لوغاريثم نماذج انحدار باريتو الاسية

إعداد

خلود معيوض مبطي البلادي

تحت إشراف

د. لمياء عبدالباسط باحارث

أستاذ مشارك في الإحصاء الحيوي (مشرف رئيسي)

د. هديل صالح كلكتاوي

أستاذ مساعد في الإحصاء الرياضي (مشرف مساعد)

المستخلص

في هذه الرسالة تم اقتراح ودراسة توزيع جديد وهو توزيع باريتوالرابع الاسي الغريب. ايضاً, تم اقتراح كلاً من توزيعي لوغاريتم باريتو الاسي ولوغاريتم باريتوالرابع الاسي الغريب. كما تم إيجاد بعض الخصائص لهذان التوزيعان ورسم المنحنى لبعض الدوال عند قيم مختلفة للمعالم. بالإضافة إلى ذلك, تم اقتراح تموذجي انحدار وهما نموذج انحدار لوغاريتم باريتو الاسي و نموذج انحدار لوغاريتم باريتوالرابع الاسي الغريب اعتماداً على التوزيعين السابقين.

تم استخدام طريقة الإمكان الاعظم لتقدير معالم توزيع باريتوالرابع الاسي الغريب. وايضاً, تم تقدير معالم نموذجي الانحدار السابقين باستخدام طريقة الإمكان الاعظم وطريقة الجاكنيف اعتماداً على عينات المراقبة. ايضاً, تم ايجاد مصفوفة التباين والتغاير والتي تم استخدامها في ايجاد فترات الثقة لمعالم النموذجين المقترحة.

تم استخدام مقياس (generalized Cook distance) ومقياس (likelihood distance) للكشف عن المشاهدات المؤثرة. ايضاً , تم استخدام بواقي مارتينجال (martingale residual) وبواقي ديفينس (deviance residual) للتحقق من افتراضات النموذجين واكتشاف القيم المتطرفة وصحة البيانات.

تم استخدام محاكاة مونتي كارلو (Monte Carlo simulation) لتقييم معالم توزيع باريتوالرابع الاسي الغريب باستخدام طريقة الامكان الاعظم. ايضاً, تم استخدام محاكاة مونتي كارلو (Monte Carlo simulation) للمقارنة بين مقدرات الامكان الاعظم ومقدرات الجاكنيف لعينات مختلفة الحجم ونسب رقابة مختلفة للنموذجين. وقد أظهرت نتائج المحاكاة ان مقدرات الجاكنيف اكثر كفاءة من مقدرات الامكان الاعظم بناءاً على متوسط الخطأ المعياري (standard error). كما تم حساب خصائص كل من بواقي مارتينجال وبواقي ديفينس للعينات المختلفة ونسب رقابة مختلفة وتمثيلها بيانياً.

اخيراً, تم استخدام بيانات قوة الالياف الزجاجية (strength of glass fibre) و بيانات ثنيات الجلد (skin folds) لمعرفة اداء توزيع باريتوالرابع الاسي الغريب. ايضاً, تم التحقق من مرونة نموذجي الانحدار السابقين من خلال تطبيقات حياه حقيقية. أشارت النتائج الى أن النماذج المقترحه مناسبة لتطبيقها في نماذج أزمنة البقاء .

Log Exponential-Pareto Regression Models

By

Kholod Mayoud AL-Beladi

Supervised by

Dr. Lamya A. Baharith

Associate Professor of Biostatistics (Advisor)

Dr. Hadeel S. Klakattawi

Assistant Professor of Mathematical Statistics (Co-Advisor)

Abstract

In this thesis, some extended exponential-Pareto models are considered. Particularly, a new odds exponential-Pareto IV distribution is defined and investigated. Also, log exponential-Pareto and log odds exponential-Pareto IV distributions are introduced. Some of their properties are derived. Moreover, two new location-scale regression models are constructed; one based on the log exponential-Pareto and the other based on the log odds exponential-Pareto IV distributions.

The maximum likelihood method is applied to estimate the parameters of the odds exponential-Pareto IV distribution. Also, the parameters of the proposed regression models are estimated using two different methods for censored data. First, the maximum likelihood estimators are obtained along with their asymptotic variance covariance matrices and confidence intervals. Additionally, jackknife method is considered as another method to estimate the unknown parameters. The generalized Cook distance and likelihood distance measures are applied to detect the possible influential observations. Also, martingale and deviance residuals are explored to assess the adequacy of the proposed model and to detect outliers.

The efficiency of maximum likelihood estimators for the parameters of odds exponential -Pareto IV distribution is investigated using some Monte Carlo simulation studies. Also, the performance of the maximum likelihood estimate is compared with the jackknife estimate through some Monte Carlo simulation studies for different sample sizes and censoring levels for the proposed regression models. The simulation results showed that the jackknife estimates are more efficient than the maximum likelihood estimates based on average of standard error. The martingale and deviance residuals properties are computed and the empirical distributions are plotted at different sample sizes and censoring levels.

Finally, the strength of glass fiber and skin folds data sets are analyzed under the odds exponential-Pareto IV distribution to show the potentiality of this new distribution in modelling real-life data. Also, the flexibility, usefulness and suitability of the log exponential-Pareto and log odds exponential-Pareto IV regression models are illustrated by mean of some real-life applications. The results demonstrated the superior performance of the proposed models comparing to some other competitive models in terms of some goodness-of-fit criteria.