وزارة التعليم العالي والبحث العلمي
جامعة الموصل
كلية علوم الحاسوب والرياضيات
قسم الإحصاء والمعلوماتية

التقدير اللامعلمي لدالة التوزيع التراكمي باستخدام عينات المجموعات المرتبة

رسالة مقدمة

إلى مجلس كلية علوم الحاسوب والرياضيات في جامعة الموصل
كجزء من متطلبات نيل شهادة ماجستير علوم في الإحصاء

من قبل

رامي سعد غريب كوريال

بإشراف

أ.م.د. ريكان عبد العزيز أحمد عبد الرحمن

١٤٤٦هـ ٢٠٢٥م

المستخلص

تسعى هذهِ الرسالة إلى تقدير دالة التوزيع التراكمي CDF و قياس نسبة الاحتمالية Odds Ratio فضلاً عن تقدير موثوقية القوة والإجهاد Stress And Strength Reliability بالاعتماد على طريقة انحدار متعدد الحدود المحلي Local Polynomial Regression (LPR) من الدرجة الأولى ومقارنتها مع تقدير الإمكان الأعظم اللامعلمي Non-Parametric Maximum Likelihood (NPML), من خلال متوسط مربعات الخطأ MSE و قيم التحيز Bias, بالاعتماد على عينات المجموعات المرتبة Ranked Sets Sampling و عينات المجموعات المرتبة الوسيطية الفردية و الزوجية Median Ranked Sets Sampling Odd and Even ($MRSS_o, MRSS_e$), درست خواص مقدر دالة التوزيع التراكمي CDF وقياس نسبة الاحتمالات الذي تمثل بـ التوقع Expectation و قيمة التحيز Bias و قيمة متوسط مربعات الخطأ MSE بالاعتماد على انحدار متعدد الحدود المحلي LPR, فضلاً عن ذلك أظهر تحليل المحاكاة بالاعتماد على حجم عينة 500 و تكرار هذه العينة 500 مرة لكل مشاهدة بأن الدرجة الأولى من انحدار متعدد الحدود المحلي ($P = 1$) تعطي أفضل النتائج مقارنة ببقية الدرجات فضلاً عن ذلك تمت دراسة درجات مختلفة من عرض النطاق الترددي Bandwidth وتبين لنا إن عرض النطاق الترددي من الدرجة 0.8 هو أفضل من بقية الدرجات. كما تم حساب الكفاءة النسبية Relative Efficiency إلى كل من الطرائق $NPML$, LPR وتبين لنا تفوق LPR على طريقة $NPML$ لتقدير دالة التوزيع التراكمي CDF, في مخلتف النواة (Normal (Gaussian), Epanechinkov) وأيضاً اجريت مقارنة بين الطرائق التي تم العتماد عليها في سحب العينات وهي طريقة RSS و $MRSS$ (الفردية و الزوجية). طبقت افضل الطرائق التي توصلنا اليها في المحاكاة على بيانات تمثلث بمتوسط انتاجية أشجار النخيل.

Ministry of Higher Education and
Scientific Research
University of Mosul
College of Computer Science and Mathematics
Department of Statistics and Informatics

Nonparametric Estimation of Cumulative Distribution Function Using Ranked Set Sampling

A Thesis Submitted to the Council of the College of Computer Science and Mathematics

University of Mosul

as a Partial Fulfillment of Requirements for the Degree of Master of Science in Statistics

By

Rami Saad Ghareeb Goryal

Supervised by

Assist.Prof.Dr. Rekan Abdulaziz Ahmed Abdulrahman

1446 A.H. **2025 A.D.**

ABSTRACT

The study in this thesis refers to the estimation of the cumulative distribution CDF Estimating the odds ratio and the reliability stress and strength. Utilizing the Local polynomial regression LPR With degree 2, and comparing it with the non-parametric estimation based on non-parametric maximum likelihood, for calculation mean square error MSE and Bias Dependence on Ranked Set Sampling RSS and Median Ranked Set Sampling Odd and Even $MRSS_O$, $MRSS_e$. We also studied the properties of the cumulative distribution function, Odds Ratio, and reliability stress and strength which is the expectation with bias and mean square error, depending on local polynomial regression. Using a simulation study, we found that the second degree-of LPR gives the best results compared to the other degrees. We then conducted a study, of different bandwidths, we concluded that the bandwidth of degree 0.8 is better than the rest of the bandwidth degrees. Then we calculated the relative efficiency of each method (LPR, NPML) it became clear to us that the efficiency of LPR is higher than the rest of the methods for estimating the cumulative distribution function CDF, in different kernels (Normal (Gaussian), epanechinkov) and depending on a sample size of 500 and repeating this sample 500 times for each observation. We also made a comparison between the methods we relied on in pulling the samples, which are the RSS and $MRSS$ methods (odd and even), it became clear to us that the proposed estimators based on the median rank set sample $MRSS$ (odd and even) gave better results than in the ranked sets sample RSS.
In practical application, applied the best methods we reached in the simulation to data represented by the average productivity of date palm trees.