

# Chapter 16

## Logistic Cusp Catastrophe Regression for Binary Outcome: Method Development and Empirical Testing



(Din) Ding-Geng Chen and Xinguang Chen

**Abstract** Cusp catastrophe models are unique to advance life sciences, psychology and behavioral studies. Extensive progresses have been made to utilize this modeling technique for continuous outcome and there is no development for binary data. To fill this gap, this chapter is then aimed to develop a cusp catastrophe modelling method for binary outcome. Building upon our previous research on the nonlinear regression cusp (RegCusp) catastrophe model for continuous outcome, we propose a logistic cusp catastrophe regression (LogisticCusp). LogisticCusp is based on the principles of logistic regression for binary outcome variable  $y$  (yes/no) being expressed as a latent binary variable  $Y$  through a logit link. This latent regression provides a mathematical connection between an observed outcome variable as a binomially distributed random variable and the deterministic cusp catastrophe at its equilibrium. By connecting the two,  $Y$  in the LogisticCusp is considered as one of the true roots of the deterministic cusp catastrophe model determined using the Maxwell or Delay conventions. We validate the method using a 5-step Monte-Carlo simulation with two predictors and three parameters for both bifurcation and asymmetry control variables. We further tested the method with binge drinking behavior in youth with data from the Monitoring the Future Study. Results from 5000 Monte-Carlo simulations indicate that the parameter estimates obtained through LogisticCusp are unbiased and efficient using maximum likelihood estimation with quasi-Newton numerical search algorithm. Results from empirical testing with real data are consistent with those estimated using other methods. LogisticCusp adds a new tool for researchers to examine many issues in

---

(Din) D.-G. Chen (✉)

School of Social Work, University of North Carolina, Chapel Hill, NC, USA  
e-mail: [dinchen@email.unc.edu](mailto:dinchen@email.unc.edu)

X. Chen

Department of Epidemiology, College of Public Health and Health Professions, College of Medicine, University of Florida, Gainesville, FL, USA

Global Health Institute, Wuhan University, Wuhan, China  
e-mail: [jimax.chen@ufl.edu](mailto:jimax.chen@ufl.edu)

© Springer Nature Switzerland AG 2020

X. Chen, (Din) D.-G. Chen (eds.), *Statistical Methods for Global Health and Epidemiology*, ICSA Book Series in Statistics,  
[https://doi.org/10.1007/978-3-030-35260-8\\_16](https://doi.org/10.1007/978-3-030-35260-8_16)

psychology, life sciences, and behavioral studies, particularly, issues in medicine and public health with the powerful cusp catastrophe modeling for binary outcome.

**Keywords** Cusp catastrophe model · Logistic cusp catastrophe regression · Bifurcation · Asymmetry · Binary outcome

## 16.1 Background

Up to date, the statistical models commonly used to examine medical, health, psychological, and socio-behavioral outcomes depends on the linear regression and continuous change approach (Chen & Chen, 2015, 2019; Chen, Stanton, Chen, & Li, 2013). However, in the real world, these outcomes are rarely linear and continuous because of the nature of the medical, health, and behavioral outcomes and the multiple, complex influences of environmental, behavioral, psychological, and biological factors (Chen, Lin, Chen, Tang, & Kitzman, 2014; Chen et al., 2010; Witkiewitz, van der Maas, Hufford, & Marlatt, 2007; Xu & Chen, 2016). What might appear to be small and inconsequential changes in one of these factors can lead to abrupt and sudden changes in an outcome (Thom, 1975). Under these conditions, a linear and continuous approach seriously limits the predictability of the influence of hypothesized factors on a particular outcome variable (Chen & Chen, 2015, 2019; Chen, Wang, & Chen, 2019) and therefore a new paradigm to incorporate nonlinear and discrete behaviors is needed to fill this knowledge gap.

### 16.1.1 Cusp Catastrophe for Nonlinear Discrete Systems

To account for nonlinearity and discrete characteristics in low-dimensional scenarios, researchers often turn to natural extensions of a linear regression model, including the kernel regression or regression/smoothing splines (Berk, 2008; Faraway, 2009; Guastello & Gregson, 2011). In addition to these nonparametric methods, other techniques for use with high-dimensional data include additive models, multivariate adaptive regression splines, random forests, neural networks, and support vector machine. These techniques have been discussed extensively elsewhere (Chen & Chen, 2017; Faraway, 2009). Despite much strength, these nonparametric methods do not have a mechanism to identify and incorporate a medical, health and behavior outcomes with sudden and discrete changes and multimodes. Cusp catastrophe model is one that is capable to quantify such a mechanism.

As a complement to many traditional analytical approaches, the cusp catastrophe model offers distinct advantages given its capacity to not only simultaneously handle complex linear and nonlinear relationships in a high-order probability density function but also to incorporate sudden jumps in outcome measures, as outlined in Zeeman (Zeeman, 1977) and Gilmore (Gilmore, 1981). Catastrophe

theory was proposed in the 1970s (Thom, 1975) to understand a complicated set of behaviors that included gradual, continuous changes as well as sudden and discrete or catastrophic changes in general. The cusp catastrophe model has been used extensively in a wide range of research fields, including the modeling of tobacco use (Xu & Chen, 2016), adolescent alcohol use (Clair, 1998), changes in adolescent substance use (Mazanov & Byrne, 2006), binge drinking among high school and college students (Chen et al., 2019; Guastello, Aruka, Doyle, & Smerz, 2008) adult population (White, Tapert, & Shukla, 2017) and problem drinking among persons living with HIV (Witkiewitz et al., 2007), sexual initiation and condom use among young adolescents (Chen et al., 2010, 2013), nursing turnover (Wagner, 2010), HIV prevention (Xu, Chen, Yu, Joseph, & Stanton, 2017), therapy and program evaluation (Guastello, 1982), health outcomes (Chen et al., 2014), and accident process (Guastello, 1989; Guastello & Lynn, 2014).

### 16.1.2 Established Methods for Cusp Catastrophe Modeling

Historically, three main implementation approaches have been established for data analysis to conduct cusp catastrophe modeling.

The first method test the outcome variable if it follows cusp catastrophe by inserting regression coefficients into the deterministic cusp model and the method was operationalized by Guastello using a *polynomial regression approach* (Guastello, 1982; Guastello et al., 2008). This method is straight forward to understand and the analysis can be completed using any software packages with regression analysis functionality (Guastello & Gregson, 2011).

The second method uses a *stochastic differential equation* from Cobb and his colleagues (Cobb, 1981; Cobb & Zacks, 1985; Grasman, van der Maas, & Wagenmakers, 2009) with likelihood estimation implemented in an R package “cusp”. Since the method was established by Cobb and implemented through Grasman’s work, this approach has been named as Cobb-Grasman cusp modeling (Chen et al., 2019).

The third method takes a different approach to solve the deterministic cusp catastrophe model with a statistical approach. Different from the Cobb-Grasman’s approach described above, in this method, the deterministic cusp catastrophe is directly casted into the classical multiple regression with the outcome variable being measured with a latent variable and the two control variables each being measured as linear combination. In this modeling approach, method for estimation of the cusp region is also provided (Chen & Chen, 2017). This Chen-Chen method has been used in modeling harm perception and social influence on binge drinking among high school students in the United States (Chen et al., 2019). In Chap. 15 of this book, this method was used to model prostate-specific antigen (PSA), a biomarker of prostate cancer in men.

### 16.1.3 *Need for Methods to Model Binary Data*

All the methods described above for cusp catastrophe modeling are for continuous outcome variables, and no one method is available for other types of outcomes, to the best of our knowledge. To fill this methodology gap, in this chapter we attempted a method to analyze binary outcome with cusp catastrophe model. In our previous research, we developed a regression-based approach to solve for cusp catastrophe model for continuous outcomes (Chen & Chen, 2017) and used it in analyzing binge drinking among youth (Chen et al., 2019). We used the same regression-based approach in this new method with the continuous outcome being replaced by binary outcome for cusp catastrophe modeling of binary data in the framework of statistical logistic regression.

## 16.2 **An Overview of the Cusp Catastrophe Model**

Catastrophe theory was proposed in the 1970s by Thom (1975) and popularized over the next two decades by several leading researchers (Cobb, 1981; Cobb & Ragade, 1978; Cobb & Watson, 1980; Cobb & Zacks, 1985; Gilmore, 1981; Thom & Fowler, 1975; Zeeman, 1977). Thom (1975) originally proposed the catastrophe theory to understand complicated phenomena that included both gradual, continuous change and sudden, discontinuous or catastrophic change.

### 16.2.1 *Deterministic Cusp Model*

To apply this model in research, the deterministic cusp catastrophe model can be specified with three components: two control factors (i.e.,  $\alpha$  and  $\beta$ ) and one outcome variable (i.e.,  $y$ ). This model is defined by a dynamic system:

$$\frac{dy}{dt} = -\frac{dV(y; \alpha, \beta)}{dy} \quad (16.1)$$

where  $V$ , commonly called the potential function, is defined as

$$V(y; \alpha, \beta) = -\alpha y - \frac{1}{2}\beta y^2 + \frac{1}{4}y^4 \quad (16.2)$$

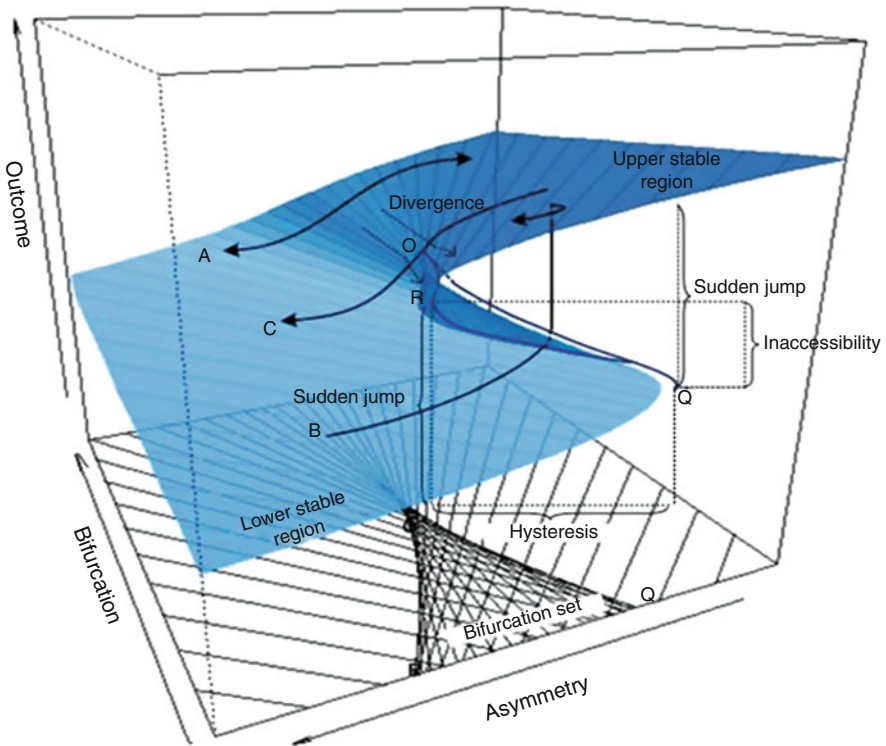
In this potential function  $V$ ,  $\alpha$  is the asymmetry or normal control factor, and  $\beta$  is the bifurcation or splitting control factor. Both  $\alpha$  and  $\beta$  are linked to determine the outcome variable  $y$  in a three-dimensional response surface. When the right side of Eq. (16.1) moves toward zero, change in the outcome  $y$  also tends toward zero with change in time; this status is called equilibrium. In general, the behavior of the

outcome  $y$  (i.e., how  $y$  changes with time  $t$ ) is complicated, but all subjects tend to move toward equilibrium the surface.

### 16.2.2 Characteristics of the Cusp Catastrophe Model

Figure 16.1 graphically depicts the equilibrium surface that reflects the response plan of the outcome measure ( $y$ ) at various combinations of the asymmetry control factor (the measure of  $\alpha$  in Fig. 16.1) and the bifurcation control factor (the measure of  $\beta$  in Fig. 16.1).

As shown in Fig. 16.1, dynamic changes in  $y$  have two stable regions (attractors), which are the lower area in the front left (lower stable region) and the upper area in the front right (upper stable region). Beyond these stable regions,  $y$  becomes sensitive to changes in  $\alpha$  and  $\beta$ . The unstable region can be projected to the control plane ( $\alpha, \beta$ ) as the *cusp region*. The cusp region is characterized by line OQ (the ascending threshold) and line OR (the descending threshold) of the equilibrium



**Fig. 16.1** Cusp catastrophe model for outcome ( $y$ ) in the equilibrium plane with an asymmetry control variable (the measure of  $\alpha$ ) and a bifurcation control variable (the measure of  $\beta$ )

surface. In this region,  $y$  becomes highly unstable with regard to changes in  $\alpha$  and  $\beta$ , jumping between the two stable regions when  $(\alpha, \beta)$  approaches the two threshold lines OQ and OR. In Fig. 16.1, paths A, B, and C depict three typical but distinct pathways of change in the health outcome measure ( $y$ ). Path A shows that in situations where  $y < 0$ , a smooth relation exists between  $y$  and  $\alpha$ . Path B shows that in situations when  $y > 0$ , if  $\alpha$  increases to reach and pass the ascending threshold link OQ,  $y$  will suddenly jump from the low stable region to the upper stable region of the equilibrium plane. Path C shows that a sudden drop occurs in  $y$  as  $\alpha$  declines to reach and pass the descending threshold line OR.

The cusp catastrophe model can be used as both a qualitative and a quantitative analytical method in research to investigate the relationship between predictors and outcome variables (e.g., behaviors or health outcomes). The qualitative approach focuses on identifying the five catastrophe elements (i.e., catastrophe flags) outlined by Gilmore (1981), whereas the quantitative approach uses numerical data to statistically fit the model.

## 16.3 Implementation of a Cusp Catastrophe Model

As described in the Introduction, since the introduction of the cusp catastrophe model, three quantitative approaches have been developed and used to implement the model for data analysis: Guastello's polynomial regression, Cobb-Grasman stochastic differential equation implemented in an R package "cusp", and Chen-Chen approach to cast the cusp catastrophe model into the nonlinear regression.

### 16.3.1 Guastello's Polynomial Approach

Specifically, as the first implementation, Guastello's approach is derived by reformulating the cusp dynamic system in Eq. (16.1) in the differential equation form into a difference equation system as outlined in Guastello (1982), Guastello et al., (2008). Since its first publication, this approach has been widely used in analyzing research data because this approach can be implemented in common statistical software packages, including SAS, SPSS, STATA, and R. This approach makes it possible the first time for many researchers to modeling social and behavioral issues with cusp catastrophe modeling. Guastello's approach is suitable for longitudinal data with outcome variables measured at two time points that are not vary far from each other.

### 16.3.2 Cobb-Grasman's Approach

As the second approach in implementing the cusp catastrophe model is Cobb-Grasman's stochastic differential equation method (named thereafter as "SDE-Cusp"). In this SDECusp approach, the deterministic cusp model in Eq. (16.1) is first extended with a probabilistic/stochastic Wiener process. With this extension, the modeling process incorporates measurement errors in the outcome variable. Using this approach, the response surface of cusp catastrophe is modeled as a probability density function where the bimodal nature of the outcome corresponds to the two states of outcome variable. Mathematically, Cobb and his colleagues (Cobb & Ragade, 1978; Cobb & Watson, 1980; Cobb & Zacks, 1985; Hartelman, van der Maas, & Molenaar, 1998; Honerkamp, 1994) cast the deterministic cusp model in Eq. (16.1) into a stochastic differential equation (SDE) as follows:

$$dz = \frac{\partial V(z, \alpha, \beta)}{\partial z} dt + dW(t) \quad (16.3)$$

where  $dW(t)$  is a white noise Wiener process with variance  $\sigma^2$ .

This extension is in fact a special case of general stochastic dynamical systems modeling with a constant diffusion function defined by  $dW(t)$ . Since the model Eq. (16.2) cannot be solved analytically, computational implementation of this stochastic model is limited. However, at the equilibrium state when time ( $t$ ) approaches the infinity, it is easier to estimate the probability density function of the corresponding limiting stationary stochastic processes. In other words, the probability density function of the outcome measure ( $y$ ) can be expressed as follows:

$$f(y) = \frac{\Psi}{\sigma^2} \exp \left[ \frac{\alpha(y - \lambda) + \frac{1}{2}\beta(y - \lambda)^2 - \frac{1}{4}(y - \lambda)^4}{\sigma^2} \right] \quad (16.4)$$

where the parameter  $\Psi$  is a normalizing constant and  $\lambda$  is used to determine the origin of  $y$ .

With this probability density function, the regression predictors  $\alpha$  and  $\beta$  can be incorporated as linear combinations to replace the canonical asymmetry factor (i.e.,  $\alpha$ ) and bifurcation factor (i.e.,  $\beta$ ). Note that as a distribution for a limiting stationary stochastic process, this probability density function in Eq. (16.3) is independent from time  $t$ , thus it can be used to model cross-sectional relationship with the advantage to detect and quantify its potential cusp nature comprising both sudden and continuous states. Moreover, the probability density function allows the well-known statistical theory of maximum likelihood to be used for model parameter estimation and statistical inference. R Package "cusp" has been developed to implement this SDECusp (Grasman et al., 2009). This SDECusp model with R package "cusp" is extremely well-suited for use with cross-sectional data. We have used this SDECusp model extensively for research and publications (Chen, Lin, et

al., 2014; Chen et al., 2013; Diks & Wang, 2016; Katerndahl, Burge, Ferrer, Wood, & Becho, 2015; Xu & Chen, 2016; Xu et al., 2017; Yu et al., 2018).

### 16.3.3 Chen-Chen’s Cusp Regression Approach

As the third approach, Chen and Chen (2017) developed a cusp catastrophe nonlinear regression model (“RegCusp”) for continuous data as a conceptual model that is guided by the statistical theory of nonlinear regression models (Seber & Lee, 2003). Following Eq. (16.1), the RegCusp model can be formulated as following:

$$y_i = Y_i + \varepsilon_i, \tag{16.5}$$

where  $y_i$  ( $i = 1, \dots, n$ ) are the observed outcome values and  $\varepsilon_i$  are the residuals from  $n$  observations, and are assumed to be normally distributed as  $\varepsilon_i \sim N(0, \sigma^2)$ .

Mathematically, it can be seen that the latent variable  $Y_i$  in Eq. (16.5) is one of the real roots of the deterministic cusp catastrophe equation:

$$\alpha_i + \beta_i Y_i - Y_i^3 = 0, \tag{16.6}$$

where  $\alpha_i$  and  $\beta_i$  are two control variables which is discussed later in the section of cusp catastrophe conventions. For any observed data with  $p$  independent variables ( $x_1, \dots, x_p$ ) and the outcome variable  $y_i$ , the variables  $\alpha_i$  and  $\beta_i$  are the control variables for  $i$ th subject.

In modeling analysis, these two control variables  $\alpha_i$  and  $\beta_i$  are modeled in a way similarly to the Cobb-Grasman (Cobb & Zacks, 1985; Grasman et al., 2009):

$$\alpha_i = a_0 + a_1 x_{1i} + \dots + a_p x_{pi} = \sum_{j=0}^p a_j x_{ji} \tag{16.7}$$

$$\beta_i = b_0 + b_1 x_{1i} + \dots + b_p x_{pi} = \sum_{j=0}^p b_j x_{ji} \tag{16.8}$$

With the formulations of Eqs. (16.5)–(16.8), a nonlinear regression method can be used to estimate the model parameters of  $\mathbf{a} = (a_0, a_1, \dots, a_p)$ ,  $\mathbf{b} = (b_0, b_1, \dots, b_p)$  from Eqs. (16.7) and (16.8). The model parameters can be estimated using maximum likelihood estimation with the likelihood function formulated as follows:

$$L(\mathbf{a}, \mathbf{b}, \sigma^2 | data) = \left( \frac{1}{\sqrt{2\pi\sigma}} \right)^n \exp \left( -\frac{\sum_{i=1}^n (z_i - Z_i)^2}{2\sigma^2} \right) \tag{16.9}$$



With the likelihood function defined in Eq. (16.10), the theory of likelihood estimation can be readily applied to estimate RegCusp parameters as well as the associated statistical inferences on parameter significance and model selection.

## 16.4 Cusp Catastrophe Modeling of Binary Data

To establish the logistic cusp catastrophe regression model, we start with the binary data structure, then introduce logistic cusp catastrophe regression, conventions and algorithm for parameter estimation, and the method for cusp region estimation.

### 16.4.1 The Binary Data Structure

Suppose data from  $n$  participants are available as  $data = (y_i, x_{1i}, \dots, x_{pi})$  ( $i = 1, \dots, n$ ) where  $y_i$  is observed binary outcome with 0/1 from the  $i$ th participants,  $x_{1i}, \dots, x_{pi}$  are the corresponding  $p$ -independent variables. Then  $y_i$  will be binary distributed as:

$$y_i \sim \text{Binary}(p_i) \quad (16.10)$$

where  $p_i = \Pr(y_i = 1)$  is the probability to observe category 1.

### 16.4.2 The Binary Cusp Catastrophe Model

We make use of the logistic type of regression to model the logit of  $p_i$  to the latent variable  $Y_i$ , such that

$$p_i = \frac{\exp(Y_i)}{1 + \exp(Y_i)} \quad (16.11)$$

is one of the real roots of the deterministic cusp catastrophe equation:

$$\alpha_i + \beta_i Y_i - Y_i^3 = 0, \quad (16.12)$$

where  $\alpha_i$  and  $\beta_i$  are two control variables which is discussed later in the section of cusp catastrophe conventions.

The two control variables of  $\alpha_i$  and  $\beta_i$  are modeled in a way similarly to SDECusp with the linear combination of multiple independent variables in Eqs. (16.7) and (16.8)

### 16.4.3 Maximun Likelihood Estimation

With the formulations of Eqs. (16.10) to (16.12), a maximum likelihood procedure can be developed to estimate the model parameters of  $\mathbf{a} = (a_0, a_1, \dots, a_m)$ ,  $\mathbf{b} = (b_0, b_1, \dots, b_m)$  from Eq. (16.7) and (16.8) as well as the associated statistical inferences on parameter significance and model selection. Based on the theory of maximum likelihood estimation, the parameters of  $\mathbf{a} = (a_0, a_1, \dots, a_m)$ ,  $\mathbf{b} = (b_0, b_1, \dots, b_m)$  are estimated by solving the system of gradient equations and their associated variances can be obtained by the Fisher information matrix or Hessian matrix.

Specifically, we construct the likelihood function from Eq. (16.10) as follows:

$$L(\mathbf{a}, \mathbf{b} | data) = \prod_{i=1}^n p_i^{y_i} (1 - p_i)^{1 - y_i} \quad (16.13)$$

To maximize the likelihood function defined in Eq. (16.13) is equivalent to maximize the log-likelihood function as follows:

$$\begin{aligned} \log L(\mathbf{a}, \mathbf{b} | data) &= \sum_{i=1}^n [y_i \log(p_i) + (1 - y_i) \log(1 - p_i)] \\ &= \sum_{i=1}^n [y_i Y_i + \log(1 - p_i)] \end{aligned} \quad (16.14)$$

### 16.4.4 Cusp Catastrophe Conventions

The cusp catastrophe model is not the traditional statistical model in which each combination of independent variables is associated with one and only one outcome value. In fact, the RegCusp model formulated from Eq. (16.6) and the LogisticCusp model formulated from Eq. (16.12) could have one, two, or three roots for each  $\alpha_i$  and  $\beta_i$  combinations depending on the locations on the control plan, defined by Eqs. (16.7) and (16.8). There three roots can be solved analytically as follows:

$$\begin{aligned} Y_1 &= \frac{1}{6} \frac{\nabla^{2/3} + 12\beta}{\nabla^{1/3}} \quad Y_2 = \frac{1}{12} \frac{\sqrt{3}I\nabla^{2/3} - 12\sqrt{3}I\beta - \nabla^{2/3} - 12\beta}{\nabla^{1/3}}, \quad \text{and} \\ Y_3 &= -\frac{1}{12} \frac{\sqrt{3}I\nabla^{2/3} - 12\sqrt{3}I\beta + \nabla^{2/3} + 12\beta}{\nabla^{1/3}} \end{aligned} \quad (16.15)$$

where  $I = \sqrt{-1}$  as the imaginary unit,  $\nabla = 108\alpha + 12\sqrt{3\Delta}$  and  $\Delta = 27\alpha^2 - 4\beta^3$  is the well-known *Cardan discriminant*. Which one to choose to fit the likelihood function in Eq. (16.13) for the latent variable  $Y$  in Eq. (16.12) would have to be determined using the Cardan discriminant.

From Eq. (16.15), it can be derived that when  $\Delta > 0$ , Eq. (16.12) has one real root; but when  $\Delta \leq 0$ , Eq. (16.12) has three real roots. Among these three roots, there are three cases: (a) if  $\alpha = \beta = \Delta = 0$ , the three roots are the same, which is referred as the *cusp point* (labeled O in Fig. 16.1); (b) if  $\Delta = 0$ , but  $\alpha \neq 0$  or  $\beta \neq 0$ , two roots are the same, which are the two lines OQ and OR forming the boundary for the cusp region (Fig. 16.1); and (c) if  $\Delta < 0$ , and  $\alpha \neq 0$  or  $\beta \neq 0$ , the three roots are distinct, which characterizes the cusp region between OQ and OR also indicated in Fig. 16.1. Therefore, this LogisticCusp model is no longer within the traditional domain of mathematical and statistical modeling. Further investigation is needed to identify the statistical properties of this LogisticCusp model.

To select the correct root for the cusp catastrophe model described by Eq. (16.12), we used two modeling conventions: *delay convention* and *Maxwell convention*. The delay convention is used to select the root from the cusp surface of  $\frac{dV(y;\alpha,\beta)}{dy} = 0$  in Eq. (16.1) that are close to the observed  $y$ . The Maxwell convention is used to select the roots on the cusp surface of  $\frac{dV(y;\alpha,\beta)}{dy} = 0$  in Eq. (16.1) corresponding to the minimum of the associated potential function  $V(y; \alpha, \beta) = \alpha y + \frac{1}{2}\beta y^2 - \frac{1}{4}y^4$ .

### 16.4.5 Cusp Region Estimation

Based on the discussion above, the boundary of the cusp region depicted in Fig. 16.1 can be constructed from  $\Delta = 0$ . Since  $\Delta = 27\alpha^2 - 4\beta^3$ , this can be solved at  $\beta = \sqrt[3]{27\alpha^2/4}$ . Therefore for the asymmetric parameter  $\alpha$  from a range of lower limit (say,  $\alpha_{Lower\ Limit}$ ) to upper limit (say,  $\alpha_{Upper\ Limit}$ ),  $\beta$  can be calculated by at  $\beta = \sqrt[3]{27\alpha^2/4}$  which would correspond to the two lines OQ and OR forming the boundary for the cusp region (Fig. 16.1).

When  $\alpha = \beta = 0$ , then  $\Delta = 0$  which would be the cusp point as commonly referred as the *cusp point* (labeled O in Fig. 16.1). When  $\Delta < 0$ , the values of  $(\alpha, \beta)$  are within the cusp region and when  $\Delta > 0$ , the values of  $(\alpha, \beta)$  are outside the cusp region.

This cusp region under  $(\alpha, \beta)$  coordinate system can be easily transformed into the original data coordinate system of the interest based on the estimated Eqs. (16.7) and (16.8). For example, if the interest is for  $(x_1, x_2)$ , we can plug the estimated Eqs. (16.7) and (16.8) with  $x_1$  and  $x_2$  varying and the other  $x$ s fixed into  $\beta = \sqrt[3]{27\alpha^2/4}$  and solve for  $x_2$  as a function of  $x_1$ . This is illustrated in the real data analysis in Sect. 4.

### 16.4.6 Numeric Search Algorithms for Parameter Estimates

There are several methods to be used to maximize the log-likelihood function in Eq. (16.13). We make use of R function “optim”. The default method is an implementation of that of Nelder and Mead (1965) which uses only function values and is robust but relatively slow. It will work reasonably well for non-differentiable functions. Another commonly-used method is a quasi-Newton method (also known as a variable metric algorithm), specifically that published simultaneously in 1970 by Broyden, Fletcher, Goldfarb and Shanno which is named as BFGS. The BFGS uses function values and gradients for the optimization. Specifically, with the log-likelihood function in Eq. (16.14), the parameters of  $\mathbf{a} = (a_0, a_1, \dots, a_p)$ ,  $\mathbf{b} = (b_0, b_1, \dots, b_p)$  are estimated by solving the system of  $2p + 2$  gradients as:

$$\begin{aligned} & \begin{pmatrix} \frac{\partial \log L}{\partial a} \\ \frac{\partial \log L}{\partial b} \end{pmatrix}_{(2p+2) \times 1} \\ &= \left( \frac{\partial \log L}{\partial a_0}, \frac{\partial \log L}{\partial a_1}, \dots, \frac{\partial \log L}{\partial a_j}, \dots, \frac{\partial \log L}{\partial a_p}, \frac{\partial \log L}{\partial b_0}, \frac{\partial \log L}{\partial b_1}, \dots, \frac{\partial \log L}{\partial b_j}, \dots, \frac{\partial \log L}{\partial b_p} \right)' \\ &= 0 \end{aligned} \tag{16.16}$$

where  $(\cdot)'$  in Eq. (16.16) denotes the vector transpose and the partial derivatives

in Eq. (16.16) can be derived as  $\begin{cases} \frac{\partial \log L}{\partial a_j} = \sum_{i=1}^n \left[ y_i \frac{\partial Y_i}{\partial a_j} - \frac{1}{1-p_i} \frac{\partial p_i}{\partial a_j} \right] \\ \frac{\partial \log L}{\partial b_j} = \sum_{i=1}^n \left[ y_i \frac{\partial Y_i}{\partial b_j} - \frac{1}{1-p_i} \frac{\partial p_i}{\partial b_j} \right] \end{cases}$  for all  $j = 0, 1, \dots, p$ .

In addition, the partial derivatives of  $\frac{\partial Y_i}{\partial a_j}$  and  $\frac{\partial Y_i}{\partial b_j}$  in the gradients can be derived from Eq. (16.12) as  $\frac{\partial Y_i}{\partial a_j} = -\frac{x_{ji}}{\beta_i - 3Y_i^2}$  and  $\frac{\partial Y_i}{\partial b_j} = -\frac{x_{ji} Y_i}{\beta_i - 3Y_i^2}$ . Also the partial derivatives of  $\frac{\partial p_i}{\partial a_j}$  and  $\frac{\partial p_i}{\partial b_j}$  in the gradients can be derived from Eq. (16.11) as  $\frac{\partial p_i}{\partial a_j} = p_i (1 - p_i) \frac{\partial Y_i}{\partial a_j}$  and  $\frac{\partial p_i}{\partial b_j} = p_i (1 - p_i) \frac{\partial Y_i}{\partial b_j}$ .

Equation (16.16) is highly complicated and it's obvious that there are no analytical solutions to solve the  $2p + 2$  gradients from Eq. (16.16) to estimate the  $2p + 2$  parameters of  $\mathbf{a} = (a_0, a_1, \dots, a_p)$  and  $\mathbf{b} = (b_0, b_1, \dots, b_p)$ . Therefore, a numerical iterative search algorithm has to be used to obtain the parameter estimators from Eq. (16.16). We make use of Newton's method (Nocedal & Wright, 1999) to solve Eq. (16.8) iteratively using following iterative scheme with a large number of iterations of  $s = 1, \dots, S$  (i.e.  $S > 1000$ ):

$$\begin{pmatrix} \mathbf{a}^{(s+1)} \\ \mathbf{b}^{(s+1)} \end{pmatrix} = \begin{pmatrix} \mathbf{a}^{(s)} \\ \mathbf{b}^{(s)} \end{pmatrix} - \begin{pmatrix} \frac{\partial^2 \log L}{\partial \mathbf{a}^2}, \frac{\partial^2 \log L}{\partial \mathbf{a} \partial \mathbf{b}} \\ \frac{\partial^2 \log L}{\partial \mathbf{a} \partial \mathbf{b}}, \frac{\partial^2 \log L}{\partial \mathbf{b}^2} \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{a}^{(s)} \\ \mathbf{b}^{(s)} \end{pmatrix} \begin{pmatrix} \frac{\partial \log L}{\partial \mathbf{a}} \\ \frac{\partial \log L}{\partial \mathbf{b}} \end{pmatrix} \begin{pmatrix} \mathbf{a}^{(s)} \\ \mathbf{b}^{(s)} \end{pmatrix} \tag{16.17}$$

Note that in the right side of Eq. (16.17),  $\left(\frac{\partial \log L}{\partial a}, \frac{\partial \log L}{\partial b}\right)$  and  $\left(\frac{\partial^2 \log L}{\partial a^2}, \frac{\partial^2 \log L}{\partial a \partial b}, \frac{\partial^2 \log L}{\partial b^2}\right)$  are the gradient vector in Eq. (16.16) and the Hessian matrix evaluated at the  $s$ th iteration of the parameters  $\begin{pmatrix} a^{(s)} \\ b^{(s)} \end{pmatrix}$ . The Hessian matrix is a  $(2p + 2) \times (2p + 2)$  matrix with its elements of the associated second derivatives. Specifically, in the Hessian matrix,

- The upper-left matrix  $\frac{\partial^2 \log L}{\partial a^2}$  is a  $(p + 1) \times (p + 1)$  matrix with diagonal elements as  $\frac{\partial^2 \log L}{\partial a_j^2} = \sum_{i=1}^n \left[ y_i \frac{\partial^2 Y_i}{\partial a_j^2} - \frac{1}{1-p_i} \frac{\partial^2 p_i}{\partial a_j^2} + \frac{1}{(1-p_i)^2} \left( \frac{\partial p_i}{\partial a_j} \right)^2 \right]$  for all  $j = 1, \dots, p$ , and the off-diagonal elements as  $\frac{\partial^2 \log L}{\partial a_j \partial a_k} = \sum_{i=1}^n \left[ y_i \frac{\partial^2 Y_i}{\partial a_j \partial a_k} - \frac{1}{1-p_i} \frac{\partial^2 p_i}{\partial a_j \partial a_k} + \frac{1}{(1-p_i)^2} \frac{\partial p_i}{\partial a_j} \frac{\partial p_i}{\partial a_k} \right]$  for all  $j, k = 1, \dots, p$  and  $j \neq k$ .
- The upper-right matrix  $\frac{\partial^2 \log L}{\partial a \partial b}$  is the same as the lower-left matrix which is a  $(p + 1) \times (p + 1)$  matrix with elements as  $\frac{\partial^2 \log L}{\partial a_j \partial b_k} = \sum_{i=1}^n \left[ y_i \frac{\partial^2 Y_i}{\partial a_j \partial b_k} - \frac{1}{1-p_i} \frac{\partial^2 p_i}{\partial a_j \partial b_k} + \frac{1}{(1-p_i)^2} \frac{\partial p_i}{\partial a_j} \frac{\partial p_i}{\partial b_k} \right]$  for all  $j, k = 1, \dots, p$ .
- The lower-right matrix  $\frac{\partial^2 \log L}{\partial b^2}$  is a  $(p + 1) \times (p + 1)$  matrix with diagonal elements as  $\frac{\partial^2 \log L}{\partial b_j^2} = \sum_{i=1}^n \left[ y_i \frac{\partial^2 Y_i}{\partial b_j^2} - \frac{1}{1-p_i} \frac{\partial^2 p_i}{\partial b_j^2} + \frac{1}{(1-p_i)^2} \left( \frac{\partial p_i}{\partial b_j} \right)^2 \right]$  for all  $j = 1, \dots, p$ , and the off-diagonal elements as  $\frac{\partial^2 \log L}{\partial b_j \partial b_k} = \sum_{i=1}^n \left[ y_i \frac{\partial^2 Y_i}{\partial b_j \partial b_k} - \frac{1}{1-p_i} \frac{\partial^2 p_i}{\partial b_j \partial b_k} + \frac{1}{(1-p_i)^2} \frac{\partial p_i}{\partial b_j} \frac{\partial p_i}{\partial b_k} \right]$  for all  $j, k = 1, \dots, p$  and  $j \neq k$ .
- In addition, all the second-order derivatives  $\frac{\partial^2 Y_i}{\partial a_j^2}, \frac{\partial^2 Y_i}{\partial a_j \partial a_k}, \frac{\partial^2 Y_i}{\partial b_j^2}, \frac{\partial^2 Y_i}{\partial b_j \partial b_k}, \frac{\partial^2 p_i}{\partial a_j^2}, \frac{\partial^2 p_i}{\partial a_j \partial a_k}, \frac{\partial^2 p_i}{\partial b_j^2}$  and  $\frac{\partial^2 p_i}{\partial b_j \partial b_k}$  in the above calculations of Hessian matrix can be similarly obtained using the first-order derivatives from the calculations in Eq. (16.16).

We name the above estimation process as ‘‘LogisticCusp’’ with respect to the ‘‘RegCusp’’ in Chen and Chen (2017).

## 16.5 Test the Logistic Cusp Catastrophe Model Through Monte-Carlo Simulation

As the first step to examine the logistic cusp regression method described above, we conducted Monte Carlo simulation studies with known parameters.

### 16.5.1 Model Settings for Simulation

To conduct Monte Carlo simulation, surrogate data are generated using Eqs. (16.10) to (16.12) with the number of observations  $n = 300$ . Two (i.e.,  $p = 3$ ) independent variables  $x_1$  and  $x_2$  are simulated independently from the standard normal distribution.

To test whether the novel model can correctly distinguish and determine the model variables, we make use of the true parameters of  $\mathbf{a} = (2, 2, 0)$ ,  $\mathbf{b} = (2, 0, 2)$  from Eqs. (16.3) and (16.4) where  $a_2 = 0$  in Eq. (16.3) to represent the correct model selection of  $x_1$  from Eq. (16.3) and  $b_1 = 0$  to represent the correct model selection of  $x_2$  from Eq. (16.4).

### 16.5.2 Steps of Simulation Study

The simulation is an iterative process, and it was completed in the following seven consecutive steps:

- Step 1: With  $n = 300$ , simulate  $x_1$  and  $x_2$  from the standard normal distribution;
- Step 2: With the true parameters  $\mathbf{a} = (2, 2, 0)$  and  $\mathbf{b} = (2, 0, 2)$  and the  $x_1$  and  $x_2$  from Step 1, calculate  $\alpha_i$  and  $\beta_i$  from Eqs. (16.7) and (16.8);
- Step 3: With the  $\alpha_i$  and  $\beta_i$  from Step 2, solve Eq. (16.12) to obtain  $Y_i$  and select the one root corresponding to the Maxwell convention, or the minimum of the associated potential function  $V(Y_i, \alpha_i, \beta_i)$ ;
- Step 4: With the selected  $Y_i$  from Step 3, generate the outcome variable  $y_i$  using Eq. (16.10);
- Step 5: Using the data generated from Steps 1 through 4, the objective function can be formed to estimate the parameters  $\mathbf{a}$  and  $\mathbf{b}$  based on Eq. (16.13) using maximum likelihood estimation.
- Step 6: Repeated Steps 1 to 5 for a large number of simulations (we used 5000 times) and record the estimated parameters

Following the steps described above, we first investigated the default Nelder and Mead optimization and we found that the estimation from the maximum likelihood is unbiased, but lack of efficiency of the Fisher information matrix for variance estimation. We further investigated the gradients and Hessian matrix from Eqs. (16.16) and (16.17) with the quasi-Newton (BFGS) and we found that BFGS produced very satisfactory variance estimation. As a routine, we run the simulation for 100,000 times to obtain the modeling results.

**Table 16.1** Summary of the result for BFGS from 100,000 simulations

Parameter	True	Mean	Median	ECP
$a_0$	2.0000	2.0412	2.0191	0.8491
$a_1$	2.0000	2.0486	2.0231	0.7649
$a_2$	0.0000	0.0129	0.0170	0.7176
$b_0$	2.0000	2.0199	2.0114	0.8341
$b_1$	0.0000	0.0032	0.0121	0.7721
$b_2$	2.0000	2.0370	2.0180	0.6823

### 16.5.3 Results and Interpretation

Table 16.1 summarizes the main results from the simulation analysis. It can be seen from Table 16.1 that the parameters are estimated unbiased (i.e., the “Mean” and “Median” of the 100,000 estimated parameters are close to the “True” values) and the empirical coverage probabilities (ECP) are very reasonable with more than 70%. We also investigated this BFGS estimation with 200,000 simulations, similar conclusions are found.

Results from the simulation studies indicate that the LogisticCusp performed quit well to estimate the known parameters of  $a$ s and  $b$ s for the asymmetry and the bifurcation control variables, including the intercept and the slope with small differences between the known values and estimates. For example, the true value for  $b_1$  is 2.0000, and the mean estimate is 2.0370.

## 16.6 Modeling Analysis with Real Data: Binge Drinking

We have known from the above simulation studies that the logistic cusp catastrophe regression works well. To further demonstrate the utility of the newly established method, we analyze real data using the logistic cusp regression method that validated from the Monte-Carlo simulations.

### 16.6.1 Data Sources and Variables

Data used for empirical testing were 1122 youth lifetime drinkers derived from the 2015 Monitoring the Future Study: A Continuing Study of American Youth (12th-Grade Survey) (ICPSR 36408, URL: <https://www.icpsr.umich.edu/icpsrweb/ICPSR/studies/36408>). Of the total sample, 48.6% were male, and 50.1% were White and 24.4% were Black, 39.8% less than 18 years of age and 60.2% were older than 18. The response variable in this study is the number of drinks (denoted by “y”) in binge drinking. Based on self-reported data, 848 (75.6%) did not engage in binge drinking in the past month, 130 (11.6%) engaged once, 72 (6.4%) engaged

twice and 72 (6.4%) engaged in three or more times. A binary variable if binge drinking (y/n) (denoted by “ $y_2$ ”) was created for modeling with participants who engaged in binge drinking at least once in the past month as yes; otherwise no.

Perception of alcohol harm was modeled as the asymmetry variable (denoted by “ $x_1$ ”). The variable was measured using responses to the question: “How much do you think people risk harming themselves (physically or in other ways), if they: (1) Take one or two drinks nearly every day? (2) Take four or five drinks nearly every day? (3) Have five or more drinks once or twice each weekend? Answer options to these questions were: 0 (*no risk*), 1 (*slight risk*), 2 (*moderate risk*), 3 (*great risk*). Items were reverse coded and mean scores (range: 0–3) were computed for analysis such that 0 (*most risk or highest level of harm*) and 3 (*least risk or lowest level of harm*). This measure was used in MTF’s research (Johnston, O’Malley, Miech, Bachman, & Schulenberg, 2017) and reported studies indicate perceived harm is a significant predictor of alcohol use in adolescents (Pedersen, Fjaer, & Gray, 2016).

Frequency of drinking in social settings was modeled as the bifurcation variable (denoted by “ $x_2$ ”) based on the responses to the question: “When you used alcohol during the last year, how often did you use it in each of the following situations?” (1) With 1 or 2 other people; and (2) at a party. Answer options to the questions were 0 (*not at all*), 1 (*few times*), 2 (*sometimes*), 3 (*most times*), and 4 (*every time*). The highest frequency (range: 0–4) at either of the two settings was used for modeling analysis. Social setting has been reported as an influential factor for alcohol use in high school and college students (Weitzman, Nelson, & Wechsler, 2003).

### 16.6.2 Modeling Analysis

Modeling analysis was conducted using the R program we developed and used in the simulation studies presented in Sect. 4. For comparison purposes, we analyzed the same data with Cobb-Grasman’s SDECusp and Chen-Chen’s RegCusp. In the modeling analysis the asymmetry variable is the perceive alcohol as less risk, and the bifurcation variable is the social setting for drinking. We consider two types of outcome variable of binge drinking,  $y$ , as continuous variable and  $y_2$ , as the binary variable. Using continuous outcome,  $y$ , we can fit the typical multiple linear regression (“Linear Regression”), the stochastic cusp catastrophe model (“SDECusp”) and the regression cusp (“RegCusp”) catastrophe model. With the binary outcome,  $y_2$ , we can fit the LogisticCusp model in this chapter.

### 16.6.3 Parameter Estimates and Comparison

Results in Table 16.2 summarizes the parameter estimates and their associated standard errors with standardized data on  $y$ ,  $x_1$  and  $x_2$ . Parameter estimates from the linear regression and the three cusp catastrophe modeling methods are all



**Table 16.2** Results from linear and 3 cusp regression modeling methods

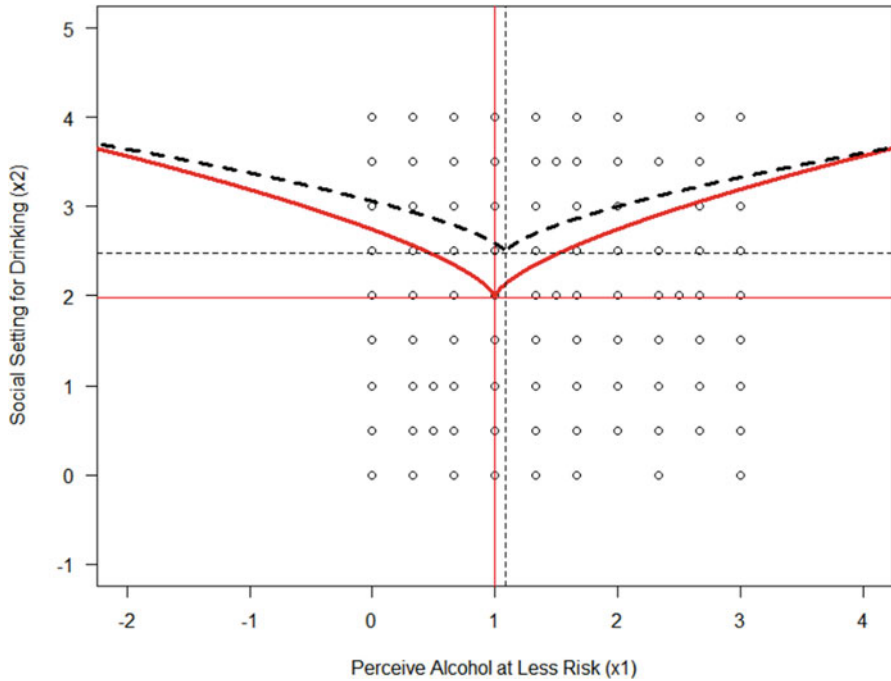
Model	Binge drinking (outcome, $y$ )	Perceived alcohol as less risk ( $x_1$ )	Social setting for drinking ( $x_2$ )	Cusp point estimated
Linear regression	Continuous ( $y$ )	$a_1 = 0.187$ (0.028)***	$b_1 = 0.299$ (0.028)***	NA
SDECusp	Continuous ( $y$ )	$a_0 = -0.595$ (0.031)*** $a_1 = 0.169$ (0.023)***	$b_0 = 3.739$ (0.087)*** $b_1 = -0.221$ (0.043)***	(-1.545, 15.950)
RegCusp	Continuous ( $y$ )	$a_0 = -0.002$ (0.035) $a_1 = 0.132$ (0.034)***	$b_0 = -0.821$ (0.033)*** $b_1 = 0.916$ (0.035)***	(1.082, 2.483)
LogisticCusp	Dichotomous ( $y_2$ )	$a_0 = 0.129$ (0.002)*** $a_1 = 1.252$ (0.037)***	$b_0 = -0.886$ (0.008)*** $b_1 = 2.948$ (0.017)***	(0.996, 1.982)

Note: values in the parenthesis are standard error

statistically highly statistically significant at  $p$ -value  $< 0.001$ , except the  $a_0$  from the RegCusp that is not ( $p > 0.05$ ).

### 16.6.4 Comparison of the Estimated Cusp Regions

With SDECusp, RegCusp and LogisticCusp models, we can estimate the cusp point in the cusp region as denoted by point O in Fig. 16.1. This can be done by setting the estimated  $\alpha$  and  $\beta$  in Eqs. (16.7) and (16.8) to be zero and solving for the corresponding values of  $x_1$  and  $x_2$  which would be the estimated cusp point as described in Section “Cusp Region Estimation”. As seen in Table 16.2, the cusp point is estimated at (-1.545, 15.950) for SDECusp catastrophe model which is out of data region. The estimated cusp point using RegCusp and BinaryCusp are (1.082, 2.483) and (0.996, 1.982), respectively. The cusp point estimated using these two method are reasonable compared to the cusp point estimated with the SDECusp method. The estimated cusp point from the SDECusp was far off the data range of the two predictor variables with  $x_1$  ranging from 0 to 3 and  $x_2$  from 0 to 4. According to the cusp point estimated with the RegCusp, sudden changes in binge drink behavior would occur only when  $x_1$ , the perceived alcohol harm was slightly greater than 1 (somewhat harmful); and  $x_2$ , the frequency of drinking in social settings was about in the middle between 2 (sometimes) and 3 (most times). If results from LogisticCusp is used, the values in the two control variables reduced a bit. Sudden changes in binge drinking would occur when  $x_1$  is approaching 1.0 (perceive alcohol use as “somewhat harmful”) and  $x_2$  is approaching 2 (sometimes)



**Fig. 16.2** Estimated cusp point along with cusp region for both RegCusp (black line) and LogisticCusp (red line) models

drinking in social settings). In another word, with a binary outcome, the estimated sudden change becomes more sensitive than with a continuous outcome.

Figure 16.2 graphically illustrates the estimated cusp points and the associated cusp regions for both RegCusp and LogisticCusp models. As seen in Fig. 16.2, the dashed lines are for RegCusp model where the estimated cusp point is at ((1.082, 2.483) and the solid lines are for LogisticCusp model where the estimated cusp point is at (0.996, 1.982).

### 16.7 Discussion and Conclusions

In this chapter, we report our research in successfully establishing the LogisticCusp method for modeling binary outcome variables. The method is grounded on the well-established logistic regression to solve for high-order cusp catastrophe models. The innovative use of a latent binary variable creates a mathematical bridge linking the deterministic cusp catastrophe with a statistical logistic regression. By application of the log likelihood method and numerical search approach with either Maxell or delayed convention, unbiased parameter estimates can be obtained; and

by application of the bootstrapping, correct model variances can also be estimated. In addition to validation through simulation, we empirically test the method using data from national probability sample of youth with a binary variable for binge drinking. Binary variables are more common than continuous variables in research. The LogisticCusp provide a new and only tool, the first time for researchers to examine challenge questions with binary outcome variables. Binary variables are widely used by researchers in almost all scientific fields in addition to life sciences, psychology and behavioral studies.

There are several advantages with the LogisticCusp method we developed. First, all binary variables suitable for logistic regression can be used for cusp catastrophe modeling to nonlinearity and discreteness of a phenomenon. Second, both the asymmetry and bifurcation variables in a logistic cusp regression can be modeled as either a single or multi-variate variable, greatly enhancing the flexibility for modeling analysis. Research from this and previous analysis (Chen et al., 2019) also indicate adequate validity of the estimated the cusp point, and the corresponding cusp region and the two threshold lines with the LogisticCusp method. In addition to assessing the validity of the estimated parameters, determination of the threshold lines provide important data guiding practice to avoid sudden changes moving toward unfavorable outcomes and to promote sudden changes leading to favorable outcomes. Third, as in other method,  $R^2$  or the variances explained by a cusp model can also be estimated as in the traditional regression analysis, facilitating model comparisons to help determine whether a study variable is nonlinear discrete or linear and continuous. Last, the method can be executed in R, free of charge.

There are a couple of limitations to the LogisticCusp method. Like many statistical methods with numerical search for parameter solutions, the LogisticCusp method is sensitive to initial values. Several measures can be used to help determine initial values: a) Generate initial values using parameter estimates for the same data but using other methods such as linear regression, logistic regression, RegCusp, and SDECusp. (b) Check if the estimated cusp point, cusp region and the two threshold lines are within the data range with a meaningful interpretation. Another limitation is the variance estimation. Like in RegCusp, the estimated variances tended to be too small for LogisticCusp. Despite that the bootstrapping provides as a remedy to this issue, we will conduct further research to understand this issue.

Despite these limitations, the establishment of the regression-based approach, including the LogisticCusp in this study and the RegCusp in our previous studies (Chen & Chen, 2017; Chen, Chen, & Zhang, 2016) provide an innovative and highly needed approach for researchers to solve for a deterministic cusp catastrophe model with a statistical method capable of handling sampling and measurement errors. In addition, the accurate estimation of the cusp point, cusp research and the threshold lines advanced cusp catastrophe modeling from qualitatively detecting the cusp to quantitatively describing cusp catastrophe. It is our anticipation that the application of the regression-based cusp catastrophe modeling methods we established will provide a set of great analytics to advance medical, health, social and behavioral studies.

**Acknowledgments** This research was support in part by National Institute of Health (NIH) Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD, R01HD075635, PIs: Chen X and Chen D).

## References

- Berk, R. A. (2008). *Statistical learning from a regression perspective*. Berlin: Springer. (ISBN 978-0-387-77500-5).
- Chen, D., & Chen, X. (2017). Cusp catastrophe regression and its application in public health and behavioral research. *International Journal of Environmental Research and Public Health*, 14(10), 1220. <https://doi.org/10.3390/ijerph14101220>
- Chen, D., Chen, X., Lin, F., Tang, W., Lio, Y. L., & Guo, T. Y. (2014). Cusp catastrophe polynomial model: Power and sample size estimation. *Open Journal of Statistics*, 4(10), 803–813. <https://doi.org/10.4236/ojs.2014.410076>
- Chen, D., Chen, X., & Zhang, K. (2016). *An exploratory statistical cusp catastrophe model*. Paper presented at the 2016 IEEE International Conference on Data Science and Advanced Analytics Montreal, QC, Canada.
- Chen, D., Lin, F., Chen, X., Tang, W., & Kitzman, H. (2014). Cusp catastrophe model: A nonlinear model for health outcomes in nursing research. *Nursing Research*, 63(3), 211–220. <https://doi.org/10.1097/NNR.0000000000000034>
- Chen, X., & Chen, D. (2015). Cusp catastrophe modeling in medical and health research. In D. G. Chen & J. Wilson (Eds.), *Innovative statistical methods for public health data*. Cham: Springer International.
- Chen, X., & Chen, D. (2019). Cognitive theories, paradigm of quantum behavior change, and cusp catastrophe modeling in social behavioral research. *Journal of the Society for Social Work and Research*, 10(1), 127–159. <https://doi.org/10.1086/701837>
- Chen, X., Lunn, S., Harris, C., Li, X. M., Deveaux, L., Marshall, S., . . . Stanton, B. (2010). Modeling early sexual initiation among young adolescents using quantum and continuous behavior change methods: Implications for HIV prevention. *Nonlinear Dynamics, Psychology, and Life Sciences*, 14(4), 491–509.
- Chen, X., Wang, Y., & Chen, D. (2019). Nonlinear dynamics of binge drinking among U.S. high school students in grade 12: Cusp catastrophe modeling of national survey data. *Nonlinear Dynamics, Psychology, and Life Sciences*, 23(4), 465–490. *Revised submission*.
- Chen, X. G., Stanton, B., Chen, D., & Li, X. M. (2013). Intention to use condom, cusp modeling, and evaluation of an HIV prevention intervention trial. *Nonlinear Dynamics, Psychology, and Life Sciences*, 17(3), 385–403.
- Clair, S. (1998). A cusp catastrophe model for adolescent alcohol use: An empirical test. *Nonlinear Dynamics, Psychology, and Life Sciences*, 2, 217–241.
- Cobb, L. (1981). Parameter-estimation for the cusp catastrophe model. *Behavioral Science*, 26(1), 75–78. <https://doi.org/10.1002/bs.3830260107>
- Cobb, L., & Ragade, R. K. (1978). Applications of catastrophe theory in the behavioral and life sciences. *Behavioral Science*, 23, 291–419.
- Cobb, L., & Watson, B. (1980). Statistical catastrophe theory: An overview. *Mathematical Modelling*, 1(4), 311–317. [https://doi.org/10.1016/0270-0255\(1980\)90041-X](https://doi.org/10.1016/0270-0255(1980)90041-X)
- Cobb, L., & Zacks, S. (1985). Applications of catastrophe-theory for statistical modeling in the biosciences. *Journal of the American Statistical Association*, 80(392), 793–802. <https://doi.org/10.2307/2288534>
- Diks, C., & Wang, J. X. (2016). Can a stochastic cusp catastrophe model explain housing market crashes? *Journal of Economic Dynamics & Control*, 69, 68–88. <https://doi.org/10.1016/j.jedc.2016.05.008>
- Faraway, J. J. (2009). *Linear model with R*. Abingdon: Taylor & Francis.

- Gilmore, R. (1981). *Catastrophe theory for scientists and engineers*. New York: Wiley.
- Grasman, R. P., van der Maas, H. L., & Wagenmakers, E. J. (2009). Fitting the cusp catastrophe in R: A cusp package primer. *Journal of Statistical Software*, 32(8), 1–27.
- Guastello, S. J. (1982). Moderator regression and the cusp catastrophe: Application of two-stage personnel selection, training, therapy, and policy evaluation. *Behavioral Science*, 27(3), 259–272.
- Guastello, S. J. (1989). Catastrophe modeling of the accident process—Evaluation of an accident reduction program using the occupational hazards survey. *Accident Analysis and Prevention*, 21(1), 61–77. [https://doi.org/10.1016/0001-4575\(89\)90049-3](https://doi.org/10.1016/0001-4575(89)90049-3)
- Guastello, S. J., Aruka, Y., Doyle, M., & Smerz, K. E. (2008). Cross-cultural generalizability of a cusp catastrophe model for binge drinking among college students. *Nonlinear Dynamics, Psychology, and Life Sciences*, 12(4), 397–407.
- Guastello, S. J., & Gregson, A. M. (2011). *Nonlinear dynamical systems analysis for the behavioral sciences using real data*. Boca Raton, FL: CRC/Taylor & Francis Group.
- Guastello, S. J., & Lynn, M. (2014). Catastrophe model of the accident process, safety climate, and anxiety. *Nonlinear Dynamics, Psychology, and Life Sciences*, 18(2), 177–198.
- Hartelman, P. A. I., van der Maas, H. L. J., & Molenaar, P. C. M. (1998). Detecting and modelling developmental transitions. *British Journal of Developmental Psychology*, 16, 97–122. <https://doi.org/10.1111/j.2044-835X.1998.tb00751.x>
- Honerkamp, J. (1994). *Stochastic dynamical system: Concepts, numerical methods*. New York, NY: VCH.
- Johnston, L. D., O'Malley, P. M., Miech, R. A., Bachman, J. G., & Schulenberg, J. E. (2017). *Monitoring the Future national survey results on drug use, 1975–2016: Overview, key findings on adolescent drug use*. Ann Arbor: Institute for Social Research, The University of Michigan.
- Katerndahl, D. A., Burge, S. K., Ferrer, R. L., Wood, R., & Becho, J. (2015). Modeling outcomes of partner violence using cusp catastrophe modeling. *Nonlinear Dynamics, Psychology, and Life Sciences*, 19(3), 249–268.
- Mazanov, J., & Byrne, D. G. (2006). A cusp catastrophe model analysis of changes in adolescent substance use: Assessment of behavioural intention as a bifurcation variable. *Nonlinear Dynamics, Psychology, and Life Sciences*, 10(4), 445–470.
- Nelder, J. A., & Mead, R. (1965). A simplex algorithm for function minimization. *Computer Journal*, 7, 308–313.
- Nocedal, J., & Wright, S. J. (1999). *Numerical optimization*. Berlin: Springer. ISBN 0-387-98793-2.
- Pedersen, W., Fjaer, E. G., & Gray, P. (2016). Perceptions of harms associated with tobacco, alcohol, and cannabis among students from the UK and Norway. *Contemporary Drug Problems*, 43, 47–61.
- Seber, G. A., & Lee, A. J. (2003). *Linear regression analysis* (2nd ed.). New York, NY: Wiley-Science.
- Thom, R. (1975). *Structural stability and morphogenesis*. New York, NY: Benjamin-Addison-Wesley.
- Thom, R., & Fowler, D. H. (1975). *Structural stability and morphogenesis: An outline of a general theory of models*. New York, NY: W. A. Benjamin.
- Wagner, C. M. (2010). Predicting nursing turnover with catastrophe theory. *Journal of Advanced Nursing*, 66(9), 2071–2084. <https://doi.org/10.1111/j.1365-2648.2010.05388.x>
- Weitzman, Nelson, T. F., & Wechsler, H. (2003). Taking up binge drinking in college: The influences of person, social group, and environment. *The Journal of Adolescent Health*, 32(1), 26–35.
- White, A. M., Tapert, S., & Shukla, S. D. (2017). Binge drinking, predictors, patterns and consequences. *The Journal of the National Institute on Alcohol Abuse and Alcoholism*, 39(1), e1–e3.
- Witkiewitz, K., van der Maas, H. L. J., Hufford, M. R., & Marlatt, G. A. (2007). Nonnormality and divergence in posttreatment alcohol use: Reexamining the Project MATCH data “another way”. *Journal of Abnormal Psychology*, 116(2), 378–394. <https://doi.org/10.1037/0021-843X.116.2.378>

- Xu, Y., & Chen, X. (2016). Protection motivation theory and cigarette smoking among vocational high school students in China: A cusp catastrophe modeling analysis. *Global Health Research and Policy, 1*, 3. <https://doi.org/10.1186/s41256-016-0004-9>
- Xu, Y., Chen, X., Yu, B., Joseph, V., & Stanton, B. (2017). The effects of self-efficacy in bifurcating the relationship of perceived benefit and cost with condom use among adolescents: A cusp catastrophe modeling analysis. *Journal of Adolescence, 61*, 31–39. <https://doi.org/10.1016/j.adolescence.2017.09.004>
- Yu, B., Chen, X., Stanton, B., Chen, D. D., Xu, Y., & Wang, Y. (2018). Quantum changes in self-efficacy and condom-use intention among youth: A chained cusp catastrophe model. *Journal of Adolescence, 68*, 187–197. <https://doi.org/10.1016/j.adolescence.2018.07.020>
- Zeeman, E. C. (1977). *Catastrophe theory: Selected papers 1972–1077*. Boston, MA: Addison-Wesley.