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

استخدام بعض طرق المعاينة المرتبة

في تقدير معلمتي توزيع لوماكس مع التطبيق على بيانات مرض سرطان المثانة

رسالة مقدمة

الى مجلس كلية علوم الحاسوب والرياضيات في جامعة الموصل

كجزء من متطلبات نيل شهادة ماجستير علوم في الاحصاء

من قبل

موسى محمد موسى فرحان الجبوري

بإشراف

أ.م.د. بان غانم عمر العاني

1445هـ **2024م**

المستخلص

لتحقيق الأهــداف التي يسعى الى تحقيقهــا الباحث في بحثـــهُ .يجب جمع البيانات والمعلومات عن ذلك البحث، وبسبب إن بعض البيانات يصعب الحصول عليها او لا يمكن الحصول عليها لأسباب تتعلق بتكلفة والجهد والوقت فلا بد لهُ من إن يختار أسلوب المعاينة الذي يضمن لهُ تحقيق اهداف البحث بأقل وقت وتكلفة وجهد ، او قد تكون في كثير من المسائل التطبيقية لا يمكن الحصول على القياسات الحقيقية للمتغير قيد الأهتمام مكلفة وبذلك تكون مضيعة للوقت ولحل هذهِ المشاكل يجب على الباحث استخدام اسلوب المعاينة الذي يضمن تقليل الوقت والجهد والتكلفة في الحصول على البيانات في هذهِ الحالة تبين ان طريقة المعاينة المجموعة المرتبة Ranked Set Sampling(RSS) مهم للحصول على المعلومات التي تفي بالغرض والتي يحتاجها الباحث. لذلك جاء الهدف الرئيسي لهذهِ الرسالة هو إيجاد الدالة الاحتمالية لتقدير معلمة الشكل ومعلمة القياس، فضلاً عن تقدير دالة الموثوقية لتوزيع لوماكس (Lomax distribution) بناءً على بعض تصاميم معاينة المجموعات المرتبة التي تشمل معاينة المجموعات المرتبة Ranked Set Sampling (RSS) ومعاينة المجموعات المرتبة المزدوجة Double Ranked Set Sampling (DRSS) باستخدام طريقتي للتقدير وهي طريقة الإمكان الأعظم(MLE) Maximum Likelihood Estimators، وطريقة الحد الأقصى لمقدرات التباعد Maximum Product Spacing (MPS). وباستخدام أسلوب المحاكاة مونت – كارلو لتوليد بيانات المتغير العشوائي الذي يتبع توزيع لوماكس لغرض المقارنة بين طرائق التقدير المختلفة لدالة الموثوقية، واستخدام المعيار الإحصائي متوسط مربعات الخطأ (MSE) تم التوصل إلى أن طريقة الحد الأقصى لمقدرات التباعد تعطي أفضل تقدير لدالة الموثوقية لتوزيع لوماكس عند استخدام أسلوب معاينة المجموعات المرتبة (RSS) بأحجام مجموعات مختلفة وعدد دورات مختلفة. وهذا ما تم إثباته من خلال النتائج التجريبية إذ كلما زاد حجم العينة كانت نتائج التقدير أفضل، إذ حققت طريقة الحد الأقصى لمقدرات التباعد نسبة أفضلية (93.333 %)، بينما عند استخدام طريقة معاينة المجموعات المرتبة المزدوجة (DRSS) لتقدير دالة الموثوقية لتوزيع لوماكس، كانت طريقة (MLE) هي الأفضل بنسبة (72.223%)،

ولبيان أهمية الموضوع من الناحية التطبيقية والواقعية، تم استخدام بيانات حقيقية عن مرضى سرطان المثانة، يمثل فيها المتغير العشوائي أزمنة سكون المرض بالأشهر (بداية الشفاء التام)

والمسجلة منذ دخول المرضى المستشفيات لغاية خروجهم منها، إذ تم اختبار أن البيانات تتبع توزيع لوماكس باستخدام اختبارات جودة المطابقة (أندرسون–دارلنك، كولموكوروف–سميرنوف، مربع كاي). تم استخدام طريقة (RSS) في سحب عينة بحجم (96) مريضاً من العينة الكلية البالغة (128) مريض وفقاً لعدد دورات (k=16) وعدد الوحدات في كل مجموعة (m=6) وباستخدام أفضل طريقة لتقدير دالة الموثوقية وهي (MPS)، أما عند استخدام طريقة (DRSS)، لغرض الحصول على عينة بحجم 96 من العينة الكلية تم سحب m=4 مجاميع وتم تكرار العملية 24 مرة وباستخدام أفضل طريقة لتقدير دالة الموثوقية وهي (MLE). لوحظ بان الموثوقية للشفاء من مرض سرطان المثانة تتناقص بزيادة أزمنة بقاء المرضى بالمشفى عند استخدام بروتوكول العلاج والتعافي.

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

Using Some Ranked Set Sampling Methods to Estimate The Two Parameters of The Lomax Distribution With Application to Bladder Cancer Data

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

Musa Mohammad Musa Farhan Al-Jaburi

Supervised by
Assis. Prof. Ban Ghanim Omar Al-Ani

2024 A.D **1445 A.H**

Abstract

To achieve the goals that the researcher seeks to achieve in his research, it is necessary to collect data and information for that research, and given the difficulty of obtaining some data or the inability to obtain it for reasons related to cost, effort and time, he must choose the best option. A sampling method that guarantees achieving the research objectives with the least time, cost and effort, or in many applied cases it may not be possible to obtain real measurements of the variable of interest, which is expensive and thus a waste of time. To solve these problems, the researcher must use a sampling method that ensures reducing the time, effort, and cost of obtaining data. In this case, it turns out that the Ranked Set Sampling (RSS) method is important for obtaining information that meets the purpose and that the researcher needs..In this thesis, the probability function was found to estimate the shape and scale parameters, and estimating the reliability function for the Lomax distribution, based on some designs of Ranked Set Sampling (RSS) and Double Ranked Set Sampling (DRSS) using four estimation methods, which are the Maximum Likelihood Estimators (MLE), Maximum Product of Spacings (MPS), By using the Monte-Carlo simulation method to generate random variable data that follows the Lomax distribution for the purpose of comparing different estimation methods for the reliability function, and using Mean Square Error (MSE) as the statistical criterion, it was concluded that the (MPS) method gives the best estimate of the reliability function for the Lomax distribution when using (RSS) with different group sizes and different number of cycles. This has been proven through experimental results, where the larger the sample size, the better the estimation results, as the method of (MPS) achieved a preference rate of (93.333%). While when using the (DRSS) method to estimate the reliability function of the Lomax distribution, the (MLE) method was the best with a percentage of (72.223).

To demonstrate the importance of the topic from a practical and realistic perspective, real data on bladder cancer patients was used, in which the random variable represents the disease remission times in

months (beginning of complete recovery) recorded from the time the patients entered the hospital until they left them. It was tested that the data follows a Lomax distribution using goodness-of-fit tests like Anderson-Darling (AD), Kolmogorov-Smirnov (K-S), and Chi-square. The (RSS) method was used to draw a sample size of (96) patients from the total sample of (128) patients according to the number of cycles (k=6) and the number of units in each group (m=16) and using the best method to estimate the reliability function, which is (MPS). While when using the (DRSS) method, a sample size of (96) patients was drawn according to the number of courses (48) and using the best method for estimating the reliability function, which is (MLE). It was observed that the reliability of recovery from bladder cancer decreases with increasing patient stay times in the hospital when using the treatment and recovery protocol.