مجلة كلية التربية الاساسية
كلية التربية الاساسية – الجامعة المستنصرية
Journal of the College of Basic Education Vol.31 (NO. 132) 2025, pp. 874-885

تقدير معالم توزيع باريتو المعمم ذو الثلاث معالم بأستعمال طريقتي العزوم و بيكاندز

اية قاسم علي ا.د. عدي طه رحيم
قسم الاحصاء/كلية الادارة و الاقتصاد/الجامعة المستنصرية
aya.ali@uomustansiriyah.edu.iq audaytaha73@gmail.com

مستخلص البحث:-
عند دراسة و تقييم الاحتمالات للاحداث النادرة ،سنستعمل توزيع باريتو المعمم (GPD) الذي يحتوي على التوزيعات الاسية. باريتو و بيتا كحالات خاصة. و لذلك سنركز على توزيع باريتو المعمم.و قد قام الباحث بتطبيقه حسب حجم العينة على عدة تطبيقات و من ضمنها التطبيق لبيانات مدة بقاء مرضى سرطان المعدة على قيد الحياة وتم التوصل الى ان البيانات تتبع توزيع باريتو المعمم ذو الثلاث معالم بأستعمال اختبار كولموكروف سميرنوف Kolmogorov-Smirnov(KS) وايضا اختبار اندرسون دارلينك Anderson-Darling(AD) و اختبار مربع كاي لايجاد اختبار حسن المطابقة.بعد تطبيق معيار متوسط مربعات الاخطاء M.S.E لطريقة العزوم و طريقة بيكاندز و توصل الباحث الى ان طريقة بيكاندز هي اول افضل طريقة تقدير للمعالم الثلاث (α, k, μ) و طريقة العزوم هي ثاني افضل طريقة تقدير للمعلمات الثلاث (α, k, μ).

الكلمات المفتاحية: توزيع باريتو المعمم ، طريقة العزوم ، طريقة بيكاندز ،سرطان المعدة،اختبار حسن المطابقة .

*** بحث مستل من رسالة الماجستير.**

(1) المقدمة Introduction

يعتبر توزيع باريتو المعمم احد التوزيعات المهمة في نمذجة القيم المتطرفة، و يعد توزيع باريتو المعمم ذات اهمية كبيرة في حيث يتم تطبيقه في الدراسات البيئية،و المالية،و مخاطر العمليات،و التأمين، و دراسات المعولية كتوزيع وقت الفشل حيث يعتبر توزيع باريتو المعمم احد التوزيعات المهمة في نمذجة القيم المتطرفة، و يعد توزيع باريتو المعمم ذات اهمية كبيرة حيث يتم تطبيقه في الدراسات البيئية، و المالية،و مخاطر العمليات،و التأمين،و كذلك تشمل تطبيقاته مستويات الاوزون في الغلاف الجوي العلوي،و الاحداث البيئية المتطرفة،و التقلبات الكبيرة في البيانات المالية و المطالبات الكبيرة في التأمين، ودراسات دالة البقاء او دالة المعولية كتوزيع لفترة البقاء او وقت الفشل،كما انه يلعب دورا مهما في نمذجة الاحداث المتطرفة،على سبيل المثال لتحليل بيانات هطول الامطار،و في تحليل تواتر الفيضانات و في تحليل بيانات اكبر ارتفاعات الامواج او مستويات سطح البحر،و الحد الاقصى لاحمال الرياح على المباني،وفي تحليل الحد الاقصى لسقوط الامطار،و في تحليل اكبر قيم الفيضانات السنوية،و قوة كسر المواد،و احمال الطائرات،و ما الى ذلك.[6][5][3]

(2) مشكلة البحث Research problem

الصعوبة في التعامل مع توزيع باريتو المعمم تكمن في تحديد شكله وتقدير معالمه، وذلك لأن هذا التوزيع يأتي بعدة أشكال تختلف حسب عدد المعلمات، والتي قد تكون اثنتين أو ثلاثاً أو أربعاً أو أكثر. وكلما زاد عدد المعلمات، زادت صعوبة تقديرها بدقة. كما أن الطرق التقليدية للتقدير غالبًا لا تؤدي إلى حلول مباشرة، بل تتطلب استخدام أساليب عددية أو تكرارية للوصول إلى التقديرات المناسبة. ومن التحديات الإضافية أيضًا كيفية اختيار البيانات التي يمكن اعتبارها تتبع توزيع باريتو المعمم، مما يزيد من تعقيد الدراسة والتحليل.

(3) هدف البحث The Goal of The Reseaech

دراسة توزيع باريتو المعمم ذو الثلاث معالم (Generalized Pareto Distribution - GPD) مع التركيز على الخصائص والمميزات له، بالإضافة إلى دراسة العزوم. كما تم استعراض طرق تقدير معلمات التوزيع، وتحديد الطريقة الأنسب باستخدام معيار مناسب للمقارنة. وتقدير معالمه باستخدام طريقتي تقدير هما: طريقة العزوم وطريقة بيكاندز، ثم تمت مقارنة النتائج لتحديد الطريقة الأفضل.

(4) منهجية البحث

في هذا البحث، تم استعمال توزيع باريتو المعمم ذو الثلاث معالم بسبب المرونة التي يتمتع بها في تمثيل و تحليل البيانات، و تم تطبيق بيانات حقيقية تخص مدة بقاء مرضى سرطان المعدة على قيد الحياة لـ(200) حالة، كذلك قمنا بتحديد شكل التوزيع لهذه البيانات. بأستعمال اختبار حسن المطابقة Goodness of fit و ذلك بأستخدام البرنامج الجاهز Easy Fit 5.5 حيث يتم استعمال اختبار مربع كاي Chi-square وكذلك اختبار كولموكروف سميرنوف Kolmogorov-Smirnov(KS) وايضا اختبار اندرسون دارلينك Anderson-Darling(AD) . في الجانب النظري تم عرض طريقتي تقدير هما طريقة العزوم و طريقة بيكاندز . و في الجانب التطبيقي تم تطبيق طريقة العزوم و طريقة بيكاندز بأستخدام برنامج iXplore و من اجل معرفة ايي طريقة تكون هي الافضل قمنا بأستعمال معيار متوسط مربعات الخطأ mse و كذلك قمنا بأستعمال معيار الانحراف المعياري و التباين و معيار التفلطح كذلك لتحديد اي طريقة الافضل.

5- الجانب النظري

(1-5) توزيع باريتو المعمم ذو الثلاثة معالم

Three-Parameter Generalizad Pareto Distribution

قام بيكاندز بدراسة توزيع باريتو المعمم*(GPD)*، و كشف عن اهميته الكبيرة في نظرية القيم المتطرفة.حيث يتكون من ثلاثة معالم(α, k, μ) هما معلمة المقياس و معلمة الشكل و معلمة الموقع على التوالي*[10]*.

تم تطويره ايضا ليصبح عائلة من التوزيعات المعروفة بتوزيع باريتو المعمم*(GPD)*،

يتعلق توزيع باريتو المعمم بثلاث معلمات :معلمة المقياس (التي يرمز لها ب α) ،معلمة الشكل (التي يرمز لها ب k) ،و معلمة الموقع (التي يرمز لها ب μ).يكون لدى المتغير العشوائي X توزيع باريتو المعمم اذا كانت دالة التوزيع التراكمي له بالصيغة الاتية: *[8]*

$$F(x) = \begin{cases} 1 - \left(1 - \dfrac{k(x-\mu)}{\alpha}\right)^{1/k} & , if\ \ k \neq 0 \\ 1 - exp\left(-\dfrac{x-\mu}{\alpha}\right) & , if\ k = 0 \end{cases}$$

و دالة الكثافة الاحتمالية تكون بالصيغة الاتية :

$$f(x) = \begin{cases} (1/\alpha)\left(1 - \dfrac{k(x-\mu)}{\alpha}\right)^{\frac{1}{k}-1} , & if\ \ k \neq 0 \\ (1/\alpha)exp\left(-\dfrac{x-\mu}{\alpha}\right) , & if\ \ k = 0 \end{cases}$$

حيث ان دالة الكثافة الاحتمالية تكون موجبة عندما $\mu \geq x$ اذا كان $k \leq 0$ ، او عندما $\mu \leq x \leq \alpha/k$ اذا كان $k > 0$.

حالة خاصة من توزيع باريتو المعمم تحدث عندما تعتبر معلمة الشكل k صفرا ،و النتيجة تكون معادلة للتوزيع الاسي.

α : معلمة المقياس $\alpha \in (0,\infty)$

K : معلمة الشكل $K \in (-\infty,\infty)$

μ : معلمة الموقع $\mu \in (-\infty,\infty)$

X : متغير عشوائي او قيمة متطرفة $\mu \leq x \leq \alpha/k$

حيث ان دالة البقاء تكون بالشكل الاتي: [7]

$$S(x) = 1 + k\left(\frac{x-\mu}{\alpha}\right)^{\frac{1}{k}}$$

و دالة المعولية تكون بالشكل التالي [2]:

$$f(x) = \begin{cases} \left(1 + \dfrac{k(x-\mu)}{\alpha}\right)^{-1/k}, & k \neq 0 \\ exp\left(-\dfrac{(x-\mu)}{\alpha}\right), & k = 0 \end{cases}$$

و يمكن كتابة صيغة المتوسط كالاتي [7]:

$$E(x) = \mu + \frac{\alpha}{1-k} \quad , \quad k > 0$$
$$E(x) = \mu + \alpha \quad , \quad k = 0$$

و يمكن كتابة صيغة التباين كالاتي[7] :

$$var(x) = \frac{\alpha^2}{(1-k)^2(1-2k)} \quad , \quad k > 0$$
$$var(x) = \alpha^2 \quad , \quad k = 0$$

(2-5) طرائق التقدير

(1-2-5) طريقة العزوم [9]

تم اشتقاق طريقة العزوم لتوزيع باريتو المعمم بواسطة هوسكينغ واليس(1987).لاحظ ان $E\left(1 - \frac{k(x-\mu)}{\alpha}\right)^r = \frac{1}{(1+kr)}$ $if\ 1 + rk > 0.$ ،العزم r ل x موجود اذا كان $k > -1/r$. بشرط ان تكون موجودة،فأن طريقة العزوم هي

$$\bar{X} = \mu + \frac{\alpha}{1+k} \quad ...(1)$$

$$S^2 = \frac{\alpha^2}{(1+k)^2(1+2k)} \quad ...(2)$$

$$G = \frac{2(1-k)(1+2k)^{0.5}}{1+3k} \quad \ldots (3)$$

حيث ان $\bar{X}$ و S^2 و G هي المتوسط و الانحراف المعياري و الالتواء للعينة على التوالي. اولا،يتم الحصول على طريقة العزوم ل k عن طريق حل المعادلة (3).يتم توضيح العلاقة بين G و k .بعد حساب k،يتم الحصول على α و μ من المعادلتين (1) و(2) كما يلي

$$\alpha = S(1+k)(1+2k)^{0.5} \quad \ldots (4)$$

$$2\mu = \bar{x} - \frac{\alpha}{\alpha+k} \quad \ldots (5)$$

(2-2-5)طريقة بيكاندز Pickands methods[4]

$$F(x;k,\alpha) = \begin{cases} 1-(1-kx/\alpha)^{1/k}, & k \neq 0, \alpha > 0 \\ 1-exp(-x/\alpha), & k = 0, \alpha > 0 \end{cases} \quad \ldots (6)$$

على الرغم من ان هذا الطرح له بعض المزايا الا انه لا يزال من الممكن تقديم اسباب عملية و نظرية لعدم فرض قيود على المعلمات.

اولا: هناك قيم ل $k > 1/2$ يتم مواجهتها فس التطبيقات العملية. علاوة على ذلك، فأن التوزيع المنتظم، الذي يستعمل في العديد من التطبيقات الاحصائية هو حالة خاصة من توزيع باريتو المعمم (*GPD*) عندما يكون $k=1$.امثلة اخرى على البيانات الواقعية حيث تقع تقديرات k خارج نطاق(-5,5) تلاحظ بشكل شائع في البيانات ذات الذيل الثقيل و البيانات المتقطعة(مثل البيانات التي يتم الحصول عليها بأستخدام اجراءات ا لاقتطاع).

ثانيا: دالة التوزيع التراكمي$F(x;k,\alpha)$في معادلة(2-6) هي دالة توزيع احتمالي صحيحة لكل القيم التي يكون فيها $\alpha > 0$ و $-\infty < k < \infty$ ،و بالتالي، من المثير للاهتمام، على الاقل من منظور نظري، تقدير التوزيع لجميع القيم الممكنة لمعاملاته. عندما $k \neq 0$.يكزن من الاكثر ملائمة احيانا العمل مع نسخة معاد تهيئتها من توزيع باريتو المعمم (*GPD*) في هذه الحالة يمكن كتابة *GPD* بالصيغة الاتية :

$$F(x) = 1-(1-x/\delta)^{1/k} \quad ,k \neq 0\,,\delta k > 0 \quad \ldots (7)$$

حيث $\delta = \alpha/k$

عندما يكون $K=0$، فأن توزيع *GPD* يتحول الى توزيع اسي بمتوسط α. تقدير αفي هذه الحالة سهل ،لان التقدير بطريقة الامكان الاعظم (*MLE*) يكون فعالا . من الان فصاعدا، سنركز على الحالة الاكثر صعوبة عندما يكون $K \neq 0$.

فكرة طريقة ال *EPM* هي الاستفادة بالكامل من المعلومات الموجودة في احصاءات الرتب عن طريق معادلة الدالة للتوزيع التراكمي (*CDF*) عند احصاءات الرتب التي يتم حسابها من العينة(*observed order statistics to their*) مع قيمها المئوية المقابلة لها، ثم استخدام المعادلات الناتجة كأساس للحصول على تقديرات اولية للمعلمات(القسم*1-3*) ثم يتم دمج هذه التقديرات بطريقة مناسبة للحصول على التقديرات النهائية للمعلمات(القسم *2-3*) .

3-1التقديرات الاولية

لتكن $x_{i:n}$ و $x_{j:n}$ هما احصائيتان من احصاءات الرتب المميزة في عينة عشوائية من الحجم n من التوزيع $F(x;k,\alpha)$. بمعادلة دالة التوزيع التراكمي *(CDF)* عند احصاءات الرتب الملاحظة مع قيمها المئوية المقابلة ، نحصل على :

$$F(x_{i:n};k,\alpha) = P_{i:n}$$

و

$$F(x_{j:n};k,\alpha) = P_{j:n} \quad \dots(8)$$

حيث

$$P_{i:n} = \frac{i - \Upsilon}{n + \beta} \quad \dots(9)$$

هو موضع رسم مناسب. في دراسة المحاكاة ،جربنا عدة قيم □ و β و وجدنا ان 0 = □ و 1 = β تعطي نتائج افضل للطريقة المقترحة .

عند استبدال(6) في (7) و اخذ اللوغاريتم، نحصل على :

$$\log(1 - x_{i:n}/\delta) = kC_i$$

و

$$\log(1 - x_{j:n}/\delta) = kC_j \quad \dots(10)$$

حيث ان $C_i = \log(1 - P_{i:n}) < 0$. يمكن ملاحظة ان المعادلة(8) هي نظام مكون من معادلتين في مجهولين، δ و k . ازالة k، نحصل على :

$$C_i \log(1 - x_{j:n}/\delta) = C_j \log(1 - x_{i:n}/\delta) \quad \dots(11)$$

بينما عند ازالة من δ، نحصل على :

$$x_{i:n}\left[1-(1-P_{j:n})^k\right] = x_{j:n}\left[1-(1-P_{i:n})^k\right] \quad \dots(12)$$

كل من المعادلتين (11)و (12) هي دالة لمتغير واحد فقط، و بالتالي يمكن حلها بسهولة بأستعمال طريقة القسمة كما هو موضح في النظرية 1 و الخوارزمية 1. و بالتالي، بأستخدام الخوارزمية 1، يمكن للمرء حل المعادلة (11)،على سبيل المثال، للحصول على δ و الحصول على مقدر ل δ . و الذي يمكن نرمز له ب $\hat{\delta}(i,j)$، ثم يتم استبدال هذا التقدير في احدى المعادلتين في (10) للحصول على تقدير مقابل ل K، و هو $\hat{k}(i,j)$ ، و الذي يعطى بواسطة :

$$\hat{k}(i,j) = \log\left(1 - x_{i:n}/\hat{\delta}(i,j)\right)/C_i \quad \dots(13)$$

يمكن بعد ذلك حساب مقدرل αعلى النحو التالي :

$$\hat{\alpha}(i,j) = \hat{k}(i,j)\hat{\delta}(i,j)$$

و بالمثل، يمكن حل المعادلة(12) للحصول على $\hat{k}(i,j)$ ثم استبداله ب $\hat{k}(i,j)$ في المعادلة (10) للحصول على $\hat{\delta}(i,j)$

تم الحصول على المقدرات التي اقترحها بيكاندز (1975)عن طريق تعيين $i = n/2$ و $j = 3n/4$. و بالتالي ،فهي حالة خاصة من المقدرات المقترحة بالنسبة لهذه القيم من i,j،يكون لنظام المعادلات في

(10) على حل مغلق، الذي يتم الحصول عليه كما يلي عند استبدال هذه القيم ل i,j في معادلة (10) نحصل على :

$$\log\left(1 - x_{n/2:n}/\delta\right) = k\log\left(1 - \frac{1}{2}\right)$$

و

$$\log\left(1 - x_{3n/4:n}/\delta\right) = k\log\left(1 - \frac{3}{4}\right)$$

مما يؤدي الى

$$\log\left(1 - x_{n/2:n}/\delta\right) - \log\left(1 - x_{3n/4:n}/\delta\right) = k\log(2)$$

من هذه المعادلة يمكننا التعبير عن k كدالة ل δ

$$K = \frac{1}{\log(2)}\log\left(\frac{\delta - x_n/2:n}{\delta - x_{3n}/4:n}\right) \quad \ldots(14)$$

للحصول على تعبير ل δ، تكتب المعادلة

$$2\log\left(1 - x_{n/2:n}/\delta\right) = \log\left(1 - x_{3n/4:n}/\delta\right)$$

مما يعني

$$\left[1 - x_{n/2:n}/\delta\right]^2 = 1 - x_{3n/4:n}/\delta$$

مما يترتب على ذلك ان

$$\hat{\delta} = \frac{x^2_{n/2:n}}{2x_{n/2:n} - x_{3n/4:n}} \quad \ldots.(15)$$

وهذا هو مقدر بيكاندز δ عند استبدال المعادلة (15) في المعادلة (14) نحصل على

$$\hat{k} = \frac{1}{\log(2)}\log\left(\frac{x_{n/2:n}}{x_{3n/4:n} - x_{n/2:n}}\right) \quad \ldots(16)$$

و هذه مقدرات بيكاندز ل k .

نعرض الان ان المقدرات السابقة معرفة بشكل جيد، اي ان المعادلة (11)تحتوي على حل اضافي بجانب الحل البديهي(*trivial solution*)ل $\delta = \pm\infty$ او ، بشكل مكافئ ، المعادلة(12)تحتوي على حل اضافي بجانب الحل البديهي عند *K=0*.

6- الجانب التطبيقي

(1-6)مقدمة :-

في هذا الفصل سيتم استعمال بيانات حقيقية و المتمثلة في عدة تطبيقات مدة بقاء مرضى سرطان المعدة على قيد الحياة في تنفيذ الجانب التطبيقي و اجراء اختبار حسن المطابقة Goodness of fit وذلك بأستعمال البرنامج الجاهزEasy Fit 5.5 حيث يتم استعمال اختبار مربع كاي Chi-square و كذلك اختباركولموكروف سميرنوفKolmogorov-Smirnov(K-S) و ايضا اختبار اندرسون دارلينك Anderson-Darling(A-D)للبيانات لمعرفة مدى ملائمة البيانات لتوزيع (GPD)،كما سيتم حساب مقدرات معلمات توزيع(GPD) بالاعتماد على جميع طرائق التقدير الواردة بالجانب النظري و من ثم تطبيقه على برنامج iXploRe ومن ثم تحديد اي من الطرق كانت هي التي تحقق نتائج افضل.

(2-6)جمع البيانات

في هذا التطبيق، تم الحصول على البيانات الحقيقية من وزارة الصحة العراقية، تمثل هذه البيانات على حالات الإصابة المسجلة خلال عامي (2021 و2022). وقد شملت هذه البيانات على الفترة الزمنية من تاريخ دخول المرضى إلى المستشفى وحتى تاريخ الخروج، سواء بالشفاء أو الوفاة. يوضح الجدول أدناه توزيع مدة البقاء على قيد الحياة، معبّراً عنها بالأشهر، وهو ما يُسهم في تخمين فترة البقاء لـ(200) مريض[1].

جدول (1) بيانات مدة بقاء مرضى سرطان المعدة على قيد الحياة.

n	S	n	S	n	S	n	S
1	8.367	51	2.800	101	3.300	151	1.867
2	0.267	52	0.633	102	4.433	152	0.667
3	0.700	53	1.000	103	1.067	153	1.167
4	1.567	54	1.033	104	2.533	154	0.500
5	6.167	55	4.700	105	0.467	155	2.467
6	7.100	56	3.667	106	5.800	156	6.400
7	0.600	57	2.167	107	2.367	157	0.433
8	0.800	58	0.900	108	11.500	158	4.267
9	1.867	59	1.133	109	2.067	159	3.967
10	1.600	60	2.700	110	2.700	160	5.133
11	2.067	61	11.333	111	7.867	161	0.567
12	5.133	62	3.167	112	0.267	162	1.167
13	4.833	63	5.967	113	1.933	163	0.567
14	6.567	64	1.600	114	2.433	164	0.500
15	10.733	65	0.533	115	0.900	165	0.467
16	0.800	66	0.533	116	3.100	166	0.900
17	0.367	67	2.600	117	3.933	167	2.100
18	3.133	68	0.733	118	0.767	168	4.867
19	9.367	69	1.167	119	0.800	169	0.367
20	5.167	70	2.400	120	1.833	170	0.267
21	1.033	71	2.000	121	3.500	171	2.200
22	4.733	72	4.800	122	0.467	172	0.600
23	2.200	73	0.567	123	2.867	173	6.100
24	1.900	74	1.733	124	7.133	174	2.400
25	10.133	75	4.733	125	0.667	175	7.933
26	0.967	76	1.400	126	4.733	176	1.233
27	2.633	77	1.300	127	0.400	177	0.467
28	5.100	78	0.733	128	3.533	178	1.500
29	8.567	79	1.633	129	1.333	179	5.600
30	4.167	80	0.533	130	1.500	180	2.700

31	3.767	81	2.900	131	1.300	181	0.900
32	1.500	82	0.333	132	0.467	182	0.633
33	3.900	83	1.200	133	1.033	183	2.300
34	2.100	84	0.567	134	6.233	184	0.300
35	0.500	85	4.033	135	1.233	185	0.867
36	1.967	86	5.567	136	8.433	186	7.500
37	2.233	87	1.233	137	2.367	187	4.000
38	1.033	88	0.667	138	5.133	188	0.633
39	1.267	89	1.433	139	0.300	189	1.433
40	4.400	90	4.067	140	1.700	190	1.067
41	2.933	91	9.200	141	0.467	191	0.800
42	1.967	92	2.833	142	1.533	192	1.700
43	2.267	93	1.267	143	0.500	193	1.300
44	0.533	94	1.167	144	0.867	194	2.367
45	5.500	95	9.367	145	1.133	195	0.800
46	8.000	96	0.933	146	0.867	196	10.333
47	5.100	97	0.900	147	0.267	197	11.900
48	1.067	98	0.400	148	4.433	198	1.200
49	1.867	99	8.500	149	1.600	199	0.400
50	1.900	100	2.567	150	1.500	200	0.433

(3-6)وصف البيانات

في البيانات تمثل مدة بقاء مرضى سرطان المعدة في المستشفى للمدة الزمنية (2021 و 2022)، تمثل هذه لـ(200) مريض.

(4-6)تحليل البيانات الاحصائية

بعد ان تم تهيئة البيانات الخاصة بمدة بقاء مرضى سرطان المعدة و الموجودة في الجدول (1). سنقوم بتحديد شكل التوزيع لهذه البيانات. و للتوصل الى شكل التوزيع و لهذه البيانات تم استعمال اختبار حسن المطابقة Goodness of fit و ذلك بأستخدام البرنامج الجاهز Easy Fit 5.5 حيث يتم استعمال اختبار مربع كاي Chi-square وكذلك اختبار كولموكروف سميرنوف Kolmogorov-Smirnov(KS) وايضا اختبار اندرسون دارلينك Anderson-Darling(AD) و ان الصيغة المستخدمة لاختبار مربع كاي لايجاد اختبار حسن المطابقة هي :

$$x^2_{call} = \frac{\sum_{i=1}^{n}(O_i - E_i)^2}{O_i}$$

$$x^2_{(200,)} =$$

اما القيم الحقيقية للمعلمات فأن البرنامج الجاهز Easy Fit 5.5 جهز لنا قيم حقيقية للمعلمات (k,μ,α). لتوزيع باريتو المعمم ذو الثلاث معالم

و ذلك اثناء اجراء اختبار حسن المطابقة لهذا التوزيع.و الجدول (2) يبين القيم الحقيقية الى (k,μ,α) و لكل متغير كالاتي.

مجلة كلية التربية الاساسية
كلية التربية الاساسية – الجامعة المستنصرية
Journal of the College of Basic Education Vol.31 (NO. 132) 2025, pp. 874-885

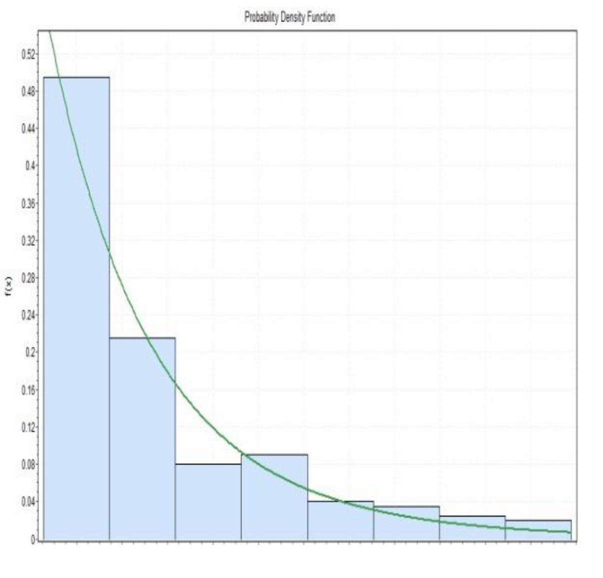
جدول (2) يبين القيم الحقيقية لمعلمات توزيع باريتو المعمم ذو الثلاث معالم (k,μ,α) حسب برنامج Easy Fit 5.5

المتغير	α	μ	k
بقاء مرضى سرطان المعدة	2.3953	0.16014	0.07434

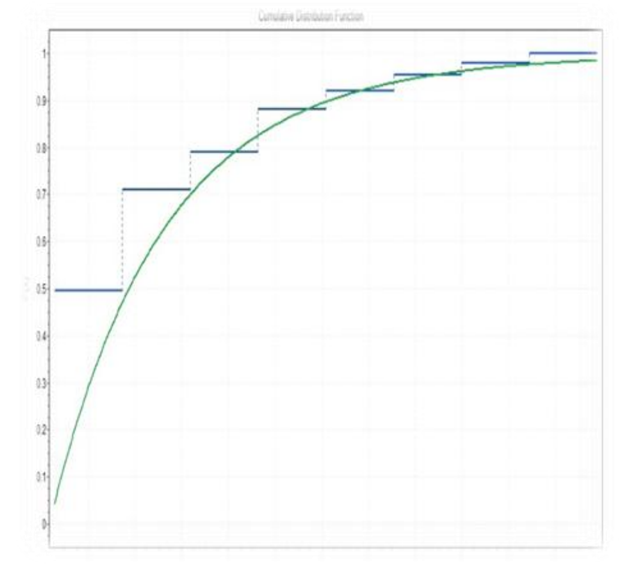
جدول (3) يبين اختبار حسن المطابقة لمتغيرات الجدول (1).

المتغير	K-S	R	A-D	R	Chi-square	R	Test
بقاء مرضى سرطان المعدة	0.04995	7	0.90219	6	4.8791	4	Yes

و نلاحظ من الجدول (3) بأن المتغيرات قبلت توزيع باريتو المعمم ذو الثلاث معالم. حيث قبلت اختبار حسن المطابقة الذي اظهرت نتائجه باستخدام البرنامج الجاهز Easy Fit 5.5.و الذي يظهر نتيجة الاختبار لحسن المطابقة لتوزيع باريتو المعمم ذو الثلاث معالم.



أ- دالة الكثافة الاحتمالية(P.D.F) لتوزيع باريتو المعمم ذو الثلاث معالم



ب- دالة التوزيع التراكمي (C.D.F)لتوزيع باريتو المعمم ذو الثلاث معالم.

شكل (1)

دالة الكثافة الاحتمالية P.D.F و دالة التوزيع التراكمية C.D.F لتوزيع باريتو المعمم ذو الثلاث معالم .

بعد ان تم تحليل البيانات الموجودة في الجدول (1) و معرفة بأنها تتوزع توزيع باريتو المعمم ذو الثلاث يمكننا الان تقدير المعالم لهذا التوزيع بأستخدام طرائق التقدير المحددة في الجانب النظري

فعند تقدير المعالم (K, μ, α) حسب طريقة العزوم و طريقة بيكاندز يمكننا استخدم المعادلات الموجودة في الجانب النظري اما بالنسبة لطريقة بيكاندز فلتقدير المعالم (K, μ, α) ايضا نقوم باستخدام المعادلات الموجودة في الجانب النظري و لكل طريقة و كالاتي :

جدول(4) يبين تقدير المعالم (K, μ, α)
لتوزيع باريتو المعمم ذو الثلاث معالم.

المعلمة	moment	Pickands
α	3.4769	25.646
µ	-0.4523	-19.303
K	-0.11604	-0.83172

و لتحديد اي من طرق التقدير تكون هي الافضل تم حساب التباين و الانحراف المعياري و التفلطح لكل طريقة و الجدول (5) يبين ذلك.

جدول (5) يبين التباين و الانحراف المعياري و معامل التفلطح لكل طريقة

الطريقة	Variance	St.d	Kurtosis
Moment	9.705	1.6156	-0.22121
Pickands	176.03	0.25647	-0.037115

من الجدول (5) و بمقارنة الانحراف المعياري للطريقتين .نلاحظ بأن افضل طريقة هي طريقة Pickands و التي امتلكت. اقل انحراف معياري و قيمته ((st.d=0.25647)) و ايضا كان لها معامل التفلطح و قيمته ((-0.037115)) و ايضا كان لها تباين و قيمته ((176.03)).

و الطريقة الاخيرة كانت طريقة العزوم التي كان لها اكبر انحراف معياري و قيمته((st.d=1.6156)) ومعامل التفلطح كانت قيمته ((-0.22121)) اما تباينها كانت قيمته ((9.705)).

جدول (6) يوضح معيار متوسط مربعات الخطأ

المعلمة	MSE(M.O.M)	MSE(Picckands)	Best MSE
α	13.438	1.153	Pick, mom
µ	8.273	0.757	Pick,mom
K	0.208	0.136	Pick, mom

(5-6)الاستنتاجات

1- بعد مقارنة الانحراف المعياري للطريقتين نلاحظ ان طريقة بيكاندز و التي امتلكت اقل انحراف معياري و كانت قيمته (0.25647) و تاتي بعدها طريقة العزوم و التي امتلكت اكبر انحراف معياري و كانت قيمته (1.6156).

2- بعد ان تم تحديد الطريقة الافضل باستخدام معيار متوسط مربعات الخطأM.S.E للطريقتين للمعالم الثلاثة فاستنتج الباحث ان طريقة بيكاندز هي الافضل للمعالم الثلاثة و تاتي بعدها طريقة العزوم للمعالم الثلاثة.

3- نلاحظ انه عند اختبار حسن المطابقة Goodness of fit تمكنا من الحصول على قيمة لمربع كاي وكذلك اختبار كولموكروف سميرنوف Kolmogorov-Smirnov(KS) وايضا اختبار اندرسون دارلينك Anderson-Darling(AD) .

Reference المصادر

1- مزهر،احمد سلام،(2024)،"استعمال الطريقة البيزية لتقدير بعض نماذج انحدار البقاء مع التطبيق"رسالة ماجستير مقدمة الى مجلس كلية الادارة و الاقتصاد،الجامعة المستنصرية.

2- Abd Elfattah, A. M., Elsherpieny, E. A., & Hussein, E. A. (2007). A new generalized Pareto distribution. *Interstat*, *12*, 1-6.

3- Abdulali, B. A. A., Abu Bakar, M. A., Ibrahim, K., & Mohd Ariff, N. (2022). Extreme value distributions: an overview of estimation and simulation. *Journal of Probability and Statistics*, *2022*(1), 5449751

4- Castillo, E., & Hadi, A. S. (1997). Fitting the generalized Pareto distribution to data. *Journal of the American Statistical Association*, *92*(440), 1609-1620.

5- Dupuis, D. J., & Tsao, M. (1998). A hybrid estimator for generalized Pareto and extreme-value distributions. *Communications in Statistics-Theory and Methods*, *27*(4), 925-941.

6- Kang, S., & Song, J. (2017). Parameter and quantile estimation for the generalized Pareto distribution in peaks over threshold framework. *Journal of the Korean Statistical Society*, *46*(4), 487-501.

7- Muraleedharan, G., & Soares, C. G. (2014). Characteristic and moment generating functions of generalised Pareto (GP3) and Weibull distributions. *Journal of Scientific Research and reports*, *3*(14), 1861-1874.

8- Salahuddin, N., Khalil, A., Mashwani, W. K., Alrajhi, S., Al-Marzouki, S., & Shah, K. (2021). On the Properties of the new generalized Pareto distribution and its applications. *Mathematical Problems in Engineering*, *2021*(1), 6855652.

9- Singh, V. P., & Guo, H. (1995). Parameter estimation for 3-parameter generalized Pareto distribution by the principle of maximum entropy (POME). *Hydrological sciences journal*, *40*(2), 165-181.

10-. Zhou, C. R., Chen, Y. F., Huang, Q., & Gu, S. H. (2017, August). Higher moments method for generalized Pareto distribution in flood frequency analysis. In *IOP Conference Series: Earth and Environmental Science* (Vol. 82, No. 1, p. 012031). IOP Publishing.

مجلة كلية التربية الاساسية
كلية التربية الاساسية – الجامعة المستنصرية
Journal of the College of Basic Education Vol.31 (NO. 132) 2025, pp. 874-885

Estimating the parameters of the three-parameter Generalized Pareto Distribution using the Method of Moments and Pickands' Method

P.Dr. Auday Taha Raheem **Aya Qasim Ali**
College of Administration and Economics - Department of Statistics Al- Mustansiriyah University
audaytaha73@gmail.com aya.ali@uomustansiriyah.edu.iq

Abstract:
When studying and evaluating the probabilities of rare events, the Generalized Pareto Distribution (GPD) is used, which includes the exponential, Pareto, and Beta distributions as special cases. Therefore, the focus is on the Generalized Pareto Distribution. The researcher applied the GPD according to the sample size in several applications, including the analysis of survival times of stomach cancer patients. It was found that the data follow a three-parameter Generalized Pareto Distribution based on the Kolmogorov-Smirnov (K-S) test, the Anderson-Darling (A-D) test, and the Chi-square goodness-of-fit test.
After applying the Mean Squared Error (MSE) criterion to both the Method of Moments and Pickands' method, the researcher concluded that Pickands' method is the best estimation method for the three parameters (α, k, μ), while the Method of Moments is the second-best estimation approach for these parameters.
Keywords: Generalized Pareto distribution, Method of Moments, Method of Pickands, Stomach Cancer, Goodness-of-Fit Test.