

اقتراح توزيع Topp-Leone المحول بصيغة Transmuted lower Records type Topp-Leone distribution proposal for the transmuted lower records type with practical application

أ.م.د صدى فايدن محمد

Ass. Prof. Dr. Sada Fayidh Mohammed

sada.f@uokerbala.edu.iq

كلية الادارة والاقتصاد جامعة كربلاء، كربلاء

College of Administration and Economics, University of Karbala

دعا احمد رسول

Duaa Ahmed Rasol

duaa.ahmed@s.uokerbala.edu.iq

كلية الادارة والاقتصاد جامعة كربلاء، كربلاء

College of Administration and Economics, University of Karbala

أ.د شروق عبد الرضا السباح

Prof. Dr Shrook A.S AL-Sabbah

shorouq.a@uokerbala.edu.iq

كلية الادارة والاقتصاد جامعة كربلاء، كربلاء

College of Administration and Economics, University of Karbala

المستخلص:

يهدف البحث إلى اقتراح وتطبيق توزيع احتمالي جديد يُعرف بتوزيع Transmuted Lower Record Type (TLRT-Topp-Leone)، والذي تم تطويره لتعزيز مرونة توزيع Topp-Leone الكلاسيكي عبر إدخال تحويل رياضي يزيد من قدرته على تمثيل البيانات ذات السلوك اللامتناظر أو الطويلة الذيل. يتلاؤ الجانب النظري استعراضًا مكملاً للخصائص الرياضية للتوزيع الجديد، بما في ذلك دالة الكثافة الاحتمالية ودالة التوزيع التراكمي ودالة البقاء، إضافةً إلى الاشتغال الرياضي لطريقة الإمكان الأعظم (MLE) لتقيير معلمات النموذج. وفي الجانب التطبيقي، تم تحليل بيانات فعلية لعدد بقاء 150 مريضًا مصابًا بفيروس COVID-19 في مستشفى الحسين بمحافظة كربلاء للفترة (2020–2024). جرى تقييم ملاءمة النموذج المقترن بالتوزيع الأصلي باستخدام اختباري Kolmogorov–Smirnov و Anderson–Darling، فضلًا عن معايير المقارنة AIC و BIC و CAIC. وقد بينت النتائج أن التوزيع المحول TLRT–Topp–Leone أكثر كفاءة ودقة في توصيف البيانات التجريبية، مما يجعله إضافة مهمة إلى عائلة التوزيعات الاحتمالية المرنة، ويفتح المجال أمام تطبيقات واسعة في تحليل بيانات البقاء الطبي والإحصاء التطبيقي.

الكلمات الافتتاحية : توزيع Topp–Leone ، تحويل ، بيانات البقاء ، تقدير الإمكان الأعظم (MLE) ، اختبارات الملاعة الإحصائية .

Abstract

This study proposes and applies a new probability distribution, referred to as the Topp–Leone distribution with Transmuted Lower Record Type transformation (TLRT–Topp–Leone). The proposed model extends the flexibility of the classical Topp–Leone distribution by incorporating a mathematical transformation that enhances its ability to capture asymmetric and heavy-tailed data. The theoretical framework provides a comprehensive derivation of the main statistical properties, including the probability density function, cumulative distribution function, and survival function, in addition to the formulation of the maximum likelihood estimation (MLE) method for parameter estimation. On the empirical side, real survival data of 150 COVID-19 patients admitted to Al-Hussein Hospital in Karbala during the period 2020–2024 were analyzed. The performance of the proposed model was evaluated against the original Topp–Leone distribution using the Kolmogorov–Smirnov and Anderson–Darling goodness-of-fit tests, as well as the information criteria AIC, BIC, and CAIC. The results demonstrated that the TLRT–Topp–Leone distribution provides a superior and more accurate fit to the observed data. Consequently, the proposed model constitutes a significant contribution to the family of flexible probability distributions and offers wide potential for applications in medical survival analysis and applied statistics.

Keywords: Topp-Leone distribution, Transmuted Lower Record Type, survival data, maximum likelihood estimation (MLE), goodness-of-fit tests

1. المقدمة

يمثل توزيع Topp-Leone أحد التوزيعات الاحتمالية التي نالت اهتماماً ملحوظاً في مجال تحليل بيانات الأعمار والموثوقية، نظراً لما يتمتع به من مرونة وقدرة على تمثيل البيانات غير المتماثلة والتي تتسم بسلوك غير خطى.

وقد ظهرت النواة الأولى لهذا التوزيع في منتصف القرن العشرين، وتحديداً في عام 1955، حين اقترحه Top وآخرون بهدف معالجة محدودية التوزيعات التقليدية في توصيف البيانات ذات السلوك المتغير. ومنذ ذلك الحين، شهد هذا التوزيع سلسلة من التطويرات والتوسعات، سواء من خلال إدخال معلمات جديدة، أو عبر دمجه مع توزيعات أخرى، أو من خلال تطبيق تحويلات احتمالية متقدمة هدفت إلى تحسين قدرته على التكيف مع أنماط البيانات المعقدة.

وقد تعددت الاتجاهات البحثية التي تناولت هذا التوزيع، فبعضها ركز على الجوانب النظرية والبنوية من خلال تحليل الخصائص الرياضية الأساسية له، مثل دوال الكثافة، التوزيع التراكمي، ودالة الخط. بينما اتجهت دراسات أخرى نحو تطوير عائلات توزيعية مشقة منه، من بينها توزيعات Transmuted G وExponentiated Topp-Leone، والتي جاءت لتلبية متطلبات تطبيقية أكثر تخصصاً في تحليل بيانات البقاء.

2. منهجية البحث

2.1 مشكلة البحث

تمثل مشكلة البحث في أن التوزيعات الاحتمالية الكلاسيكية تقصر إلى المرونة الكافية لتمثيل بيانات البقاء وبيانات الأعمار ذات السلوك غير المتماثل أو المحتوية على قيم متطرفة، مما يضعف دقة التقدير وكفاءة النمنجة. كما أن الأدبيات الإحصائية ما زالت محدودة في توظيف تحويلات من نوع Transmuted Lower Record Type على توزيعات مرنة مثل Topp-Leone، الأمر الذي يستدعي تطوير نماذج أكثر قرابة علىمحاكاة هذه الخصائص.

2.2 هدف البحث

يهدف البحث إلى تطوير نموذج احتمالي مرن يعتمد على توزيع Topp-Leone من خلال تطبيق تحويل إحصائي بصيغة Transmuted Lower Records Type، وذلك من خلال صياغة توزيع إحصائي جديد يُعرف بتوزيع Topp-Leone . واعتماد طريقة الإمكان الأعظم (MLE) لتقدير معلمات التوزيع المقول بصيغة Transmuted Lower Records Type المقترن.

2.3 المتغيرات العشوائية المستمرة ودوالها الاحتمالية

المتغير العشوائي هو متغير يتخذ قيمًا عدديًا تعتمد على نتائج تجربة عشوائية. وينقسم إلى متغيرات متقطعة تأخذ قيمًا محددة، ومتغيرات مستمرة تأخذ قيمًا غير منتهية ضمن مدى معين. في تحليل بيانات البقاء، غالباً ما يُستخدم المتغير العشوائي المستمر لتمثيل الزمن حتى وقوع حدث معين، مثل الوفاة أو الفشل أو الشفاء. ويرتبط هذا المتغير بعدد من الدوال الاحتمالية التي تصف خصائصه الرياضية، وتساعد في بناء النماذج وتقدير المعلمات الإحصائية بدقة [1] .

أولاً : دالة الكثافة الاحتمالية (Probability Density Function - PDF)

تمثل دالة الكثافة الاحتمالية $f(t)$ الوظيفة التي تعبّر عن مدى تركز الاحتمال عند كل نقطة زمنية t ، وتحُرّف بأنها المشقة الأولى لدالة التوزيع التراكمي $F(t)$:

$$f(t) = \frac{d}{dt} F(t) \quad (1)$$

تتمتع هذه الدالة بعدة خصائص أساسية، $0 \leq f(t) \leq 1$ لجميع قيم t . التكامل الكلي على المجال يساوي واحداً $\int_{-\infty}^{\infty} f(t) dt = 1$. تستخدم لحساب احتمال وقوع المتغير في فترة زمنية محددة:

$$P(a \leq T \leq b) = \int_a^b f(t) dt \quad (2)$$

تُستخدم دالة الكثافة أيضاً في طرق تقدير المعلمات مثل طريقة الاحتمالية العظمى، وفي اشتقاق باقي الدوال الاحتمالية كدالة البقاء ودالة الخطر [2].

ثانياً : دالة التوزيع التراكمي (Cumulative Distribution Function - CDF)

تُعرف دالة التوزيع التراكمي $F(t)$ بأنها الدالة التي تمثل احتمال التراكمي لأن يكون المتغير العشوائي T أصغر من أو يساوي الزمن t ، ونكتب رياضياً كالتالي:

$$F(t) = P(T \leq t) = \int_{-\infty}^t f(x) dx \quad (3)$$

حيث إنها دالة غير متناقصة، يعني أنه إذا $t_1 < t_2$ فإن $F(t_1) \leq F(t_2)$. تبدأ من الصفر وتنتهي بالواحد $F(-\infty, +\infty) = 0,1$. ويمكن اشتقاق دالة الكثافة منها:

$$\frac{d}{dt} F(t) = f(t) \quad (4)$$

في تحليل الفترات الزمنية، وفي تحديد خصائص البيانات التراكمية، كما تُعد مرجعاً أساساً لدالة البقاء.

ثالثاً : دالة البقاء (Survival Function - S(t))

تعبر دالة البقاء $S(t)$ عن احتمال أن تبقى الوحدة قيد الدراسة دون أن تتعرض للفشل حتى اللحظة الزمنية t ، وتُعرف على أنها متممة دالة التوزيع التراكمي:

$$S(t) = P(T > t) = 1 - F(t) \quad (5)$$

وهي دالة متناقصة، تبدأ بقيمة $1 = S(0)$ وتتناقص تدريجياً إلى أن تؤول إلى الصفر. كما يمكن التعبير عنها أيضاً من خلال التكامل:

$$S(t) = \int_t^{\infty} f(x) dx \quad (6)$$

هذه الدالة في تقدير الاحتمال المتبقى للبقاء، وفي المقارنة بين المجموعات المختلفة باستخدام اختبارات مثل Log-rank ، كما تُعد أساسية في بناء منحنيات Kaplan-Meier وفي تحليل الرقابة اليمنى [1].

2.4 توزيع Top-Leone - الخصائص والسلوك

يُعد توزيع Leone-Topp Leone من التوزيعات الاحتمالية المستمرة التي ظهرت لأول مرة في الأدبيات الإحصائية على يد Leone عام 1955، حيث طُور لتوفير نموذج قادر على تمثيل البيانات ذات السلوك المتتصاعد في دالة الخطر عبر الزمن [3-14]. ويتميز هذا التوزيع ببنائه الرياضية البسيطة نسبياً ومونته في تمثيل ظواهر الحياة أو الفشل التي تتسم بزيادة احتمال الحدث مع مرور الوقت، مما يجعله ملائماً لتحليل بيانات البقاء ودراسات الموثوقية [4-11].

2.4.1 الصيغة الرياضية للتوزيع

تُعرف الدالة الكثافية لاحتمالية توزيع Leone-Topp (PDF) على الفترة $(0,1)$ كما يأتي:

$$f(x; v) = 2v(1-x)[2x-x^2]^{v-1}, \quad 0 < x < 1, \quad v > 0 \quad (7)$$

في حين تُعطى دالة التوزيع التراكمي (CDF) بالشكل:

$$F(x; v) = [2x - x^2]^v, \quad 0 < x < 1 \quad (8)$$

أما دالة البقاء (Survival Function) فهي:

$$S(x; v) = 1 - [2x - x^2]^v \quad (9)$$

وتشتت دالة الخطر (Hazard Function) كما يلي:

$$h(x; \alpha) = \frac{f(x; v)}{S(x; v)} = \frac{2v(1-x)[2x-x^2]^{v-1}}{1-[2x-x^2]^v} \quad (10)$$

2.4.2 التحليل السلوكي والهندسي للتوزيع

يمتاز توزيع Topp-Leone بكونه توزيعاً غير متماثل، حيث تتأثر كثافته الاحتمالية بدرجة المعلمة v . إذ أن زيادة قيمة v تؤدي إلى زيادة تركز الاحتمال في القيم القريبة من 1، مما يعكس سلوكاً تراكمياً متزايداً يعبر عن زيادة احتمال وقوع الحدث مع مرور الوقت. أما في الحالات التي تكون فيها $v < 1$ ، فإن الكثافة تكون أكبر عند القيم القريبة من الصفر، مما يشير إلى سلوك تنازلي في الخطر. ويُظهر التحليل الهندسي للتوزيع أن له شكلاً مرتباً يتميز بذروة واحدة، ويُظهر ذيولاً تختلف في انحدارها باختلاف المعلمة، مما يتبع له تمثيل حالات مختلفة من بيانات الفشل [5-12].

2.4.3 تطبيقات توزيع Topp-Leone

لقد استخدم توزيع Topp-Leone في عدة مجالات تطبيقية، أبرزها تحليل عمر المنتجات في الدراسات الصناعية، وتحليل زمن الفشل في الأنظمة الهندسية، وتقدير زمن البقاء في الدراسات الطبية [6]. وتكم فائدته في أنه يُظهر سلوكاً خطراً متزايداً، وهو ما يتاسب مع الظواهر التي يكون فيها احتمال الفشل أعلى كلما طالت فترة المراقبة. ومع ذلك، فإن محدودية الدعم في الفترة $(0,1)$ تتطلب أحياناً تحويل البيانات أو تطوير توزيع أكثر مرونة ليشمل المجال $(0,\infty)$ [7-15].

2.5 تحويل Transmuted Lower Record Type

يجمع تحويل Transmuted Lower Record Type بين طريقتين رياضيتين قويتين: الأولى هي تحويل G الذي يضيف معاملًا لضبط شكل التوزيع، والثانية هي تحويل L.R. التي تُعيد تشكيل دالة التوزيع باستخدام خصائص القيم الأدنى المسجلة. ويهدف هذا الدمج إلى بناء توزيع من قدر قادر على تمثيل البيانات ذات الانحراف العالي أو التي تتسم بسلوك غير متماثل في دالة الخطر، لا سيما في دراسات البقاء أو البيانات البيئية.

2.5.1 الصيغة الرياضية العامة Transmuted Lower Record Type (TLRT)

تُعد قاعدة التحويل المعروفة باسم Transmuted Lower Record Type ، والمُشار إليها اختصاراً بـ TLRT ، من بين المناهج الإحصائية الحديثة التي تعتمد على التوزيع الأصلي كدالة تراكمية $G(x)$ لتوليد نماذج احتمالية جديدة ذات مرونة أعلى. وقد تم تناول هذه القاعدة بصورة منهاجية في دراسة حديثة لـ Balakrishnan and He (2021)، حيث استندت إلى خصائص السجلات الدنيا(Lower Records)، التي غالباً ما تهمل في التحليل الإحصائي التقليدي رغم أهميتها في تمثيل السلوك الطرفي للبيانات [8-13].

يقوم هذا التحويل على تعديل دالة التوزيع التراكمي الأصلية $G(x)$ لتوليد دالة تراكمية جديدة $F(x)$ باستخدام معامل تحويل $p \in (0,1)$ ، وذلك وفق الصيغة:

$$G(x) = F(x)[1 - p \log(F(x))], \quad 0 < p \leq 1 \quad (11)$$

ويُشتق من هذه الصيغة الشكل التحليلي لدالة الكثافة الاحتمالية الجديدة $(x) f$ كما يلي:

$$g(x) = f(x)[1 - p \log(f(x))], \quad 0 < p \leq 1 \quad (12)$$

يمتاز هذا النموذج بعدة خصائص نظرية، منها:

أن التوزيع الناتج $(x) G$ يُعرف باسم توزيع السجلات الدنيا المحوّلة، ويُعد أكثر قدرة على تمثيل البيانات ذات الانحراف العالي أو السلوك الطرفي المعقّد مقارنة بالتوزيعات الكلاسيكية.

عندما تكون $0 = F(x)$ ، يصبح التحويل غير معرف، وهو ما يُحدّد من إمكانية استخدامه في الجزء الأيسر من نطاق البيانات. يمكن من خلال هذا التحويل توليد عائلات توزيعية جديدة عبر تعديل دوال التوزيع الأصلية، فيما يُعرف باسم توزيعات السجلات المحوّلة (Record-Based Transmuted Distributions).

كما اقترح الباحثان صيغة بديلة لنفس التحويل تأخذ الشكل التالي:

$$G(x) = F(x) + pF(x)\log(F(x)), \quad 0 < p \leq 1 \quad (13)$$

وينتج عنها دالة الكثافة الاحتمالية بالشكل:

$$g(x) = f(x)[1 - p - p\log(F(x))] \quad (14)$$

تُظهر هذه الصيغة مدى مرونة تحويل TLRT في تعديل التوزيع الأصلي بما يسمح بالسيطرة الدقيقة على شكل النموذج الناتج، مما يُعزّز دقتها في تمثيل بيانات الأعمار وبيانات البقاء ذات السلوك غير المتماثل أو التفرط العالي [8].

2.6 تطوير توزيع Topp -Leone بصيغة Transmuted Lower Record Type

باعتبار دالة التوزيع التراكمي (c.d.f) لتوزيع Topp-Leone والمُعطاة في المعادلة (8) فإن تعويض هذه الدالة في صيغة تحويل السجلات الدنيا المحوّل (Transmuted Lower Record Type) الواردة في المعادلة (13) يعطي دالة توزيع تراكمي جديدة بالشكل:

$$G(x; v, p) = F(x)[1 - p \log(F(x))] = [2x - x^2]^v[1 - p \log([2x - x^2]^v)] \quad (15)$$

وبتطبيق خاصية اللوغاريتم على القوس الداخلي، نحصل على:

$$G(x; v, p) = [2x - x^2]^v[1 - p v \log(2x - x^2)] \quad (16)$$

وباستخدام القاعدة العامة لاشتقاق دالة الكثافة الاحتمالية الناتجة عن تحويل TLRT ، والتي تأخذ الشكل:

$$g(x) = f(x)[1 - p(1 + \log(F(x)))] \quad (17)$$

وبما أن دالة الكثافة الأصلية (p.d.f) لتوزيع Topp-Leone هي:

$$f(x; v) = 2v(1 - x)[2x - x^2]^{v-1} \quad (18)$$

وأن $v^v F(x; v) = [2x - x^2]$ ، فإن دالة الكثافة المعدلة الناتجة عن تحويل TLRT تأخذ الشكل:

$$g(x; v, p) = 2v(1-x)[2x-x^2]^{v-1}[1-p(1+v \log(2x-x^2))] \quad (19)$$

وذلك ضمن المجال $0 < x < 1$ ، $v > 0$ ، $0 < p < 1$

للتأكد من أن الدالة $g(x)$ تمثل دالة كثافة احتمالية صحيحة، يجب أن تتحقق شرط التكامل:

$$\int_0^1 g(x; v, p) dx = 1 \quad (20)$$

نبدأ بتعويض معادلة (19) داخل التكامل:

$$= \int_0^1 2v(1-x)[2x-x^2]^{v-1}[1-p(1+v \log(2x-x^2))] dx \quad (21)$$

نقوم بتوسيع القوس داخل التكامل:

$$= \int_0^1 2v(1-x)[2x-x^2]^{v-1}dx - p \int_0^1 2v(1-x)[2x-x^2]^{v-1}dx - pv \int_0^1 2v(1-x)[2x-x^2]^{v-1} \log(2x-x^2) dx \quad (22)$$

الحد الأول هو تكامل دالة الكثافة الأصلية $f(x)$ لتوزيع Topp-Leone ، والذي يتحقق $\int_0^1 f(x) dx = 1$ وعليه فإن التعبير الكامل يصبح: $\int_0^1 g(x) dx = 1 - p + p = 1$. بذلك تكون قد أنشأنا توزيعاً جديداً يُعرف باسم Transmuted Lower Record Type شرط $\int_0^1 g(x) dx = 1$. يتميز هذا التوزيع بمرنة عالية في تمثيل السلوك الطرفي والانحراف في البيانات، ويُعد مناسباً لمنطقة بيانات الأعمار والبقاء الواقعية ضمن المجال $(0,1)$ كما تم التحقق رياضياً من أن دالة الكثافة الناتجة تحقق خصائص التوزيع الاحتمالي بشكل دقيق.

2.7 طريقة الإمكان الأعظم (Maximum Likelihood Estimation)

تُعد طريقة الإمكان الأعظم من أكثر طرائق التقدير شيوعاً في الإحصاء لما تتميز به من خصائص رياضية تجعلها وسيلة فعالة لتقدير معلمات النماذج الاحتمالية. فهي توفر مقدرات كافية لاحتفاظ جميع المعلومات الإحصائية في العينة، وتتميز بكونها غير متغيرة بحيث يساوي متوسطها القيمة الحقيقة للمعلمة، إضافةً إلى ثباتها الذي يضمن دقة أكبر مع زيادة حجم العينة. كما تمتاز مقدرات الإمكان الأعظم بالكافأة من خلال تحقيقها أصغر تباين ممكن، وبخاصية عدم التغير التي تسمح بتحويل التقديرات والحصول على مقدرات جديدة بنفس الكفاءة. لذلك، فإن استعمال هذه الطريقة في تقدير معلمات التوزيع المقترن يتيح الحصول على تقديرات قوية خاصة عندما يتتوفر حجم عينة كافٍ. ولتقدير معلمات توزيع TLRT-Topp-Leone باستخدام طريقة الإمكان الأعظم نفرض لدينا x_1, x_2, \dots, x_n تمثل عينة عشوائية مستقلة مأخوذة من التوزيع المقترن $TLRT - TL(v, p)$ ، فإن دالة الكثافة الاحتمالية تعطي كما في المعادلة (21) . وبالتالي، تكون دالة الإمكان :

$$L(v, p) = \prod_{i=1}^n g(x_i; v, p) = \prod_{i=1}^n 2v(1-x_i)[2x_i-x_i^2]^{v-1}[1-p(1+v \log(2x_i-x_i^2))] \quad (23)$$

وبأخذ اللوغاريتم الطبيعي للطرفين نحصل على دالة لوغاريتيم الإمكان:

$$\log L(v, p) = \sum_{i=1}^n [\log(2v) + \log(1-x_i) + (v-1)\log(2x_i-x_i^2) + \log(1-p(1+v \log(2x_i-x_i^2)))] \quad (24)$$

لحساب القيم التقديرية المثلى للمعلمتين v و p ، نأخذ المشتقات الجزئية لدالة لوغاريتيم الإمكان:

بالنسبة لـ v :

$$\frac{\partial \log L}{\partial v} = \sum_{i=1}^n \left[\frac{1}{v} + \log(2x_i - x_i^2) - \frac{p \log(2x_i - x_i^2)}{1-p(1+v \log(2x_i - x_i^2))} \right] = 0 \quad (25)$$

بالنسبة لـ p :

$$\frac{\partial \log L}{\partial p} = \sum_{i=1}^n \left[\frac{-1-v \log(2x_i - x_i^2)}{1-p(1+v \log(2x_i - x_i^2))} \right] = 0 \quad (26)$$

المعادلتان (25) و(26) غير خطيتين، ولا يمكن حلّهما تحليلياً، لذا يتم استعمال الطرق العددية المناسبة (مثل طريقة نيوتن-رافسون) للحصول على القيم التقديرية المثلث للمعلمات، باستعمال هذه القيم التقديرية، يمكن استبدالها في دالة البقاء:

$$S(x; \hat{v}, \hat{p}) = 1 - [2x - x^2]^{\hat{v}} [1 - \hat{p}]^{\hat{v}} \log(2x - x^2) \quad (27)$$

للحصول على القيم التقديرية لدالة البقاء الخاصة بالتوزيع TLRT-TL.

3. الجانب التطبيقي Application

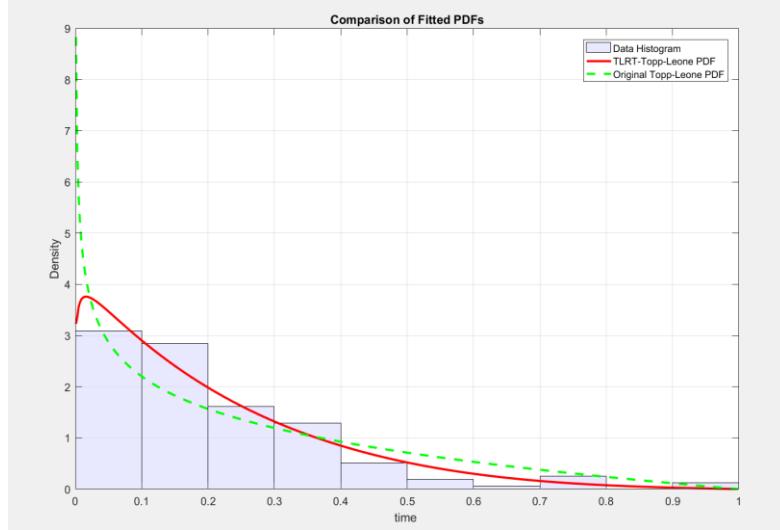
تم الحصول على بيانات حقيقة من مستشفى الحسين التعليمي في محافظة كربلاء، تتعلق بمدة بقاء المرضى المصابين بفيروس كورونا داخل المستشفى، وذلك خلال الفترة الممتدة من عام 2020 إلى عام 2024. وقد تضمنت البيانات (150) مشاهدة تمثل عدد الساعات التي قضاها كل مريض في المستشفى منذ لحظة الإدخال وحتى الخروج. تم تسجيل هذه المعلومات بدقة ضمن السجلات الطبية الرسمية، واعتمدت كأساس للتحليل الإحصائي في هذا الجانب التطبيقي من الدراسة.

لا اختبار ملاءمة التوزيع (Goodness-of-Fit Test) تم إجراء اختبار كولموغروف-سميرنوف (K-S) واختبار أندرسون-دارلينغ (AD) لقياس مدى ملاءمة البيانات الفعلية للتوزيعين احتماليين Topp-Leone و TLRT-Topp-Leone .

جدول (1) قيم الاختبار الاحصائي لحسن المطابقة

dist	parameters	K-S	P_value	AD	P_value
TLRT-Topp-Leone	V=1.3	0.0766	0.3272	1.2299	0.032
Toop_Leone	P= 0.9	0.1836	0.0002	8.9079	0.00001

حيث تم ثبيت قيمة المعلمة $v=1.3$ في كلا التوزيعين، وكانت قيمة معلمة التحويل $p=0.9$ في التوزيع المحول. أظهرت نتائج اختبار K-S أن قيمة الإحصاء بلغت 0.0766 مع قيمة احتمالية (P-value) تساوي 0.3272 ، مما يشير إلى قبول الفرضية القائلة بأن البيانات تتبع التوزيع المحول، لكن القيمة الاحتمالية أكبر من 0.05. أما بالنسبة لاختبار AD فقد بلغت القيمة 1.2299 بقيمة احتمالية 0.032 ، وهي مقبولة نسبياً ضمن بعض مستويات المعنوية المرنة. في المقابل، عند تطبيق نفس الاختبارات على توزيع Topp-Leone الأصلي، أظهرت النتائج ضعف الملاءمة؛ إذ بلغت قيمة K-S 0.1836 مع P-value صغيرة جدًا 0.0002 ، مما يشير إلى رفض الفرضية. كذلك، سُجل اختبار AD قيمة مرتفعة جداً بلغت 8.9079 مع P-value مساوية لـ 0.00001 ، مما يعزز ضعف ملاءمة هذا التوزيع للبيانات الفعلية. بناءً على ذلك، يمكن القول إن توزيع Topp-Leone المحول يمثل النموذج الأنسب لتمثيل بيانات مدة البقاء في المستشفى، مقارنةً للتوزيع التقليدي.



شكل (1) دالة الكثافة الاحتمالية للتوزيعين ويبين مدى أفضلية التوزيع المحول

يتضح من الشكل أعلاه أن دالة الكثافة الاحتمالية للتوزيع Topp-Leone المحول (الخط الأحمر) تُظهر توافقًا واضحًا مع سلوك البيانات الفعلية مقارنةً بالتوزيع الأصلي (الخط الأخضر المقطوع). حيث يقترب المنحنى الأحمر من شكل الهستوغرام في جميع النقاط، خصوصًا في بداية ونهاية التوزيع، بينما يبدو التوزيع الأصلي أكثر انحرافًا عند القيم الصغيرة. مما يؤكد أن التوزيع المحول يوفر ملائمة أفضل لتمثيل بيانات مدة البقاء في المستشفى.

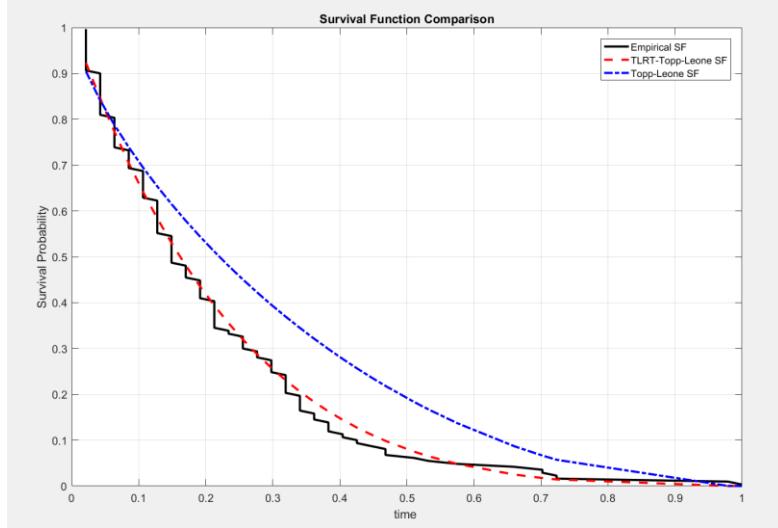
جدول (2) قيم المعايير

dist	-LL	AIC	BIC	CAIC
TLRT-Topp-Leone	-7297.69519	14595.79039	14596.39907	14595.99039
Toop_Leone	-851.0552759	17023.1055	17026.1489	17024.1055

في ضوء النتائج الإحصائية المعروضة في الجدول (2)، تم إجراء مقارنة بين توزيع Topp-Leone والتوزيع المحول TLRT-Topp-Leone باستخدام مجموعة من معايير المفضلة المعتمدة في تحليل النماذج الاحتمالية، وهي: دالة السجل الاحتمالية السالبة (-LL)، معيار معلومات أكايكي (AIC)، معيار معلومات بيز (BIC)، ومعيار أكايكي المصحح (CAIC).

وقد أظهرت النتائج أن التوزيع المحول TLRT-Topp-Leone حقق أقل القيم لجميع المعايير مقارنة بالتوزيع الأصلي، إذ كانت قيمة تساوي (-LL=-7297.695)، بينما بلغت قيمة (AIC=145957.904)، و (CAIC=145963.991)، و (BIC=145963.991). أما توزيع Topp-Leone فقد سجل قيمًا أعلى بكثير، إذ بلغت (-LL=-851.055)، و قيمة (AIC=145959.904)، و قيمة (CAIC=17024.106)، و (BIC=17026.149)، و (AIC=17023.106).

تشير هذه النتائج إلى أن التوزيع المحول يتمتع بقدرة أكبر على تمثيل البيانات الحقيقية بكفاءة ودقة، مما يؤكد تفوقه على التوزيع الأصلي من حيث ملائمة النموذج الإحصائي.



شكل (2) قيمة دالة البقاء التجريبية ودالة البقاء لكلا التوزيعين وبين التقارب بين التوزيع المحول المنحني التجربى

الشكل اعلاه يوضح مقارنة بين دوال البقاء التجريبية والنظرية لكل من التوزيع الأصلي Topp-Leone والتوزيع المحول TLRT-Topp-Leone، وذلك لغرض فحص مدى ملاءمة كل نموذج للبيانات الفعلية الخاصة بمدة البقاء في المستشفى.

من خلال الشكل، نلاحظ أن المنحني الأسود يمثل دالة البقاء التجريبية (Empirical SF) المستخرجة من البيانات الفعلية، في حين يمثل المنحني الأحمر المتقطع دالة البقاء المشتقة من التوزيع المحول، والمنحني الأزرق المتقطع يمثل دالة البقاء للتوزيع الأصلي. يُظهر الشكل أن التوزيع المحول TLRT-Topp-Leone يقدم تقريباً أكثر دقة لدالة البقاء التجريبية، حيث يتلقاط معها في عدد من النقاط ويتبع سلوكها عن قرب، بخلاف التوزيع الأصلي الذي يميل إلى الابتعاد بشكل ملحوظ عنها. يعزز هذا التقارب بين التوزيع المحول والمنحني التجربى من كفاءة التوزيع المقترن في تمثيل خصائص بيانات البقاء الواقعية بدقة أعلى من التوزيع التقليدى .

4. الاستنتاجات و التوصيات

4.1 الاستنتاجات

أظهرت نتائج البحث أن التوزيع المقترن Transmuted Lower Record Type Topp-Leone المحول بصيغة يمتلك بمرنة عالية وقدرة أكبر على تمثيل البيانات ذات السلوك غير المتماثل أو الطويلة الذيل مقارنة بالتوزيع الأصلي. أكدت اختبارات الملاءمة Anderson-Darling و Kolmogorov-Smirnov و CAIC ، AIC ، BIC ومعايير المقارنة أن النموذج الجديد يوفر تمثيلاً أدق وأكثر كفاءة لبيانات البقاء الواقعية. كما بيّنت التحليلات التطبيقية على بيانات مرضى COVID-19 أن التوزيع المحول يقترب بشكل كبير من المنحنيات التجريبية، مما يعزز موثوقيته في توصيف الظواهر الطبية. وبذلك يشكل هذا التوزيع إضافة مهمة إلى عائلة التوزيعات الاحتمالية المرنة، ويفك جوى استخدام تحويلات TLRT في تطوير النماذج الإحصائية.

4.2 التوصيات

توسيع استخدام التوزيع المقترن في مجالات تطبيقية أخرى مثل البيانات البيئية والمالية التي تتميز بعدم التماثل ووجود قيم متطرفة. دراسة طرق تقدير بديلة إلى جانب الإمكان الأعظم، مثل الطرائق البايزية أو المربعات الصغرى، لمقارنة دقة النتائج في حالات العينات الصغيرة أو البيانات المراقبة.

اختبار أداء التوزيع TLRT-Topp-Leone باستخدام بيانات محاكاة ذات سيناريوهات مختلفة للتأكد من كفاءته تحت ظروف متعددة. دمج التوزيع المقترن ضمن نماذج بقاء متعددة المتغيرات أو نماذج المخاطر النسبية لتوسيع قدرته التفسيرية في الدراسات

الطيبة. تشجيع الباحثين على استكشاف تحويلات إحصائية مشابهة لتحسين توزيعات أخرى، بهدف بناء نماذج أكثر مرنة وتوافقاً مع البيانات المعقّدة.

المصادر

1. Kleinbaum, D. G., & Klein, M. (2012). *Survival analysis: A self-learning text* (3rd ed.). Springer.
2. Kartsonaki, C. (2016). *Survival analysis*. Oxford University Press.
3. Topp, C. W., & Leone, F. C. (1955). A family of J-shaped frequency functions. *Journal of the American Statistical Association*, 50(269), 209–219.
4. Al-Babtain, A. A., Ibrahim, N. A., & El-Gohary, A. (2017). Exponential Topp-Leone distribution: Properties and applications. *Journal of Statistical Theory and Applications*, 16(2), 189–202.
5. Kotz, S., & Seier, E. (2007). Kurtosis of the Topp–Leone distribution and comparison with other distributions. *Journal of Applied Statistical Science*, 15(3), 187–193.
6. Pourdarvish, A., Momenzadeh, M., & Jafari, A. A. (2015). The exponentiated Topp–Leone distribution: Properties and applications. *Journal of Statistical Computation and Simulation*, 85(16), 3261–3273.
7. Al-Zahrani, B. M. (2012). Goodness-of-fit tests for the Topp-Leone distribution with estimated parameters. *Journal of Applied Statistical Science*, 19(2), 105–116.
8. Balakrishnan, N., & He, M. (2020). A record-based transmuted family of distributions. In *Advances in Statistics-Theory and Applications: Honoring the Contributions of Barry C. Arnold in Statistical Science* (pp. 3–24). Springer International Publishing.
9. Akaike, H. (1973). Maximum likelihood identification of Gaussian autoregressive moving average models. *Biometrika*, 60(2), 255–265.
10. Massey Jr, F. J. (1951). The Kolmogorov-Smirnov test for goodness of fit. *Journal of the American Statistical Association*, 46(253), 68–78.
11. Alizadeh, M., Ganjali, M., & Tahmasebi, S. (2016). The Topp-Leone-G family of distributions. *REVSTAT–Statistical Journal*, 14(1), 79–104.
12. Tahir, M. H., Alzaatreh, A., & Mansoor, M. (2015). Topp–Leone-G family of distributions. *Journal of the Egyptian Mathematical Society*, 23(2), 349–358.
13. Tanış, V. N. (2018). A new transmuted lower record type power function distribution: Properties and applications. *Hacettepe Journal of Mathematics and Statistics*, 47(6), 1381–1393.
14. Nyamajiwa, T., Balakrishnan, N., & Zografos, K. (2024). The Gamma Tipo Dos Half Logistic Topp-Leone-G distribution: Properties and applications. *Journal of Statistical Distributions and Applications*, 11(1), 1–22.
15. El-Gohary, A., Alshamrani, A., & Al-Otaibi, A. N. (2013). The Topp-Leone Inverted Exponential distribution. *Life Science Journal*, 10(2), 3512–3518.