كلية التجارة
قسم الاحصاء التطبيقى والتأمين

تحليل نماذج التحويل شبه المعلمية للبيانات المبتورة

Semiparametric Analysis of Transformation Models with Censored Data

تحت اشراف

أ.د/فاطمة على محمد عبد العاطى
أستاذ الاحصاء التطبيقي
كلية التجارة-جامعة المنصورة

أ.د/ هشام محمد المنجى
أستاذ الإحصاء المساعد
كلية التجارة _ جامعة المنصورة

اعداد

آية محمود عبد العزيز الزينى

المستخلص

غالبًا ما تظهر البيانات المبتورة من ناحية اليسار في دراسات الأوبئة والمتابعة الفردية بسبب سحب العينات المتحيزة ، لأن الأشخاص الذين لديهم فترات بقاء أقصر تميل هذه الدراسات إلى استبعادهم من العينة. علاوة على ذلك ، غالبًا ما يخضع وقت بقاء الأشخاص محل الدراسة للرقابة من ناحية اليمين .

في هذا البحث ، تمت دراسة فئة عامة من نماذج التحويل شبه المعلمية التي تتضمن نموذج المخاطر النسبية ونموذج الإحتمالات النسبية كحالات خاصة لتحليل البيانات المبتورة من ناحية اليسار والخاضعة للرقابة من ناحية اليمين .

وتم اقتراح استخدام طريقة الاحتمال الشرطي و تم تطويرها لتناسب طريقة الإمكان الأعظم الشرطية(CMLE) بهدف تقدير معالم الانحدار و دالة الخطر التراكمية لهذه النماذج والمعادلات المشتقة لمعامل الانحدار، وتُظهر طريقة الإمكان الأعظم الشرطية لتكون متسقة وطبيعية بشكل متقارب .

وأخيرا ، تم تقدير نموذج بوكس كوس والنموذج اللوغارتيمي لبيانات تم مراقبتها من ناحية اليمين وتحديد المتغيرات التى لها دلالة إحصائية .

الكلمات المفتاحية :

دالة الإمكان الأعظم الشرطية – نموذج الخطر النسبي – نموذج التحويل شبه المعلمي – البيانات المقتطعة .

Abstract:

Left–truncated data often arise in epidemiology and individual follow–up studies due to a biased sampling plan where subjects with shorter survival times tend to be excluded from the sample. Moreover, the survival time of recruited subjects are often subject to right censoring .

In this article, a general class of semi–parametric transformation models that include proportional hazards model and proportional odds model as special cases is studied for the analysis of left–truncated and right–censored data.

We propose a conditional likelihood approach and develop the conditional maximum likelihood estimators (CMLE) for the regression parameters and cumulative hazard function of these models. The derived score equations for regression parameter. The (CMLE) is shown to be consistent and asymptotically normal.

Finally, on the Application side, both the Box–Cox model and the logarithmic model for right –censored data are estimated and the variables with the most statistical significance are determined.

Keywords:–Maximum conditional likelihood– Proportional hazards model – Semiparametric transformation model –Truncation data.

1. المقدمة

تحليل البقاء هو فرع من فروع التحليل الإحصائى يستخدم في تحليل البيانات التي تتكون من الملاحظات التي تمت على متغيرات عشوائية ذات قيمة غير سلبية . و يركز الإهتمام بهذا الموضوع علي مجموعة من الأفراد لكل منهم فترة محددة من الوقت عند حدوث أحداث معينة ويسمى هذا الحدث (الفشل أو "البقاء".) .

كما تأتى الأهمية العملية فى تحديد المتغيرات الأكثر تأثيرا فى المرض لما له من أهمية عظمى فى المجال الطبى ويهدف أيضا تقدير نموذج التحويل شبه المعلمى فى إطار البيانات الخاضعة للرقابة من ناحية اليمين عادةً ما يشار إلى البيانات التي تم جمعها من هذه التجارب على أنها بيانات البقاء على قيد الحياة ، أو بيانات وقت وقوع الحدث ، أو بيانات وقت الفشل ويتم تمثيلها بواسطة متغير عشوائي غير سلبي (T) .

(T): الدالة الرئيسية التي تُستخدم بشكل شائع في تحليل البقاء على قيد الحياة هي دالة البقاء على قيد الحياة ، ويعرف إحتمال البقاء على قيد الحياة على النحو التالي :

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

حيث أن : $S(t)$ دالة البقاء و تمثل احتمال وقوع الحدث.

كما أن دالة التوزيع التراكمي (CDF) هى:

$$F(t) = P(T \leq t) \qquad (2)$$

وهناك دالة أخرى شائعة الاستخدام في تحليل البقاء هي دالة الخطر أو معدل الفشل اللحظي (T) وتعرف كالتالى :-

$$\lambda(t) = \lim_{h \to 0} \frac{P(t \leq T < t+h / T \geq t)}{h} \qquad (3)$$

ومن السهل إيجاد أن :

$$\lambda(t) = \frac{f(t)}{s(t)} = -\frac{d}{dt} \log S(t) \qquad (4)$$

وتُعرف دالة الخطر التراكمية لـ (T) كمايلى:

$$\Lambda(t) = \int_0^t \lambda(s)ds \quad (5)$$

من..(4) و(5) نحصل على:

$$S(t) = e^{-\int_0^t \lambda(s)ds} = e^{-\Lambda(t)} \quad (6)$$

هناك نوع آخر من المتغيرات في تحليل البقاء يسمى المتغيرات المشتركة وتمثل تأثير بعض المعالجات (أو المتغيرات التفسيرية) مثل العمر والجنس وما إلى ذلك على وقت بقاء الفرد. لنفترض أن لكل فرد متجه يتكون من عدد q من الأبعاد للمتغيرات المشتركة Z وفي حالة المتغيرات غير المشتركة (non-covariate) يمكننا الحصول على دالة البقاء المشروط لـ T بالنظر إلى المتغير المشترك $z = z$ حيث

$$S(t/z) = e^{-\int_0^t \lambda(s/z)ds} = e^{-\Lambda(t/z)} \quad (7)$$

حيث $\lambda(t/z)$ و $(\Lambda t/z)$ هي دالة الخطر الشرطي والخطر التراكمي لدالة T حيث $Z = z$ لنحصل على:

$$\lambda(t/z) = \lim_{h\to 0} \frac{P(t \le T < t+h / T \ge t, z)}{h} \quad (8)$$

$$،\Lambda(t/z) = \int_0^t \lambda(s/z)ds \quad (9)$$

وعادة ما يتكون النموذج من فئة من التوزيعات أو الكثافات التي نعتقد أنها قد تكون قد وُلدت من البيانات المرصودة. وبشكل أساسي ، هناك نوعان من النماذج: نموذج معلمي ونموذج غير معلمي.

فالنموذج المعلمي هو نموذج يمكن معرفته بواسطة متجهة q من الأبعاد ، أما النموذج الغير معلمي (نموذج غير محدود الأبعاد) هو نموذج يحتوي على جميع P.d.f في حجم العينة بشكل مطلق ومستمر مع قياس مجال σ . [1]

في النموذج المعلمي يُعرف توزيع النموذج بالمعلمات غير المعروفة بينما في النموذج اللامعلمي تكون المعلمة غير المعروفة هي توزيع النموذج نفسه .

لا يزال لدينا نوع آخر من النماذج وهي حالة وسيطة بين النموذجين المذكورين أعلاه ، تسمى هذه الحالة **بالنموذج الشبه معلمي** ، وهو نموذج يحتوي على توزيعات احتمالية موصوفة من خلال معلمة تحتوي على مكون متناهي الأبعاد ومكوّن غير متناهي الأبعاد.

النموذج المعلمي يشار اليه:

$$\mathcal{P} = \{\mathcal{P}(x,\theta): \theta \epsilon \Theta \subset \mathbb{R}^q\} \qquad (10)$$

حيث أن $\mathcal{P}$ هو نموذج معلمى.

وعلى سبيل المثال قد يتم تعريف النموذج المعلمى الطبيعى على أنه:

$$\mathcal{P} = \{N(\theta_1,\theta_2): (\theta_1,\theta_2) \in \Theta \subset R \times R^+\} \qquad (11)$$

أما النموذج الغير معلمي يشار اليه كالتالى:

$$\{\mathcal{P}: \mathcal{P}(x) \geq 0 \ for\ all x \in \chi, \int_x \mathcal{P}(x) dv(x) = 1\} \qquad (12)$$

عادة ما يتم الإشارة إلى النموذج الشبه معلمي كالتالى:

$$\mathcal{P} = \{\mathcal{P}(x,\beta,\eta): \beta \in \Theta \subset \mathcal{R}^q, \eta \in \mathcal{H}\} \qquad (13)$$

حيث أن :

q-dimension parameter : β

η : معلمة غير متناهية الأبعاد .

$\mathcal{H}$: فراغ الدالة مثل Hilbert space.

وقد تواجه هذه الدراسة عدة مشاكل منها (على سبيل المثال لا الحصر):-

- تقدير معاملات الأبعاد غير المحدودة في ظل وجود معلمة أبعاد محدودة في نموذج تحويل الشبة معلمي ، وجود صعوبة فى التعامل مع البيانات المبتورة .

2. الاستعراض المرجعي

تناولت دراسة Zeng & Lin(2006) فئة من نماذج التحويل الشبه معلمي لتصف تأثيرات المتغيرات المشتركة بمرور الوقت من خلال دوال الكثافة لعمليات العد ،وتشمل هذه الفئة الكثافة الاحتمالية ونماذج التحويل الشبه معلمي

وقد تناولت دراسة(2016) Zhang التقدير الشبه معلمي لنموذج التحويل الخطي جزئيا وذلك بتحديد الكمية الشرطية مع عدم وجود معلمة مفروضه إما من خلال نموذج دالة الارتباط أو توزيع الخطأ .

وتناولت دراسة(2017) Sokullu & Stouli دراسة البيانات الخاضعة لرقابة الفترات الزمنية من خلال نموذج مكون من جزئين الأول نموذج الانحدار اللوجستيي والثانى النموذج الشبه معلمي وتم تطوير هذا النموذج عن طريق خوارزمية من نوع EM وتحديد اتساقها وكفائتها.

كما تناولت Li et al (2018) تحليل الإنحدار للبيانات الخاضعة للرقابة المزدوجة وذلك من خلال نماذج التحويل شبه المعلمية وتم فيها دراسة الخصائص للاحتمالات العظمى بما فيها ..الاتساق _ الكفاءه الشبه معلمية .

وأوضحت الدراسات مدى اهتمامها بنموذج التحويل شبه المعلمي لذلك يهدف البحث الى تقدير نموذج التحويل الشبه معلمي للبيانات المبتورة من ناحيه اليمن وبالتحديد دراسة نموذج بوكس كوكس والنموذج اللوغاريتيمي من خلال دالة الإمكان الأعظم الشرطية .

3. النماذج محل الدراسة

يُعتبر نموذج التحويل الخطي نموذجًا شبه معلمي يتكون من معلمة غير معروفة من أبعاد محدودة (finite dimensional) ومن أجزاء ذات أبعاد غير محدودة (infinite dimensional)

نفرض أن T ← وقت البقاء على قيد الحياة ، z ← يكون متجه q من الأبعاد للمتغيرات المشتركة لتأثيرات نموذج Z على الاستجابة لـ T .

لذلك نفرض أن نموذج التحويل الخطي(2006) Lu and Tsiatis .

$$H(t) = -\beta^T Z + \epsilon \qquad (14)$$

حيث أن:

H : دالة أحادية متزايدة غير معروفة .

β : هي معلمة انحدار ذات q للأبعاد.

ϵ : مصطلح الخطأ الذي يُفترض أنه يتبع دالة توزيع معروفة F_ϵ خالية من المتغير Z . كما أن دالة البقاء الشرطية لـ T من Z تمثل الشكل

$$S(T/Z) = \mathbb{P}(T > t/Z)$$

$$= \mathbb{P}(H(T) + \beta^T Z > H(t) + \beta^T Z/Z) \quad (15)$$

$$S_T(t/Z) = e^{-\Lambda_\epsilon(H(t)+\beta^T Z/Z)} \quad (16)$$

حيث $\Lambda_\epsilon(./Z)$ هى دالة الخطر التراكمي الشرطى من ∈ المعطى ل Z ، ويوجد حالتين لنموذج التحويل الخطي : نموذج الخطر النسبي (نموذج كوكس) ونموذج الاحتمالات التناسبية (proportional odd model) ، بافتراض أن ϵ يتبع توزيع extreme value distribution. ومن المعروف أن دالة البقاء على قيد الحياة لـ ϵ هي $S_\epsilon(x)=e^{-e^x}$ ومن نموذج التحويل الخطي (3) يختصر إلى [1]

$$\Lambda_T(t/z) = \Lambda_{T0}(t)e^{\beta^T z} \quad (17)$$

وهو نموذج الخطر النسبي $S_{T0}(.)$ ، $\Lambda_{T0}(.)$ تُعرف باسم بقاء خط الأساس ودالة الخطر التراكمي على التوالي.

ثانيا :النماذج الشبه معلمية

نماذج التحويل لعمليات العد وهو يرتبط فئة نماذج التحويل الخطي بتحويل غير معروف خطيا لوقت الفشل T إلى متجه الوقت (time invariant) للمتغيرات المشتركة z

$$H(T) = -\beta^T Z + \varepsilon \quad (18)$$

H(.) دالة تزايدة .

β هي مجموعة من معالم الانحدار غير المعروفة .

ε خطأ عشوائى .

فى النموذج يكون التحويل H(.) يكون إفتراضى arbitrary، ولا ينبغى أن يفرض ε بشكل مقيد ، فهناك دائمًا تحول مثل " ε " له أي توزيع معين ، وقام بتوسيع المعادلة للسماح بمتغيرات زمنية متغيرة وأحداث متكررة [1]، لذا سوف يتم فرض أن:

N^* (t) تمثل عملية العد التي تسجل عدد الأحداث التي حدثت بحلول الوقت t

z (.) يكون متجهًا للمتغيرات المشتركة المتغيرة بمرور الوقت .

لذا نحدد دالة الكثافة الاحتمالية لـ N^*t بشرط أن $\{Z(s), s \leq t\}$ وتأخذ الشكل

$$\Lambda(t/Z) = G\left[\int_0^t R^* \exp\{\beta^T Z(s)\} d\Lambda(s)\right] \qquad (19)$$

G هي دالة قابلة للتفاضل باستمرار وتتزايد بشكل ثابت

$R^*(.)$مؤشر عملية العد ، β هو متجه لمعلمات انحدار غير معروفة

$\Lambda(.)$ هي دالة متزايدة غير محددة.

بالنسبة لبيانات البقاء $R^*(t) = I(T \geq t)$ حيث$I(.)$ يمثل مؤشر الدالة ، بالنسبة للأحداث المتكررة $R^* = 1$ ...ومن المفيد الأخذ في الاعتبار نموذج بوكس كوكس التحويلي التالي :

$$G(x) = \frac{(1+x)^{\theta}-1}{\theta}, \qquad \theta \geq 0 \qquad (20)$$

حيث $\rho = 0$ بالمقابله $G(x) = \log(1+x)$.

و نموذج التحويل اللوغاريتمي .

$$G(x) = \frac{\log(1+rx)}{r}, \qquad r \geq 0 \qquad (21)$$

حيث r=0 بالمقابلة لـ $G(x) = x$ واختيار $G(x) = x$ ينتج عنه نموذج الخطر النسبي المألوف أو نموذج الكثافة . (Cox, 1972; Andersen and Gill, 1982) . إذا كان $N^*(.)$ لديه قفزة واحدة في وقت البقاء على قيد الحياة T و Z هو الوقت الثابت ، ثم المعادلة (2) تختزل إلى المعادلة (1)

ثالثا : نموذج بوكس كوكس التحويلى(Box–cox transformation model)

يتم استخدامه لتعديل الشكل التوزيعي لمجموعة من البيانات ليتم توزيعها بشكل أكثر طبيعية بحيث يمكن استخدام الاختبارات وحدود الثقة التي تتطلب التوزيع الطبيعي بشكل مناسب .ولكن تقنية هذا النموذج قد لا تتناسب مع القيم المتطرفة .

ويتبع الشكل الرياضى التالى :-

$$Y = (X + \delta)^{\lambda} \qquad (22)$$

حيث λ هو الأس (القوة) و هو مقدار التحول الذي يتم إضافته عندما تكون X صفراً أو سالبة. عندما تكون λ صفرًا ، يتم استبدال التعريف أعلاه بـ

$$Y = \ln(X + \delta) \qquad (23)$$

عادةً ما يتم فحص قيم λ القياسية 2، 1.5، 1، 0.5، 0، -0.5، -1، -1.5، -2،

وقام بحل القيمة المثلى لـ λ باستخدام تقدير الاحتمالية القصوى. يقوم أيضًا بحساب حدود الثقة حول القيمة المثلى. . بالاعتماد على القيمة المعيارية الأكثر ملاءمة بين حدود الثقة.على سبيل المثال ، إذا كانت حدود الثقة من 0.4 إلى 1.1 ، فسيتم تعيين على القيمة القياسية "1" (بدون تحويل) لأن هذا هو الأكثر ملاءمة. يجب توخي الحذر عند استخدام حدود الثقة ، لأنها تعتمد بشكل كبير على حجم العينة.

4. دالة الإمكان الأعظم الشرطية تتم على خطوتين

الخطوة الأولى: تعمل على تعظيم دالة الاحتمال اللوغاريتمى الهامشي.

الخطوة الثانية: تعمل الخطوة الثانية على تعظيم دالة احتمالية اللوغاريتم الشرطي .

5. طريقة التقدير

1- يتم أولا تقدير المعادلات لنموذج التحويل الشبه معلمي باستخدام طريقة martingale arguments واستخدامها Chen et al (2002) لتحليل بيانات غير متحيزة من ناحية اليمين .

2- يتم استخدام طريقة الإمكان الأعظم الشرطية (CMLE) لتقدير نماذج التحويل شبه المعلمية حيث أن دالة الكثافة التراكمية للنموذج

$$\Lambda(t \mid Z) = \{G \quad R(t)\exp(\beta^{T} Z)\} \qquad (24)$$

حيث يتم تقدير *R(.)* ، *β* عندما تكون *Z* منفصلة .

G هى دالة التحويل ، *R(t)* دالة الخطر التراكمية الأساسية.

أولا : تقدير المعادلات EE

نفرض أن $(i = 1, \dots, n)$ $(X_i, V_i, Z_i\delta_i)$ وتكون العينه المبتورة التي يتم مراقبتها تعرف كالآتى $Y_i(t) = I_{[V_i \le t \le X_i]}$ ، $N_i(t) = I_{[X_i \le t\,\delta_i = 1]}$ ونفرض أيضا أن $\mathcal{F}(t)$ *σ-field*

ونفرض أن $P(Z_i) = P(V \le T/Z_i)$. وتشير الى القيم الصحيحة β, h, β_0, h_0 على التوالى .

لذلك نفرض أن :

$$M_i(t) = N_i(t) - \int_0^t Y_i d\Lambda_\epsilon(\beta^T Z_i + h_0(s)) \qquad (25)$$

حيث أن $\Lambda_\epsilon(u) = G(e^u)$ هي دالة الخطر التراكمية لـ ϵ .

$$E[Y_i(t) / Z_i] = P(V_i \le t \le X_i / Z_i)$$

$$= P(V \le t \le X / V \le T, Z_i)$$

$$=P(Z_i)^{-1}P(V \le t \le C/Z_i)P(T \ge t/Z_i) \quad (26)$$

$$E[N_i(t) / \mathcal{F}(t-)] = Y_i d\Lambda_\epsilon(\beta_0^T Z + h_0(t)) \quad (27)$$

كما اعتبر (2011) Shen) المعادلتين التاليتين تقديرا لكلا من β و $h(t)$

$$U_1(B,h) = \sum_{i=1}^{n}[dN_i(t) - Y_i(t)d\Lambda_\epsilon(\beta^T z_i + h(t))] = 0 \quad (28)$$

$$U_2(\beta,h) = \sum_{i=1}^{n}\int_0^{\tau_c} Z_i[dN_i(t) - Y_i(t)d\Lambda_\epsilon(\beta^T Z_i + h(t))] = 0 \quad (29)$$

ثانيا : دالة الإمكان الأعظم الشرطية CML

بافتراض أن $Q(x,v,z)=P(C\le x/V = v, Z = z)$ ويشار إلى التوزيع الشرطى C اذا كانت $V=v$ ، $Z=z$ وحيث دالة الاحتمال $L(F,K,Q)$ كما يلى :

$$L = \prod_{i=1}^{n}\{d(F(X_i/Z_i)dk(V_i/Z_i)[1 - Q(X_i/V_i,Z_i]/p(Z_i))\}^{\delta_i}$$

$$\times \prod_{i=1}^{n}\{dQ(X_i/V_i,Z_i)dK(V_i/Z_i)[1 - Q(X_i/V_i,Z_i] / p(Z_i\}^{1-\delta_i} \quad (30)$$

كما في *Wang(1991)* ، نحلل L إلى ثلاثة عوامل تنتج :-

$$L = \prod_{i=1}^{n}\left\{\frac{[dF(X_i/Z_i)]^{\delta_i}[1 - F(X_i / Z_i]^{1-\delta_i}}{1 - F(V_i -/ Z_i)}\right\}\prod_{i=1}^{n}\left\{\frac{dK(V_i/Z_i)[1 - F(V_i - Z_i)]}{p(Z_i)}\right\}$$

$$\times \prod_{i=1}^{n}\{[1 - Q(X_i/V_i,Z_i]^{\delta_i}[dQ(X_i/V_i,Z_i]^{1-\delta_i}]\} = L_1 L_2 L_3 \quad (31)$$

L_1, L_2, L_3 ← تمثل دوال الاحتمال . وبالنظر الى دالة النموذج*(24)* ينتج الآتى :-

$$P(Z_i)\int_{a_k}^{b_k} exp\{-G\{R(v)e^{\beta^T z_i}\} dK(v) \quad (32)$$

6. الجانب التطبيقى :

تتناول الدراسة التطبيقية بيانات طبية لعينة من الأطفال المرضى المصابون بمرض التهاب الدماغ ، والتى تم علاجهم من هذا المرض بمستشفى حميات المنصورة ، وقد تم تجميع البيانات وعددها 133 طفل أصيبوا بهذا المرض وكان عدد الحالات التى توفت خلال فترة الدراسة هي 90 حالة بينما عدد الحالات التى ظلت على قيد الحياة حتى نهايه الفترة عددها 43 حالة وتمثل الحالات المبتورة وهو مايعتمد عليه فى هذا البحث ،. تم استخدام برنامج R والإعتماد على Package library(survival) – library(nleqslv)- library(boot)

❖ وصف المتغيرات المستخدمة فى البحث

اسم المتغير بالانجليزية	اسم المتغير باللغة العربية	التعريف
Gender	النوع	ويأخذ القيمة 1 وتعنى ذكر والقيمة صفر وتعنى أنثى .
Age	العمر	ويتم قياس الأعمار هؤلاء المرضى بالشهور من شهر الى 120 شهرا .
Date Diagnose	تاريخ التشخيص	وهو وقت دخول المريض إلى المستشفى وتشخيصة .
End date	تاريخ الخروج	تاريخ أخر مشاهدة لكل مريض .
Survival time	وقت المراقبة	الفرق ما بين تاريخ التشخيص وتاريخ الخروج .
Censoring	المراقبة	ويأخذ القيمة صفر تعنى أن المريض ظل على قيد الحياة والقيمة 1 وتعنى أن المريض قد توفى
Job	الوظيفة	وتنقسم إلى جزئين إما طالب أو دون سن المدرسة .
ICP	قياس ضغط الدم	يمثل درجات قياس ضغط الدم .

ثم قياس متغيرات تفسيرية (النوع _العمر _الوظيفة _قياس ضغط الدم _ وقت البقاء) ومتغير الاستجابة (البتر).تم تقدير كلا من النموذج اللوغاريتمي ونموذج بوكس كوكس بالإعتماد على ثلاثة متغيرات (العمر – النوع – الوظيفة) .

الحالة الأولى بالاعتماد على ثلاثة متغيرات

أولا : بالنسبة لمتغير العمر

من النتائج المتحصل عليها في الجدول(1) والخاصة بتقديرات النموذج شبه المعلمي لطريقة تقدير Box-Cox وطريقة التقدير اللوغاريتمية (Logarithm) يمكن تلخيص هذه النتائج عن طريق مقارنة نتائج التقديرات (EST) والإنحراف المعياري المقارب لها (ASE) والدلالة المعنوية (P-value) مع زيادة قيمة الـ θ وهي معلمة النموذج حيث افترضنا قيمة لـ θ=0.25,0.5,0.75,1,1.25,1.5,2.

(أ) بالنسبة للتقديرات شبه المعلمية لنموذج *Box-cox* يمكن استنتاج الملاحظات التالية:

يلاحظ أنه عندما كانت قيمة θ=0.25 أن معامل التقدير شبه المعلمي للمتغير العمر (Age) يساوي-0.0431 بإنحراف معياري مقارب قدره 0.01038وقيمة المعنوية المحسوبة (p-value) أقل من 0.001 مما يعني وجود دلالة معنوية لمتغير العمر وذلك عند مستوي الدلالة الإحصائية (0.05).

بصفة عامة نجد أنه عند جميع قيم θ فإن قيمة التقدير شبه المعلمي باستخدام نموذج Box-Cox لمعامل متغير العمر (Age) كلها كانت سلبية وفي تناقص مع زيادة قيمة الـ θ حتي تصل إلي قيمة θ=1 ثم بعد ذلك تتجه قيمة التقدير معامل متغير العمر للتزايد مع زيادة قيمة الـ θ ثم تتجه للتناقص مرة أخري، أما بالنسبة للإنحراف المعياري المقارب فنجد أنه مع زيادة قيمة الـ θ فإن الإنحراف المعياري المقارب يقل حتي تصل إلي قيمة θ=1 ثم بعد ذلك تتجه قيمة تقدير الإنحراف المعياري المقارب للتزايد مع زيادة قيمة الـ θ ثم تتجه للتناقص مرة أخري ، أما بالنسبة للمعنوية المحسوبة فإن متغير العمر له دلالة معنوية عند جميع قيم الـ θ مع ملاحظة أن قيمة المعنوية المحسوبة تقل كلما زادت قيمة الـ θ .

(ب)بالنسبة لتقديرات شبه المعلمية للنموذج اللوغاريتمي *Logarithm* يمكن استنتاج الملاحظات التالي

يلاحظ نجد أنه عندما كانت قيمة θ=0.5 معامل التقدير شبه المعلمي لمتغير العمر (Age) يساوي - 0.0208 بإنحراف معياري مقارب قدره 0.0096 وقيمة المعنوية المحسوبة (p-value) أقل من 0.0307مما يعني وجود دلالة معنوية لمتغير العمر وذلك عند مستوي الدلالة الإحصائية (0.05).بالنسبة لتقديرات معامل متغير العمر.

بصفة عامة نجد أنه عند جميع قيم θ فإن قيمة التقدير شبه المعلمي باستخدام نموذج Logarithm لمعامل متغير العمر (Age) كانت سلبية حتي قيمة $\theta=1$ ثم ايجابية بعد قيمة الـ θ أكبر من الواحد الصحيح وبالتالي فإن العلاقة طردية بين قيمة الـ θ والتقدير لمعامل متغير العمر، أما بالنسبة للإنحراف المعياري المقارب فنجد أنه مع زيادة قيمة الـ θ فإن الإنحراف المعياري المقارب يزداد ومنه نستنتج أن العلاقة طردية بين الإنحراف المعياري المقارب للتقدير وقيمة الـ θ، أما بالنسبة للمعنوية المحسوبة فإن متغير العمر له دلالة معنوية عند قيم الـ $\theta=0.25,0.5,2$ مع ملاحظة أن قيمة المعنوية المحسوبة تزداد كلما زادت قيمة الـ θ حتي تصل $\theta=0.75$ ثم بعد ذلك تتجه قيمة المعنوية المحسوبة للتناقص مع زيادة قيمة الـ θ.

جدول (1) التقديرات بالنسبه لمتغير العمر

	Transformation	Box-cox			Logarithm		
		EST	ASE	P-Value	EST	ASE	P-Value
age	θ=0.25	-0.0431	0.01038	0.0000	-2.7369	0.0058	2.47e-6
	θ=0.50	-0.0303	0.0070	1.55e-5	-0.0208	0.0096	0.0307
	θ=0.75	-0.0221	0.0041	5.33e-8	-0.0112	0.0139	0.4227
	θ=1	-0.0171	0.0052	0.0011	0.0188	0.0256	0.4620
	θ=1.25	-1.4180	0.0280	4.22e-7	0.0327	0.0251	0.1920
	θ=1.5	-0.0122	0.0025	7.11e-7	0.04086	0.0223	0.0675
	θ=2	-0.1041	0.0018	5.73e-9	0.0533	0.0228	0.0197

(أ) بالنسبة لمتغير النوع الموضح بالجدول (2) نجد أنه عندما كانت قيمة $\theta=0.5$ فإن معامل التقدير شبه المعلمي باستخدام نموذج Box-Cox لمتغير النوع (*Gender*) يساوى 0.4002- بإنحراف معياري مقارب قدره 0.8170 وقيمة المعنوية المحسوبة 0.6242 مما يعني عدم معنوية متغير النوع في العلاقة المفترضة حيث أن المعنوية المحسوبة أكبر من مستوي الدلالة الإحصائية (0.05). يلاحظ أنه عند جميع قيم θ فإن قيمة التقدير شبه المعلمي باستخدام نموذج Box-Cox لمعامل متغير النوع (*Gender*) كلها كانت سلبية وفي تناقص مع زيادة قيمة الـ θ، أما بالنسبة للإنحراف المعياري المقارب فنجد أنه مع زيادة قيمة الـ θ فإن الإنحراف المعياري المقارب يقل ومنه نستنتج أن العلاقة عكسية بين الخطأ المعياري للتقدير وقيمة الـ θ، أما بالنسبة للمعنوية المحسوبة فإن متغير النوع ليس له دلالة معنوية عند جميع قيم الـ θ مع ملاحظة أن قيمة المعنوية المحسوبة تقل كلما زادت قيمة الـ θ .

ب)بالنسبة للتقديرات شبه المعلمية للنموذج اللوغاريتمي Logarithm يمكن استنتاج الملاحظات التالية:

يلاحظ أنه عندما كانت قيمة $\theta=0.5$ فإن معامل التقدير شبه المعلمي باستخدام نموذج *Logarithm* لمتغير النوع (*Gender*) يساوي 0.7534- بإنحراف معياري مقارب قدره 5078.1 وقيمة المعنوية المحسوبة (p-value) 0.6173 مما يعني عدم معنوية متغير النوع في العلاقة المفترضة حيث أن المعنوية المحسوبة أكبر من مستوي الدلالة الإحصائية (0.05). بصفة عامة نجد أنه عند جميع قيم θ فإن قيمة التقدير شبه المعلمي باستخدام نموذج Logarithm لمعامل متغير النوع (Gender) كلها كانت سلبية ماعدا عند θ=2 حيث كان التقدير موجب وفي تناقص مع زيادة قيمة الـ θ مع ملاحظة أن التقدير يقل حتي تصل إلي قيمة θ=0.75 ثم بعد ذلك تتجه قيمة التقدير للتزايد مع زيادة قيمة الـ θ مرة أخري ومنه نستنتج أن العلاقة متغيرة بين التقدير وقيمة الـ θ بمعني أنها عكسية ثم طردية، أما بالنسبة للإنحراف المعياري المقارب فنجد أنه مع زيادة قيمة الـ θ فإن الإنحراف المعياري المقارب يزداد ومنه نستنتج أن العلاقة طردية بين الإنحراف المعياري المقارب للتقدير وقيمة الـ θ أما بالنسبة للمعنوية المحسوبة فإن متغير النوع ليس له دلالة معنوية عند جميع قيم الـ θ مع ملاحظة أن قيمة المعنوية المحسوبة تزداد كلما زادت قيمة الـ θ.

جدول (2) التقديرات بالنسبة لمتغير العمر

	Transformation	Box-cox			Logarithm		
		EST	ASE	P-Value	EST	ASE	P-Value
Gender	θ=0.25	-0.4002	0.8170	0.6242	-0.4661	0.7027	0.5071
	θ=0.50	-0.2657	0.3830	0.4879	-0.7534	1.5078	0.6173
	θ=0.75	-0.1843	0.2785	0.5080	-1.1831	1.9909	0.5523
	θ=1	-0.1380	0.2180	0.5265	-0.6502	3.3098	0.8442
	θ=1.25	-0.1109	0.1764	0.5296	-0.4117	3.5403	0.9074
	θ=1.50	-0.0935	0.1501	0.5333	-0.1926	3.9664	0.9612
	θ=2	-0.0749	0.1162	0.5190	0.3131	5.8273	0.9571

أ)بالنسبة للتقديرات شبه المعلمية لنموذج Box-cox يمكن استنتاج الملاحظات التالية نجد أن معامل التقدير شبه المعلمي لمتغير الوظيفة (Job) يساوي0.5193 بإنحراف معياري مقارب قدره 0.4364 وقيمة المعنوية المحسوبة 0.2340 (p-value) مما يعني عدم الدلالة المعنوية لمتغير الوظيفة حيث أن المعنوية المحسوبة أكبر من مستوي الدلالة الإحصائية (0.05). بصفة عامة نجد أنه عند جميع قيم θ فإن قيمة التقدير شبه المعلمي باستخدام نموذج Box-Cox لمعامل متغير الوظيفة (Job) كلها كانت سلبية ماعدا عند قيمة الـ 0.25θ= وفي تناقص مع زيادة قيمة الـ θ أما بالنسبة للإنحراف المعياري المقارب فنجد أنه مع زيادة قيمة الـ θ فإن الإنحراف المعياري المقارب يقل ومنه نستنتج أن العلاقة عكسية بين الخطأ المعياري للتقدير وقيمة الـ θ، أما بالنسبة للمعنوية المحسوبة فإن متغير الوظيفة ليس له دلالة معنوية عند جميع قيم الـ θ ماعدا عند قيمة θ=2 حيث كانت دلالة المتغير معنوية مع ملاحظة أن قيمة المعنوية المحسوبة تقل كلما زادت قيمة الـ θ .

ب بالنسبة لتقديرات شبه المعلمية للنموذج اللوغاريتمي Logarithm يمكن استنتاج الملاحظات التالية:- نجد أن معامل التقدير شبه المعلمي لمتغير الوظيفة (*Job*) يساوي 1.8267 بخطأ معياري قدره 1.4988 وقيمة المعنوية المحسوبة 0.2229 (p-value) مما يعني عدم الدلالة المعنوية .

ثالثا بالنسبة لتقديرات معامل متغير الوظيفة (*Job*):

بصفة عامة نجد أنه عند جميع قيم θ فإن قيمة التقدير شبه المعلمي باستخدام نموذج Logarithm لمعامل متغير الوظيفة (Job) كلها كانت موجبة وفي تزايد مع زيادة قيمة الـ θ، أما بالنسبة للإنحراف المعياري فنجد أنه مع زيادة قيمة الـ θ فإن الإنحراف المعياري المقارب يقل حتي تصل إلي قيمة θ=0.75 ثم بعد ذلك تتجه قيمة تقدير الإنحراف المعياري المقارب للتزايد مع زيادة قيمة الـ θ مرة أخري، أما بالنسبة للمعنوية المحسوبة فإن متغير الوظيفة ليس له دلالة معنوية عند جميع قيم الـ θ مع ملاحظة أن قيمة المعنوية المحسوبة تقل حتي تصل عند قيمة θ=0.75 ثم بعد ذلك تتجه للزيادة مرة أخري مع زيادة قيمة الـ θ المحسوبة أكبر من مستوي الدلالة الإحصائية (0.05).

جدول (3) بالنسبة لتقديرات متغير الوظيفه

	Transformation	Box-cox			Logarithm		
		EST	ASE	P-Value	EST	ASE	P-Value
job	θ=0.25	0.5193	0.4364	0.2340	0.0045	0.7468	0.9952
	θ=0.50	-0.6320	0.2766	0.0223	1.8267	1.4988	0.2229
	θ=0.75	-0.1843	0.2785	0.5080	-1.1831	1.9909	0.5523
	θ=1	-0.1380	0.2180	0.5265	-0.6502	3.3098	0.8442
	θ=1.25	-0.1109	0.1764	0.5296	-0.4117	3.5403	0.9074
	θ=1.50	-0.0935	0.1501	0.5333	-0.1926	3.9664	0.9612
	θ=2	-0.0749	0.1162	0.5190	0.3131	5.8273	0.9571

الحالة الثانية : بالإعتماد على متغير واحد فقط (العمر) والموضحة فى جدول (4)

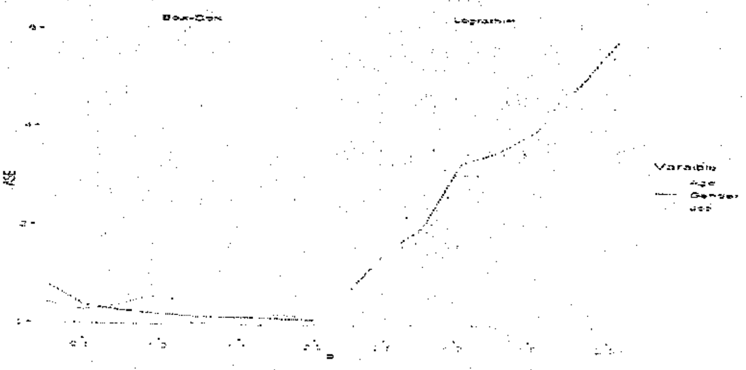
أولاً: بالنسبة للتقديرات شبه المعلمية لنموذج Box-cox والنموذج اللوغاريتمي يمكن استنتاج الملاحظات التالية عند θ = 0.5 نلاحظ أن القيمة المعنوية المحسوبة (p-value) أقل من 0.001 مما يعني معنوية متغير العمر في العلاقة المفترضة حيث أن المعنوية المحسوبة أقل من مستوي الدلالة الأحصائية (0.05). بينما في التقدير شبه المعلمي باستخدام نموذج Box-Cox لثلاث المتغيرات (النوع والعمر والوظيفة) نجد أن معامل التقدير شبه المعلمي للمتغير العمر (Age) يساوي -0.0431 بإنحراف معياري مقارب قدره 0.01038 وقيمة المعنوية المحسوبة (p-value) أقل من 0.001 مما يعني وجود دلالة معنوية

لمتغير العمر وذلك عند مستوي الدلالة الأحصائية (0.05). ونلاحظ أن الأشارة مازالت سالبة وان الدلالة المعنوية موجودة في الحالتين. وقيمة المعنوية المحسوبة (p-value) أقل من 0.001 مما يعني وجود دلالة معنوية لمتغير العمر وذلك عند مستوي الدلالة الأحصائية (0.05). ونلاحظ أن الأشارة مازالت سالبة وان الدلالة معنوية فى الحالتين ،في الرسم التوضيحي للتقديرات والإنحراف المعيارى المعياري يمكن التحقق من الملاحظات السابقة لطريقة التقدير شبه المعلمي باستخدام النموذجين

جدول (4) تقدرات النماذج بالإعتماد على متغير العمر فقط

θ	Transformation					
	Box-Cox			Logarithm		
	EST	ASE	p-value	EST	ASE	p-value
0.25	-0.0340	0.0061	0.00000	-0.0283	0. 0050	0.0000
0.50	-0.027	0.0047	0.00000	-0.0332	0. 0096	0.0006
0.75	-0.0183	0.0040	0.00001	-0.0192	-	-
1	-0.0142	0.0038	0.00017	0.0040	0. 0249	0.8723
1.25	-0.0118	0.0026	0.00000	0.0172	0. 0295	0.5606
1.5	-0.0102	0.0021	0.0000	0.0273	0.0325	0.4016
2	-	-	-	-	-	-

في الرسم التوضيحي للتقديرات والإنحراف المعيارى المعياري يمكن التحقق من الملاحظات السابقة لطريقة التقدير شبه المعلمي بأستخدام النموذجين



شكل(1)مقارنة نماذج Box–Cox النماذج اللوغاريتمية لمتوسط الإنحراف المعياري

7. النتائج

يمكن إيجاز النتائج التى تم التوصل إليها فى هذا البحث فى النقاط التالية :

1) طريقة الإمكان الأعظم الشرطية CMLE تعد اختيارا جيدا من حيث الكفاءة والقوة عندما يكون التوزيع الأساسى لوقت الاقتطاع غير معروف .

2) أظهرت نتائج (CMLE) أنها أكثر كفاءة من النتائج التى تعتمد على تقدير المعادلات (EE).

3) عندما تكون متجة المتغيرات المشتركة Z منفصلة فإن β أكثر فاعلية .

(5) من خلال الجدول التالى يتضح أن نموذج بوكس كوكس هو النموذج الأفضل باستخدام الإنحراف المعيارى المقارب .

جدول (5)

النموذج الافضل	Logarithm	Box-Cox	Transformation	
	ASE	ASE	θ	المتغير
Box-Cox	0.0096	0.007	0.5	AGE
	0.0139	0.0041	0.75	
	0.0256	0.0052	1	
	0.0223	0.0025	1.5	
	0.0228	0.0018	2	
	ASE	ASE		المتغير
Box-Cox	1.5078	0.383	0.5	GEND ER
	1.9909	0.2785	0.75	
	3.3098	0.218	1	
	3.9664	0.1501	1.5	
	5.8273	0.1162	2	
	ASE	ASE		المتغير
Box-Cox	1.4988	0.2766	0.5	JOB
	1.7392	0.3948	0.75	
	3.0814	0.6172	1	
	3.7179	0.2446	1.5	
	5.5625	0.1585	2	

8. التوصيات

(1) استخدام طريقة CMLE لأنها من أكثر المقدرات كفاءة (يقال أن الإحصاء $\hat{\theta}$ كاف لـ θ إذا كان $\hat{\theta}$ يحتوي على جميع المعلومات حول θ المتوفرة في متغير البيانات بأكمله x ؛ رسميًا $\hat{\theta}$ يكفي لـ θ إذا كان التوزيع الشرطي لـ X المعطى $\hat{\theta}$ لا يعتمد على θ).

(2) تقييم مدى كفاية نماذج التحويل الشبه معلمية لضمان صحة الاستدلال ودقة التنبؤ. توصى الدراسة بتوسيع هذا المنهج على البيانات المبتورة من ناحية اليسار والخاضعة للرقابة من ناحية اليمين (LTRC)

(3) توصى الدراسة باستخدام اختبار جودة المطابقة لتحديد النموذج الأفضل فى نماذج التحويل شبه المعلمية.

8. Zeng, D., & Lin, D. Y. (2007). Maximum likelihood estimation in semiparametric regression models with censored data. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 69(4), 507-564

9. Zeng, D., & Lin, D. Y. (2006). Efficient estimation of semiparametric transformation models for counting processes. Biometrika, 93(3), 627-640

9. المراجع

1. Al-Mosawi, R. R., & Aziz, A. (2020). On Parametric Linear Transformation Model with Left-Truncated and Interval-Censored Data. Journal of Education for Pure Science-University of ThiQar, 10(1), 100-111
2. Chen, C. M., & Shen, P. S. (2018). Conditional maximum likelihood estimation in semiparametric transformation model with LTRC data. Lifetime data analysis, 24(2), 250-272
3. Chen, K., Jin, Z., & Ying, Z. (2002). Semiparametric analysis of transformation models with censored data. Biometrika, 89(3), 659-668.
4. Li, S., Hu, T., Wang, P., & Sun, J. (2018). A class of semiparametric transformation models for doubly censored failure time data. Scandinavian Journal of Statistics, 45(3), 682-698
5. Sokullu, S., & Stouli, S. (2017). Cross-Validation Selection of Regularisation Parameter (s) for Semiparametric Transformation Models. Annals of Economics and Statistics/Annales d'Économie et de Statistique, (128), 67-108
6. Shen, P. S. (2011). Semiparametric analysis of transformation models with left-truncated and right-censored data. Computational Statistics, 26(3), 521-537.
7. Zhang, Z. (2016). Semiparametric estimation of partially linear transformation models under conditional quantile restriction. Econometric Theory, 32(2), 458-497

1- Introduction

Recently, firms have increasingly been competing in business environments characterized by short product life cycles, globally extended supply chains, and volatile demand patterns. Supply chains no longer have only to deal with the risks associated with the uncertainty of demand or supply variability, but also now have to deal with economic and environmental disruption events (Chiang et al., 2012, Snyder, 2012) Given such developments, the cultivation of supply chain agility has been suggested as an effective response strategy (Braunscheidel and Suresh, 2009; Swafford et al., 2006).

Supply chain agility has been considered a lot recently as a way for organizations to rapidly reply to changing business environment and improve their customer service levels (Mehralian et al., 2015). Supply chain agility is widely considered to be the most critical success factor in today's competitive marketplace (Goldman, Nagel, & Preiss, 1995), as an agile supply chain enables its member firms to be more market-sensitive, more capable of synchronizing supply with demand, and better able to achieve shorter cycle times (Chan et. al., 2016).

Furthermore, Trends towards greater interconnectivities, interdependencies, and cost-efficient philosophies has increased the potential impact of supply chain disruptions. This means that supply chains are increasingly vulnerable to disruption and at risk of losses and supply chain breakdown, which compromise supply chain survivability. Thus, supply chain capabilities for agility and resilience represents an emerging supply chain concern alongside longstanding recognition of the need for supply chains to strategically develop and/or combine their capabilities to