirical distributions of BLN samples and LN samples. The actual distance between BLN and LN is hard to evaluate, which is controlled by a bound in Theorem 2.2. For this reason, we evaluated the distance in an indirect way. Suppose $\boldsymbol{y}$ and $\boldsymbol{y}'$ are samples from BLN and LN distributions, respectively. Let D^{KS} be the Kolmogorov-Smirnov distance defined by (7) between empirical distributions of $\boldsymbol{y}$ and $\boldsymbol{y}'$. We repeated sampling from BLN and LN and evaluating D^{KS}, then we calculated the average distance $\overline{D^{KS}}$. We picked six combinations of (μ, σ^2): $(0, .5)$, $(0, 3)$, $(0, 8)$, $(1, .5)$, $(1, 3)$ and $(1, 8)$ and computed $\overline{D^{KS}}$ for increasing levels of n. The sample size was 1000 and the number of replications was 100. Figure 3 shows that $\overline{D^{KS}}$ decays as n increases. The distance is obviously affected by the values of μ and σ^2.

The second simulation study is comparing three candidate models. Suppose the BLN model is our underlying true model. We simulated data for both case group (with $\mu = 0$ and $\sigma^2 = .5$) and control group (with $\mu = 1$ and $\sigma^2 = 1$). The sample size was 100 for each group. We simulated 4 datasets for $n = 10$, 20, 50 and 100. We considered three candidate models: the BLN model, the continuous LN model and the proportional odds (PO) model. To compare all three models, we computed the corrected Akaike information criterion (AICc) and the Bayesian information criterion (BIC). We repeated this procedure for 100 times and calculated the average AICc and BIC. Results are reported in Table 1. It is not surprising that both AICc and BIC prefer the BLN model which is the underlying model. As n goes large, the difference becomes small between the BLN model and the LN model. In other words, it is acceptable to treat a dense score continuously. On the other hand, the PO model becomes less preferred as n goes large. This is because the PO model is penalized for involving too many cut-points. It is known that the AICc is penalized less strongly than does the BIC. Therefore, the difference of BIC is greater than of AICc.

Table 1. Simulation Study: The average AICc and BIC for BLN, LN and PO models. Δ in the table is the average difference.

	$\overline{\text{AICc}}$					$\overline{\text{BIC}}$				
n	BLN	LN	PO	Δ_{B-L}	Δ_{B-P}	BLN	LN	PO	Δ_{B-L}	Δ_{B-P}
10	868	900	902	-32	-34	881	913	936	-32	-55
25	1179	1211	1250	-32	-71	1192	1224	1321	-32	-129
50	1432	1446	1498	-14	-66	1445	1459	1615	-14	-170
100	1699	1702	1830	-3	-131	1712	1715	1985	-3	-273

The simulation studies illustrate the fact of convergence proven in Section 2. The second simulation also shows that the BLN model could be more effective in some scenarios. In particular, if there is a good reason to believe that the data are from an underlying parametric distribution, it would be wise to consider a parametric treatment.

4. Data Analysis

The BLN model (9) was fitted for the data shown in Figure 1. Estimates and fitted models are plotted in Figure 4. The model in general fits well for "emotional" ($n = 25$) and "physical" ($n = 20$) except the "peak inflation" part on "emotional". Both shapes are sketched by a smooth curve, however, discretized. For "role emotional" ($n = 3$) and "role physical" ($n = 4$), the model fitting is not as satisfactory as for the other two. It seems unnecessary to use the parametric BLN distribution when n is small. Similar findings are reported in Table 2. AICc and BIC were compared for the BLN model and the PO model. The PO model is strongly preferred for small-n scores, but less preferred for large-n scores. Note, the "emotional" score here is an exception due to its "peak inflation". In general, if scores are modeled separately, we may consider using different models for different scores. However, if joint modeling is needed, using the BLN setting will be beneficial because it is ready for a multivariate extension, for which references are Coull and Agresti (2000) and Rabe-Hesketh and Skrondal (2001).

Table 2. Model Comparison for QOL Data: AICc and BIC for BLN and PO models. Δ in the table is the difference.

Score	n	AICc(BLN)	AICc(PO)	Δ_{B-P}	BIC(BLN)	BIC(PO)	Δ_{B-P}
role emotional	3	4745	4568	177	4767	4591	176
role physical	4	6313	6138	175	6336	6166	170
emotional	25	12323	12283	40	12346	12426	-80
physical	20	12341	12371	-30	12364	12491	-127

To test whether there is a difference between case and control groups, a likelihood ratio test can be performed on $H_0 : \mu_0 = \mu_1, \sigma_0^2 = \sigma_1^2$. P-Values are .59, $\ll .01$, .81 and $\ll .01$

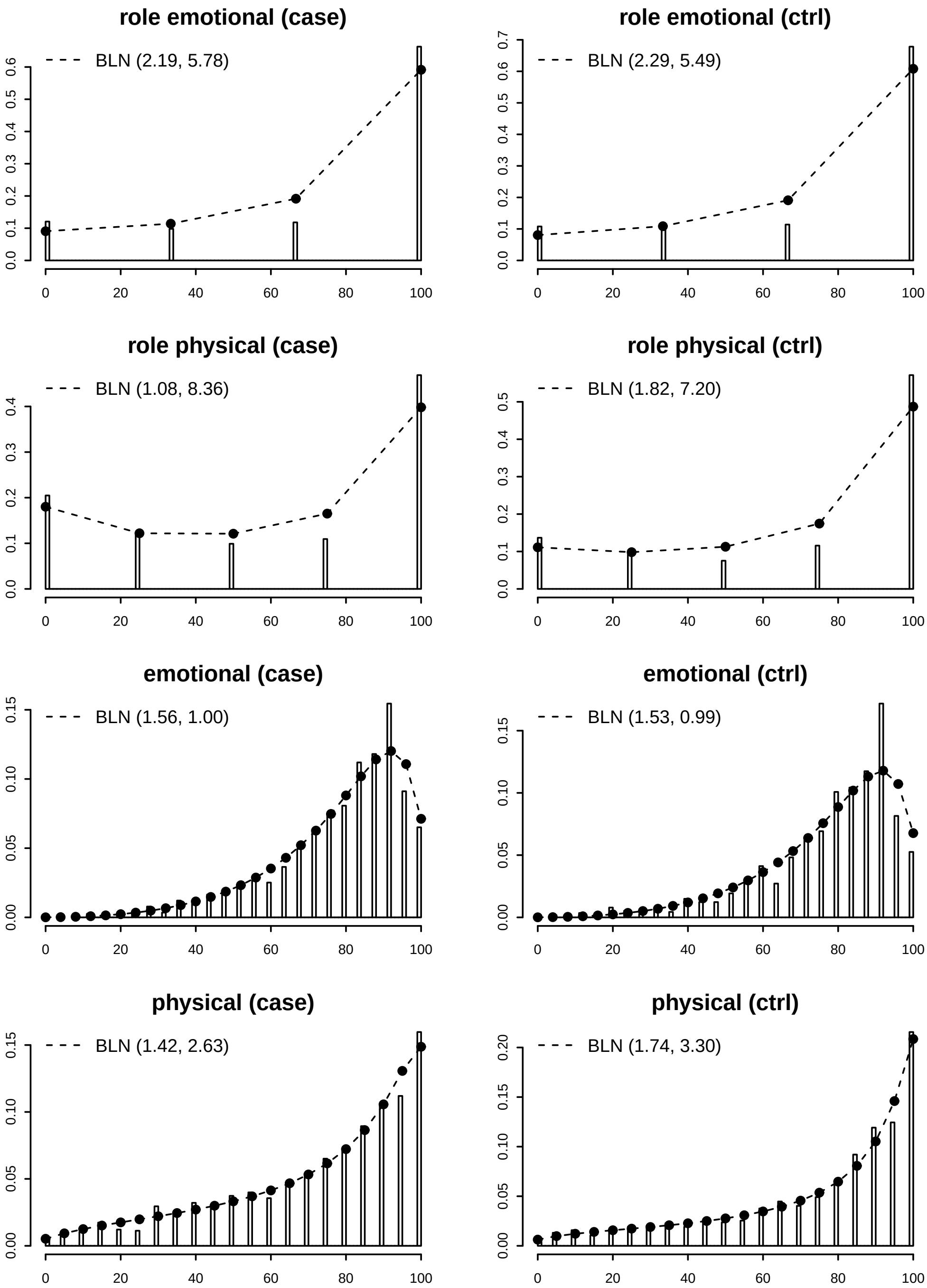

Figure 4. Comparison of Data and Fitted BLN Models

for "role emotional", "role physical", "emotional" and "physical", respectively. The test results indicate that, for "role emotional" and "emotional", there is little difference between case group and control group, while for "role physical" and "physical", the difference is significant. This result is well caught by the estimates in Figure 4. The interpretation of this finding also makes sense. The physical conditions of patients with breast cancer should be different from those of normal people (worse, in fact). However, their emotional conditions seem to be unaffected comparing with normal people. More interesting studies could be conducted by introducing other explanatory variables into the model.

5. Discussion

This paper provides an alternate model for the bounded outcome score which is a special kind of ordinal data. The ordinal data has ordered categories where the distance between categories is meaningless. A bounded outcome score, usually, is on an ordered discrete-scale and hence is ordinal. However, it is debatable whether the distance should matter in this case. On the other hand, treating a bounded outcome score categorical is inefficient when it is dense (large n). Without grouping, the BLN model parameterizes the shape of a score, and naturally handles the situation with large n. One drawback of the BLN parameterization is that the class of shapes sometimes might be too limited with only two free parameters. This problem could be alleviated by considering more generalized parameterizations. Mead (1965) introduced a generalized LN distribution by adding a power parameter, $Y = [1/(1+e^{-Z})]^{1/\theta}$ where Z is normal. This leads to a generalized BLN distribution which has a richer class of shapes. On extreme cases when the shape cannot be parameterized or is not important, one may still consider categorizing or grouping data. Nevertheless, the BLN approach has its unique advantage on certain datasets for which it should be considered as a candidate model.

Appendix A. Proof of Theorem 2.1

Let $X_n \sim binomial(n,p)$. By the Weak Law of Large Numbers,

$$\lim_{n\to\infty} P_X\left(\frac{X_n}{n} \le q\right) = \begin{cases} 1 & \text{if } p < q, \\ 0 & \text{if } p > q. \end{cases}$$

Suppose $Y_n \sim BLN_n(\mu, \sigma^2)$, then by Fubini's Theorem and the Dominated Convergence Theorem,

$$\begin{aligned}
\lim_{n\to\infty} P_Y\left(\frac{Y_n}{n} \le q\right) &= \lim_{n\to\infty} \sum_{y=0}^{[nq]} \int_0^1 \frac{n!}{y!(n-y)!} p^y (1-p)^{n-y} f_{LN}(p \mid \mu, \sigma^2) dp \\
&= \int_0^1 \lim_{n\to\infty} \sum_{y=0}^{[nq]} \frac{n!}{y!(n-y)!} p^y (1-p)^{n-y} f_{LN}(p \mid \mu, \sigma^2) dp \\
&= \int_0^1 \lim_{n\to\infty} P_X\left(\frac{X_n}{n} \le q\right) f_{LN}(p \mid \mu, \sigma^2) dp \\
&= \int_0^q f_{LN}(p \mid \mu, \sigma^2) dp,
\end{aligned}$$

which implies that Y_n/n converges in law to $LN(\mu, \sigma^2)$.

Appendix B. Proof of Theorem 2.2

Let $X_n \sim binomial(n,p)$ and $\Phi(\cdot)$ be the CDF of $N(0,1)$. By the Berry-Esseen Theorem,

$$\sup_q \left| P_X\left(\frac{X_n}{n} \le q\right) - \Phi\left(\frac{\sqrt{n}(q-p)}{\sqrt{p(1-p)}}\right) \right| \le \frac{C}{\sqrt{n}} \left\{p^2 + (1-p)^2\right\},$$

where C is a constant.

Suppose $Y_n \sim BLN_n(\mu, \sigma^2)$ and $Y \sim LN(\mu, \sigma^2)$. Let $G_n(\cdot)$ be the CDF of Y_n/n and $G(\cdot)$ be the CDF of Y.

$$\begin{aligned}
d_n &= \sup_q |G_n(q) - G(q)| \\
&= \sup_q \left| \int_0^1 P_X\left(\frac{X_n}{n} \le q\right) f_{LN}(p \mid \mu, \sigma^2) dp - \int_0^q f_{LN}(p \mid \mu, \sigma^2) dp \right| \\
&\le \sup_q \left| \int_0^1 \left\{ P_X\left(\frac{X_n}{n} \le q\right) - \Phi\left(\frac{\sqrt{n}(q-p)}{\sqrt{p(1-p)}}\right) \right\} f_{LN}(p \mid \mu, \sigma^2) dp \right| \\
&\quad + \sup_q \left| \int_0^1 \Phi\left(\frac{\sqrt{n}(q-p)}{\sqrt{p(1-p)}}\right) f_{LN}(p \mid \mu, \sigma^2) dp - \int_0^q f_{LN}(p \mid \mu, \sigma^2) dp \right|.
\end{aligned}$$

Let the first term of above be T_1 and the second term be T_2. Use f_{LN} for $f_{LN}(p \mid \mu, \sigma^2)$ for simplicity. Then

$$T_1 \le \int_0^1 \frac{C}{\sqrt{n}} \left\{p^2 + (1-p)^2\right\} f_{LN} dp \le \frac{C}{\sqrt{n}}$$

and

$$T_2 = \sup_q \left\{ \int_0^q \left\{ 1 - \Phi\left(\frac{\sqrt{n}(q-p)}{\sqrt{p(1-p)}}\right) \right\} f_{LN} dp + \int_q^1 \Phi\left(\frac{\sqrt{n}(q-p)}{\sqrt{p(1-p)}}\right) f_{LN} dp \right\}.$$

Let $\alpha \in (0, 1/2)$. Since $\sqrt{p(1-p)} \le 1/2$,

$$\begin{aligned}
T_{22} &= \int_q^1 \Phi\left(\frac{\sqrt{n}(q-p)}{\sqrt{p(1-p)}}\right) f_{LN} dp \\
&\le \int_q^1 \Phi\left(2\sqrt{n}(q-p)\right) f_{LN} dp \\
&= \int_{q+n^{-\alpha}}^1 \Phi\left(2\sqrt{n}(q-p)\right) f_{LN} dp + \int_q^{q+n^{-\alpha}} \Phi\left(2\sqrt{n}(q-p)\right) f_{LN} dp \\
&\le \Phi\left(-2n^{\frac{1}{2}-\alpha}\right) \int_{q+n^{-\alpha}}^1 f_{LN} dp + \Phi(0) \int_q^{q+n^{-\alpha}} f_{LN} dp.
\end{aligned}$$

Similarly,

$$\begin{aligned}
T_{21} &= \int_0^q \left\{ 1 - \Phi\left(\frac{\sqrt{n}(q-p)}{\sqrt{p(1-p)}}\right) \right\} f_{LN} dp \\
&\le \Phi\left(-2n^{\frac{1}{2}-\alpha}\right) \int_0^{q-n^{-\alpha}} f_{LN} dp + \Phi(0) \int_{q-n^{-\alpha}}^q f_{LN} dp.
\end{aligned}$$

Hence

$$
\begin{aligned}
T_2 &\leq \Phi\left(-2n^{\frac{1}{2}-\alpha}\right)\int_0^1 f_{LN}dp + \Phi(0)\sup_q\left\{\int_{q-n^{-\alpha}}^{q+n^{-\alpha}} f_{LN}dp\right\} \\
&= \Phi\left(-2n^{\frac{1}{2}-\alpha}\right) + \frac{1}{2}\sup_q\left\{\int_{q-n^{-\alpha}}^{q+n^{-\alpha}} f_{LN}dp\right\}.
\end{aligned}
$$

It is not difficult to show that the density function $f_{LN}(p \mid \mu, \sigma^2)$ is bounded in $(0,1)$. Therefore the mode of $LN(\mu, \sigma^2)$ exists. Denote the mode $M(\mu, \sigma^2)$. Then

$$
\sup_q\left\{\int_{q-n^{-\alpha}}^{q+n^{-\alpha}} f_{LN}dp\right\} \leq 2M(\mu, \sigma^2)n^{-\alpha}.
$$

Finally,

$$
\begin{aligned}
d_n &\leq T_1 + T_2 \\
&\leq \frac{C}{\sqrt{n}} + \Phi\left(-2n^{\frac{1}{2}-\alpha}\right) + M(\mu, \sigma^2)n^{-\alpha},
\end{aligned}
$$

which completes the proof.

Acknowledgements

This research was supported in part by NSF grants DMS-1007874 and SES-1024080, and the NIH grant R01-CA112159. We thank an anonymous referee for the valuable comments.

References

Aitchison, J. and Begg, C., 1976. Statistical diagnosis when cases are not classified with certainty. Biometrika, 63, 1-12.

Coull, B.A. and Agresti, A., 2000. Random effects modeling of multiple binomial responses using the multivariate binomial-logit-normal distribution. Biometrics, 56, 73-80.

Johnson, N., 1949. Systems of frequency curves generated by methods of translation. Biometrika, 36, 149-176.

Johnson, N.L. and Kotz, S. and Balakrishnan, N., 1995. Continuous univariate distributions. New York: John Wiley & Sons, Inc.

Korolev, V. and Shevtsova, I., 2010. On the upper bound for the absolute constant in the Berry-Esseen inequality. Theory of Probability and its Applications, 54, 638-658.

Lesaffre, E. and Rizopoulos, D and Tsonaka, R., 2007. The logistic transform for bounded outcome scores. Biostatistics, 8, 72-85.

McCullagh, P., 1980. Regression models for ordinal data (with discussion). Journal of the Royal Statistical Society, Series B, 42, 109-142.

Mead, R., 1965. A generalised logit-normal distribution. Biometrics, 21, 721-732.

Rabe-Hesketh, S. and Skrondal, A., 2001. Parameterization of multivariate random effects models for categorical data. Biometrics, 57, 1256-1264.

Ware, J.E.Jr. and Sherbourne, C. D., 1992. The MOS 36-Item Short-Form Health Survey (SF-36): I. Conceptual Framework and item Selection. Medical Care, 30, 473-483.