



# A Flexible Skewed Link Model for Ordinal Outcomes: An Application to Infertility

Mohammad Chehrazi<sup>1</sup>, Seyed Hassan Saadat<sup>2</sup>, Mahmoud Hajjahmadi<sup>1\*</sup>, Mirko Spiroski<sup>3</sup>

<sup>1</sup>Department of Biostatistics and Epidemiology, School of Public Health, Babol University of Medical Sciences, Babol, Iran; <sup>2</sup>Behavioral Sciences Research Center, Lifestyle Institute, Baqiyatallah University of Medical Sciences, Tehran, Iran; <sup>3</sup>Scientific Foundation SPIROSKI, Skopje, Republic of Macedonia

## Abstract

**BACKGROUND:** An important issue in modeling categorical response data is the choice of the links. The commonly used complementary log-log link is inclined to link misspecification due to its positive and fixed skewness parameter.

**AIM:** The objective of this paper is to introduce a flexible skewed link function for modeling ordinal data with some covariates.

**METHODS:** We introduce a flexible skewed link model for the cumulative ordinal regression model based on Chen model.

**RESULTS:** The main advantage suggested by the proposed links is the skewed link provide much more identifiable than the existing skewed links. The propriety of posterior distributions under proper and improper priors is explored in detail. An efficient Markov chain Monte Carlo algorithm is developed for sampling from the posterior distribution.

**CONCLUSION:** The proposed methodology is motivated and illustrated by ovary hyperstimulation syndrome data.

**Edited by:** Slavica Hristomanova-Mitkovska

**Citation:** Chehrazi M, Saadat SH, Hajjahmadi M, Spiroski M. A Flexible Skewed Link Model for Ordinal Outcomes: An Application to Infertility. Open Access Maced J Med Sci. 2020 Apr 05; 8(A):119-124. https://doi.org/10.3889/oamjms.2020.4386

**Keywords:** Latent variable; Cumulative regression; Markov chain Monte Carlo; Ordinal data; Skewed link function

\***Correspondence:** Mahmoud Hajjahmadi, Department of Biostatistics and Epidemiology, School of Public Health, Babol University of Medical Sciences, Babol, Iran. Tel.: +98-9111116135. Fax: +98-1132191973. E-mail: hajjahmadi@yahoo.com

**Received:** 03-Feb-2020

**Revised:** 01-Mar-2020

**Accepted:** 03-Mar-2020

**Copyright:** © 2020 Mohammad Chehrazi, Seyed Hassan Saadat, Mahmoud Hajjahmadi, Mirko Spiroski

**Funding:** This research did not receive any financial support

**Competing Interests:** The authors have declared that no competing interests exist

**Open Access:** This is an open-access article distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0)

## Introduction

When you face categorical response data, you initially and unconsciously think about logistic or probit regression to model. Other commonly used models are the student t link and complementary log-log model for modeling binary outcome data. The usual way to model the binomial response is to use a generalized linear model, where model the latent probability of “success” by a linear function of covariates through a link function [1]. However, these popular links do not always provide the best fit for a given data set. Chen *et al.* (1999) used the rates at which the probability of a given binary response approaches 0 and 1 to describe a link [2]. By their notation, the logit and the probit links are symmetric link functions since the response function  $p(x)$  approaches 0 at the same rate as it approaches 1, so the links do not always provide the best fit available for a given dataset, however. In this case, the link could be misspecified, which can yield substantial bias in the mean response estimates [3]. On the other hand, the complementary log-log (cloglog) link is positively skewed with the response curve approaching 0 fairly

slowly but approaching one quite sharply. However, the cloglog link has a fixed negative skewness. As a result, it lacks both the flexibility to let the data tell how much skewness should be incorporated and the ability to allow for positive skewness [4]. In short, binomial data might often be better modeled with flexible link functions that allow for both positive and negative skew and that allow the data to determine the amount of skewness required.

Several authors have proposed different link functions. Stukel (1988) suggested a class of generalized logistic models for modeling binary data, but, in the presence of covariates, Stukel’s model yields improper posterior distributions for many types of noninformative improper prior, including the improper uniform prior for the regression coefficients [5], [6]. Aranda-Ordaz (1981) proposed two separate one-parameter models for additional flexibility in the logistic model [7]. Guerrero and Johnson (1982) used Box-Cox transformation on the odds ratio to form a more flexible class of model [8]. Jones (2004) proposed a family of flexible distributions based on the distribution of order statistics [9]. Using a latent variable approach of Albert and Chib (1993) [10], Chen *et al.* (1999) proposed a

class of skewed link models, where the underlying latent variable has a mixed-effects model structure [2]. However, the model proposed by Chen *et al.* (1999) has the limitation that the intercept term is confounded with the skewness parameter. Kim *et al.* (2008) proposed a class of generalized skewed t-link models using a latent variable approach, which achieves proper posteriors for regression coefficients under uniform priors [6]. Unfortunately, the range of the skewness for generalized skewed t-link is limited due to a constraint on the shape parameter required for the identifiability of the model. More recently, Wang and Dey (2010) propose the generalized extreme value link function to allow more flexible skewness controlled by the shape parameter, but the standard logistic and probit links are not among the special cases of this family [11].

In this article, we generalize the Chen model to a class of skewed link models, including intercept for ordinal response data and apply it to ovary hyperstimulation syndrome (OHSS) data.

The rest of the paper is organized as follows. We introduce the OHSS data in Section 2. In Section 3, we propose a general class of skewed link functions for ordinal outcome data using the latent variable approach. Section 4 discusses the prior specification and posterior proprieties of the parameter in the proposed model under a fully Bayesian framework. Section 5 clarifies some computational issues in the model as well as the criteria for model comparisons. Several comprehensive simulation studies are reported in Section 6 with detailed discussions. Finally, in Section 7, we fit the proposed model on the OHSS data. We conclude our paper in Section 8 and all the proofs of the theorems are deferred to the Appendices.

## Method

First, we establish a general notation. Suppose that  $(y_{i1}, \dots, y_{ij})$  be binary indicators of the response for subject  $i$ , where

$$\begin{aligned}
 Y_i = j &\sim \text{Bernolli}(\pi_j(x_i)) \\
 \sum_{j=1}^J \pi_j(x_i) &= 1 \\
 \pi_j(x_i) &= P(Y_i \leq j) - P(Y_i \leq j-1) \\
 g(P(Y_i \leq j)) &= \alpha_j + \beta^{*'} x_i \quad (1)
 \end{aligned}$$

Let  $X_i = (x_{i1}, x_{i2}, \dots, x_{ip})'$  be the corresponding  $p$ -dimensional vector of covariates, let  $\beta_t' = (\beta_1, \beta_2, \dots, \beta_p)$  be a  $p$ -dimensional vector of regression coefficients and, let  $\{\alpha_j\}$  be increasing in  $j$ . Each cumulative link has its own intercept. This model has the same effects  $\beta^{*}$  for each link.

Suppose that  $-\infty = \alpha_j < \alpha_{j+1} < \dots < \alpha_j = +\infty$  are cut points of a continuous scale. Based on the latent variable approach of Albert and Chib (1993), an ordinal regression model can be described as follows. Let  $Y_i^*$  be a latent variable such that

$$Y_i = \int_0^j \alpha_{j-1} < Y_i^* \leq \alpha_j \quad \text{O.W} \quad (2)$$

Where  $Y_i^* = \beta^{*'} x_i + \delta^* z_i + \varepsilon_i^*$  and  $\varepsilon_i^* \sim F = (\mu_{\varepsilon^*}, \sigma_{\varepsilon^*}^2)$  are symmetric and  $|\mu_{\varepsilon^*}| \geq |\delta^*|$  and  $F$  is a cumulative distribution function. We suppose  $z_i \sim G$  is the cumulative distribution function of a skewed distribution and independent of  $\varepsilon_i^*$ .

The model has several nice properties. First, the underlying latent variable has a mixed-effects model structure. Second,  $\delta$  is the skewness parameter and, when  $G$  has support of  $R^+$  and the intercept is known, the model is positively skewed if  $\delta > 0$  and negatively skewed if  $\delta < 0$ . Third, the model reduces to symmetric link model when  $\delta = 0$  or  $G$  is a degenerate distribution at 0. Fourth, it facilitates the easy implementation of the Gibbs sampling algorithm. However, the model has the limitation that the intercept and  $\delta$  in (5) are confounded with each other. One way of handling this problem is to exclude the intercept from the model, as done in Chen *et al.* (1999). Without an intercept in the model,  $\delta$  plays a dual role. For example, when  $G$  has  $R^+$  as its support and  $\delta$  is negative, we may not be able to tell whether this negative value comes from a negative intercept or from the negative skewness of the link. In addition, this approach may lead to a loss in goodness-of-fit. Another possible solution is to leave an intercept in the model but fix  $\delta = 1$ , but then we may lose flexibility in controlling the amount of skewness in the link.

To overcome the identifiability problem in the model in (5), we propose the following skewed generalized t-link model. We begin with the following model:

$$Y_i' = \frac{Y_i^*}{\mu_{\varepsilon^*}}, \beta' = \frac{\beta^{*'}}{\mu_{\varepsilon^*}}, \mu' = \frac{\mu^*}{\mu_{\varepsilon^*}} \quad (3)$$

After reparameterization, the proposed link model takes the form:

$$Y_i' = \beta' x_i + \frac{\delta^*}{\mu_{\varepsilon^*}} z_i + \varepsilon_i' \quad (4)$$

Where

$$\left| \delta' = \frac{\delta^*}{\mu_{\varepsilon^*}} \right| \leq 1, \varepsilon_i' \sim \left(1, \frac{\sigma_{\varepsilon^*}^2}{\mu_{\varepsilon^*}^2}\right), \mu_{\varepsilon^*}^2 \neq 0, \frac{\sigma_{\varepsilon^*}^2}{\mu_{\varepsilon^*}^2} = \sigma_{\varepsilon'}^2$$

Finally, the probabilities for  $Y_i \leq j$  can be written as follow:

$$\begin{aligned} P(Y_i \leq j) &= P(Y_i' \leq \alpha_j) = P(\beta' x_i + \delta' z_i + \varepsilon_i' \leq \alpha_j) = \\ &= P(\varepsilon_i' \leq \alpha_j - \beta' x_i - \delta' z_i) = \int_{-\infty}^{+\infty} F(\alpha_j - \beta' x_i - \delta' z_i) dG(z_i) \quad (5) \end{aligned}$$

### Special case

We consider a special case which showing the marginal distribution of  $Y_i'$  is skewed normal when  $\varepsilon_i'$  is standard normal and  $z_i$  is half-standard normal.

$$\varepsilon_i' \sim N(0,1)$$

$$z_i \sim HN^+(0,1)$$

The characteristic function of half-standard normal and s normal IS:

$$\begin{aligned} \varphi_{z_i}(t) &= \exp\left(\frac{-t^2}{2}\right) \left[ 1 + i \sqrt{\frac{2}{\pi}} \int_0^t \exp\left(\frac{x^2}{2}\right) dx \right] \\ \varphi_{\varepsilon_i'}(t) &= \exp(it\mu_{\varepsilon_i'} - \frac{1}{2}\sigma_{\varepsilon_i'}^2 t^2) \\ \varphi_{Y_i'}(t) &= E\left(\exp\left(it(\beta' x_i + \delta' z_i + \varepsilon_i')\right)\right) = \\ &= \exp\left(it(\beta' x_i)\right) \cdot \exp\left(it(\delta' z_i)\right) \cdot \exp\left(it\mu_{\varepsilon_i'}\right) = \\ &= \exp\left(it(\beta' x_i)\right) \cdot \exp\left(it\left(0 - \frac{1}{2}t^2\right)\right) \cdot \exp\left(\frac{-\delta'^2 t^2}{2}\right) \\ &\quad \left[ 1 + i \sqrt{\frac{2}{\pi}} \int_0^{\delta' t} \exp\left(\frac{x^2}{2}\right) dx \right] \\ &= \exp\left(it(\beta' x_i) - \frac{1}{2}t^2 - \frac{\delta'^2 t^2}{2}\right) \cdot \left[ 1 + i \sqrt{\frac{2}{A}} \int_0^{\delta' t} \exp\left(\frac{x^2}{2}\right) dx \right] = \\ &= \exp\left(it(\beta' x_i) - \frac{1}{2}t^2(1 + \delta'^2)\right) \cdot \\ &\quad \left[ 1 + i \sqrt{\frac{2}{\pi}} \int_0^{\frac{\delta'}{\sqrt{1+\delta'^2}} t} \frac{\delta'}{\sqrt{1+\delta'^2}} \exp\left(\frac{x^2}{2}\right) dx \right] \quad (6) \end{aligned}$$

The last equation is a characteristic function of skewed normal.

$$\text{Where } Y_i' \sim SN(\beta' x_i, \sqrt{1 + \delta'^2}, \delta')$$

### The prior and posterior distributions

The likelihood function for the model is given by:

$$\begin{aligned} L(\beta', \delta', \alpha_j, \sigma_{\mu}^2 | y, X, n, J) &= \\ \prod_{i=1}^n \prod_{j=1}^J \int_{-\infty}^{+\infty} [F(\alpha_j - \beta' x_i - \delta' z_i) & \\ - F(\alpha_{j-1} - \beta' x_i - \delta' z_i)]^{y_{ij}} [g(z_i)]^{y_{ij}} dz_i & \quad (7) \end{aligned}$$

Following Gibbs sampling, we should define a Bayesian framework for hypothetical data which the outcome has four categories and having a couple of independent variables. First of all, using a simple transformation on  $Y_i'$  residuals will be normalized.

Therefore, it does not need estimating  $\sigma_{\mu}^2$ .

$$h_i = \frac{Y_i' - 1}{\sigma_{\mu}^2} = \beta'' x_i + \delta'' z_i + \frac{\varepsilon_i' - 1}{\sigma_{\varepsilon}^2}, \frac{\varepsilon_i' - 1}{\tilde{A}_5^2} = \zeta_i \quad (8)$$

We assume that  $\delta'$ ,  $\beta'$ , and  $\alpha$  are independent a priori. Thus, the proper prior for the parameters are defined as follows.

$$\delta' \sim \text{unif}(-1, 1)$$

$$\beta' \sim N_p(0, kI_p), \quad k = \text{a known constant}$$

$$\alpha = (\alpha_1, \alpha_2, \alpha_3) \sim N_3(0, kI_3)$$

We also considered positive standard half-normal and normal for  $z$  and  $F$ , respectively.

Hence, the posterior distribution was obtained in below:

$$\begin{aligned} p(\alpha, \beta', \delta', h, z | h) &\propto \prod_{i=1}^n \prod_{j=1}^3 \\ &\left\{ 1(y_i = j) \cdot 1(\alpha_{j-1} < h_i \leq \alpha_j) \cdot \exp\left(\frac{-1}{2} z_i^2\right) \right\} \\ &\cdot \exp\left\{\frac{-1}{2}(h_i - \alpha_j - \beta' x_i - \delta' z_i)^2\right\} \\ &\cdot \exp\left\{\frac{-\alpha' I_3 \alpha}{2k}\right\} \cdot \exp\left\{\frac{-\beta' I_p \beta}{2k}\right\} \quad (9) \end{aligned}$$

Let  $\tau_i = 1$  if  $Y_i = j$  otherwise  $= \infty$ . Let  $X$  denote the  $n \times p$  known design matrix with rows  $X_i'$  having full column rank associated with all  $n$  observations and define  $X^*$  as the matrix with rows  $\tau_i X_i'$ , where  $p$  is the dimension of  $\beta$ .

**Theorem:** Suppose prior distribution is improper for,  $p(\beta) \propto 1$ , and the following conditions are satisfied:

(C1)  $X$  is of full rank.

(C2) There exists  $a > 0$  and  $\varepsilon \in R^n$  such that  $X^* a = 0$ .

(C3)  $(\alpha_j)^p - (\alpha_{j-1})^p < \infty$

and  $p(\alpha)$  would be proper, then the posterior is proper (the integral of the likelihood times the prior is finite).

Although the analytical evaluation of the joint posterior distribution of  $(\alpha, \beta', \delta', h, z)$  given in (9) does not appear possible, the random-effects structure of the skewed link model allows us to develop an efficient Markov chain Monte Carlo sampling algorithm to sample from this joint posterior distribution. For ease of presentation, we consider  $G$  equal exponential distribution only as the Markov chain Monte Carlo sampling algorithms for other choices of  $G$  are similar.

$$\alpha \sim N_2 \left[ U^{-1} (H - x'B - \delta z), U^{-1} \right], U = I - \frac{I}{k}$$

$$z_i \sim HN^+ \left( \frac{\delta k}{\delta^2 + 1}, \frac{1}{\delta^2 + 1} \right)$$

$$\delta \sim TN \left( A, \frac{1}{z_i} \right), A = \frac{h_i - \alpha_j - x_i'B}{z_i}$$

$$h_i \sim TN(\alpha_j - x_i'B - \delta z_i, 1, \alpha_{j-1}, \alpha_j)$$

$$B \sim N_2 \left[ \Sigma^{-1} X' (H - \alpha_j - \delta z), \Sigma^{-1} \right], \Sigma = \frac{I}{k} + X'X$$

$$\Gamma = \{ \gamma > 0, \gamma(\theta) = (\gamma\delta, \gamma B, \gamma S) \}$$

$$p_S(\gamma | w, z, x, y) \propto \gamma^{a_0} \exp \left\{ -\gamma \left( S b_0 + \frac{S A^2}{2} \right) \right\}$$

$$p_\delta(\delta | S, w, z, x, y) \propto \exp$$

$$\left\{ -\frac{S^2 \delta^2 Z'Z}{2} \left( 3 - \frac{(Z'Z)^{-1} Z'(W - XB)}{\delta} \right)^2 \right\}$$

$$p_B(\gamma | \delta, S, w, z, x, y) \propto \exp \left\{ -\frac{1}{2} \left( \gamma - \frac{S^2 B^{-1} X'W}{\frac{I}{C_0} + X'X} \right)^2 \right. \\ \left. - \frac{1}{2} \left( \frac{I}{C_0} + X'X \right) \left( \gamma - \frac{S^2 B^{-1} X'W}{\frac{I}{C_0} + X'X} \right)^2 \right\}$$

### Analysis of the data

We consider data from a randomized controlled trial (RCT) study of women randomly assigned to one of two treatment groups Albumin or Cabergoline,  $n = 138$ , which is a subset of the data published in Tehraninejad *et al.* [12]. The ordinal response variable

( $y$ ) is OHSS, which takes the values 0, 1, and 2 and 3, where 0 denotes that the OHSS has not occurred and 1–3 indicates the severity of OHSS (mild, moderate, severe). We consider six factors: Treatment group, number of MII, HCG dose, ascites, HCT (hematocrit), and insemination. All variables are continuous except treatment group which is binary, taking the values 0 and 1, where 1 denotes Albumin and 0 indicates Cabergoline. In addition, 69 and 56 patients developed OHSS in Albumin and Cabergoline groups, respectively. In Markov chain Monte Carlo sampling, we standardized all covariates.

We use the DIC measure proposed by Spiegelhalter *et al.* [13] to compare ordinal regression models under different links. Let  $\theta$  denote the collection of all parameters involved in the model under consideration. For example, for the skewed link model in (2), we have  $\theta = (\alpha, \beta', \delta', h, z)$ . Then,

$$DIC = D(\theta) + 2pD \tag{10}$$

where  $D(\theta)$  is a deviance function and  $\theta = E(\theta | y, X)$  is the posterior mean of  $\theta$ . In (10),  $pD$  is the effective number of model parameters, which is calculated as  $pD = E\{D(\theta) | y, X\} - D(\theta)$ . For the proposed skewed link model, we take the deviance function to be of form  $D(\theta) = -2 \log p(y | X, \beta, \delta, v1)$ , where  $p(y | X, \beta, \delta, v1)$  is the likelihood function based on the observed data given in Guerrero and Johnson [8]. In Cowles and Carlin [14],  $D(\theta)$  is a Bayesian measure of lack of fit or adequacy and  $2pD$  is the complexity penalty term. The smaller the DIC value, the better the model fits the data. The other properties of the DIC can be found in Spiegelhalter *et al.* (2002). Note that it is important to integrate out all latent variables in the deviance calculation, as this yields a more appropriate penalty term  $2pD$ .

Table 1 shows the values of DIC for six covariate-based models under symmetric and skewed links. With the model, the effective number of model parameters  $pD$  is about 6 for all cases. For example, the values of  $pD$  are 6.25 when  $F$  is normal and  $G = E$ , 6.12 when  $F$  is Cauchy and  $G =$  negative exponential (NE) and 5.69 when  $F$  is normal and  $G$  is degenerated at zero.

**Table 1: OHSS data. Values of DIC for models with all sets of covariates under symmetric and skewed links**

F	G	DIC
Normal	E	701.9
	HN	696.7
Logistic	E	702.2
	HN	702.0
Cauchy	E	695.2
	HN	691.9
Complementary log-log	-	723.6

Covariates (TG, Acite, no.MII, Insemination, dose HCG, HCT).

We can see some interesting patterns in the values of DIC in Table 1. For any set of covariates, the skewed link model and  $G = NE$  fits the data best. Except for the probit link, a negatively skewed link fits the data better than the positively skewed and symmetric links within each class of Cauchy links. When  $F$  is normal,



the DIC values for different G's are very similar under the models that include the six covariates. For any given G, the proposed link achieves a better fit than the probit and Cauchy links. Within the class of the proposed link, the models that include the treatment group greatly improve the fit for both the symmetric and skewed links over the models with HCG dose. When ascites is included in the model, the number of MII leads to much further improvement under the symmetric link but not much under the skewed links. The symmetric logit link with the covariates fits the data much better than the negatively skewed link with the covariates. Therefore, it is important that the choice of links should always be done in conjunction with the selection of covariates.

Table 2 shows the posterior means, the posterior standard deviations, and the 95% highest posterior density intervals of the parameters of the regression model with covariates treatment group, number of MII, dose of HCG, ascites, HCT, and insemination under the symmetric probit link and the skewed probit link. Except for the treatment group and ascites, the posterior estimates of all regression coefficients are positive, which implies that the cumulative probability of OHSS is an increasing function of the number of MII, dose of HCG, HCT, and insemination. In addition, all six covariates are significant under both links, as all interval estimates fail to contain 0. Note that, under the skewed probit link model, the posterior means of  $\delta$  is 0.43.

**Table 2: OHSS data. Posterior estimates under symmetric and proposed skewed link models**

Model	Variable	Estimate	Std Dev	95% HPD Interval	
Probit link	Intercept 1	2.25	0.99	(0.30, 4.41)	
	Intercept 2	2.82	0.38	(2.07, 3.74)	
	Trt Group				
	Albumin	Ref		Ref	
	Cabergoline	-5.01	1.21	(-7.34, -2.65)	
	No.MII	0.12	0.05	(0.04, 0.21)	
	Dose HCG	1.90	0.59	(0.77, 3.02)	
	Ascites	-4.48	0.89	(-6.11, -2.59)	
	HCT	0.38	0.09	(0.20, 0.57)	
	Insemination	0.12	0.04	(0.04, 0.19)	
	Skewed probit link	Intercept 1	2.49	1.10	(0.31, 4.49)
		Intercept 2	3.08	0.43	(2.06, 3.89)
		Trt Group			
		Albumin	Ref		
Cabergoline		-4.89	1.29	(-7.05, -2.45)	
No.MII		0.13	0.05	(0.04, 0.20)	
Dose HCG		1.94	0.62	(0.80, 3.00)	
Ascites		-4.03	0.94	(-6.01, -2.3)	
HCT		0.41	0.10	(0.21, 0.59)	
Insemination		0.13	0.05	(0.03, 0.24)	
		0.43	0.15	(0.12, 0.78)	

Std Dev, standard deviation; Trt: Treatment; HCT: Hematocrit; HPD.

In all the computations, there are 150,000 Gibbs samples, but we only used 30,000 iterations, obtained from every 5<sup>th</sup> iteration, to compute all quantities of interest, using a burn-in of 2000 iterations. The computer codes were written in R 3.4.4 with double precision accuracy. The convergence of the Gibbs sampler was checked using several diagnostic procedures, as recommended by Cowles and Carlin [14]. Approximate convergence is reached after 2000 iterations.

## Discussion

This article has presented a class of new skewed link models for analyzing ordinal response data with covariates. Chen *et al.* suggested that an asymmetric link model for binary outcomes may be more appropriate than symmetric link model when the number of 1's is much different than the number of 0's [2]. In this paper, the outcome is ordinal and there has not been any considerable evaluation in this regard. In the example, the data were generated from the asymmetric complementary log-log model. We showed that the skewed logit model with a negative skewness parameter six fits the data better than the symmetric logit model. However, the counts of 1's and 0's only give an indication for the choice of links, and the other factors, such as the distribution of covariates, may also affect the choice of links for a given dataset. Although the proposed skewed link model does not include the complementary log-log link model, our large-scale simulation study has shown that the complementary log-log link model can be well approximated by a skewed probit model with a half-normal distribution for G and a skewness parameter roughly equal to 0.55. Compared to the complementary log-log link model, the advantage of the proposed skewed link model is that it allows more flexibility to model skewness; that is, the skewness can be determined by the data. There are, however, a few aspects that warrant further discussion. The most important issue is the choice of F and G in the model defined by Chen *et al.* [2]. In this article, we took F to be a normal or logistic cdf and G to have the density of a standard half-normal cdf. In practice, the choice of F is usually restricted to a standard cdf, such as a normal or logistic, whereas there are many plausible choices of G. The choice of G corresponding to a half-normal cdf is analytically tractable and easily implementable from a computational perspective. Nonparametric mixture models such as Dirichlet process mixture models can be implemented in modeling F as well as G. Some related work in this direction has been done by Newton *et al.* [15], Basu and Mukhopadhyay [16]. Following, one can develop model diagnostic strategies for modeling ordinal response data. In addition, a more general skewed link model in this regard could be produced.

## Acknowledgment

We gratefully thank Professor Ming Hui Chen and Dipak Day for their valuable comments to improve and enrich this paper.

### References

1. McCullagh P, Nelder JA. Generalized Linear Models. Boca Raton, Florida: CRC Press; 1989.
2. Chen MH, Dey DK, Shao QM. A new skewed link model for dichotomous quantal response data. *J Am Stat Assoc.* 1999;94:1172-86.
3. Czado C, Santner TJ. The effect of link misspecification on binary regression inference. *J Stat Plann Inference.* 1992;33:213-31. [https://doi.org/10.1016/0378-3758\(92\)90069-5](https://doi.org/10.1016/0378-3758(92)90069-5)
4. Jiang X, Dey DK, Prunier R, Wilson AM, Holsinger KE. A new class of flexible link functions with application to species co-occurrence in cape floristic region. *Ann Appl Stat.* 2013;7:2180-204. <https://doi.org/10.1214/13-aos663>
5. Stukel TA. Generalized logistic models. *J Am Stat Assoc.* 1988;83:426-31.
6. Kim S, Chen MH, Dey DK. Flexible generalized t-link models for binary response data. *Biometrika.* 2008;95:93-106. <https://doi.org/10.1093/biomet/asm079>
7. Aranda-Ordaz FJ. On two families of transformations to additivity for binary response data. *Biometrika.* 1981;68:357-63. <https://doi.org/10.1093/biomet/68.2.357>
8. Guerrero VM, Johnson RA. Use of the box-cox transformation with binary response models. *Biometrika.* 1982;69:309-14. <https://doi.org/10.1093/biomet/69.2.309>
9. Jones M. Families of distributions arising from distributions of order statistics. *Test.* 2004;13:1-43. <https://doi.org/10.1007/bf02602999>
10. Albert JH, Chib S. Bayesian analysis of binary and polychotomous response data. *J Am Stat Assoc.* 1993;88:669-79.
11. Wang X, Dey D. A Flexible Skewed Link Function for Binary Response Data. University of Connecticut, Department of Statistics, Technical Report; 2008.
12. Tehraninejad ES, Hafezi M, Arabipoor A, Azimineko E, Chehrizi M, Bahmanabadi A. Comparison of cabergoline and intravenous albumin in the prevention of ovarian hyperstimulation syndrome: A randomized clinical trial. *J Assist Reprod Genet.* 2012;29:259-64. <https://doi.org/10.1007/s10815-011-9708-4>
13. Spiegelhalter DJ, Best NG, Carlin BP, Van Der Linde A. Bayesian measures of model complexity and fit. *J R Stat Soc.* 2002;64:583-639. <https://doi.org/10.1111/1467-9868.00353>
14. Cowles MK, Carlin BP. Markov chain Monte Carlo convergence diagnostics: A comparative review. *J Am Stat Assoc.* 1996;91(434):883-904. <https://doi.org/10.1080/01621459.1996.10476956>
15. Newton MA, Czado C, Chappell R. Bayesian inference for semiparametric binary regression. *J Am Stat Assoc.* 1996;91:142-53. <https://doi.org/10.1080/01621459.1996.10476671>
16. Basu S, Mukhopadhyay S. Binary response regression with normal scale mixture links. *Biostat Basel.* 2000;5:231-42.

### Appendix

**Proof:** Let  $\varepsilon_1, \dots, \varepsilon_n$  be independent random variables with common distribution function  $F$ . We have

$$\left[ F(\alpha_j - \beta' X_i - \delta' z_i) - F(\alpha_{j-1} - \beta' x_i - \delta' z_i) \right]^{y_{ij}} = E \left\{ \begin{matrix} 1(\varepsilon_i \leq \tau_i(\alpha_j - \beta' x_i - \delta' z_i)) \\ -1(\varepsilon_i \leq -\tau_i(-\alpha_{j-1} + \beta' x_i + \delta' z_i)) \end{matrix} \right\}$$

Let  $z^* = (\tau_1 z_1, \dots, \tau_n z_n)'$  and  $\alpha = (\alpha_1, \dots, \alpha_{j-1})$  using Fubini's theorem and a result of Lemma 4.1 of Chen and Shao (2001) that there exists a constant  $K$  depending only on  $X$  such that  $\|-\beta\| \leq K \|-\alpha\|$  whenever  $-\alpha^* \beta \leq -u$ , we have

$$\int_{-\infty}^{+\infty} \int_{R^{p-1}} \int_{-1}^{+1} \left[ \prod_{i=1}^n \prod_{j=1}^J \int_{-\infty}^{+\infty} \left[ F(\alpha_j - \beta' X_i - \delta' z_i) \right]^{y_{ij}} \left[ -F(\alpha_{j-1} - \beta' x_i - \delta' z_i) \right]^{y_{ij}} \right] p(\delta') p(\alpha) d\delta' dB d\alpha$$

$$= \int_{-\infty}^{+\infty} \int_{R^{nJ}} E \left\{ \begin{matrix} \int_{R^p} \left( \prod_{i=1}^n \prod_{j=1}^J (dG(z_i))^{y_{ij}} \right) \int_{R^p} \left[ 1(\varepsilon_i \leq \alpha_j + \tau_i(-\beta' x_i - \delta' z_i)) \right. \\ \left. 1(\varepsilon_i \leq \alpha_{j-1} - \tau_i(\beta' x_i + \delta' z_i)), \leq i \leq n \right] \end{matrix} \right\} dB$$

$$p(\delta) p(\alpha) d\delta d\alpha$$

$$= \int_{-\infty}^{+\infty} \int_{R^{nJ}} E \left\{ \begin{matrix} \int_{R^{nJ}} \left( \prod_{i=1}^n \prod_{j=1}^J (dG(z_i))^{y_{ij}} \right) \int_{R^p} \left[ 1(-\beta' X^* \leq \alpha_j - \varepsilon + \delta' Z^*) \right. \\ \left. -1(-\beta' X^* \leq \alpha_{j-1} - \varepsilon + \delta' Z^*) \right] \end{matrix} \right\} dB$$

$$p(\delta) p(\alpha) d\delta d\alpha$$

$$< K \int_{-\infty}^{+\infty} \int_{R^{nJ}} E \left\{ \begin{matrix} \int_{R^{nJ}} \prod_{i=1}^n \prod_{j=1}^J (dG(z_i))^{y_{ij}} \times \|\mu^*\|^{p+1} \\ +(-\delta \|z^*\|)^p + \|\alpha_j\|^p - \|\varepsilon^*\|^p \\ +(-\delta \|z^*\|)^p + \|\alpha_{j-1}\|^p \end{matrix} \right\}$$

$$p(\delta) p(\alpha) d\delta d\alpha =$$

$$K \int_{-\infty}^{+\infty} \int_{R^{nJ}} E \left\{ \begin{matrix} \int_{R^{nJ}} \prod_{i=1}^n \prod_{j=1}^J (dG(z_i))^{y_{ij}} \\ \times \left( (\alpha_j)^p - (\alpha_{j-1})^p \right) \end{matrix} \right\}$$

$$p(\delta) p(\alpha) d\delta d\alpha$$

The last statement could be finite if:  $(\alpha_j)^p - (\alpha_{j-1})^p < \infty$  and  $p(\alpha)$  would be a proper prior.