ASYMPTOTIC PROPERTIES FOR A GENERAL EXTREME-VALUE REGRESSION MODEL

(ESTIMADORES CORRIGIDOS PARA UM MODELO GUMBEL DE REGRESSÃO)

WAGNER BARRETO DE SOUZA

Orientador: Klaus Leite Pinto Vasconcellos

Área de concentração: Estatística Matemática

Dissertação submetida como requerimento parcial para obtenção do grau de Mestre em Estatística pela Universidade Federal de Pernambuco.

Recife, fevereiro de 2009

Souza, Wagner Barreto de

Asymptotic properties for a general extreme-value regression model / Wagner Barreto de Souza - Recife : O Autor, 2009.

51 folhas : il., fig., tab.

Dissertação (mestrado) – Universidade Federal de Pernambuco. CCEN. Estatística, 2009.

Inclui bibliografia e apêndice.

1. Estatística Matemática. I. Título.

519.9 CDD (22.ed.) MEI2009-030

Universidade Federal de Pernambuco
Pós-Graduação em Estatística

17 de fevereiro de 2009
(data)

Nós recomendamos que a dissertação de mestrado de autoria de

Wagner Barreto de Souza

intitulada

"Estimadores Corrigidos para um Modelo Gumbel de Regressão"

seja aceita como cumprimento parcial dos requerimentos para o grau de Mestre em Estatística.

Klaus Leite Pinto Vasconcellos
Coordenador da Pós-Graduação em Estatística

Banca Examinadora:

Klaus Leite Pinto Vasconcellos
Klaus Leite Pinto Vasconcellos orientador

Enrico Antonio Colosimo (UFMG)

Raydonal Ospina M.
Raydonal Ospina

Este documento será anexado à versão final da dissertação.

Agradecimentos

Primeiramente, agradeço à minha mãe, Roseane Barreto de Souza, meu pai, Marcos Batista Nogueira de Souza e minha irmã Michelle Maria Barreto de Souza pelo amor, carinho, educação, incentivo, ajuda... Com certeza, são as pessoas mais importantes da minha vida e devo minha vida a eles. Também agradeço à minha tia Sheila, meu tio Gildo e minha vó Helena por todo carinho e ajuda.

Ao Rodrigo Bernardo, Saulo Roberto, Alexandre Simas e Alessandro Henrique pela amizade desde os primeiros períodos da graduação, ajuda e divertimento. Eles são como irmãos para mim e nunca vou esquecer o que já fizeram por mim. Nunca esquecerei das idas à sinuca (que faziam com que a gente chegasse atrasado nas aulas) e à Passira bar. Também agradeço ao Leonardo José e Ennio Cesar pela amizade e confiança.

À Andrea Rocha, Vanessa dos Santos e Juliana Kátia pela amizade conquistada durante minha graduação, pelas ajudas, apoio e divertimento. Também agradeço a Maíra, Syntia e Rosangela por suas brincadeiras e carinho.

Agradeço à Alice Lemos por tudo que passamos juntos durante o mestrado, amizade, confiança e carinho. Também agradeço-a por suas ajudas, em particular, por ter ajudado-me com a entrega desta dissertação.

Também agradeço aos meus colegas de mestrado, Olga, Cícero, Lídia, Manoel, Juliana e Wilton.

Ao professor Klaus Leite Pinto Vasconcellos por todo aprendizado em Estatística desde a graduação, por ter sido meu orientador de mestrado e por ter dado-me uma boa base em Estatística.

Ao professor Gauss Cordeiro pelos ensinamentos em Estatística, pelas pesquisas, pelo apoio acadêmico e ajudas.

Aos meus professores do ensino fundamental Joana D'arc, Fábio e Jorge por acreditarem em mim e por suas várias ajudas.

Aos funcionários Cícero, Valéria, Cândido e Maurício por toda ajuda e brincadeiras.

À CAPES pelo apoio financeiro e à Universidade Federal de Pernambuco por ter concedido-me uma grande oportunidade de obter os títulos de bacharel e mestre em Estatística.

Por fim, agradeço ao professores Raydonal Ospina e Enrico Colosimo por terem aceitado participar da minha banca de mestrado.

ASYMPTOTIC PROPERTIES FOR A GENERAL EXTREME-VALUE REGRESSION MODEL

Resumo

Nesta dissertação, introduzimos um modelo geral de regressão de valor-extremo e obtemos a partir de Cox e Snell (1968) fórmulas gerais para o viés de segunda ordem das estimativas de máxima verossimilhança (EMV) dos parâmetros. Apresentamos fórmulas que podem ser computadas por meio de regressões lineares ponderadas. Além disso, obtemos a assimetria de ordem $n^{-1/2}$ dos EMVs dos parâmetros usando a fórmula de Bowman e Shenton (1998). Casos especiais deste modelo e um estudo de simulação com os resultados obtidos com o uso da fórmula de Cox e Snell (1968) são apresentados. Um uso prático deste modelo e das fórmulas obtidas para correção de viés é apresentado.

Palavras-chave: Modelo de regressão de valor-extremo; Covariáveis de dispersão; Estimativas de máxima verossimilhança; Correção de viés; Assimetria.

ASYMPTOTIC PROPERTIES FOR A GENERAL EXTREME-VALUE REGRESSION MODEL

Abstract

In this thesis we introduce a general extreme-value regression model and derive Cox and Snell's (1968) general formulae for second-order biases of maximum likelihood estimates (MLEs) of the parameters. We present formulae which can be computed by means of weighted linear regressions. Furthermore, we give the skewness of order $n^{-1/2}$ of the MLEs of the parameters by using Bowman and Shenton's (1998) formula. Special cases of this model and a simulation study with results obtained with the use of Cox and Snell's (1968) formulae are presented. A practical use of this model and of the derived formulae for bias correction is also presented.

Keywords: Extreme-value regression model; Dispersion covariates; Maximum likelihood estimates; Bias correction; Skewness.

Contents

List of Tables

Chapter 1

Introduction

The extreme-value distribution, named after Emil Julius Gumbel (1891-1966), is perhaps the most widely applied statistical distribution for climate modelling and has found some applications in engineering. It is also known as Gumbel distribution and log-Weibull distribution.

The extreme-value distribution is very useful in predicting the chance that an extreme earthquake, flood or other natural disaster will occur. Some applications as accelerated life testing through earthquakes, horse racing, rainfall, queues in supermarkets, sea currents, wind speeds, tracking race records and others are listed in a recent book by Kotz and Nadarajah (2000).

The extreme-value regression model is one of the most common models in analyzing lifetime data. Nelson (1982) and Meeker and Escobar (1998) have discussed this regression model extensively, and also illustrated its application.

In this thesis we have two aims. First, we introduce a general extreme-value regression model. We suppose that the location and dispersion parameters vary across observations through non-linear regression models. The systematic components of the general extreme-value regression model introduced by us follows as in the generalized nonlinear models with dispersion covariates (Cordeiro and Udo, 2008); the latter extends the generalized linear models (GLMs) that were introduced by Nelder and Weddeburn (1972) and the generalized nonlinear models (GNLMs) introduced by Cordeiro and Paula (1989). In contrast with the GNLMs, our model has the dispersion parameter defined by a nonlinear regression structure, as we mentioned earlier. Other similar regression models have been introduced in the literature. Dey et al (1997) introduced the overdispersed generalized linear models (OGLMs) and Cordeiro and Botter (2001) considered an additional regression model for a scale parameter in the OGLMs which is incorporated in the variance function. Jϕrgensen (1987) introduced a general class of models called exponential dispersion models. This class of models contain the GNLMs as a particular case and the location and scale parameters are orthogonal. In contrast to the exponential dispersion models, in the general extreme-value regression

models, the parameters corresponding to location and dispersion are not orthogonal (Cox and Reid, 1987). Ferrari and Cribari-Neto (2004) proposed a class of regression models for modelling rates and proportions assuming the response has a beta distribution. Azzalini and Capitanio (1999) examined further probabilistic properties of the multivariate skew-normal distribution, which extends the class of normal distributions and, in particular, they discussed the skew-normal linear regression models.

Our second aim is to provide formulae for the second-order biases of the maximum likelihood estimates (MLEs) of the parameters of our model. When the sample size is large, the biases of the MLEs are negligible, since, in general, their order is $O(n^{-1})$, while the asymptotic standard errors are of order $O(n^{-1/2})$. However, the bias correction is important when the sample size is small. In the literature, many authors have derived expressions for the second-order biases of the MLEs in a variety of regression models. Box (1971) gives a general expression for the n^{-1} bias in multivariate nonlinear models with known covariance matrices. Young and Bakir (1987) show the usefulness of bias correction in generalized log-gamma regression models, which have as particular case the extreme-value linear regression model. Cordeiro and McCullagh (1991) give general matrix formulae for bias correction in GLMs. Cordeiro and Vasconcellos (1997) obtain general matrix formulae for bias correction in multivariate nonlinear regression models with normal errors. This result is extended in Vasconcellos and Cordeiro (1997) to cover heteroscedastic models, while Cordeiro, Vasconcellos and Santos (1998) derive the second-order bias for univariate nonlinear Student-t regression. Vasconcellos and Cordeiro (2000) provide an expression for the second-order bias in multivariate nonlinear Student-t regression model and show it can be computed from a weighted linear regression. Cordeiro and Botter (2001) and Cordeiro and Uto (2008) derive general formulae for second-order biases of MLEs in OGLMs and GNLMs with dispersion covariates, respectively. Ospina et al (2006) provide the second-order biases of MLEs in beta linear regression models. Recently, Simas et al. (2008) obtained the second-order bias for a general beta regression model.

The rest of the thesis is organized as follows. In Chapter 2 we introduce the model and obtain the score function and Fisher's information matrix. With this, we discuss estimation by maximum likelihood and provide asymptotic confidence intervals for the MLEs of the parameters. In Chapters 3 and 4 formulae for the second-order biases and skewness of the MLEs are derived for the parameters. Some special cases are presented and discussed in Chapter 5 as the extreme-value linear regression model, the extreme-value linear regression with dispersion covariates and the extreme-value nonlinear regression model. Monte Carlo simulations are presented in Section 6.1 of Chapter 6 to evaluate the performance of Cox and Snell's (1968) general formula. An application to a real data set is presented in Section 6.2 of Chapter 6 and we conclude the thesis in Chapter 7.

Introdução

A distribuição de valor extremo, assim denominada por Emil Julius Gumbel (1891-1966), é talvez uma das mais utilizadas para modelar dados climáticos e também tem sido aplicada em alguns problemas de engenharia. Esta distribuição também é conhecida como distribuição Gumbel e distribuição log-Weibull.

A distribuição de valor extremo é muito usada para predizer a chance de que um grande terremoto ou outro desastre natural aconteça. Algumas aplicações como teste de vida acelerado através de testes sismos, corridas de cavalos, chuva, filas nos supermercados, correntes marítimas, velocidade do vento, monitoramento de raça e outros registros estão listadas no livro de Kotz e Nadarajah.

O modelo de regressão de valor extremo é um dos mais comuns em análise de dados de sobrevivência. Nelson (1982) e Meeker e Escobar (1998) discutem este modelo extensivamente e também ilustram sua aplicação.

Nesta dissertação, temos dois objetivos. Primeiro, introduzimos um modelo geral de regressão de valor extremo. Supomos que os parâmetros de locação e escala variam com as observações através de duas estruturas não-lineares de regressão. As componentes sistemáticas do modelo geral de regressão de valor extremo introduzido por nós seguem como nos modelos não-lineares generalizados com covariáveis de dispersão (Cordeiro e Udo, 2008); estes últimos extendem os modelos lineares generalizados (MLGs) que foram introduzidos por Nelder e Weddeburn (1972) e o modelo não linear generalizado (MNLG) introduzido por Cordeiro e Paula (1989). Em contraste com os MNLGs, nosso modelo tem o parâmetro de dispersão definido por uma estrutura não-linear de regressão, como mencionado anteriormente. Outros modelos similares de regressão tem sido introduzidos na literatura. Dey et al. (1997) introduziu o modelo linear generazalido com superdispersão (MLGS) e Cordeiro e Botter (2001) consideraram uma estrutura adicional de regressão para o parâmetro de escala no MLGS que é incorporado na função de variância. Jϕrgensen (1987) introduziu uma classe geral de modelos chamada modelos exponenciais de dispersão. Esta classe de modelos contém os MLGs como caso particular; seus parâmetros de locação e escala são ortogonais. Em contraste com os modelos exponenciais de dispersão, no modelo geral de regressão de valor extremo, os parâmetros de locação e escala não são ortogonais (Cox e Reid, 1987). Ferrari e Cribari-Neto (2004) propuseram uma classe de modelos de regressão para modelar

razões e proporções assumindo que a resposta tem uma distribuição beta. Azzalini e Capitanio (1999) examinaram propriedades probabilísticas da distribuição normal assimétrica multivariada, que extende a distribuição normal multivariada e, em particular, eles discutem o modelo de regressão linear normal assimétrico.

Nosso segundo objetivo é obter fórmulas para o vieses de segunda ordem das estimativas de máxima verossimilhança (EMVs) dos parâmetros do nosso modelo. Quando o tamanho da amostra é grande, os vieses das EMVs são negligíveis, uma vez que que, em geral, sua ordem é $O(n^{-1})$, enquanto os erros padrão assintóticos são de ordem $O(n^{-1/2})$. Como sempre, porém, a correção de viés é importante quando o tamanho da amostra é pequeno. Na literatura, muitos autores tem obtido expressões para os vieses de segunda ordem das EMVs em uma variedade de modelos de regressão. Box (1971) obtem uma expressão geral para o viés n^{-1} em modelos não-lineares multivariados com matrizes de covariância conhecidas. Young e Bakir (1987) mostra a utilidade da correção de viés para o modelo de regressão log-gamma generalizado, que tem como caso particular o modelo linear de regressão de valor extremo. Cordeiro e McCullagh (1991) obtem uma fórmula geral matricial para a correção de viés nos MLGs. Cordeiro e Vasconcellos (1997) forneceram fórmulas matriciais para o viés em modelos não-lineares multivariados com erros seguindo distribuição normal. Este resultado é estendido por Vasconcellos e Cordeiro (1997) para cobrir modelos heteroscedásticos, enquanto Cordeiro, Vasconcellos e Santos (1998) calculam o viés de segunda ordem para o modelo não-linear de regressão t-Student univariado. Vasconcellos e Cordeiro (2000) obtiveram uma expressão para o viés de segunda ordem para o modelo de regressão não-linear multivariado t-Student e mostraram que o viés pode ser computado a partir de uma regressão linear ponderada. Cordeiro e Botter (2001) e Cordeiro e Uto (2008) obtiveram fórmulas gerais para o viés de segunda ordem das EMVs em MLGS e MNLGs com covariáveis de dispersão, respectivamente. Ospina et al. (2006) calcularam o viés de segunda ordem das EMVs para o modelo de regressão linear beta. Recentemente, Simas et al. (2008) obtiveram os vieses de segunda ordem para um modelo geral de regressão beta.

O resto da dissertação está organizado como segue. No Capítulo 2, introduzimos o modelo e obtemos a função score e matriz de informação de Fisher. Com isto, discutimos estimação por máxima verossimilhança e obtemos intervalos de confiança assintóticos para as EMVs dos parâmetros. Nos Capítulos 3 e 4, fórmulas para os vieses de segunda ordem e assimetria das EMVs são obtidas. Alguns casos especiais são apresentados e discutidos no Capítulo 5 como o modelo de regressão linear, linear com covariáveis de dispersão e não-linear de valor extremo. Simulações de Monte Carlo são apresentadas na Seção 6.1 do Capítulo 6 para avaliar o desempenho das fórmulas gerais de Cox e Snell (1968). Uma aplicação a dados reais é apresentada na Seção 6.2 do Capítulo 6 e concluímos a dissertação no Capítulo 7.

Chapter 2

General extreme-value regression model

Resumo

Neste capítulo introduzimos um modelo geral de regressão de valor extremo definindo duas estruturas não-lineares de regressão para os parâmetros de locação e escala. Este modelo generaliza o modelo de regressão linear de valor extremo. Discutimos estimação por máxima verossimilhança e obtemos a matriz de informação de Fisher. Com isso, encontramos a distribuição assintótica dos estimadores de máxima verossimilhança e damos intervalos de confiança assintóticos. Além disso, apresentamos os testes da razão de verossimilhança, Score e Wald para verificar hipóteses no modelo geral de regressão de valor extremo.

2.1 The model

Let Y be a random variable with extreme-value distribution. Then, the probability density function of Y is given by

$$g(y;\mu,\phi) = \phi^{-1}\exp\left(-\frac{y-\mu}{\phi}\right)\exp\left(-\exp\left(-\frac{y-\mu}{\phi}\right)\right), \quad y \in R, \tag{2.1}$$

where $\mu \in R$ and $\phi > 0$ are the location and dispersion parameters, respectively.

The moment generation function of Y is given by $E(e^{tY}) = e^{t\mu}\Gamma(1-\phi t)$, with $|t| < \phi^{-1}$. Hence, we obtain the mean and variance of Y:

$$E(Y) = \mu + \gamma\phi \tag{2.2}$$

and

$$Var(Y) = \frac{\pi^2}{6}\phi^2, \tag{2.3}$$

respectively, where γ is the Euler constant.

Let $Y = (Y_1, \ldots, Y_n)^\top$ be a random sample, where each Y_i has probability density function (pdf) as (2.1) with location parameter μ_i and dispersion parameter ϕ_i. We assume that the components on both parameter vectors $\mu = (\mu_1, \ldots, \mu_n)^\top$ and $\phi = (\phi_1, \ldots, \phi_n)^\top$ vary across observations through nonlinear regression models.

The extreme-value regression model with dispersion covariates is defined by (2.1) and by two systematic components which are parameterized as

$$g_1(\mu) = \eta_1 = f_1(X; \beta), \quad g_2(\phi) = \eta_2 = f_2(Z; \theta), \tag{2.4}$$

where $g_1(\cdot)$ and $g_2(\cdot)$ are known strictly monotonic and twice differentiable link functions that map R and R^+, respectively, $f_1(\cdot; \beta)$ and $f_2(\cdot; \theta)$ are continuously twice differentiable nonlinear functions, $\beta = (\beta_1, \ldots, \beta_p)^\top$ ($\beta \in R^p$) and $\theta = (\theta_1, \ldots, \theta_q)^\top$ ($\theta \in R^q$) are sets of unknown parameters to be estimated and X and Z are $n \times p$ and $n \times q$ matrices with columns representing different covariates and ranks p and q, respectively; X and Z are not necessarily different.

2.2 Estimation and inference

The log-likelihood function of Y with observed value y is

$$\ell \equiv \ell(\beta, \phi) = -\sum_{i=1}^{n} \log(\phi_i) - \sum_{i=1}^{n} \frac{y_i - \mu_i}{\phi_i} - \sum_{i=1}^{n} \exp\left(- \frac{y_i - \mu_i}{\phi_i} \right), \tag{2.5}$$

with μ_i and ϕ_i defined by (2.4).

The score function is defined by $U \equiv U(\beta, \theta) = (\partial \ell / \partial \beta^\top, \partial \ell / \partial \theta^\top)^\top$. Let $Y^* = (Y_1^*, \ldots, Y_n^*)^\top$, $y^* = (y_1^*, \ldots, y_n^*)^\top$, $\mu^* = E(Y^*) = (\mu_1^*, \ldots, \mu_n^*)^\top$ and $v = (v_1, \ldots, v_n)^\top$, where $Y_i^* = \exp(-Y_i/\phi_i)$, $y_i^* = \exp(-y_i/\phi_i)$, $\mu_i^* = \exp(-\mu_i/\phi_i)$ and $v_i = -1 + (y_i - \mu_i)/\phi_i[1 - \exp(-(y_i - \mu_i)/\phi_i)]$, for $i = 1, \ldots, n$. Hence, we have that

$$U_j(\beta, \theta) = \frac{\partial \ell}{\partial \beta_j} = -\sum_{i=1}^{n} (y_i^* - \mu_i^*) \frac{1}{\mu_i^* \phi_i} \frac{d\mu_i}{d\eta_{1i}} \frac{\partial \eta_{1i}}{\partial \beta_j}, \quad j = 1, \ldots, p,$$

and

$$U_J(\beta, \theta) = \frac{\partial \ell}{\partial \theta_J} = \sum_{i=1}^{n} v_i \frac{1}{\phi_i} \frac{d\phi_i}{d\eta_{2i}} \frac{\partial \eta_{2i}}{\partial \theta_J}, \quad J = 1, \ldots, q.$$

In matrix notation, we have that

$$\frac{\partial \ell}{\partial \beta} = \widetilde{X}^\top \Omega^{-1} M_1 (y^* - \mu^*), \qquad \frac{\partial \ell}{\partial \theta} = \widetilde{S}^\top \Phi^{-1} M_2 v, \tag{2.6}$$

where we define the $n \times p$ matrix $\widetilde{X} = (\partial\eta_{1i}/\partial\beta_j)_{i,j}$, the $n \times q$ matrix $\widetilde{S} = (\partial\eta_{2i}/\partial\theta_j)_{i,j}$ and the diagonal matrices $\Omega = -\text{diag}(\mu_1^*\phi_1, \ldots, \mu_n^*\phi_n)$, $\Phi = \text{diag}(\phi_1, \ldots, \phi_n)$ $M_1 = \text{diag}(d\mu_i/d\eta_{1i})$ and $M_2 = \text{diag}(d\phi_i/d\eta_{2i})$.

The maximum likelihood estimators (MLEs) of the parameters β and θ are obtained by solving the nonlinear system $U = 0$ and do not have closed-form. Therefore, nonlinear optimization algorithms such as a Newton algorithm or a quasi-Newton algorithm are needed to find the MLEs of the parameters; for details, see Nocedal and Wright (1999).

We now obtain an expression for Fisher's information matrix. First, we define the quantities $W_{\beta\beta} = \text{diag}\{(d\mu_i/d\eta_{1i})^2/\phi_i^2\}$, $W_{\theta\theta} = \text{diag}\{[\Gamma^{(2)}(2) + 1](d\phi_i/d\eta_{2i})^2/\phi_i^2\}$ and $W_{\beta\theta} = \text{diag}\{(\gamma - 1)(d\mu_i/d\eta_{1i})(d\phi_i/d\eta_{2i})/\phi_i^2\}$. With this, it is shown in Appendix A that the Fisher's information matrix is given by

$$K = K(\beta, \theta) = \begin{pmatrix} K_{\beta\beta} & K_{\beta\theta} \\ K_{\theta\beta} & K_{\theta\theta} \end{pmatrix} = \begin{pmatrix} \widetilde{X}^\top W_{\beta\beta}\widetilde{X} & \widetilde{X}^\top W_{\beta\theta}\widetilde{S} \\ \widetilde{S}^\top W_{\theta\beta}\widetilde{X} & \widetilde{S}^\top W_{\theta\theta}\widetilde{S} \end{pmatrix}.$$

Define the matrices $\mathbb{X}$ and $\widetilde{W}$ with dimensions $2n \times (p+q)$ and $2n \times 2n$, respectively, as $\mathbb{X} = \begin{pmatrix} \widetilde{X} & 0 \\ 0 & \widetilde{S} \end{pmatrix}$ and $\widetilde{W} = \begin{pmatrix} W_{\beta\beta} & W_{\beta\theta} \\ W_{\theta\beta} & W_{\theta\theta} \end{pmatrix}$. Then, we can write the Fisher's information matrix as

$$K = \mathbb{X}^\top \widetilde{W} \mathbb{X}. \tag{2.7}$$

We have $W_{\beta\theta} \neq 0$, thus, the parameters β and θ are not orthogonal in contrast with GLMs (Nelder and Wedderburn, 1972). Let $\tau = (\beta^\top, \theta^\top)^\top$ be the parameter vector. Under usual regularity conditions for maximum likelihood estimation and assuming $J(\tau) = \lim_{n\to\infty} K(\tau)/n$ exists and is not singular, we have that

$$\sqrt{n}(\widehat{\tau} - \tau) \xrightarrow{d} N_{p+q}(0, J^{-1}(\tau)),$$

where $\widehat{\tau}$ is the maximum likelihood estimator of τ and "$\xrightarrow{d}$" denotes convergence in distribution. Hence, denoting τ_r the rth component of τ, it comes $(\widehat{\tau}_r - \tau_r)[K^{rr}(\widehat{\tau})]^{-1/2} \xrightarrow{d} N(0,1)$, where $K^{rr}(\widehat{\tau})$ is the rth diagonal element of $K^{-1}(\widehat{\tau})$. Thus, an asymptotic confidence interval for τ_r is given by $\widehat{\tau}_r \pm z_{1-\alpha/2}[K^{rr}(\widehat{\tau})]^{1/2}$ with asymptotic coverage $100(1-\alpha)\%$, where α is the significance level and z_ξ is the ξ quantile of the standard normal distribution.

We now discuss inference for large sample in the extreme-value regression model with dispersion covariates. Consider the partition $\tau = (\tau_1^\top, \tau_2^\top)^\top$, where τ_1 has dimension g. Suppose we are interested in testing the hypothesis $H_0 : \tau_1 = \tau_1^{(0)}$ (null hypothesis) versus $H_1 : \tau_1 \neq \tau_1^{(0)}$ (alternative hypothesis). The statistic of the likelihood ratio (LR) test is

$$\omega_1 = 2\{\ell(\widehat{\tau}) - \ell(\widetilde{\tau})\},$$

where ℓ is the log-likelihood function defined in (2.5) and $\widetilde{\tau}$ is the maximum likelihood estimator of τ under the null hypothesis.

The score test can also be performed. Let $U_{\tau 1}$ and $K^{\tau_1\tau_1}$ be respectively, the vector and the matrix containing the components of the vector score (2.6) and of the matrix (2.7), corresponding to τ_1. The Rao's statistic is given by

$$\omega_2 = \widetilde{U}_{\tau_1}^{\top}\widetilde{K}^{\tau_1\tau_1}\widetilde{U}_{\tau_1},$$

where tildes indicate that quantities are evaluated under null hypotheses. Another test that can be used is Wald's test. The statistic of the test is

$$\omega_3 = (\widehat{\tau}_1 - \tau_1^{(0)})^{\top} K^{\tau_1^{(0)}\tau_1^{(0)}} (\widehat{\tau}_1 - \tau_1^{(0)}),$$

where $\widehat{\tau}_1$ is the maximum likelihood estimator of τ_1.

Under H_0 and some regularity conditions, we have that $\omega_i \xrightarrow{d} \chi^2_g$, for $i = 1, 2, 3$, where g is the number of estimated parameters under H_1 minus the number of estimated parameters under H_0. Hence, the decision rule for the ith test is as follows. We reject the null hypothesis if $\omega_i > \chi^2_{g,1-\alpha/2}$, where α is the significance level and $\chi^2_{g,\,\xi}$ represents the ξ quantile of the chi-squared distribution with g degrees of freedom.

Chapter 3

Bias correction of the MLEs

Resumo

Neste capítulo calculamos os vieses de segunda ordem das EMVs dos parâmetros para o modelo geral de regressão de valor extremo introduzido por nós, utilizando a fórmula de Cox e Snell (1968). O viés de segunda ordem é dado pela fórmula (3.6). Esta fórmula mostra que o viés pode ser calculado através de uma regressão linear ponderada. Também propomos estatísticas de testes modificadas para teste de hipóteses; essas estatísticas são baseadas nas EMVs corrigidas usando a fórmulas do viés encontrada. Além disso, as fórmulas (3.7) e (3.8) dão os vieses de segunda ordem das EMVs dos parâmetros de locação e escala, respectivamente.

3.1 Formulae for the biases

In this Chapter we obtain an expression for the second-order biases of the MLEs of the parameters of the general extreme-value regression model using Cox and Snell's (1968) general formula. First, we introduce some notation. The partial derivatives of the log-likelihood (2.5) with respect to the unknown parameters of the vectors β and θ are indicated by indices $\{j, l, \ldots\}$ and $\{J, L, \ldots\}$, respectively. Hence, we denote $U_j = \partial\ell/\partial\beta_j$, $U_J = \partial\ell/\partial\theta_J$, $U_{jL} = \partial^2\ell/\partial\beta_j\partial\theta_L$, $U_{jlM} = \partial^3\ell/\partial\beta_j\partial\beta_l\partial\theta_M$ etc. To denote the moments of the partial derivatives above, we use the notation introduced by Lawley (1956): $\kappa_{jl} = E(U_{jl})$, $\kappa_{j,l} = E(U_jU_l)$, $\kappa_{jl,M} = E(U_{jl}U_M)$, etc., where all κ's refer to a total over the sample and are, in general, of order $O(n)$. The derivatives of these moments are denoted as $\kappa_{jl}^{(m)} = \partial\kappa_{jl}/\partial\beta_m$, $\kappa_{jl}^{(M)} = \partial\kappa_{jl}/\partial\theta_M$ etc. Not all the κ's are functionally independent. For example, $\kappa_{jl,m} = \kappa_{jlm} - \kappa_{jl}^{(m)}$ gives the covariance between the first partial derivative of the log-likelihood (2.5) with respect to β_m and mixed partial second derivative with respect to β_j, β_l. Note also that $\kappa_{j,l} = -\kappa_{jl}$ and $\kappa_{L,M} = -\kappa_{LM}$ are typical elements of the information matrices $K_{\beta\beta}$ and $K_{\theta\theta}$ for β and θ, respectively. Let $\kappa^{j,l} = -\kappa^{jl}$ and $\kappa^{J,L} = -\kappa^{JL}$ be the corresponding elements of their inverses $K^{\beta\beta}$ and $K^{\theta\theta}$, which are $O(n^{-1})$.

Cox and Snell's (1968) formula can be used to obtain the second-order bias of the MLE for the ath component of the parameter vector $\widehat{\tau} = (\widehat{\tau}_1, \ldots, \widehat{\tau}_{p+q})^\top = (\widehat{\beta}^\top, \widehat{\theta}^\top)^\top$:

$$
\begin{aligned}
B(\widehat{\tau}_a) \;=\; & \sum_{j,l,m} \kappa^{aj}\kappa^{lm}\left\{\kappa_{jl}^{(m)} - \frac{1}{2}\kappa_{jlm}\right\} + \sum_{J,l,m} \kappa^{aJ}\kappa^{lm}\left\{\kappa_{Jl}^{(m)} - \frac{1}{2}\kappa_{Jlm}\right\} + \\
& \sum_{j,L,m} \kappa^{aj}\kappa^{Lm}\left\{\kappa_{jL}^{(m)} - \frac{1}{2}\kappa_{jLm}\right\} + \sum_{j,l,M} \kappa^{aj}\kappa^{lM}\left\{\kappa_{jl}^{(M)} - \frac{1}{2}\kappa_{jlM}\right\} + \\
& \sum_{J,L,m} \kappa^{aJ}\kappa^{Lm}\left\{\kappa_{JL}^{(m)} - \frac{1}{2}\kappa_{JLm}\right\} + \sum_{J,l,M} \kappa^{aJ}\kappa^{lM}\left\{\kappa_{Jl}^{(M)} - \frac{1}{2}\kappa_{Jlm}\right\} + \\
& \sum_{j,L,M} \kappa^{aj}\kappa^{LM}\left\{\kappa_{jL}^{(M)} - \frac{1}{2}\kappa_{jLM}\right\} + \sum_{J,L,M} \kappa^{aJ}\kappa^{LM}\left\{\kappa_{JL}^{(M)} - \frac{1}{2}\kappa_{JLM}\right\}.
\end{aligned}
\tag{3.1}
$$

The parameters β and θ are not orthogonal, thus, the entries of the matrix $W_{\beta\theta}$ are not all zero. Hence, all terms in (3.1) must be considered, which makes the derivation cumbersome. In Appendix B we calculate all cumulants used in (3.1). After tedious algebra, we obtain an expression for the second-order bias of $\widehat{\beta}$ in matrix form:

$$
\begin{aligned}
B(\widehat{\beta}) \;=\; & K^{\beta\beta}\widetilde{X}^\top\left[W_1 Z_{\beta d} + W_2 D_\beta + (W_3 + W_5) Z_{\beta\theta d} + W_4 D_\theta + W_7 Z_{\theta d}\right] 1_{n\times 1} + \\
& K^{\beta\theta}\widetilde{S}^\top\left[W_3 Z_{\beta d} + W_4 D_\beta + (W_6 + W_7) Z_{\beta\theta d} + W_8 Z_{\theta d} + W_9 D_\theta\right] 1_{n\times 1},
\end{aligned}
\tag{3.2}
$$

where the matrices $W_1, \ldots, W_9$ are defined in Appendix B, $1_{n\times 1}$ denotes a vector $n \times 1$ of ones, $Z_{\beta d} = \mathrm{diag}(\widetilde{X}K^{\beta\beta}\widetilde{X}^\top)$, $Z_{\beta\theta d} = \mathrm{diag}(\widetilde{X}K^{\beta\theta}\widetilde{S}^\top)$, $Z_{\theta d} = \mathrm{diag}(\widetilde{S}K^{\theta\theta}\widetilde{S}^\top)$, $D_\beta = \mathrm{diag}\{d_{1\beta}, \ldots, d_{n\beta}\}$, $D_\theta = \mathrm{diag}\{d_{1\theta}, \ldots, d_{n\theta}\}$ with $d_{i\beta} = \mathrm{trace}\left(\widetilde{\widetilde{X}}_i K^{\beta\beta}\right)$, $d_{i\theta} = \mathrm{trace}\left(\widetilde{\widetilde{S}}_i K^{\theta\theta}\right)$, $\widetilde{\widetilde{X}}_i = (\partial^2\eta_{1i}/\partial\beta_j\partial\beta_l)_{j,l}$ and $\widetilde{\widetilde{S}}_i = (\partial^2\eta_{2i}/\partial\theta_J\partial\theta_L)_{J,L}$, for $i = 1, \ldots, n$.

In a similar way, we have an expression for the second-order bias of $\widehat{\theta}$ in matrix form:

$$
\begin{aligned}
B(\widehat{\theta}) \;=\; & K^{\theta\beta}\widetilde{X}^\top\left[W_1 Z_{\beta d} + W_2 D_\beta + (W_3 + W_5) Z_{\beta\theta d} + W_4 D_\theta + W_7 Z_{\theta d}\right] 1_{n\times 1} + \\
& K^{\theta\theta}\widetilde{S}^\top\left[W_3 Z_{\beta d} + W_4 D_\beta + (W_6 + W_7) Z_{\beta\theta d} + W_8 Z_{\theta d} + W_9 D_\theta\right] 1_{n\times 1}.
\end{aligned}
\tag{3.3}
$$

We define the $(2n \times 1)$-vectors δ_1 and δ_2 as

$$
\delta_1 = \begin{pmatrix} \left[W_1 Z_{\beta d} + (W_3 + W_5) Z_{\beta\theta d} + W_7 Z_{\theta d}\right] 1_{n\times 1} \\ \left[W_3 Z_{\beta d} + (W_6 + W_7) Z_{\beta\theta d} + W_8 Z_{\theta d}\right] 1_{n\times 1} \end{pmatrix}
$$

and

$$
\delta_2 = \begin{pmatrix} \left[W_2 D_\beta + W_4 D_\theta\right] 1_{n\times 1} \\ \left[W_4 D_\beta + W_9 D_\theta\right] 1_{n\times 1} \end{pmatrix},
$$

and the $p\times(p+q)$ upper and $q\times(p+q)$ lower blocks of the matrix $K(\tau)^{-1}$ by $K^{\beta *}=(K^{\beta\beta}K^{\beta\theta})$ and $K^{\theta *}=(K^{\theta\beta}K^{\theta\theta})$, respectively. With these definitions, we can write the second-order biases of $\widehat{\beta}$ and of $\widehat{\theta}$ as

$$B(\widehat{\beta})=K^{\beta *}\mathbb{X}^{\top}(\delta_1+\delta_2) \tag{3.4}$$

and

$$B(\widehat{\theta})=K^{\theta *}\mathbb{X}^{\top}(\delta_1+\delta_2), \tag{3.5}$$

respectively. Hence, from (3.4) and (3.5) we conclude that the second order-bias of the MLE of the joint vector $\tau=(\beta^{\top},\theta^{\top})^{\top}$ has the form

$$B(\widehat{\tau})=K^{-1}\mathbb{X}^{\top}(\delta_1+\delta_2)=(\mathbb{X}^{\top}\widetilde{W}\mathbb{X})^{-1}\mathbb{X}^{\top}(\delta_1+\delta_2).$$

Defining $\epsilon_1=\widetilde{W}^{-1}\delta_1$ and $\epsilon_2=\widetilde{W}^{-1}\delta_2$, it follows

$$B(\widehat{\tau})=(\mathbb{X}^{\top}\widetilde{W}\mathbb{X})^{-1}\mathbb{X}^{\top}\widetilde{W}(\epsilon_1+\epsilon_2). \tag{3.6}$$

Formula (3.6) shows that the second-order bias of $\widehat{\tau}$ is easily obtained as the vector of regression coefficients in the formal linear regression of $\epsilon_1+\epsilon_2$ on the columns of $\mathbb{X}$ with $\widetilde{W}$ as weight matrix. We can express (3.6) as $B(\widehat{\tau})=B_1(\widehat{\tau})+B_2(\widehat{\tau})$, with $B_1(\widehat{\tau})=(\mathbb{X}^{\top}\widetilde{W}\mathbb{X})^{-1}\mathbb{X}^{\top}\widetilde{W}\epsilon_1$ and $B_2(\widehat{\tau})=(\mathbb{X}^{\top}\widetilde{W}\mathbb{X})^{-1}\mathbb{X}^{\top}\widetilde{W}\epsilon_2$. If $\epsilon_2=0$, formula (3.6) gives the second-order bias for the extreme-value linear regression with linear dispersion covariates. Therefore, the bias $B_1(\widehat{\tau})$ and $B_2(\widehat{\tau})$ may be regarded as the linearity and nonlinearity terms in the total bias.

Hence, we can define a corrected estimator $\bar{\tau}$ as $\bar{\tau}=\widehat{\tau}-\widehat{B}(\widehat{\tau})$, where $\widehat{B}(\widehat{\tau})$ denotes the bias (3.6) with the unknown parameters replaced by their MLEs. The corrected estimator $\bar{\tau}$ is expected to have better sampling properties than the uncorrected $\widehat{\tau}$ (see Section 6.1). The asymptotic distribution of $\sqrt{n}(\bar{\tau}-\tau)$ is $N_{p+q}(0,J^{-1}(\tau))$, where, as before, we assume that $J(\tau)=\lim_{n\to\infty}K(\tau)/n$ exists and is nonsingular. Hence, an asymptotic confidence interval for τ_r (the rth element of τ) is given by $\bar{\tau}_r\pm z_{1-\alpha/2}[K^{rr}(\bar{\tau})]^{1/2}$ with asymptotic coverage $100(1-\alpha)\%$, where α is the significance level and z_{ξ} is the ξ quantile of the standard normal distribution.

3.2 Modified statistics of the test

We now propose modified statistics of the test based on the corrected MLEs of the parameters. Consider again the partition $\tau=(\tau_1^{\top},\tau_2^{\top})^{\top}$, where τ_1 has dimension g and suppose we are interested in testing the hypothesis $H_0:\tau_1=\tau_1^{(0)}$ (null hypothesis) versus $H_1:\tau_1\neq\tau_1^{(0)}$ (alternative hypothesis). The modified statistic of the likelihood ratio (LR) test is

$$\omega_1^{*}=2\{\ell(\bar{\tau})-\ell(\bar{\tau}^{*})\},$$

where ℓ is the log-likelihood function defined in (2.5) and $\bar{\tau}$ and $\bar{\tau}^*$ are the corrected MLEs of τ under alternative and null hypothesis, respectively.

In a similar form, we propose the Rao's modified statistic of the test given by

$$\omega_2^* = \bar{U}_{\tau_1}^\top \bar{K}^{\tau_1\tau_1} \bar{U}_{\tau_1},$$

where 'bars' indicate the quantities evaluated under null hypotheses and with the corrected MLEs. As before, $U_{\tau 1}$ and $K^{\tau_1\tau_1}$ are the vector and the matrix containing the components of the vector score (2.6) and of the matrix (2.7) respectively, corresponding to τ_1.

The Wald's modified statistic of the test is

$$\omega_3^* = (\bar{\tau}_1 - \tau_1^{(0)})^\top K^{\tau_1^{(0)}\tau_1^{(0)}} (\bar{\tau}_1 - \tau_1^{(0)}),$$

where $\bar{\tau}_1$ is the corrected MLE of τ_1.

Under H_0 and some regularity conditions, we have that $\omega_i \xrightarrow{d} \chi^2_g$, for $i = 1, 2, 3$. The decision rule for the ith test is as mentioned before. We hope that the three modified tests have better performance than the conventional tests.

3.3 Bias correction of the MLEs of μ and ϕ

We now give matrix formulae for the second-order biases of the MLEs of μ and ϕ. First, expanding the functions $\widehat{\eta}_{1i} = f_1(x_i^\top, \widehat{\beta})$ and $\widehat{\eta}_{2i} = f_2(x_i^\top, \widehat{\theta})$ given in (2.4) in Taylor's series up to second-order around the points β and θ, respectively, we obtain

$$\widehat{\eta}_{1i} - \eta_{1i} = \widetilde{X_i}^\top (\widehat{\beta} - \beta) + \frac{1}{2}(\widehat{\beta} - \beta)^\top \widetilde{\widetilde{X_i}} (\widehat{\beta} - \beta) + o_p(\|\widehat{\beta} - \beta\|^2)$$

and

$$\widehat{\eta}_{2i} - \eta_{2i} = \widetilde{s_i}^\top (\widehat{\theta} - \theta) + \frac{1}{2}(\widehat{\theta} - \theta)^\top \widetilde{\widetilde{S}}_i (\widehat{\theta} - \theta) + o_p(\|\widehat{\theta} - \theta\|^2)$$

where $\widetilde{X_i}$ and $\widetilde{S}_i$ are the ith row of the matrices $\widetilde{X}$ and $\widetilde{S}$, respectively. Hence, the second-order biases of $\widehat{\eta}_1$ and $\widehat{\eta}_2$ in notation matrix are given by

$$B(\widehat{\eta}_1) = \widetilde{X} B(\widehat{\beta}) + \frac{1}{2} D_\beta 1_{n\times 1} \quad \text{and} \quad B(\widehat{\eta}_2) = \widetilde{S} B(\widehat{\theta}) + \frac{1}{2} D_\theta 1_{n\times 1}.$$

We now expand the functions $\widehat{\mu}_i = g_1^{-1}(\widehat{\eta}_{1i})$ and $\widehat{\phi}_i = g_2^{-1}(\widehat{\eta}_{2i})$ in Taylor's series up to second-order around the points η_{1i} and η_{2i}, respectively. With this, it follows that

$$\widehat{\mu}_i - \mu_i = \frac{d\mu_i}{d\eta_{1i}}(\widehat{\eta}_{1i} - \eta_{1i}) + \frac{1}{2}\frac{d^2\mu_i}{d\eta_{1i}^2}(\widehat{\eta}_{1i} - \eta_{1i})^2 + o_p((\widehat{\eta}_{1i} - \eta_{1i})^2)$$

and

$$\widehat{\phi}_i - \phi_i = \frac{d\phi_i}{d\eta_{2i}}(\widehat{\eta}_{2i} - \eta_{2i}) + \frac{1}{2}\frac{d^2\phi_i}{d\eta_{2i}^2}(\widehat{\eta}_{2i} - \eta_{2i})^2 + o_p((\widehat{\eta}_{2i} - \eta_{2i})^2).$$

Hence, we obtain the second-order biases of $\widehat{\mu}_i$ and $\widehat{\phi}_i$:

$$B(\widehat{\mu}_i) = B(\widehat{\eta}_{1i})\frac{d\mu_i}{d\eta_{1i}} + \frac{1}{2}Var(\widehat{\eta}_{1i})\frac{d^2\mu_i}{d\eta_{1i}^2} \quad \text{and} \quad B(\widehat{\phi}_i) = B(\widehat{\eta}_{2i})\frac{d\phi_i}{d\eta_{2i}} + \frac{1}{2}Var(\widehat{\eta}_{2i})\frac{d^2\phi_i}{d\eta_{2i}^2}.$$

Defining the $n \times n$ diagonal matrices $T_1 = \text{diag}\{d^2\mu_i/d\eta_{1i}^2\}$ and $T_2 = \text{diag}\{d^2\phi_i/d\eta_{2i}^2\}$, we provide an expression for the second-order biases of MLEs of μ and ϕ in notation matrix, as it follows:

$$B(\widehat{\mu}) = \frac{1}{2}\left\{M_1[2\widetilde{X}B(\widehat{\beta}) + D_\beta 1_{n\times 1}] + Z_{\beta d}T_1 1_{n\times 1}\right\}$$

and

$$B(\widehat{\theta}) = \frac{1}{2}\left\{M_2[2\widetilde{S}B(\widehat{\theta}) + D_\theta 1_{n\times 1}] + Z_{\theta d}T_2 1_{n\times 1}\right\}.$$

For the extreme-value regression model, using (3.4) and (3.5) we obtain

$$B(\widehat{\mu}) = \frac{1}{2}\left\{M_1[2\widetilde{X}K^{\beta *}\mathbb{X}^\top(\delta_1 + \delta_2) + D_\beta 1_{n\times 1}] + Z_{\beta d}T_1 1_{n\times 1}\right\} \tag{3.7}$$

and

$$B(\widehat{\phi}) = \frac{1}{2}\left\{M_2[2\widetilde{S}K^{\theta *}\mathbb{X}^\top(\delta_1 + \delta_2) + D_\theta 1_{n\times 1}] + Z_{\theta d}T_2 1_{n\times 1}\right\}. \tag{3.8}$$

Therefore, the corrected estimators $\widetilde{\mu} = \widehat{\mu} - \widehat{B}(\widehat{\mu})$ and $\widetilde{\phi} = \widehat{\phi} - \widehat{B}(\widehat{\phi})$ of μ and ϕ, respectively, have biases of order $O(n^{-2})$, where $\widehat{B}(\cdot)$ denotes the MLE of $B(\cdot)$, i.e., the unknown parameters are replaced by their MLEs.

Chapter 4

Skewness

Resumo

Aqui, obtemos fórmulas para o terceiro cumulante das EMVs dos parâmetros usando a fórmula dada por Bowman and Shenton (1998). Os principais resultados deste capítulo são dados pelas fórmulas (4.3) e (4.4). Com isso e com a matriz de informação de Fisher, fórmulas de ordem $n^{-1/2}$ são obtidas para a assimetria das EMVs dos parâmetros.

4.1 Formulae for the skewness

We here give an asymptotic formula of order $n^{-1/2}$, where n is the sample size, for the skewness of the distribution of the MLEs of the parameters in general extreme-value regression model.

The assumption of symmetry plays a crucial role in many statistical models. The measure of skewness most well-known is Pearson's standardized third cumulant defined by $\gamma_1 = \kappa_3/\kappa_2^{3/2}$, where κ_2 and κ_3 are the second and third cumulants of the distribution, respectively; when $\gamma_1 > 0$ ($\gamma_1 < 0$) the distribution is positively (negatively) skewed. If the distribution is symmetrical, γ_1 vanishes; therefore its value will give some indication of departure from symmetry. Further, it has also been suggested and used as a possible measure of nonnormality of the distribution.

Bowman and Shenton (1998) give an asymptotic formula for the skewness of the distribution of the MLE of a parameter and discuss applications related to the MLEs of the parameters of the Poisson, binomial, normal, gamma and beta distributions. Cordeiro and Cordeiro (2001) give an asymptotic formula of order $n^{-1/2}$ for the skewness of the distribution of the maximum likelihood estimates of the linear parameters in generalized linear models. Moreover, they also give asymptotic formulae for the skewness of the distribution of the maximum likelihood estimates of the dispersion and precision parameters.

We use the same notation introduced in Section 3. Let $\kappa_3(\widehat{\tau}_a) = E[(\widehat{\tau}_a - \tau_a)^3]$ be a third cumulant of the MLE $\widehat{\tau}_a$ of τ_a, for $a = 1, \ldots, p+q$. Using the formula found by Bowman and Shenton (1998), we can write to order n^{-2}

$$\begin{aligned}
\kappa_3(\widehat{\tau}_a) &= \sum_{j,l,m} \kappa^{a,j}\kappa^{a,l}\kappa^{a,m}\{\kappa_{j,l,m} + 3\kappa_{jlm} + 6\kappa_{jl,m}\} + \sum_{J,l,m} \kappa^{a,J}\kappa^{a,l}\kappa^{a,m} \times \\
&\{\kappa_{J,l,m} + 3\kappa_{Jlm} + 6\kappa_{Jl,m}\} + \sum_{j,L,m} \kappa^{a,j}\kappa^{a,L}\kappa^{a,m}\{\kappa_{j,L,m} + 3\kappa_{jLm} + 6\kappa_{jL,m}\} + \\
&\sum_{j,l,M} \kappa^{a,j}\kappa^{a,l}\kappa^{a,M}\{\kappa_{j,l,M} + 3\kappa_{jlM} + 6\kappa_{jl,M}\} + \sum_{J,L,m} \kappa^{a,J}\kappa^{a,L}\kappa^{a,m} \times \\
&\{\kappa_{J,L,m} + 3\kappa_{JLm} + 6\kappa_{JL,m}\} + \sum_{J,l,M} \kappa^{a,J}\kappa^{a,l}\kappa^{a,M}\{\kappa_{J,l,M} + 3\kappa_{JlM} + 6\kappa_{Jl,M}\} + \\
&\sum_{j,L,M} \kappa^{a,j}\kappa^{a,L}\kappa^{a,M}\{\kappa_{j,L,M} + 3\kappa_{jLM} + 6\kappa_{JL,M}\} + \sum_{J,L,M} \kappa^{a,J}\kappa^{a,L}\kappa^{a,M} \times \\
&\{\kappa_{J,L,M} + 3\kappa_{JLM} + 6\kappa_{JL,M}\}. \qquad (4.1)
\end{aligned}$$

Plugging the Bartlett identities $\kappa_{j,l,m} = 2\kappa_{jlm} - \kappa_{jl}^{(m)} - \kappa_{jm}^{(l)} - \kappa_{lm}^{(j)}$ and $\kappa_{jl,m} = \kappa_{jl}^{(m)} - \kappa_{jlm}$ in (4.1), we obtain the third cumulant of $\widehat{\tau}_a$ as a function of the cumulants calculated in the previous Section as

$$\begin{aligned}
\kappa_3(\widehat{\tau}_a) &= -\sum_{j,l,m} \kappa^{a,j}\kappa^{a,l}\kappa^{a,m}\{\kappa_{jlm} - 5\kappa_{jl}^{(m)} + \kappa_{jm}^{(l)} + \kappa_{lm}^{(j)}\} - \sum_{J,l,m} \kappa^{a,J}\kappa^{a,l}\kappa^{a,m} \times \\
&\{\kappa_{Jlm} - 5\kappa_{Jl}^{(m)} + \kappa_{Jm}^{(l)} + \kappa_{lm}^{(J)}\} - \sum_{j,L,m} \kappa^{a,j}\kappa^{a,L}\kappa^{a,m}\{\kappa_{jLm} - 5\kappa_{jL}^{(m)} + \kappa_{jm}^{(L)} + \kappa_{Lm}^{(j)}\} \\
&- \sum_{j,l,M} \kappa^{a,j}\kappa^{a,l}\kappa^{a,M}\{\kappa_{jlM} - 5\kappa_{jl}^{(M)} + \kappa_{jM}^{(l)} + \kappa_{lM}^{(j)}\} - \sum_{J,L,m} \kappa^{a,J}\kappa^{a,L}\kappa^{a,m} \times \\
&\{\kappa_{JLm} - 5\kappa_{JL}^{(m)} + \kappa_{Jm}^{(L)} + \kappa_{Lm}^{(J)}\} - \sum_{J,l,M} \kappa^{a,J}\kappa^{a,l}\kappa^{a,M}\{\kappa_{JlM} - 5\kappa_{Jl}^{(M)} + \kappa_{JM}^{(l)} + \kappa_{lM}^{(J)}\} \\
&- \sum_{j,L,M} \kappa^{a,j}\kappa^{a,L}\kappa^{a,M}\{\kappa_{jLM} - 5\kappa_{jL}^{(M)} + \kappa_{jM}^{(L)} + \kappa_{LM}^{(j)}\} - \sum_{J,L,M} \kappa^{a,J}\kappa^{a,L}\kappa^{a,M} \times \\
&\{\kappa_{JLM} - 5\kappa_{JL}^{(M)} + \kappa_{JM}^{(L)} + \kappa_{LM}^{(J)}\}. \qquad (4.2)
\end{aligned}$$

We now define the matrices $B_{\beta\beta} = K^{\beta\beta}\widetilde{X}^\top$, $B_{\beta\theta} = K^{\beta\theta}\widetilde{S}^\top$, $B_{\theta\theta} = K^{\theta\theta}\widetilde{S}^\top$, $B_{\theta\beta} = K^{\theta\beta}\widetilde{S}^\top$, $N_\beta = (h_{\beta 1}, \ldots, h_{\beta n})$, $N_\theta = (h_{\theta 1}, \ldots, h_{\theta n})$, $N_{\beta\theta} = (h_{\beta\theta 1}, \ldots, h_{\beta\theta n})$ and $N_{\theta\beta} = (h_{\theta\beta 1}, \ldots, h_{\theta\beta n})$, with $h_{\beta i} = H_{\beta i}1_{p\times 1}$, $h_{\theta i} = H_{i\theta}1_{q\times 1}$, $h_{\beta\theta i} = H_{i\beta\theta}1_{p\times 1}$, $h_{\theta\beta i} = H_{i\theta\beta}1_{q\times 1}$, $H_{\beta i} = \text{diagonal}\{K^{\beta\beta}\widetilde{\widetilde{X}}_i K^{\beta\beta}\}$, $H_{\theta i} = \text{diagonal}\{K^{\theta\theta}\widetilde{\widetilde{S}}_i K^{\theta\theta}\}$, $H_{\beta\theta i} = \text{diagonal}\{K^{\beta\theta}\widetilde{\widetilde{S}}_i K^{\theta\beta}\}$ and $H_{\theta\beta i} =$
diagonal$\{K^{\theta\beta}\widetilde{\widetilde{X}}_i K^{\beta\theta}\}$, for $i = 1, \ldots, n$. Here, diagonal$\{A\}$ denotes a diagonal matrix with same diagonal as A.

With these definitions and after tedious algebra, which is shown with some details in Appendix C, the third cumulant of $\widehat{\beta}$ is obtained as

$$\begin{aligned}\kappa_3(\widehat{\beta}) &= [B^{(3)}_{\beta\beta}V_1 + B^{(3)}_{\beta\theta}V_8]1_{n\times 1} + B^{(2)}_{\beta\beta}[2V_3 + V_5]B^{\top}_{\beta\theta}1_{p\times 1} + B_{\beta\beta}[V_6 + 2V_7]B^{(2)\top}_{\beta\theta}1_{p\times 1} \\ &\quad +\text{diagonal}\{N_\beta[V_2B^{\top}_{\beta\beta} + V_4B^{\top}_{\beta\theta}] + N_{\beta\theta}[V_4B^{\top}_{\beta\beta} + V_9B^{\top}_{\beta\theta}]\}1_{p\times 1}, \qquad (4.3)\end{aligned}$$

where $V_1, \ldots, V_9$ are defined in the Appendix C, $A^{(2)} = A \odot A$ and $A^{(3)} = A \odot A \odot A$, with $\odot$ denoting the direct product.

In analogous way, we obtain the third cumulant of the MLE $\widehat{\theta}$ of θ as

$$\begin{aligned}\kappa_3(\widehat{\theta}) &= [B^{(3)}_{\theta\beta}V_1 + B^{(3)}_{\theta\theta}V_8]1_{n\times 1} + B^{(2)}_{\theta\beta}[2V_3 + V_5]B^{\top}_{\theta\theta}1_{q\times 1} + B_{\theta\beta}[V_6 + 2V_7]B^{(2)\top}_{\theta\theta}1_{q\times 1} \\ &\quad +\text{diagonal}\{N_{\theta\beta}[V_2B^{\top}_{\theta\beta} + V_4B^{\top}_{\theta\theta}] + N_\theta[V_4B^{\top}_{\theta\beta} + V_9B^{\top}_{\theta\theta}]\}1_{q\times 1}. \qquad (4.4)\end{aligned}$$

From the third cumulants of order n^{-2} of the MLEs of $\widehat{\beta}$ and $\widehat{\theta}$, given by (4.3) and (4.4), and the asymptotic covariance matrix $(\mathbb{X}^{\top}\widetilde{W}\mathbb{X})^{-1}$, we obtain the skewness of order $n^{-1/2}$ of the MLEs of $\widehat{\beta}$ and $\widehat{\theta}$.

Chapter 5

Some special models

Resumo

Neste capítulo, apresentamos alguns casos especiais do modelo geral de regressão de valor extremo. A matriz de informação de Fisher, vieses de segunda ordem e o terceiro cumulante das EMVs são dados para cada caso.

5.1 Extreme-value nonlinear homoskedastic regression model

The extreme-value nonlinear homoskedastic regression model is obtained by taking $g_1(\mu_i) = \eta_{1i} = f_1(x_i^\top, \beta)$ and $\phi_i = \eta_{2i} = \theta$, $i = 1, \ldots, n$ in (2.4). The elements of the score function are given by

$$\frac{\partial \ell}{\partial \beta} = \widetilde{X}^\top \Omega^{-1} M_1 (y^* - \mu^*), \qquad \frac{\partial \ell}{\partial \theta} = \theta^{-1} \sum_{i=1}^{n} v_i,$$

and Fisher's information matrix is given by

$$K = K(\beta, \theta) = \frac{1}{\theta^2} \begin{pmatrix} \widetilde{X}^\top \mathrm{diag}\{(d\mu_i/d\eta_{1i})^2\} \widetilde{X} & (\gamma - 1) \widetilde{X}^\top \mathrm{diag}\{d\mu_i/d\eta_{1i}\} 1_{n\times 1} \\ (\gamma - 1) 1_{n\times 1}^\top \mathrm{diag}\{d\mu_i/d\eta_{1i}\} \widetilde{X} & n[\Gamma^{(2)}(2) + 1] \end{pmatrix}.$$

The second-order biases of $\widehat{\beta}$ and $\widehat{\theta}$ are given by

$$B(\widehat{\beta}) = K^{\beta *} \mathbb{X}^\top (\delta_1 + \delta_2) \quad \text{and} \quad B(\widehat{\theta}) = K^{\theta *} \mathbb{X}^\top (\delta_1 + \delta_2),$$

where $\mathbb{X} = \begin{pmatrix} \widetilde{X} & 0 \\ 0 & 1_{n\times 1} \end{pmatrix}$, $\delta_1 = \begin{pmatrix} [W_1 Z_{\beta d} + (W_3 + W_5) Z_{\beta\theta d} + W_7 Z_{\theta d}]\, 1_{n\times 1} \\ [W_3 Z_{\beta d} + (W_6 + W_7) Z_{\beta\theta d} + W_8 Z_{\theta d}]\, 1_{n\times 1} \end{pmatrix}$ and $\delta_2 = \begin{pmatrix} W_2 D_\beta 1_{n\times 1} \\ W_4 D_\beta 1_{n\times 1} \end{pmatrix}$.

The second-order biases of $\widehat{\mu}$ and $\widehat{\phi}$ are given by

$$B(\widehat{\mu}) = \frac{1}{2}\left\{M_1[2\widetilde{X}K^{\beta *}\mathbb{X}^\top(\delta_1+\delta_2) + D_\beta 1_{n\times 1}] + Z_{\beta d}T_1 1_{n\times 1}\right\}$$

and

$$B(\widehat{\theta}) = M_2 1_{n\times 1}K^{\theta *}\mathbb{X}^\top(\delta_1+\delta_2).$$

The third cumulant of order n^{-2} of the MLEs of β and θ are given by

$$\begin{aligned}\kappa_3(\widehat{\beta}) &= [B^{(3)}_{\beta\beta}V_1 + B^{(3)}_{\beta\theta}V_8]1_{n\times 1} + B^{(2)}_{\beta\beta}[2V_3+V_5]B^\top_{\beta\theta}1_{p\times 1} + B_{\beta\beta}[V_6+2V_7]B^{(2)\top}_{\beta\theta}1_{p\times 1}\\ &\quad +\text{diagonal}\{N_\beta[V_2B^\top_{\beta\beta} + V_4B^\top_{\beta\theta}]\}1_{p\times 1},\end{aligned}$$

and

$$\kappa_3(\widehat{\theta}) = [B^{(3)}_{\theta\beta}V_1 + B^{(3)}_{\theta\theta}V_8]1_{n\times 1} + B^{(2)}_{\theta\beta}[2V_3+V_5]B^\top_{\theta\theta} + B_{\theta\beta}[V_6+2V_7]B^{(2)\top}_{\theta\theta},$$

respectively, with $B_{\beta\theta} = K^{\beta\theta}1_{1\times n}$ and $B_{\theta\theta} = K^{\theta\theta}1_{1\times n}$.

The extreme-value linear homoskedastic regression model is obtained with $f_1(x_i^\top,\beta) = x_i^\top\beta$, $i=1,\ldots,n$. Hence, expressions for score function, Fisher's information matrix, second-order biases and third cumulant of the MLEs of β and θ are given from above formulae with $\widetilde{X}=X$, $\delta_2 = 0_{p+q\times 1}$, $D_\beta = 0_{n\times n}$, $T_1 = 0_{n\times n}$ and $N_\beta = 0_{p\times n}$, where $0_{n_1\times n_2}$ is the matrix of zeros with dimensions $n_1\times n_2$. Further, we have the simplified quantities

$$\omega_{1i} = \frac{1}{2\theta^3}\left(\frac{d\mu_i}{d\eta_{1i}}\right)^3, \quad \omega_{2i} = -\frac{1}{2\theta^2}\left(\frac{d\mu_i}{d\eta_{1i}}\right)^2, \quad \omega_{3i} = -\frac{3-\gamma}{4\theta^3}\left(\frac{d\mu_i}{d\eta_{1i}}\right)^2,$$

$$\omega_{4i} = \frac{1-\gamma}{2\theta^2}\frac{d\mu_i}{d\eta_{1i}}, \quad \omega_{5i} = \frac{\gamma+1}{2\theta^3}\left(\frac{d\mu_i}{d\eta_{1i}}\right)^2, \quad \omega_{6i} = -\frac{1}{2\theta^3}[4(\gamma-1)-\Gamma^{(2)}(2)]\frac{d\mu_i}{d\eta_{1i}},$$

$$\omega_{7i} = \frac{\Gamma^{(2)}(2)}{2\theta^3}\frac{d\mu_i}{d\eta_{1i}}, \quad \omega_{8i} = -\left[\Gamma^{(3)}(2)/2 + \Gamma^{(2)}(2)\right]\theta^{-3}\left(\frac{d\phi_i}{d\eta_{2i}}\right)^3,$$

$$v_{1i} = 2\omega_{1i}, \quad v2i = 6\omega_{2i}, \quad v_{3i} = 4\frac{\gamma-5}{\gamma-3}\omega_{3i}, \quad v_{4i} = 6\omega_{4i}, \quad v_{5i} = 2\frac{\gamma-7}{\gamma+1}\omega_{5i},$$

$$v_{6i} = 2\frac{8(1-\gamma)+\Gamma^{(2)}(2)}{4(1-\gamma)+\Gamma^{(2)}(2)}\omega_{6i}, \quad v_{7i} = 2\frac{\Gamma^{(2)}(2)-4(1-\gamma)}{\Gamma^{(2)}(2)}\omega_{7i}$$

and

$$v_{8i} = \frac{\Gamma^{(3)}(2) - 2}{\Gamma^{(3)}(2)/2 + \Gamma^{(2)}(2)}\omega_{8i},$$

for $i = 1, \ldots, n$. In this model, we obtain the second-order biases of $\widehat{\mu}$ and $\widehat{\phi}$ given by

$$B(\widehat{\mu}) = M_1 X K^{\beta *} \mathbb{X}^\top \delta_1 \text{ and } B(\widehat{\phi}) = K^{\theta *} \mathbb{X}^\top \delta_1.$$

Furthermore, if ϕ is known and $g_1(\mu) = \mu$, it follows that

$$B(\widehat{\beta}) = \frac{\phi}{2}(X^\top X)^{-1} X^\top \text{diagonal}\{X(X^\top X)^{-1} X^\top\} 1_{n\times 1},$$

which agrees with the formula found in Cordeiro and Ramos (2006).

5.2 Extreme-value nonlinear regression model with linear dispersion covariates

We now take $g_1(\mu_i) = \eta_{1i} = f_1(x_i^\top, \beta)$ and $g_2(\phi_i) = \eta_{2i} = z_i^\top \theta$. With this, we have the extreme-value nonlinear regression model with linear dispersion covariates. From (2.6), we have the elements of the score function given by

$$\frac{\partial \ell}{\partial \beta} = \widetilde{X}^\top \Omega^{-1} M_1 (y^* - \mu^*) \quad \text{and} \quad \frac{\partial \ell}{\partial \theta} = S^\top \Phi^{-1} M_2 v.$$

For inference, it is required the Fisher's information matrix, which is $K = \mathbb{X}^\top \widetilde{W} \mathbb{X}$. Here, $\mathbb{X} = \begin{pmatrix} \widetilde{X} & 0 \\ 0 & S \end{pmatrix}$ and $\widetilde{W}$ is defined as before. From (3.6), it follows the expression for the second-order biases of the MLEs:

$$B(\widehat{\tau}) = (\mathbb{X}^\top \widetilde{W} \mathbb{X})^{-1} \mathbb{X}^\top \widetilde{W} (\epsilon_1 + \epsilon_2),$$

with $\epsilon_2 = \widetilde{W}^{-1} \begin{pmatrix} W_2 D_\beta 1_{n\times 1} \\ W_4 D_\beta 1_{n\times 1} \end{pmatrix}$. From (3.7) and (3.8), we obtain

$$B(\widehat{\mu}) = \frac{1}{2}\left\{ M_1 [2\widetilde{X} K^{\beta *} \mathbb{X}^\top (\delta_1 + \delta_2) + D_\beta 1_{n\times 1}] + Z_{\beta d} T_1 1_{n\times 1} \right\}$$

and

$$B(\widehat{\phi}) = M_2 S K^{\theta *} \mathbb{X}^\top (\delta_1 + \delta_2).$$

The third cumulants of order n^{-2} of the MLEs of β and θ are given by

$$\begin{aligned} \kappa_3(\widehat{\beta}) \;=\; & [B^{(3)}_{\beta\beta} V_1 + B^{(3)}_{\beta\theta} V_8] 1_{n\times 1} + B^{(2)}_{\beta\beta} [2V_3 + V_5] B^\top_{\beta\theta} 1_{p\times 1} + B_{\beta\beta} [V_6 + 2V_7] B^{(2)\top}_{\beta\theta} 1_{p\times 1} \\ & + \text{diagonal}\{N_\beta [V_2 B^\top_{\beta\beta} + V_4 B^\top_{\beta\theta}]\} 1_{p\times 1}, \end{aligned}$$

and

$$\kappa_3(\widehat{\theta}) \;=\; [B^{(3)}_{\theta\beta}V_1 + B^{(3)}_{\theta\theta}V_8]1_{n\times 1} + B^{(2)}_{\theta\beta}[2V_3 + V_5]B^{\top}_{\theta\theta}1_{q\times 1} + B_{\theta\beta}[V_6 + 2V_7]B^{(2)\top}_{\theta\theta}1_{q\times 1},$$

respectively.

We obtain the extreme-value linear regression model with linear dispersion covariates by making $f_1(x_i^\top, \beta) = x_i^\top\beta$ and $g_2(\phi_i) = \eta_{2i} = z_i^\top\theta$, for $i = 1, \ldots, n$. Hence, expressions for score function, Fisher's information matrix, second-order biases and third cumulant of the MLEs of β and θ are given from above formulae with $\widetilde{X} = X$, $\delta_2 = 0_{p+q\times 1}$, $D_\beta = 0_{n\times n}$,$T_1 = 0_{n\times n}$ and $N_\beta = 0_{p\times n}$.

For this model, the second-order biases of the MLEs of μ and ϕ are given by

$$B(\widehat{\mu}) = M_1XK^{\beta *}\mathbb{X}^\top\delta_1 \text{ and } B(\widehat{\phi}) = M_2SK^{\theta *}\mathbb{X}^\top\delta_1.$$

Further, we obtain the simplified quantities

$$\omega_{1i} = \frac{1}{2\phi_i^3}\left(\frac{d\mu_i}{d\eta_{1i}}\right)^3, \quad \omega_{2i} = -\frac{1}{2\phi_i^2}\left(\frac{d\mu_i}{d\eta_{1i}}\right)^2, \quad \omega_{3i} = -\frac{3-\gamma}{4\phi_i^3}\left(\frac{d\mu_i}{d\eta_{1i}}\right)^2\frac{d\phi_i}{d\eta_{2i}},$$

$$\omega_{4i} = \frac{1-\gamma}{2\phi_i^2}\frac{d\mu_i}{d\eta_{1i}}\frac{d\phi_i}{d\eta_{2i}}, \quad \omega_{5i} = \frac{\gamma+1}{2\phi_i^3}\left(\frac{d\mu_i}{d\eta_{1i}}\right)^2\frac{d\phi_i}{d\eta_{2i}},$$

$$\omega_{6i} = -\frac{1}{2\phi_i^3}[4(\gamma-1) - \Gamma^{(2)}(2)]\frac{d\mu_i}{d\eta_{1i}}\left(\frac{d\phi_i}{d\eta_{2i}}\right)^2, \quad \omega_{7i} = \frac{\Gamma^{(2)}(2)}{2\phi_i^3}\frac{d\mu_i}{d\eta_{1i}}\left(\frac{d\phi_i}{d\eta_{2i}}\right)^2,$$

$$\omega_{8i} = -\left[\Gamma^{(3)}(2)/2 + \Gamma^{(2)}(2)\right]\phi_i^{-3}\left(\frac{d\phi_i}{d\eta_{2i}}\right)^3, \quad v_{1i} = 2\omega_{1i}, \quad v2i = 6\omega_{2i},$$

$$v_{3i} = 4\frac{\gamma-5}{\gamma-3}\omega_{3i}, \quad v_{4i} = 6\omega_{4i}, \quad v_{5i} = 2\frac{\gamma-7}{\gamma+1}\omega_{5i}, \quad v_{6i} = 2\frac{8(1-\gamma)+\Gamma^{(2)}(2)}{4(1-\gamma)+\Gamma^{(2)}(2)}\omega_{6i},$$

$$v_{7i} = 2\frac{\Gamma^{(2)}(2) - 4(1-\gamma)}{\Gamma^{(2)}(2)}\omega_{7i} \quad \text{and} \quad v_{8i} = \frac{\Gamma^{(3)}(2) - 2}{\Gamma^{(3)}(2)/2 + \Gamma^{(2)}(2)}\omega_{8i},$$

for $i = 1, \ldots, n$.

Chapter 6

Numerical results

Resumo

Neste capítulo apresentamos resultados numéricos utilizando a simulação de Monte Carlo e também uma aplicação do nosso modelo a dados reais. Na seção 6.1 um estudo de simulação foi feito utilizando a simulação de Monte Carlo. O objetivo desta simulação foi comparar os desempenhos das EMVs corrigidas e não-corrigidas dos parâmetros. Usamos um modelo não-linear de regressão de valor extremo com covariáveis de dispersão linear e na tabela 6.1 observamos que as EMVs corrigidas apresentam menor viés e menor erro quadrático médio que as EMVs não-corrigidas. Além disso, as tabelas 6.2 e 6.3 mostram as EMVs corrigidas e não-corrigidas para média e desvio-padrão, respectivamente. Destas tabelas, observamos que a correção de viés é mais relevante para o desvio-padrão; com relação a média, não há muito ganho em se corrigir o viés. Além disso, a estimação intervalar é melhorada usando EMVs corrigidas em relação à estimação intervalar usando EMVs não-corrigidas, como pode ser visto nas tabelas 6.4 and 6.5. Na seção 6.2, ajustamos um modelo não-linear de regressão de valor extremo com covariáveis de dispersão linear para um conjunto de dados analisado por Nelson (1984; 1990. página 272). Apresentamos as EMVs corrigidas e não-corrigidas dos parâmetros; além disso, damos intervalos de confiança com nível de significância de 5%. Também observamos que a hipótese do modelo ser homoscedástico é rejeitada, com isso, mostrando a utilidade das covariáveis de dispersão.

6.1 Simulation

The aim of this Section is to compare the finite sample performances of the uncorrected and corrected MLEs of the parameters of the extreme-value regression model through Monte Carlo simulation. We use the extreme-value regression model with the systematic components

$$\log \mu_i = \beta_1 x_{1i} + e^{\beta_2 x_{2i}} \quad \text{and} \quad \log \phi_i = \theta_1 + \theta_2 x_{1i}, \quad i = 1, \ldots, n.$$

The values of the covariates x_1 and x_2 were obtained as random draws from the $N(-1, 1)$ and $U(0, 1)$, respectively. We use 10000 Monte Carlo replications and all simulations were

carried out using the Ox matrix programming language (Doornik, 2006).

We simulate the regression model with $\beta_1 = 1$, $\beta_2 = -1$, $\theta_1 = -0.7$, $\theta_2 = 0.5$. Table 6.1 shows the MLEs and corrected MLEs with respective mean squared error (MSE) for the sample sizes 20 and 40. We see that the biases of the MLEs were reduced, mainly those relating to dispersion covariates. The MSEs of the MLEs were also reduced with the second-order correction for all cases. We also observe that the uncorrected and corrected MLEs have their biases reduced when $n = 40$ in relation to $n = 20$, which was already expected.

$n = 20$	β_1	β_2	θ_1	θ_2
MLEs	1.0222	−1.0300	−0.89532	0.43765
	(0.022664)	(0.50081)	(0.19462)	(0.087644)
Corrected MLEs	1.0072	−1.0123	−0.73966	0.47669
	(0.021288)	(0.36390)	(0.14255)	(0.073796)
$n = 40$	β_1	β_2	θ_1	θ_2
MLEs	1.0060	−0.99541	−0.77236	0.48221
	(0.0047162)	(0.046229)	(0.040934)	(0.022347)
Corrected MLEs	1.0002	−0.99912	−0.70555	0.49453
	(0.0045484)	(0.045846)	(0.034298)	(0.020585)

Table 6.1: MLEs and corrected MLEs of the parameters for $n = 20, 40$.

Table 6.2 shows $E(Y)$, given by (2.2), its MLE $\widehat{E(Y)}$ and corrected MLE $\widetilde{E(Y)}$, for sample size 20, where Y has extreme-value distribution with the systematic components given previously. The values in parenthesis denote the MSE of the estimatives. We observe that the corrected MLE presents smaller bias and MSE than the original MLE in almost all cases.

The standard deviance of Y, say $sd(Y)$, given by square root of (2.3), its estimator $\widehat{sd(Y)}$ and corrected MLE $\widetilde{sd(Y)}$ are given in Table 6.3, for sample size 20. The values in parenthesis denote the MSE of the estimatives. The corrected MLE reduces the bias and MSE in almost all cases, in comparison to the MLE. Further, the corrected MLE gives values close to the $sd(Y)$, while the MLE does not. From this, we see the need to use the correction.

Tables 6.4 and 6.5 give uncorrected and corrected confidence interval for the MLEs of the parameters and the coverage of the intervals; we use the significance levels $\alpha = 0.01, 0.05, 0.1$. We observe that the corrected confidence interval presents a coverage closer to the true coverage than the uncorrected confidence interval. Further, the coverage of the uncorrected and corrected confidence interval approaches the true coverage when the sample size increases.

$n \rightarrow$	1	2	3	4	5
$E(Y)$	0.86957	3.9682	0.80961	0.35171	0.49377
$\widehat{E(Y)}$	0.85988	4.1385	0.78592	0.34892	0.48305
	(0.011812)	(0.76582)	(0.011488)	(0.0043071)	(0.0069011)
$\widetilde{E(Y)}$	0.86601	4.0312	0.80341	0.35305	0.49242
	(0.011361)	(0.71259)	(0.011271)	(0.0043731)	(0.0069530)
$n \rightarrow$	6	7	8	9	10
$E(Y)$	0.52817	1.0460	0.63938	0.60248	1.6144
$\widehat{E(Y)}$	0.52024	1.0233	0.62909	0.59347	1.5948
	(0.0059510)	(0.013317)	(0.0075229)	(0.0069274)	(0.026123)
$\widetilde{E(Y)}$	0.52666	1.0379	0.63654	0.60023	1.6063
	(0.0058381)	(0.013235)	(0.0073223)	(0.0067230)	(0.027142)
$n \rightarrow$	11	12	13	14	15
$E(Y)$	0.78390	0.35716	1.9231	2.1774	1.0054
$\widehat{E(Y)}$	0.77282	0.35454	1.9022	2.2015	0.99219
	(0.0099204)	(0.0041048)	(0.031911)	(0.096365)	(0.014294)
$\widetilde{E(Y)}$	0.78017	0.35840	1.9147	2.1842	1.0002
	(0.0096023)	(0.0041524)	(0.033716)	(0.095876)	(0.014088)
$n \rightarrow$	16	17	18	19	20
$E(Y)$	0.56643	0.19300	1.9698	0.24531	0.18865
$\widehat{E(Y)}$	0.55266	0.20058	1.9275	0.24936	0.19749
	(0.0076330)	(0.0029065)	(0.014780)	(0.0037163)	(0.0033845)
$\widetilde{E(Y)}$	0.56359	0.19839	1.9584	0.24971	0.19478
	(0.0076149)	(0.0027981)	(0.014002)	(0.0037258)	(0.0032117)

Table 6.2: MLEs and corrected MLEs of $E(Y)$ with its respective MSE for $n = 20$.

6.2 Application

In this Section we present an application to a real data set. The real data set gives low-cycle fatigue life data for a strain-controlled test on 26 cylindrical specimens of a nickel-base superalloy. The data were originally described and analyzed in Nelson (1984; 1990. page 272). Four of the specimens were removed from test before failure. In addition to the number of cycles, each specimen has a level of pseudostress (Young's modulus times strain). The purpose of Nelson's analysis was to estimate the curve giving the number of cycles at which 0.1% of the population of such specimens would fail, as a function of the pseudostress.

We exclude the censored data and use only the uncensored data. The uncensored data for the number of cycles (say y) are 211.629, 200.027, 155.000, 13.949, 152.680, 156.725, 56.723,

$n \rightarrow$	1	2	3	4	5
$sd(Y)$	0.43227	0.92208	0.34879	0.24416	0.27576
$\widehat{sd(Y)}$	0.38286	0.85886	0.31515	0.23553	0.25896
	(0.0097726)	(0.33017)	(0.0048163)	(0.0043465)	(0.0039319)
$\widetilde{sd(Y)}$	0.42371	0.92943	0.34461	0.24791	0.27687
	(0.0087356)	(0.32027)	(0.0044193)	(0.0044971)	(0.0040421)
$n \rightarrow$	6	7	8	9	10
$sd(Y)$	0.32533	0.43194	0.35993	0.35245	0.54282
$\widehat{sd(Y)}$	0.29679	0.38258	0.32397	0.31804	0.47835
	(0.0042313)	(0.0097424)	(0.0052048)	(0.0049358)	(0.026265)
$\widetilde{sd(Y)}$	0.32271	0.42339	0.35505	0.34803	0.53154
	(0.0040304)	(0.0087072)	(0.0047118)	(0.0045072)	(0.025138)
$n \rightarrow$	11	12	13	14	15
$sd(Y)$	0.40419	0.25467	0.58019	0.67539	0.45219
$\widehat{sd(Y)}$	0.35966	0.24327	0.51216	0.60185	0.39957
	(0.0075408)	(0.0041578)	(0.035562)	(0.071879)	(0.011765)
$\widetilde{sd(Y)}$	0.39689	0.25752	0.56881	0.66568	0.44287
	(0.0066873)	(0.0043146)	(0.034660)	(0.071866)	(0.010631)
$n \rightarrow$	16	17	18	19	20
$sd(Y)$	0.30080	0.17127	0.52654	0.18640	0.15678
$\widehat{sd(Y)}$	0.27791	0.18356	0.46386	0.19410	0.17361
	(0.0039325)	(0.0071945)	(0.022894)	(0.0063484)	(0.0081751)
$\widetilde{sd(Y)}$	0.29996	0.18152	0.51543	0.19529	0.16831
	(0.0039113)	(0.0063645)	(0.021708)	(0.0059097)	(0.0068239)

Table 6.3: MLEs and corrected MLEs of $sd(Y)$ with its respective MSE for $n = 20$.

121.075, 112.002, 43.331, 12.076, 13.181, 18.067, 21.300, 15.616, 13.030, 8.489, 12.434, 9.750, 11.865, 6.705, 5.733 and their respective levels of pseudostress (say x) are 80.3, 80.6, 84.3, 85.2, 85.8, 86.4, 87.2, 87.3, 91.3, 99.8, 100.1, 100.5, 113.0, 114.8, 116.4, 118.0, 118.4, 118.6, 120.4, 142.5, 144.5, 145.9.

We assume each number of cycles (y) follows an extreme-value distribution with pdf in the form (2.1) and systematic components parameterized as

$$\mu_i = \beta_0 + \beta_1 e^{\beta_2 x_i} \quad \text{and} \quad \log \phi_i = \theta_0 + \theta_1 x_i,$$

for $i = 1, \ldots, 22$.

Table 6.6 shows the uncorrected and corrected MLEs of the parameters with their respective standard error and asymptotic confidence interval (ACI) at significance level of 5%. We

$(1-\alpha)\times 100\%$; UN	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\theta}_0$	$\hat{\theta}_1$
99%	(0.7698;1.2746)	(−1.9393;−0.1206)	(−1.6162;−0.1744)	(−0.0537;0.9290)
	(0.89670)	(0.88870)	(0.89880)	(0.90570)
95%	(0.8302;1.2143)	(−1.7219;−0.3381)	(−1.4439;−0.3468)	(0.0637;0.8115)
	(0.80640)	(0.82220)	(0.79900)	(0.80790)
90%	(0.8610;1.1834)	(−1.6106;−0.4493)	(−1.3557;−0.4350)	(0.1239;0.7514)
	(0.73360)	(0.76570)	(0.72350)	(0.72540)
$(1-\alpha)\times 100\%$; CO	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\theta}_0$	$\hat{\theta}_1$
99%	(0.7223;1.2921)	(−2.0266;0.0019)	(−1.4605;−0.0188)	(−0.0153;0.9687)
	(0.92970)	(0.92550)	(0.94370)	(0.92870)
95%	(0.7904;1.2240)	(−1.7841;−0.2406)	(−1.2881;−0.1912)	(0.1023;0.8510)
	(0.86340)	(0.87440)	(0.86320)	(0.84450)
90%	(0.8253;1.1891)	(−1.6600;−0.3646)	(−1.2000;−0.2794)	(0.1625;0.7909)
	(0.7970)	(0.8273)	(0.7919)	(0.7707)

Table 6.4: Uncorrected (UN) and corrected (CO) confidence interval for the MLEs of the parameters with its respective coverage below in parenthesis for $n = 20$.

$(1-\alpha)\times 100\%$; UN	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\theta}_0$	$\hat{\theta}_1$
99%	(0.8554;1.1566)	(−1.5148;−0.4760)	(−1.1972;−0.3475)	(0.1590;0.8054)
	(0.96230)	(0.96740)	(0.96020)	(0.96760)
95%	(0.8914;1.1206)	(−1.3906;−0.6002)	(−1.0956;−0.4491)	(0.2363;0.7281)
	(0.89900)	(0.91470)	(0.895100)	(0.90080)
90%	(0.9098;1.1021)	(−1.3271;−0.6637)	(−1.0437;−0.5011)	(0.2758;0.6886)
	(0.83820)	(0.85870)	(0.83300)	(0.83800)
$(1-\alpha)\times 100\%$; CO	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\theta}_0$	$\hat{\theta}_1$
99%	(0.8404;1.1600)	(−1.5542;−0.4431)	(−1.1304;−0.2807)	(0.1712;0.8179)
	(0.97180)	(0.97760)	(0.97720)	(0.97360)
95%	(0.8786;1.1218)	(−1.4215;−0.5767)	(−1.0288;−0.3823)	(0.2485;0.7406)
	(0.92030)	(0.93190)	(0.91950)	(0.91420)
90%	(0.8982;1.1023)	(−1.3536;−0.6446)	(−0.9769;−0.4342)	(0.2880;0.7010)
	(0.86500)	(0.88510)	(0.86270)	(0.85360)

Table 6.5: Uncorrected (UN) and corrected (CO) confidence interval for the MLEs of the parameters with its respective coverage below in parenthesis for $n = 40$.

see that the intercept and the explanatory variable are individually significant at the significance level of 5% for μ and ϕ because the confidence intervals do not contain the value zero. In particular, we reject the hypothesis $H_0 : \theta_1 = 0$ in favour of the hypothesis $H_1 : \theta_1 \neq 0$. With this, we conclude that the model is heteroskedastic and see the usefulness of including dispersion covariates. The length of the corrected ACIs of $\widehat{\beta}_0$, $\widehat{\beta}_1$ and $\widehat{\beta}_2$ increased in relation to the uncorrected, while the length of the corrected ACIs of $\widehat{\theta}_0$ and $\widehat{\theta}_1$ decreased in relation to the uncorrected.

Then, the uncorrected fitted systematic components are given by

$$\widehat{\mu}_i = 5.8789 + 1.2137 \times 10^5 e^{-0.085241 x_i} \quad \text{and} \quad \log \widehat{\phi}_i = 9.7437 - 0.067976 x_i,$$

and the corrected systematic components are given by

$$\widetilde{\mu}_i = 5.7660 + 1.2137 \times 10^5 e^{-0.085132 x_i} \quad \text{and} \quad \log \widetilde{\phi}_i = 9.6198 - 0.065876 x_i,$$

Uncorrected	MLE	Stand. error	ACI(95%)
β_0	5.8789	0.59195	(4.7187;7.0390)
β_1	1.2137×10^5	1.6794×10^{-6}	$1.2137\times10^5 \pm 2.3788 \times 10^{-6}$
β_2	-0.085241	0.0020041	$(-0.089169;-0.081313)$
θ_0	9.7437	0.85846	(8.0612;11.426)
θ_1	-0.067976	0.0079762	$(-0.083609;-0.052343)$
Corrected	MLE	Stand. error	ACI(95%)
β_0	5.7660	0.70163	(4.3908;7.1412)
β_1	1.2137×10^5	2.0682×10^{-6}	$1.2137\times10^5 \pm 4.0536 \times 10^{-6}$
β_2	-0.085132	0.0021645	$(-0.089374;-0.080890)$
θ_0	9.6198	0.85949	(7.9353;11.304)
θ_1	-0.065876	0.0079850	$(-0.08152;-0.05022)$

Table 6.6: Uncorrected and corrected MLEs of the extreme-value nonlinear regression model with linear dispersion covariates for the analysed data set.

for $i = 1, \ldots, 22$.

Figure 6.1 shows graphics of the uncorrected and corrected mean curves for the real data set. From these graphics, we see that both curves fitted well the data. Furthermore, table 6.7 shows the uncorrected and corrected mean and standard deviance. We observe that the corrected mean gives values close to the uncorrected mean, while the corrected standard deviance has a larger difference to the uncorrected standard deviance.

Uncorrected mean					
177.05	172.94	129.73	121.05	115.60	110.42
103.89	103.10	76.335	41.538	40.694	39.600
18.379	16.724	15.439	14.308	14.047	13.920
12.860	7.1339	6.9553	6.8458		
Corrected mean					
176.86	172.79	129.96	121.34	115.93	110.77
104.28	103.50	76.831	41.996	41.147	40.047
18.607	16.923	15.613	14.459	14.193	14.062
12.978	7.0776	6.8914	6.7772		
Uncorrected standard deviance					
93.140	91.260	70.966	66.754	64.086	61.525
58.269	57.874	44.096	24.744	24.244	23.594
10.087	8.9256	8.0057	7.1807	6.9881	6.8937
6.0998	1.3580	1.1853	1.0777		
Corrected standard deviance					
94.172	92.372	72.771	68.658	66.044	63.529
60.320	59.930	46.217	26.511	25.994	25.320
11.082	9.8307	8.8364	7.9414	7.7321	7.6294
6.7642	1.5197	1.3258	1.2048		

Table 6.7: Uncorrected and corrected MLEs of the mean and standard deviance of the fitted model for the real data set.

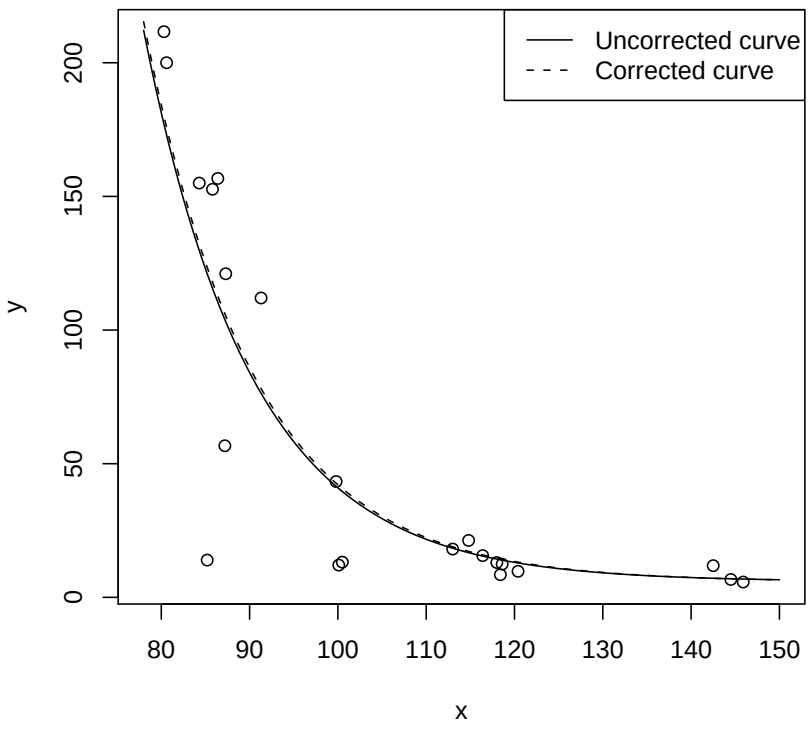

Figure 6.1: Graphics of the uncorrected and corrected mean curves for the real data set.

Chapter 7

Conclusions

In this thesis we introduced a general extreme-value regression model by considering two nonlinear regression structures for the location and scale parameters. We obtain the score function and Fisher's information matrix. With this, we discussed estimation by maximum likelihood and provided asymptotic confidence intervals for the MLEs of the parameters. We also obtained the second-order biases and skewness of the MLEs of the parameters, which were the main results of this thesis. Further, we proposed modified statistics of tests based on the corrected MLEs.

Furthermore, we used a nonlinear extreme-value regression model with linear dispersion covariates and, through Monte Carlo simulation, we observe the usefulness of the bias correction mainly for the standard deviance, as can be seen in Section 6.1. Further, interval estimation is better by using the corrected MLEs than uncorrected MLEs of the parameters.

An application to a real data was also presented. The data set gives low-cycle fatigue life data for a strain-controlled test on 26 cylindrical specimens of a nickel-base superalloy. The data were originally described and analyzed in Nelson (1984; 1990. page 272). We excluded the censored observations and fitted a nonlinear extreme-value regression model with linear dispersion covariates and observed the good fit of the uncorrected and corrected curves for the mean. The hypothesis of the model to be homoskedastic were rejected, showing the usefulness of the dispersion covariates. The fitting of the corrected standard deviance shows that the original MLE of the standard deviance was underestimating it, as happened in the simulation study.

Conclusões

Nesta dissertação, introduzimos um modelo geral de regressão de valor extremo considerando duas estruturas não-lineares de regressão para os parâmetros de locação e escala. Obtemos a função score e matriz de informação de Fisher. Com isto, discutimos estimação por máxima verossimilhança e obtemos intervalos de confiança assintóticos para as EMVs dos parâmetros. Também obtemos os vieses de segunda ordem e assimetria das EMVs dos parâmetros, que foram os principais resultados desta dissertação. Além disso, propusemos estatísticas de teste modificadas baseadas nas EMVs corrigidas.

Usamos um modelo não-linear de regressão de valor extremo com covariáveis de dispersão linear e, através de simulação de Monte Carlo, observamos a utilidade da correção de viés, principalmente para o desvio-padrão, como pode ser visto na seção 6.1. Além disso, a estimação intervalar é melhorada usando EMVs corrigidas em relação à estimação intervalar usando EMVs não-corrigidas.

Uma aplicação a dados reais foi também apresentada. Os dados foram originalmente descritos e analisados por Nelson (1984; 1990. página 272). Excluímos as observações censuradas e ajustamos um modelo não-linear de regressão com covariáveis de dispersão linear e observamos o bom ajuste das curvas para a média corrigida e não-corrigida. A hipótese do modelo ser homoscedástico foi rejeitada, mostrando assim a utilidade das covariáveis de dispersão. O desvio-padrão corrigido mostra que a EMV original subestimou o desvio-padrão, como aconteceu no estudo de simulação.

Bibliography

[1] Azzalini, A., Capitanio, A. (1999). Statistical applications of the multivariate skew-normal distribution. *Journal of the Royal Statistical Society B*, 61, 579–602.

[2] Bowman, K.O., Shenton, L.R. (1998). Asymptotic Skewness and the Distribution of Maximum Likelihood Estimators. *Communications in Statistics. Theory and Methods*, 27, 2743–2760.

[3] Box, M. (1971). Bias in nonlinear estimation (with discussion). *Journal of the Royal Statistical Society B*, 33, 171–201.

[4] Cordeiro, G.M., Botter, D.A. (2001). Second-order biases of maximum likelihood estimates in overdispersed generalized linear models. *Statistics and Probability Letters*, 55, 269–280.

[5] Cordeiro, G.M., McCullagh, P., (1991). Bias correction in generalized linear models. *Journal of the Royal Statistical Society B*, 53, 629-643.

[6] Cordeiro, G.M., Paula, G.A., (1989). Improved likelihood ratio statistics for exponential family nonlinear models. *Biometrika*, 76, 93-100.

[7] Cordeiro, G.M., Ramos, E.M.L.S., (2006). Some second-order asymptotics for extreme value linear regression models. *Communications in Statistics. Theory and Methods*, 35, 197-207.

[8] Cordeiro, G.M., Udo, M.C.T. (2008). Bias correction in generalized nonlinear models with dispersion covariates. *Communications in Statistics. Theory and Methods*, 37, 2219-2225.

[9] Cordeiro, G.M., Vasconcellos, K.L.P. (1997). Bias Correction For A Class Of Multivariate Nonlinear Regression Models. *Statistics and Probability Letters*, 35, 155–164.

[10] Cordeiro, G.M., Vasconcellos, K.L.P., Santos, M.L. (1998). On The Second-Order Bias Of Parameter Estimates In Nonlinear Regression Models With Student T Errors. *Journal of Statistical Computation and Simulation*, 60, 363–378.

[11] Cordeiro, H.H., Cordeiro, G.M. (2001). Skewness for Parameters in Generalized Linear Models. *Communications in Statistics. Theory and Methods*, 30, 1317–1334.

[12] Cox, D.R., Reid, N. (1987). Parameter orthogonality and approximate conditional inference. *Journal of the Royal Statistical Society B*, 49, 1–39.

[13] Cox, D.R., Snell, E.J. (1968). A general definition of residuals. *Journal of the Royal Statistical Society B*, 30, 248–275.

[14] Dey, D.K., Gelfand, A.E., Peng, F. (1997). Overdispersed generalized linear models. *Journal of Statistical Planning and and inference*, 64, 93–107.

[15] Doornik, J.A. (2006). An Object-Oriented Matrix Language - Ox 4. Timberlake Consultants Press, London. 5th ed.

[16] Ferrari, S. L. P., Cribari-Neto, F. (2004). Beta regression for modeling rates and proportions. *J. Appl. Statist.*, 31, 799-815.

[17] Jϕrgensen, B. (1987) Exponential dispersion models (with discussion). *Journal of the Royal Statistical Society B*, 49, 127-162.

[18] Kotz, S., Nadarajah, S. (2000). Extreme Value Distributions: Theory and Applications. Imperial College Press.

[19] Lawley , D.N. (1956). A general method for approximating to the distribution of the likelihood ratio criteria. *Biometrika*, 71, 233-244.

[20] Meeker, W.Q., Escobar, L.A. (1998). Statistical Methods for Reliability Data. John Wiley & Sons, New York.

[21] Nelder, J.A., Wedderburn, R.W.M. (1972). Generalized Linear Models. *J. Roy. Statist. Soc. A*, 135, 370-384.

[22] Nelson, W. (1982). Applied Life Data Analysis. John Wiley & Sons, New York.

[23] Nelson, W. (1985). Weibull analysis of reliability data with few or no failures. *Journal of Quality Technology*, 17, 140-146.

[24] Nelson, W. (1990). Accelerated Testing: Statistical Models. Test Plans, and Data Analyses. John Wiley & Sons, New York.

[25] Nocedal, J., Wright, S.J. (1999). Numerical optimization. Springer.

[26] Ospina, R., Cribari-Neto, F., Vasconcellos, K.L.P., (2006). Improved point and interval estimation for a beta regression model. *Computational Statistics & Data Analysis*, 51, 960–981.

[27] Simas, A.B., Barreto-Souza, W., Rocha, A.V. (2008). Improved estimators for a general class of beta regression models. (Submitted for publication. Preprint: arXiv:0809.1878v1).

[28] Vasconcellos, K.L.P., Cordeiro, G.M. (1997). Approximate Bias For Multivariate Non-linear Heteroscedastic Regressions. *Brazilian Journal of Probability and Statistics*, 11, 141–159.

[29] Vasconcellos, K.L.P., Cordeiro, G.M. (2000). Bias Corrected Estimates in Multivariate Student-t Regression Models. *Communications in Statistics B*, 29, 797–822.

[30] Young, D., Bakir, S. (1987). Bias correction for a generalized log-gamma regression model, *Technometrics*, 29, 183–191.

Appendix A

We now derive Fisher's information matrix. The second-order derivatives of (2.5) with respect to the parameters β and ϕ are given by

$$\begin{aligned} U_{j,l} = \frac{\partial^2 \ell}{\partial \beta_j \partial \beta_l} &= -\sum_{i=1}^{n} \frac{1}{\phi_i^2} \exp\left(-\frac{y_i-\mu_i}{\phi_i}\right)\left(\frac{d\mu_i}{d\eta_{1i}}\right)^2 \frac{\partial \eta_{1i}}{\partial \beta_j}\frac{\partial \eta_{1i}}{\partial \beta_l} + \\ &\quad \sum_{i=1}^{n} \frac{1}{\phi_i}\left[1-\exp\left(-\frac{y_i-\mu_i}{\phi_i}\right)\right]\left[\frac{d^2\mu_i}{d\eta_{1i}^2}\frac{\partial \eta_{1i}}{\partial \beta_j}\frac{\partial \eta_{1i}}{\partial \beta_l} + \frac{d\mu_i}{d\eta_{1i}}\frac{\partial^2 \eta_{1i}}{\partial \beta_j \partial \beta_l}\right], \end{aligned} \tag{7.1}$$

$$\begin{aligned} U_{J,L} = \frac{\partial^2 \ell}{\partial \theta_J \partial \theta_L} &= -\sum_{i=1}^{n} \frac{1}{\phi_i^2}\frac{y_i-\mu_i}{\phi_i}\left[1+\exp\left(-\frac{y_i-\mu_i}{\phi_i}\right)\left(\frac{y_i-\mu_i}{\phi_i}-1\right)\right]\left(\frac{d\phi_i}{d\eta_{2i}}\right)^2 \times \\ &\quad \frac{\partial \eta_{2i}}{\partial \theta_J}\frac{\partial \eta_{2i}}{\partial \theta_L} + \sum_{i=1}^{n}\frac{1}{\phi_i}\left\{-1+\frac{y_i-\mu_i}{\phi_i}-\exp\left(-\frac{y_i-\mu_i}{\phi_i}\right)\frac{y_i-\mu_i}{\phi_i}\right\} \times \\ &\quad \left\{\left[\frac{d^2\phi_i}{d\eta_{2i}^2}-\frac{1}{\phi_i}\left(\frac{d\phi_i}{d\eta_{2i}}\right)^2\right]\frac{\partial \eta_{2i}}{\partial \theta_J}\frac{\partial \eta_{2i}}{\partial \theta_L} + \frac{d\phi_i}{d\eta_{2i}}\frac{\partial^2 \eta_{2i}}{\partial \theta_J \partial \theta_L}\right\} \end{aligned} \tag{7.2}$$

and

$$\begin{aligned} U_{j,L} = \frac{\partial^2 \ell}{\partial \beta_j \partial \theta_L} &= -\sum_{i=1}^{n}\frac{1}{\phi_i^2}\exp\left(-\frac{y_i-\mu_i}{\phi_i}\right)\frac{y_i-\mu_i}{\phi_i}\frac{d\mu_i}{d\eta_{1i}}\frac{d\phi_i}{d\eta_{2i}}\frac{\partial \eta_{1i}}{\partial \beta_j}\frac{\partial \eta_{2i}}{\partial \theta_L} - \\ &\quad \sum_{i=1}^{n}\frac{1}{\phi_i^2}\left[1-\exp\left(-\frac{y_i-\mu_i}{\phi_i}\right)\right]\frac{d\mu_i}{d\eta_{1i}}\frac{d\phi_i}{d\eta_{2i}}\frac{\partial \eta_{1i}}{\partial \beta_j}\frac{\partial \eta_{2i}}{\partial \theta_L}. \end{aligned} \tag{7.3}$$

Under regularity conditions, it is known that the expected value of the score function vanishes. Hence, we obtain

$$E\left[\exp\left(-\frac{Y_i-\mu_i}{\phi_i}\right)\right] = 1 \quad \text{and} \quad E\left[\exp\left(-\frac{Y_i-\mu_i}{\phi_i}\right)\frac{Y_i-\mu_i}{\phi_i}\right] = \gamma - 1,$$

where γ is Euler constant and $i = 1, \ldots, n$. It is easy to show that

$$E\left[\exp\left(-\frac{Y_i-\mu_i}{\phi_i}\right)\left(\frac{Y_i-\mu_i}{\phi_i}\right)^2\right] = \Gamma^{(2)}(2),$$

where $\Gamma^{(2)}(.)$ denotes the second derivative of the gamma function. Using the results above and taking the expected values in (7.1), (7.2) and (7.3), it comes

$$\kappa_{j,l} = -\sum_{i=1}^{n} \frac{1}{\phi_i^2}\left(\frac{d\mu_i}{d\eta_{1i}}\right)^2 \frac{\partial \eta_{1i}}{\partial \beta_j}\frac{\partial \eta_{1i}}{\partial \beta_l}, \quad \kappa_{J,L} = -\sum_{i=1}^{n} \frac{1}{\phi_i^2}[\Gamma^{(2)}(2) - 1]\left(\frac{d\phi_i}{d\eta_{2i}}\right)^2 \frac{\partial \eta_{2i}}{\partial \theta_j}\frac{\partial \eta_{2i}}{\partial \theta_l}$$

and

$$\kappa_{j,L} = -\sum_{i=1}^{n} \frac{\gamma - 1}{\phi_i^2}\frac{d\mu_i}{d\eta_{1i}}\frac{d\phi_i}{d\eta_{2i}}\frac{\partial \eta_{1i}}{\partial \beta_j}\frac{\partial \eta_{2i}}{\partial \theta_l}.$$

In notation matrix, we have

$$E\left(-\frac{\partial^2 \ell}{\partial \beta \partial \beta^\top}\right) = \widetilde{X}^\top W_{\beta\beta}\widetilde{X}, \quad E\left(-\frac{\partial^2 \ell}{\partial \theta \partial \theta^\top}\right) = \widetilde{S}^\top W_{\theta\theta}\widetilde{S}$$

and

$$E\left(-\frac{\partial^2 \ell}{\partial \beta \partial \theta^\top}\right) = \widetilde{X}^\top W_{\beta\theta}\widetilde{S}.$$

Appendix B

We here present the cumulants of third order and use the notation of the Section 3 and $(jl)_i = \partial^2\eta_{1i}/\partial\beta_j\beta_l$, $(JL)_i = \partial^2\eta_{2i}/\partial\theta_J\theta_L$, $(j,l)_i = \partial\eta_{1i}/\partial\beta_j\partial\eta_{1i}/\partial\beta_l$, $(jl,m)_i = \partial^2\eta_{1i}/\partial\beta_j\partial\beta_l\eta_{1i}/\partial\beta_m$, $(jlm)_i = \partial^3\eta_{1i}/\partial\beta_j\partial\beta_l\beta_m$ and so on.

Taking the expected value of the partial derivatives of (7.1), (7.2) and (7.3), it follows that

$$\begin{aligned}\kappa_{jlm} &= -\sum_{i=1}^{n}\left[\phi_i^{-3}\left(\frac{d\mu_i}{d\eta_{1i}}\right)^3 + 3\phi_i^{-2}\frac{d\mu_i}{d\eta_{1i}}\frac{d^2\mu_i}{d\eta_{1i}^2}\right](j,l,m)_i - \sum_{i=1}^{n}\phi_i^{-2}\left(\frac{d\mu_i}{d\eta_{1i}}\right)^2 \times \\ &\quad [(lm,j)_i + (jm,l)_i + (jl,m)_i],\end{aligned}$$

$$\kappa_{jlM} = \sum_{i=1}^{n}\left[\frac{3-\gamma}{\phi_i^3}\left(\frac{d\mu_i}{d\eta_{1i}}\right)^2 + \frac{1-\gamma}{\phi_i^2}\frac{d^2\mu_i}{d\eta_{1i}^2}\right]\frac{d\phi_i}{d\eta_{2i}}(j,l,M)_i - \sum_{i=1}^{n}\frac{\gamma-1}{\phi_i^2}\frac{d\mu_i}{d\eta_{1i}}\frac{d\phi_i}{d\eta_{2i}}(jl,M)_i,$$

$$\begin{aligned}\kappa_{JLM} &= \sum_{i=1}^{n}\left[(\Gamma^{(3)}(2)+6\Gamma^{(2)}(2)+4)\phi_i^{-3}\left(\frac{d\phi_i}{d\eta_{2i}}\right)^3 - 3(\Gamma^{(2)}(2)+1)\phi_i^{-2}\frac{d\phi_i}{d\eta_{2i}}\frac{d^2\phi_i}{d\eta_{2i}^2}\right] \times \\ &\quad (J,L,M)_i - (\Gamma^{(2)}(2)+1)\sum_{i=1}^{n}\phi_i^{-2}\left(\frac{d\phi_i}{d\eta_{2i}}\right)^2[(JM,L)_i + (LM,J)_i + (JL,M)_i]\end{aligned}$$

and

$$\begin{aligned}\kappa_{JLm} &= \sum_{i=1}^{n}\left\{[4(\gamma-1)-\Gamma^{(2)}(2)]\phi_i^{-3}\left(\frac{d\phi_i}{d\eta_{2i}}\right)^2\frac{d\mu_i}{d\eta_{1i}} + (1-\gamma)\phi_i^{-2}\frac{d^2\phi_i}{d\eta_{2i}^2}\frac{d\mu_i}{d\eta_{1i}}(J,L,m)_i\right\} + \\ &\quad (1-\gamma)\sum_{i=1}^{n}\phi_i^{-2}\frac{d\mu_i}{d\eta_{1i}}\frac{d\phi}{d\eta_{2i}}(JL,m)_i.\end{aligned}$$

The partial derivatives of the elements of Fisher's information matrix are given by

$$\kappa_{jl}^{(m)} = -2\sum_{i=1}^{n}\phi_i^{-2}\frac{d\mu_i}{d\eta_{1i}}\frac{d^2\mu_i}{d\eta_{1i}^2}(j,l,m)_i - \sum_{i=1}^{n}\phi_i^{-2}\left(\frac{d\mu_i}{d\eta_{1i}}\right)^2[(jm,l)_i + (lm,j)_i],$$

$$\kappa_{jl}^{(M)} = 2\sum_{i=1}^{n}\phi_i^{-3}\left(\frac{d\mu_i}{d\eta_{1i}}\right)^2\frac{d\phi_i}{d\eta_{2i}}(j,l,M)_i, \quad \kappa_{JL}^{(m)} = 0,$$

$$\kappa_{JL}^{(M)} = 2(\Gamma^{(2)}(2)+1)\sum_{i=1}^{n}\left[\phi_i^{-3}\left(\frac{d\phi_i}{d\eta_{2i}}\right)^3 - \phi_i^{-2}\frac{d\phi_i}{d\eta_{2i}}\frac{d^2\phi_i}{d\eta_{2i}^2}\right](J,L,M)_i - (\Gamma^{(2)}(2)+1)\times \sum_{i=1}^{n}\phi_i^{-2}\left(\frac{d\phi_i}{d\eta_{2i}}\right)^2[(JL,M)_i + (LM,J)_i],$$

$$\kappa_{jL}^{(m)} = (1-\gamma)\sum_{i=1}^{n}\phi_i^{-2}\left[\frac{d^2\mu_i}{d\eta_{1i}^2}(j,L,m)_i + \frac{d\mu_i}{d\eta_{1i}}(jm,L)_i\right]\frac{d\phi_i}{d\eta_{2i}}$$

and

$$\kappa_{jL}^{(M)} = (\gamma-1)\left\{\sum_{i=1}^{n}\left[2\phi_i^{-3}\left(\frac{d\phi_i}{d\eta_{2i}^2}\right) - \phi_i^{-2}\frac{d^2\phi_i}{d\eta_{2i}^2}\right]\frac{d\mu_i}{d\eta_{1i}}(j,L,M)_i - \sum_{i=1}^{n}\phi_i^{-2}\frac{d\mu_i}{d\eta_{1i}}\frac{d\phi_i}{d\eta_{2i}}(LM,j)_i\right\}.$$

We define the quantities

$$\omega_{1i} = \frac{1}{2}\left[\phi_i^{-3}\left(\frac{d\mu_i}{d\eta_{1i}}\right)^3 - \phi_i^{-2}\frac{d\mu_i}{d\eta_{1i}}\frac{d^2\mu_i}{d\eta_{1i}^2}\right], \quad \omega_{2i} = -\frac{1}{2}\phi_i^{-2}\left(\frac{d\mu_i}{d\eta_{1i}}\right)^2,$$

$$\omega_{3i} = \frac{1}{2}\left[(1-\gamma)\phi_i^{-2}\frac{d^2\mu_i}{d\eta_{1i}^2} - (3-\gamma)\phi_i^{-3}\left(\frac{d\mu_i}{d\eta_{1i}}\right)^2\right]\frac{d\phi_i}{d\eta_{2i}},$$

$$\omega_{4i} = \frac{1-\gamma}{2}\phi_i^{-2}\frac{d\mu_i}{d\eta_{1i}}\frac{d\phi_i}{d\eta_{2i}}, \quad \omega_{5i} = \frac{1}{2}\left[(\gamma+1)\phi_i^{-3}\left(\frac{d\mu_i}{d\eta_{1i}}\right)^2 + (\gamma-1)\phi_i^{-2}\frac{d^2\mu_i}{d\eta_{1i}^2}\right]\frac{d\phi_i}{d\eta_{2i}},$$

$$\omega_{6i} = -\frac{1}{2}\left\{[4(\gamma-1) - \Gamma^{(2)}(2)]\phi_i^{-3}\left(\frac{d\phi_i}{d\eta_{2i}}\right)^2 + (1-\gamma)\phi_i^{-2}\frac{d^2\phi_i}{d\eta_{2i}^2}\right\}\frac{d\mu_i}{d\eta_{1i}},$$

$$\omega_{7i} = \frac{1}{2}\left[\Gamma^{(2)}(2)\phi_i^{-3}\left(\frac{d\phi_i}{d\eta_{2i}}\right)^2 + (1-\gamma)\phi_i^{-2}\frac{d^2\phi_i}{d\eta_{2i}^2}\right]\frac{d\mu_i}{d\eta_{1i}},$$

$$\omega_{8i} = -\left[\frac{\Gamma^{(3)}(2)}{2} + \Gamma^{(2)}(2)\right]\phi_i^{-3}\left(\frac{d\phi_i}{d\eta_{2i}}\right)^3 - \frac{1}{2}\left(\Gamma^{(2)}(2)+1\right)\phi_i^{-2}\frac{d\phi_i}{d\eta_{2i}}\frac{d^2\phi_i}{d\eta_{2i}^2},$$

$$\omega_{9i} = -\frac{\Gamma^{(2)}(2)+1}{2}\phi_i^{-2}\left(\frac{d\phi_i}{d\eta_{2i}}\right)^2$$

and the diagonal matrices $W_k = \mathrm{d}iag\{\omega_{k1},\ldots,\omega_{kn}\}$ for $i = 1,\ldots,n$ and $k = 1,\ldots,9$; $\Gamma^{(3)}(\cdot)$ denotes the third derivative of the gamma function.

With the notation introduced above, we have that

$$\kappa_{jl}^{(m)} - \frac{1}{2}\kappa_{jlm} = \sum_{i=1}^{n}\omega_{1i}(j,l,m)_i - \sum_{i=1}^{n}\omega_{2i}[(jl,m)_i - (jm,l)_i - (lm,j)_i],$$

$$\kappa_{Jl}^{(m)} - \frac{1}{2}\kappa_{Jlm} = \sum_{i=1}^{n}\omega_{3i}(J,l,m)_i + \sum_{i=1}^{n}\omega_{4i}(lm,J)_i,$$

$$\kappa_{jL}^{(m)} - \frac{1}{2}\kappa_{jLm} = \sum_{i=1}^{n}\omega_{3i}(j,L,m)_i + \sum_{i=1}^{n}\omega_{4i}(jm,L),$$

$$\kappa_{jl}^{(M)} - \frac{1}{2}\kappa_{jlM} = \sum_{i=1}^{n}\omega_{5i}(j,l,U)_i - \sum_{i=1}^{n}\omega_{4i}(jl,M)_i,$$

$$\kappa_{JL}^{(m)} - \frac{1}{2}\kappa_{JLm} = \sum_{i=1}^{n}\omega_{6i}(J,L,m)_i - \sum_{i=1}^{n}\omega_{4i}(JL,m)_i,$$

$$\kappa_{Jl}^{(M)} - \frac{1}{2}\kappa_{JlM} = \sum_{i=1}^{n}\omega_{7i}(J,l,U)_i + \sum_{i=1}^{n}\omega_{4i}(JM,l)_i,$$

$$\kappa_{jL}^{(M)} - \frac{1}{2}\kappa_{jLM} = \sum_{i=1}^{n}\omega_{7i}(j,L,U)_i + \sum_{i=1}^{n}\omega_{4i}(LM,j)_i$$

and

$$\kappa_{JL}^{(M)} - \frac{1}{2}\kappa_{JLM} = \sum_{i=1}^{n} \omega_{8i}(J,L,U)_i + \sum_{i=1}^{n} \omega_{9i}[(LM,J)_i - (JL,M)_i + (JM,L)_i].$$

We now calculate the terms in (3.1) to obtain the second-order bias order for the MLEs of the extreme-value regression model with dispersion covariates. Define e_a as the ath column vector of the $p \times p$ identity matrix, $\mathbf{1_{n\times 1}}$ vector of ones with dimension $n \times 1$, $L_{1,a} = \sum_{i=1}^{n} \omega_{4i} \sum_{j,l,M} \kappa^{aj}\kappa^{lM}(jl,M)_i$ and $L_{2,a} = -\sum_{i=1}^{n} \omega_{4i} \sum_{J,L,m} \kappa^{aJ}\kappa^{Lm}(JL,m)_i$. Hence, it follows that

$$\begin{aligned}
\sum_{j,l,m} \kappa^{aj}\kappa^{lm}\left\{\kappa_{jl}^{(m)} - \frac{1}{2}\kappa_{jlm}\right\} &= \sum_{i=1}^{n} \omega_{1i} \sum_{j} \kappa^{aj}(j)_i \sum_{l,m} \kappa^{lm}(l,m)_i + \\
&\quad \sum_{i=1}^{n} \omega_{2,i} \sum_{j} \kappa^{aj}(j)_i \sum_{l,m} \kappa^{lm}(lm)_i \\
&= e_a^\top K^{\beta\beta} \widetilde{X}^\top W_1 Z_{\beta d} 1_{n\times 1} + e_a^\top K^{\beta\beta} \widetilde{X}^\top W_2 D_\beta 1_{n\times 1},
\end{aligned}$$

$$\begin{aligned}
\sum_{J,l,m} \kappa^{aJ}\kappa^{lm}\left\{\kappa_{Jl}^{(m)} - \frac{1}{2}\kappa_{Jlm}\right\} &= \sum_{i=1}^{n} \omega_{3i} \sum_{J} \kappa^{aJ}(J)_i \sum_{l,m} \kappa^{lm}(l,m)_i + \\
&\quad \sum_{i=1}^{n} \omega_{4,i} \sum_{J} \kappa^{aJ}(J)_i \sum_{l,m} \kappa^{lm}(lm)_i \\
&= e_a^\top K^{\beta\theta} \widetilde{S}^\top W_3 Z_{\beta d} 1_{n\times 1} + e_a^\top K^{\beta\theta} \widetilde{S}^\top W_4 D_\beta 1_{n\times 1},
\end{aligned}$$

$$\begin{aligned}
\sum_{j,L,m} \kappa^{aj}\kappa^{Lm}\left\{\kappa_{jL}^{(m)} - \frac{1}{2}\kappa_{jLm}\right\} &= \sum_{i=1}^{n} \omega_{3i} \sum_{j} \kappa^{aj}(j)_i \sum_{L,m} \kappa^{Lm}(L,m)_i + L_{1,a} \\
&= e_a^\top K^{\beta\beta} \widetilde{X}^\top W_3 Z_{\beta\theta d} 1_{n\times 1} + L_{1,a},
\end{aligned}$$

$$\begin{aligned}
\sum_{j,l,M} \kappa^{aj}\kappa^{lM}\left\{\kappa_{jl}^{(M)} - \frac{1}{2}\kappa_{jlM}\right\} &= \sum_{i=1}^{n} \omega_{5i} \sum_{j} \kappa^{aj}(j)_i \sum_{l,M} \kappa^{lM}(l,M)_i - L_{1,a} \\
&= e_a^\top K^{\beta\beta} \widetilde{X}^\top W_5 Z_{\beta\theta d} 1_{n\times 1} - L_{1,a},
\end{aligned}$$

$$\begin{aligned}
\sum_{J,L,m} \kappa^{aJ}\kappa^{Lm}\left\{\kappa_{JL}^{(m)} - \frac{1}{2}\kappa_{JLm}\right\} &= \sum_{i=1}^{n} \omega_{6i} \sum_{J} \kappa^{aJ}(J)_i \sum_{L,m} \kappa^{Lm}(L,m)_i + L_{2,a} \\
&= e_a^\top K^{\beta\theta} \widetilde{S}^\top W_6 Z_{\beta\theta d} 1_{n\times 1} + L_{2,a},
\end{aligned}$$

$$\begin{aligned}\sum_{J,l,M}\kappa^{aJ}\kappa^{lM}\left\{\kappa_{Jl}^{(M)}-\frac{1}{2}\kappa_{JlM}\right\} &= \sum_{i=1}^{n}\omega_{7i}\sum_{J}\kappa^{aJ}(J)_i\sum_{l,M}\kappa^{lM}(l,M)_i - L_{2,a}\\ &= e_a^\top K^{\beta\theta}\widetilde{S}^\top W_7 Z_{\beta\theta d}1_{n\times 1} - L_{2,a},\end{aligned}$$

$$\begin{aligned}\sum_{j,L,M}\kappa^{aj}\kappa^{LM}\left\{\kappa_{jL}^{(M)}-\frac{1}{2}\kappa_{jLM}\right\} &= \sum_{i=1}^{n}\omega_{7i}\sum_{j}\kappa^{aj}(j)_i\sum_{L,M}\kappa^{LM}(L,M)_i +\\ &\quad \sum_{i=1}^{n}\omega_{4i}\sum_{j}\kappa^{aj}(j)_i\sum_{L,M}\kappa^{LM}(LM)_i\\ &= e_a^\top K^{\beta\beta}\widetilde{X}^\top W_7 Z_{\theta d}1_{n\times 1} + e_a^\top K^{\beta\beta}\widetilde{X}^\top W_4 D_\theta 1_{n\times 1},\end{aligned}$$

and

$$\begin{aligned}\sum_{J,L,M}\kappa^{aJ}\kappa^{LM}\left\{\kappa_{JL}^{(M)}-\frac{1}{2}\kappa_{JLM}\right\} &= \sum_{i=1}^{n}\omega_{8i}\sum_{J}\kappa^{aJ}(J)_i\sum_{L,M}\kappa^{LM}(L,M)_i +\\ &\quad \sum_{i=1}^{n}\omega_{9i}\sum_{J}\kappa^{aJ}(J)_i\sum_{L,M}\kappa^{LM}(LM)_i\\ &= e_a^\top K^{\beta\theta}\widetilde{S}^\top W_8 Z_{\theta d}1_{n\times 1} + e_a^\top K^{\beta\theta}\widetilde{S}^\top W_9 D_\theta 1_{n\times 1}.\end{aligned}$$

Appendix C

We here calculate the cumulants of the expression (4.2). With the cumulants obtained previously, it is easy to show that

$$
\begin{aligned}
-\kappa_{jlm} + 5\kappa_{jl}^{(m)} - \kappa_{jm}^{(l)} - \kappa_{lm}^{(j)} &= \sum_{i=1}^{n} \left[\phi_i^{-3} \left(\frac{d\mu_i}{d\eta_{1i}} \right)^3 - 3\phi_i^{-2} \frac{d\mu_i}{d\eta_{1i}} \frac{d^2\mu_i}{d\eta_{1i}^2} \right] (j,l,m)_i - \\
&\quad 3 \sum_{i=1}^{n} \phi_i^{-2} \left(\frac{d\mu_i}{d\eta_{1i}} \right)^2 [(jm,l)_i + (j,lm)_i - (jl,m)_i],
\end{aligned}
$$

$$
\begin{aligned}
-\kappa_{Jlm} + 5\kappa_{Jl}^{(m)} - \kappa_{Jm}^{(l)} - \kappa_{lm}^{(J)} &= \sum_{i=1}^{n} \left[(\gamma-5)\phi_i^{-3} \left(\frac{d\mu_i}{d\eta_{1i}} \right)^2 \frac{d\phi_i}{d\eta_{2i}} + 3(1-\gamma)\phi_i^{-2} \frac{d^2\mu_i}{d\eta_{1i}^2} \frac{d\phi_i}{d\eta_{2i}} \right] \times \\
&\quad (J,l,m)_i + 3(1-\gamma) \sum_{i=1}^{n} \phi_i^{-2} \frac{d\mu_i}{d\eta_{1i}} \frac{d\phi_i}{d\eta_{2i}} (J,lm)_i,
\end{aligned}
$$

$$
\begin{aligned}
-\kappa_{jLm} + 5\kappa_{jL}^{(m)} - \kappa_{jm}^{(L)} - \kappa_{Lm}^{(j)} &= \sum_{i=1}^{n} \left[(\gamma-5)\phi_i^{-3} \left(\frac{d\mu_i}{d\eta_{1i}} \right)^2 \frac{d\phi_i}{d\eta_{2i}} + 3(1-\gamma)\phi_i^{-2} \frac{d^2\mu_i}{d\eta_{1i}^2} \frac{d\phi_i}{d\eta_{2i}} \right] \times \\
&\quad (j,L,m)_i + 3(1-\gamma) \sum_{i=1}^{n} \phi_i^{-2} \frac{d\mu_i}{d\eta_{1i}} \frac{d\phi_i}{d\eta_{2i}} (jm,L)_i,
\end{aligned}
$$

$$
\begin{aligned}
-\kappa_{jlM} + 5\kappa_{jl}^{(M)} - \kappa_{jM}^{(l)} - \kappa_{lM}^{(j)} &= \sum_{i=1}^{n} \left[(\gamma-7)\phi_i^{-3} \left(\frac{d\mu_i}{d\eta_{1i}} \right)^2 \frac{d\phi_i}{d\eta_{2i}} - 3(1-\gamma)\phi_i^{-2} \frac{d^2\mu_i}{d\eta_{1i}^2} \frac{d\phi_i}{d\eta_{2i}} \right] \times \\
&\quad (j,l,M)_i - 3(1-\gamma) \sum_{i=1}^{n} \phi_i^{-2} \frac{d\mu_i}{d\eta_{1i}} \frac{d\phi_i}{d\eta_{2i}} (jl,M)_i,
\end{aligned}
$$

$$
\begin{aligned}
-\kappa_{JLm} + 5\kappa_{JL}^{(m)} - \kappa_{Jm}^{(L)} - \kappa_{Lm}^{(J)} &= \sum_{i=1}^{n} \left\{ [8(1-\gamma) + \Gamma^{(2)}(2)]\phi_i^{-3} \frac{d\mu_i}{d\eta_{1i}} \left(\frac{d\phi_i}{d\eta_{2i}} \right)^2 - 3(1-\gamma) \times \right. \\
&\quad \left. \phi_i^{-2} \frac{d\mu_i}{d\eta_{1i}} \frac{d^2\phi_i}{d\eta_{2i}^2} \right\} (J,L,m) - 3(1-\gamma) \sum_{i=1}^{n} \phi_i^{-2} \frac{d\mu_i}{d\eta_{1i}} \frac{d\phi_i}{d\eta_{2i}} (JL,m)_i,
\end{aligned}
$$

$$-\kappa_{JlM}+5\kappa_{Jl}^{(M)}-\kappa_{JM}^{(l)}-\kappa_{lM}^{(J)} = \sum_{i=1}^{n}\left\{[-4(1-\gamma)+\Gamma^{(2)}(2)]\phi_i^{-3}\frac{d\mu_i}{d\eta_{1i}}\left(\frac{d\phi_i}{d\eta_{2i}}\right)^2+3(1-\gamma)\times \right.$$
$$\left.\phi_i^{-2}\frac{d\mu_i}{d\eta_{1i}}\frac{d^2\phi_i}{d\eta_{2i}^2}\right\}(J,l,M)+3(1-\gamma)\sum_{i=1}^{n}\phi_i^{-2}\frac{d\mu_i}{d\eta_{1i}}\frac{d\phi_i}{d\eta_{2i}}(JM,l)_i,$$

$$-\kappa_{jLM}+5\kappa_{jL}^{(M)}-\kappa_{jM}^{(L)}-\kappa_{LM}^{(j)} = \sum_{i=1}^{n}\left\{[-4(1-\gamma)+\Gamma^{(2)}(2)]\phi_i^{-3}\frac{d\mu_i}{d\eta_{1i}}\left(\frac{d\phi_i}{d\eta_{2i}}\right)^2+3(1-\gamma)\times \right.$$
$$\left.\phi_i^{-2}\frac{d\mu_i}{d\eta_{1i}}\frac{d^2\phi_i}{d\eta_{2i}^2}\right\}(j,L,M)+3(1-\gamma)\sum_{i=1}^{n}\phi_i^{-2}\frac{d\mu_i}{d\eta_{1i}}\frac{d\phi_i}{d\eta_{2i}}(j,LM)_i$$

and

$$-\kappa_{JLM}+5\kappa_{JL}^{(M)}-\kappa_{JM}^{(L)}-\kappa_{LM}^{(J)} = \sum_{i=1}^{n}\left[(-\Gamma^{(3)}(2)+2)\phi_i^{-3}\left(\frac{d\phi_i}{d\eta_{2i}}\right)^3-3(\Gamma^{(2)}(2)+1)\times \right.$$
$$\left.\phi_i^{-2}\frac{d\phi_i}{d\eta_{2i}}\frac{d^2\phi_i}{d\eta_{2i}^2}\right](J,L,M)_i-3(\Gamma^{(2)}(2)+1)\sum_{i=1}^{n}\phi_i^{-2}\left(\frac{d\phi_i}{d\eta_{2i}}\right)^2[(J,LM)_i+(JM,L)_i-(JL,M)_i].$$

We define

$$v_{1i}=\phi_i^{-3}\left(\frac{d\mu_i}{d\eta_{1i}}\right)^3-3\phi_i^{-2}\frac{d\mu_i}{d\eta_{1i}}\frac{d^2\mu_i}{d\eta_{1i}^2},\quad v_{2i}=-3\phi_i^{-2}\left(\frac{d\mu_i}{d\eta_{1i}}\right)^2,$$

$$v_{3i}=(\gamma-5)\phi_i^{-3}\left(\frac{d\mu_i}{d\eta_{1i}}\right)^2\frac{d\phi_i}{d\eta_{2i}}+3(1-\gamma)\phi_i^{-2}\frac{d^2\mu_i}{d\eta_{1i}^2}\frac{d\phi_i}{d\eta_{2i}},\quad v_{4i}=3(1-\gamma)\phi_i^{-2}\frac{d\mu_i}{d\eta_{1i}}\frac{d\phi_i}{d\eta_{2i}},$$

$$v_{5i}=(\gamma-7)\phi_i^{-3}\left(\frac{d\mu_i}{d\eta_{1i}}\right)^2\frac{d\phi_i}{d\eta_{2i}}-3(1-\gamma)\phi_i^{-2}\frac{d^2\mu_i}{d\eta_{1i}^2}\frac{d\phi_i}{d\eta_{2i}},$$

$$v_{6i}=[8(1-\gamma)+\Gamma^{(2)}(2)]\phi_i^{-3}\frac{d\mu_i}{d\eta_{1i}}\left(\frac{d\phi_i}{d\eta_{2i}}\right)^2-3(1-\gamma)\phi_i^{-2}\frac{d\mu_i}{d\eta_{1i}}\frac{d^2\phi_i}{d\eta_{2i}^2},$$

$$v_{7i}=[-4(1-\gamma)+\Gamma^{(2)}(2)]\phi_i^{-3}\frac{d\mu_i}{d\eta_{1i}}\left(\frac{d\phi_i}{d\eta_{2i}}\right)^2+3(1-\gamma)\phi_i^{-2}\frac{d\mu_i}{d\eta_{1i}}\frac{d^2\phi_i}{d\eta_{2i}^2},$$

$$v_{8i} = [2 - \Gamma^{(3)}(2)]\phi_i^{-3}\left(\frac{d\phi_i}{d\eta_{2i}}\right)^3 - 3[\Gamma^{(2)}(2)+1]\phi_i^{-2}\frac{d\phi_i}{d\eta_{2i}}\frac{d^2\phi_i}{d\eta_{2i}^2}$$

and

$$v_{9i} = -3[\Gamma^{(2)}(2)+1]\phi_i^{-2}\left(\frac{d\phi_i}{d\eta_{2i}}\right)^2,$$

and the diagonal matrices $V_k = \mathrm{d}iag\{v_{k1},\ldots,v_{kn}\}$, for $i = 1,\ldots,n$ and $k = 1,\ldots,9$.

Hence, we obtain

$$-\kappa_{jlm} + 5\kappa_{jl}^{(m)} - \kappa_{jm}^{(l)} - \kappa_{lm}^{(j)} = \sum_{i=1}^{n} v_{1i}(j,l,m)_i + \sum_{i=1}^{n} v_{2i}[(jm,l)_i + (j,lm)_i - (jl,m)_i],$$

$$-\kappa_{Jlm} + 5\kappa_{Jl}^{(m)} - \kappa_{Jm}^{(l)} - \kappa_{lm}^{(J)} = \sum_{i=1}^{n} v_{3i}(J,l,m)_i + \sum_{i=1}^{n} v_{4i}(J,lm)_i,$$

$$-\kappa_{jLm} + 5\kappa_{jL}^{(m)} - \kappa_{jm}^{(L)} - \kappa_{Lm}^{(j)} = \sum_{i=1}^{n} v_{3i}(j,L,m)_i + \sum_{i=1}^{n} v_{4i}(L,jm)_i,$$

$$-\kappa_{jlM} + 5\kappa_{jl}^{(M)} - \kappa_{jM}^{(l)} - \kappa_{lM}^{(j)} = \sum_{i=1}^{n} v_{5i}(j,l,M)_i - \sum_{i=1}^{n} v_{4i}(jl,M)_i,$$

$$-\kappa_{JLm} + 5\kappa_{JL}^{(m)} - \kappa_{Jm}^{(L)} - \kappa_{Lm}^{(J)} = \sum_{i=1}^{n} v_{6i}(J,L,m)_i - \sum_{i=1}^{n} v_{4i}(JL,m)_i,$$

$$-\kappa_{JlM} + 5\kappa_{Jl}^{(M)} - \kappa_{JM}^{(l)} - \kappa_{lM}^{(J)} = \sum_{i=1}^{n} v_{7i}(J,l,M)_i + \sum_{i=1}^{n} v_{4i}(JM,l)_i,$$

$$-\kappa_{jLM} + 5\kappa_{jL}^{(M)} - \kappa_{jM}^{(L)} - \kappa_{LM}^{(j)} = \sum_{i=1}^{n} v_{7i}(j,L,M)_i + \sum_{i=1}^{n} v_{4i}(j,LM)_i,$$

and

$$-\kappa_{JLM} + 5\kappa_{JL}^{(M)} - \kappa_{JM}^{(L)} - \kappa_{LM}^{(J)} = \sum_{i=1}^{n} v_{8i}(J,L,M)_i + \sum_{i=1}^{n} v_{9i}[(J,LM)_i + (JM,L)_i - (JL,M)_i].$$

We now calculate the terms in (4.2) to obtain the third cumulant of order n^{-2} for the MLEs of the extreme-value regression model with dispersion covariates. It follows that

$$
\begin{aligned}
-\sum_{j,l,m}\kappa^{a,j}\kappa^{a,l}\kappa^{a,m}\{\kappa_{jlm}-5\kappa_{jl}^{(m)}+\kappa_{jm}^{(l)}+\kappa_{lm}^{(j)}\} &= e_a^\top[B_{\beta\beta}^{(3)}V_1 1_{n\times 1}+ \\
&\quad \text{diagonal}\{N_\beta V_2 B_{\beta\beta}^\top\}1_{p\times 1}],
\end{aligned}
$$

$$
\begin{aligned}
-\sum_{J,l,m}\kappa^{a,J}\kappa^{a,l}\kappa^{a,m}\{\kappa_{Jlm}-5\kappa_{Jl}^{(m)}+\kappa_{Jm}^{(l)}+\kappa_{lm}^{(J)}\} &= e_a^\top B_{\beta\beta}^{(2)}V_3 B_{\beta\theta}^\top 1_{p\times 1}+ \\
&\quad \text{diagonal}\{N_\beta V_4 B_{\beta\theta}^\top\}1_{p\times 1}],
\end{aligned}
$$

$$
\begin{aligned}
-\sum_{j,L,m}\kappa^{a,j}\kappa^{a,L}\kappa^{a,m}\{\kappa_{jLm}-5\kappa_{jL}^{(m)}+\kappa_{jm}^{(L)}+\kappa_{Lm}^{(j)}\} &= e_a^\top[B_{\beta\beta}^{(2)}V_3 B_{\beta\theta}^\top 1_{p\times 1}+ \\
&\quad \text{diagonal}\{N_\beta V_4 B_{\beta\theta}^\top\}1_{p\times 1}],
\end{aligned}
$$

$$
\begin{aligned}
-\sum_{j,l,M}\kappa^{a,j}\kappa^{a,l}\kappa^{a,M}\{\kappa_{jlM}-5\kappa_{jl}^{(M)}+\kappa_{jM}^{(l)}+\kappa_{lM}^{(j)}\} &= e_a^\top[B_{\beta\beta}^{(2)}V_5 B_{\beta\theta}^\top 1_{p\times 1}- \\
&\quad \text{diagonal}\{N_\beta V_4 B_{\beta\theta}^\top\}1_{p\times 1}],
\end{aligned}
$$

$$
\begin{aligned}
-\sum_{J,L,m}\kappa^{a,J}\kappa^{a,L}\kappa^{a,m}\{\kappa_{JLm}-5\kappa_{JL}^{(m)}+\kappa_{Jm}^{(L)}+\kappa_{Lm}^{(J)}\} &= e_a^\top B_{\beta\beta}V_6 B_{\beta\theta}^{(2)\top}1_{p\times 1}- \\
&\quad \text{diagonal}\{N_{\beta\theta}V_4 B_{\beta\beta}^\top\}1_{p\times 1}],
\end{aligned}
$$

$$
\begin{aligned}
-\sum_{J,l,M}\kappa^{a,J}\kappa^{a,l}\kappa^{a,M}\{\kappa_{JlM}-5\kappa_{Jl}^{(M)}+\kappa_{JM}^{(l)}+\kappa_{lM}^{(J)}\} &= e_a^\top[B_{\beta\beta}V_7 B_{\beta\theta}^{(2)\top}1_{p\times 1}+ \\
&\quad \text{diagonal}\{N_{\beta\theta}V_4 B_{\beta\theta}^\top\}1_{p\times 1}],
\end{aligned}
$$

$$
\begin{aligned}
-\sum_{j,L,M}\kappa^{a,j}\kappa^{a,L}\kappa^{a,M}\{\kappa_{jLM}-5\kappa_{jL}^{(M)}+\kappa_{jM}^{(L)}+\kappa_{LM}^{(j)}\} &= e_a^\top[B_{\beta\beta}V_7 B_{\beta\theta}^{(2)\top}1_{p\times 1}+ \\
&\quad \text{diagonal}\{N_{\beta\theta}V_4 B_{\beta\beta}^\top\}1_{p\times 1}]
\end{aligned}
$$

and

$$
\begin{aligned}
-\sum_{J,L,M}\kappa^{a,J}\kappa^{a,L}\kappa^{a,M}\{\kappa_{JLM}-5\kappa_{JL}^{(M)}+\kappa_{JM}^{(L)}+\kappa_{LM}^{(J)}\} &= e_a^\top[B_{\beta\theta}^{(3)}V_8 1_{n\times 1}+ \\
&\quad \text{diagonal}\{N_{\beta\theta}V_9 B_{\beta\theta}^\top\}1_{p\times 1}].
\end{aligned}
$$