مجلة التجارة والتمويل

https://caf.journals.ekb.eg/

كلية التجارة – جامعة طنطا

العدد : الرابع

ديسمبر 2023

(الجزء الثانى)

التقدير اللامعلمي لدالة البقاء ثنائية المتغيرات للبيانات المراقبة من اليمين

Non-Parametric Estimation of Bivariate Survival Function for Right Censored Data

إعداد

هبة أحمد حسن أحمد

المدرس المساعد بقسم الإحصاء والرياضة والتأمين

كلية التجارة – جامعة دمنهور

hebaahmedhassan1982@gmail.com

الأستاذ الدكتور

سهير فهمي حجازي

أستاذ الإحصاء – كلية تجارة – جامعة طنطا

sohair.higazi@ commerce.tanta.edu.eg

الأستاذ الدكتور

دينا حسن عبد الهادي

أستاذ الإحصاء – كلية تجارة – جامعة طنطا

dina.abdelhady@commerce.tanta.edu.eg

المستخلص

يعد التقدير اللامعلمي لدالة البقاء ثنائية المتغيرات $S(t_x, t_y)$ في ظل وجود "**المراقبة**" ذا أهمية كبيرة في التطبيقات، واعتمدت أغلب المقدرات اللامعلمية لهذه الدالة على مقدر "**KM**" (**Kaplan and Meier, 1958**)، وأوضح **Shafiq *et al.* (2007)** أن هذا المقدر يعاني من كثير من القصور. يهدف هذا البحث بإحلال مقدر "**KM**" بمقدرات "**NA**" (**Nelson,1972; Aalen, 1978**)، "**S-VR**" (**Susarla and Van Ryzin, 1980**)، "**WKM**" (**Jan *et al.*, 2005**)، "**MWKM**" (**Shafiq *et al.*, 2007**)، "**Huang**" (**Huang, 2008**) للحصول على مقدرات جديدة **بالتعديل** على مقدر **الاستقلال** اللامعلمي الأصلي المعتمد على أسلوب "**KM**" $\widehat{S}_{KM}(t_x, t_y)$ (**Wang and Zafra, 2009**) في حالة "**استقلال**" وقت وقوع الحدثين (T_x, T_y) ووجود مراقبة "**مستقلة**" في محاولة للوصول إلى مقدرات تتلافى بعض عيوب المقدرات السابقة مثل مقدر "**DB**" (**Dabrowska, 1988**)، ومقدر "**DV**" (**Wang and Zafra, 2009**). تم إجراء دراسة المحاكاة باستخدام لغة البرمجة "**R**" عند حجمي للعينة ($n = 30$ ، **60**)، ونسبتي للمراقبة ($CR =$ **30%** ، **40%**)، وخمسة احتمالات بقاء حقيقية ($S(t_x, t_y) =$ **25%** ، **50%** ، **70%** ، **75%** ، **99%**)، وذلك لتقييم مقدر **الاستقلال** اللامعلمي الأصلي $\widehat{S}_{KM}(t_x, t_y)$ وأيضًا المقدرات المقترحة بالتعديل عليه في حالة المراقبة المستقلة واستقلال وقت وقوع الحدثين. وتمت المقارنة بين المقدرات اللامعلمية المختلفة للدالة $S(t_x, t_y)$ باستخدام مقدار التحيز التجريبي ومقدار التباين التجريبي والكفاءة النسبية التجريبية وتقديرات احتمالات البقاء ثنائية المتغيرات التجريبية. وأظهرت دراسة المحاكاة وفقًا لمعايير المفاضلة العديد من النتائج من أهمها: "**أفضلية**" التوفيق اللامعلمي لمقدر **الاستقلال** اللامعلمي المقترح $\widehat{S}_H(t_x, t_y)$ على جميع التوفيقات اللامعلمية الأخرى لمقدر **الاستقلال** اللامعلمي الأصلي $\widehat{S}_{KM}(t_x, t_y)$ حيث ثبت أنه "**الأقل تحيزًا والأقل تباينًا والأقل تقديرًا لاحتمالات البقاء ثنائية المتغيرات التجريبية والأكثر كفاءة**"، أيضًا أوضحت النتائج "**تفوق**" مقدر **الاستقلال** اللامعلمي المقترح $\widehat{S}_H(t_x, t_y)$ على جميع المقدرات اللامعلمية المقترحة الأخرى بزيادة حجم العينة. كما تم تطبيق المقدرات اللامعلمية المبحوثة على "**بيانات فعلية**" لعينة من الهيئة المعاونة (**فئة معيد**) بكلية تجارة جامعة دمنهور معرضة لمراقبة "**مستقلة**" وأحداث "**مستقلة**".

الكلمات المفتاحية: البيانات المراقبة من اليمين، تحليل البقاء ثنائي المتغيرات، دالة البقاء ثنائية المتغيرات، المراقبة المستقلة، مقدر **الاستقلال** اللامعلمي ثنائي المتغيرات.

Abstract

Nonparametric estimation of the bivariate survival function $S(t_x, t_y)$ in the presence of censoring is of great interest in many applications, and most of nonparametric estimators of this function were based on the **"KM"** estimator **(Kaplan and Meier, 1958)**, and **Shafiq *et al.* (2007)** explained show that this estimator suffer from many weaknesses. This research aims to replace the **"KM"** estimator with the **"NA" (Nelson, 1972; Aalen, 1978)**, **"S-VR" (Susarla and Van Ryzin, 1980)**, **"WKM" (Jan *et al.*, 2005)**, **"MWKM" (Shafiq *et al.*, 2007)**, and **"Huang" (Huang, 2008)** estimators to obtain new estimators by modifying the original nonparametric **Independence** estimator based on the **"KM"** method $\widehat{S}_{KM}(t_x, t_y)$ **(Wang and Zafra, 2009)** in the case of the independence of the time of occurrence of the two events (T_x, T_y) and the presence of independent censoring in an attempt to reach estimators that avoid some of the drawbacks of previously used estimators such as **"DB"** estimator **(Dabrowska, 1988)**, and **"DV"** estimator **(Wang and Zafra, 2009)**. Simulation study was conducted using the **"R"** programming language at two sample sizes **(n = 30 , 60)**, two censored ratios **(CR = 40% , 30%)**, and five true survival probabilities **($S(t_x, t_y)$ = 25% , 50% , 70% , 75% , 99%)** in order to evaluate the original nonparametric **Independence** estimator $\widehat{S}_{KM}(t_x, t_y)$ and also the proposed estimators by modifying it in the case of independent censoring and the independence of the time of occurrence of the two events.Comparison between the different nonparametric estimators of $S(t_x, t_y)$ were performed using the empirical bias, the empirical variance, empirical relative efficiency, and estimates of bivariate survival probabilities criteria. Simulation study indicate that the proposed estimator $\widehat{S}_H(t_x, t_y)$ have less bias, less variability (consistent) and more efficient than all other nonparametric fits of the original nonparametric **Independence** estimator and that $\widehat{S}_H(t_x, t_y)$ **"outperformed"** all other proposed non-parametric estimators with increasing the sample size. The studied nonparametric estimators were also applied to **"actual data"** for a sample of the assistant body (teaching assistant category) at the Faculty of Commerce at Damanhour University, exposed to **"independent"** Censoring and **"independent"** events.

Keywords: **"KM"** Method; Right Censoring; Bivariate Survival Analysis; Bivariate Survival Function; Independent Censoring; Bivariate Nonparametric **Independence** Estimator.

١. مقدمة Introduction

يرجع تطور "**تحليل البقاء**" Survival Analysis إلى القرن السابع عشر مع أول "**جدول حياة**" Life Table تم تطويره في عام **(1662)**، حيث كان "**حدث الوفاة**" هو حدث الاهتمام. ويستخدم "**مصطلح البقاء**" لجميع أنواع الأحداث وفي مجالات عديدة على سبيل المثال، حدث "**الوفاة**" (الطب الحيوي)، حدث "**تحليل المدة**" Duration Analysis (الاقتصاد)، حدث "**الزواج**" (الديموغرافيا)، حدث "**تحليل تاريخ الأحداث**" Event–History Analysis (علم الاجتماع)، وحدث "**الموثوقية وتحليل وقت الفشل**" Reliability and Failure Time Analysis (الهندسة) **(Dai et al., 2016)**.

ويواجه الباحثون العديد من "**الصعوبات**" في كثير من تطبيقات "**تحليل البقاء**"، وأهم هذه الصعوبات: انتهاء الفترة الزمنية المحددة للدراسة وانسحاب و/أو فقد بعض المشاهدات، وأيضًا الأخطار المتنافسة Competing Risks **(Emura *et al.*, 2019)**. ونتيجة لهذه الصعوبات يكون لدي الباحثون (في نهاية الدراسة) "**نوعين**" من البيانات وهما:

١. "**بيانات كاملة**" **Complete Data**: وهي معلومات كاملة عن المتجه العشوائي $(T_y، T_x)$، حيث T_x هو وقت وقوع الحدث الأول بينما T_y هو وقت وقوع الحدث الثاني وذلك لمشاهدة واحدة. ولا يواجه الباحثون أي صعوبات عند التحليل الإحصائي لهذه البيانات.

٢. "**بياتات مراقبة**" **Censored Data**: وهي تشمل بيانات جزئية عن المتجه العشوائي $(T_y، T_x)$. وتنتشر البيانات المراقبة بشكل شائع في تخصصات مختلفة مثل الاقتصاد والهندسة وعلوم الحياة **(Qi *et al.*, 2022)**.

وقد اكتشفت مشكلة "**البيانات المراقبة**" في تحليل البقاء ثنائي المتغيرات Bivariate Survival Analysis منذ بداية الثمانيات. وتظهر "**بيانات وقت الحدث ثنائية المتغيرات**" Bivariate Event Time Data عند إجراء دراسة البقاء على "**وحدات مزدوجة**" Paired Units مثل التوائم أو الأعضاء المتطابقة للمشاهدة **(Prentice and Cai, 1992)**. وفي هذه الحالة تأتي المفردات "**كأزواج**"، ويكون هناك "**قياسان زمنيان**" يجب مراعاتهما لكل مكون

في الزوج. ويرجع ظهور هذه البيانات بشكل متكرر في البحث العلمي إلى أن كل مشاهدة في الدراسة تتعرض "**لنوعين من الأحداث**" (على سبيل المثال، المصاب بفيروس نقص المناعة البشرية يصاب بالإيدز ويتوفى في النهاية) أو لأن هناك "**ارتباطًا طبيعيًا أو اصطناعيًا**" مثل إن أوقات الحدث داخل نفس الزوج تكون مرتبطة ببعضها البعض (على سبيل المثال، فقدان البصر الشديد في العين اليمنى واليسرى) **(Lin and Ying, 1993)**.

وتتميز "**بيانات البقاء**" Survival Data عن جميع أنواع البيانات الأخرى في أنها تتطور "**ديناميكيًا**" مع الزمن **(Hougaard, 2000)**، وتتسم "**بيانات البقاء**" بالآتي: يتم مشاهدة البيانات "**جزئيًا**"، وتكون بيانات "**خاصة**" تتم مشاهدتها وملاحظتها، حيث إن وقت وقوع الحدث يكون معروفًا، ولكن قد يحدث ذلك بعد وقت تقييم البيانات، وفي هذه الحالة يكون لدينا "**بيانات مراقبة**". كما أن العديد من "**بيانات البقاء**" يمكن أن تكون "**شرطية**" Conditioned على ما هو معروف حاليًا وقد يتغير هذا بمرور الزمن وهذا يبدو طبيعيًا للغاية؛ حيث إننا نأخذ في الاعتبار فقط مثلاً "**خطر الحدث**" Event Risk للمشاهدات التي لم تتعرض للحدث.

وأشار **Emura and Michimae (2022)** إلى أن تحليل البقاء (أو الموثوقية) يتعامل مع بيانات وقت الحدث غير الكاملة حيث أن إمكانية ملاحظتها بشكل كامل نادر الحدوث. وأوضح **Janssen and Veraverbeke (2023)** أن التحدي الرئيسي في تحليل بيانات البقاء هو تطوير أساليب إحصائية تأخذ في الاعتبار "**المعلومات غير الكاملة**" الموجودة في "**البيانات المراقبة**". وقد حظيت مشكلة "**البيانات المراقبة**" باهتمام كبير في أغلب الأدبيات الإحصائية. وبذل العديد من الإحصائيين جهودهم في حل هذه المشكلة. حيث قام **Lin and Ying (1993)**، وأيضًا **Wang and Wells (1997)** بتقدير دالة البقاء ثنائية المتغيرات Bivariate Survival Function $S(t_x, t_y)$ في ظل بعض آليات المراقبة الخاصة Special Censoring Mechanisms. كما اقترح **Prentice and Cai (1992)** مقدرًا لتقدير الدالة $S(t_x, t_y)$ يعتمد على دوال البقاء الهامشية Marginal Survival Functions لها ودالة التغاير Covariance Function لها. ويعد أوائل من اقترحوا "**مقدرات لامعلمية**" Nonparametric Estimators لتوفيق "**دالة بقاء المجتمع ثنائية المتغيرات**" Bivariate

Population Survival Function كل من **Dabrowska (1988)**، **Pruitt (1991)**، **van der Laan (1996)**.

ومن الجدير بالذكر، إن العديد من هذه المقدرات ليس لها أي توزيعات احتمالية حقيقية، وذات صيغ غير واضحة، ويعتمد الكثير منها (بشكل كبير) على اختيار "**المعالم الممهدة**" Smoothing Parameters. وبعض هذه المقدرات إما "**متحيزة**" Biased، أو "**غير كفء**" Inefficient أو "**غير متسقة**" Inconsistent أو "**غير محددة بالكامل في نطاق دوال البقاء المشتركة**" وبالتالي فهي غير قابلة للتطبيق العملي.

ويتم تصنيف "**البيانات المراقبة**" طبقًا "**لمعيارين**" هما:

(١) **معيار آلية الوقوع:** طبقًا لهذا المعيار صنف **Miller *et al.* (1980)** "**البيانات المراقبة**" إلى "**ثلاثة**" أنواع هي: **(أ)** "**البيانات المراقبة من اليسار**" Left Censored Data ويقصد بها بيانات المشاهدات التي "**وقع**" لها الحدث "**قبل**" بداية الدراسة، **(ب)** "**البيانات المراقبة خلال فترة**" Interval Censored Data وهي بيانات المشاهدات التي "**وقع**" لها الحدث "**خلال فترة محددة**" من الدراسة، ولكن وقت وقوع الحدث "**غير معلوم**"، وأخيرًا **(ج)** "**البيانات المراقبة من اليمين**" Right Censored Data وهي بيانات المشاهدات التي "**لم يقع**" لها الحدث "**أثناء**" فترة الدراسة.

(٢) **معيار النوع:** تنقسم "**البيانات المراقبة**" وفقًا لهذا المعيار إلى:-

أ. "**البيانات المراقبة من النوع الأول**" **Type-I Censoring Data**: تنتشر هذه البيانات في "**التطبيقات الهندسية**" وتعد "**الأكثر**" شيوعًا في الواقع، نظرًا لأن "**إيقاف**" التجربة يتم بعد مرور فترة معينة أو عند نقطة زمنية معينة. وتوصف البيانات بأنها مراقبة من النوع "**الأول**" إذا كان الوقت المراقب Censored Time "**ثابتًا**"، وكان عدد المشاهدات المعرضة للحدث "**عشوائيًا**"، بالإضافة إلى أن "**جميع**" المشاهدات تدخل إلى الدراسة "**عند نفس الوقت**" (**Ramos *et al.*, 2020**).

ب. "**البيانات المراقبة من النوع الثاني**" **Type-II Censoring Data**: تنتشر هذه البيانات في "**التطبيقات الهندسية**". وتعد البيانات مراقبة من النوع "**الثاني**" إذا كان الوقت المراقب "**عشوائيًا**"، وكان عدد المشاهدات المعرضة للحدث "**ثابتًا**"، بالإضافة إلى أن "**جميع**" المشاهدات تدخل إلى الدراسة "**عند نفس الوقت**". وبصفة عامة، فإن إيقاف التجربة يتم بعد تعرض عدد معين من المشاهدات للحدث والذي يشار إليها بالرمز "r" **(Ramos *et al.*, 2020)**.

ج. "**البيانات الموقوفة عشوائيًا**" **Random Censoring Data**: تنتشر هذه البيانات في "**المجالات الطبية**". وهناك صعوبة في تحديد هذا النوع من البيانات مما يتطلب تعريفا δ (مؤشر الحدث Event Indicator) لكل مشاهدة في الدراسة، كما أن دخول المشاهدات إلى الدراسة يكون "**غير متزامن**" **(Miller *et al.*, 1980)**.

وتندرج البيانات المراقبة من النوعين **الأول، والثاني، والعشوائية** تحت البيانات المراقبة "**من اليمين**"، نظرًا لأن المتغير الذي نهتم به يستغرق وقتًا كبيرًا لحدوثه، وبالتالي لا نستطيع مشاهدته أثناء فترة الدراسة **(Ramos *et al.*, 2020)**. ونظرًا لأن البيانات المراقبة من "**اليمين**" تعد من أكثر الأنواع شيوعًا فسوف يقتصر هذا البحث على تحليل هذا النوع من البيانات فقط.

إن "**التوصيف الخاطئ**" للبيانات باختيار "**توزيع معلمي**" قد يؤدي إلى الحصول على مقدرات لدالة البقاء ثنائية المتغيرات لا تتمتع بخصائص جيدة. خاصة وأن بعض الظواهر التي يختص بها تحليل بيانات البقاء يصعب التعرف على توزيعاتها، وبالتالي لا يكون من السهل تقدير دالة البقاء ثنائية المتغيرات "**معلمي**". وفي المقابل، يعد "**التقدير اللامعلمي**" لدالة البقاء ثنائية المتغيرات في ظل وجود "**المراقبة**" ذا أهمية كبيرة في التطبيقات، خاصة مع تعدد الأدبيات الإحصائية التي تناولت التقدير اللامعلمي ثنائي المتغيرات مقارنةً بالأدبيات التي تناولت التقدير المعلمي. حيث أصبحت الطرق اللامعلمية "**المفضلة**" في تحليل بيانات البقاء

لعدة "**أسباب**" تتمثل في: **(أ)** احتواء البيانات على بعض الخصائص التي لا يمكن الحصول عليها بسهولة بواسطة "**النماذج المعلمية**" Parametric Models، **(ب)** هناك بعض المجالات التي يكون فيها "**توفيق**" Fit النموذج المعلمي "**غير مناسب**"، وأيضًا **(ج)** تعتبر "**الطرق اللامعلمية**" هي النتيجة المنطقية في حالة "**الشرطية**" **(Hougaard, 2000)**.

وتتسم معظم "**المقدرات اللامعلمية**" لدالة البقاء ثنائية المتغيرات على سبيل المثال مقدر **Dabrowska (1988)**، ومقدر **Pruitt (1991)**، ومقدر **Prentice and Cai (1992)**، ومقدر **van der Laan (1996)** بالتحيز وعدم الاتساق وعدم الكفاءة خاصةً في ظل وجود "**مراقبة كثيفة**" Heavy Censoring. إضافةً لذلك، اعتمدت أغلب صيغ مقدرات دالة البقاء ثنائية المتغيرات على أسلوب "**KM**" الذي ينتج عنه تقديرات "**غير دقيقية ومتحيزة لأعلى**" لاحتمالات البقاء Survival Probabilities.

لهذه الأسباب، يحاول البحث الوصول إلى مقدر لامعلمي مناسب لتقدير دالة البقاء ثنائية المتغيرات للبيانات المراقبة من اليمين وذلك للتغلب على بعض عيوب المقدرات السابق استخدامها فيما يتعلق بالتحيز والتباين والكفاءة وتقديرات احتمالات البقاء، من خلال إجراء **تعديل** على مقدر **الاستقلال** اللامعلمي $\widehat{S}_{KM}(t_x, t_y)$ في حالة "**استقلال**" وقت وقوع الحدثين ووجود مراقبة "**مستقلة**"، وذلك باستبدال مقدر "**KM**" **(Kaplan and Meier, 1958)** المستخدم في تقدير "**دالة البقاء الهامشية لوقت وقوع الحدث**" $S(.)$ بكل من: **مقدر** "**NA**" **(Nelson,1972; Aalen, 1978)**، **مقدر** "**S-VR**" **(Susarla and Van Ryzin, 1980)**، **مقدر** "**WKM**" **(Jan *et al.*, 2005)**، **مقدر** "**MWKM**" **(Shafiq *et al.*, 2007)**، **وأخيرًا مقدر** "**Huang**" **(Huang, 2008)**.

ويهدف البحث إلى: **(١)** "**اقتراح**" بعض المقدرات اللامعلمية الجديدة لدالة البقاء ثنائية المتغيرات "**بالتعديل**" على مقدر **الاستقلال** اللامعلمي الأصلي $\widehat{S}_{KM}(t_x, t_y)$ في حالة "**استقلال**" وقت وقوع الحدثين ووجود مراقبة "**مستقلة**"؛ **(٢)** "**المقارنة**" بين أداء مقدر **الاستقلال** اللامعلمي الأصلي $\widehat{S}_{KM}(t_x, t_y)$ والمقدرات اللامعلمية المقترحة (خمس مقدرات) لاختيار أفضل المقدرات اعتمادًا على نتائج "**دراسة المحاكاة**"، وتتم المقارنة وفقًا **(١)** لمقدار التحيز

التجريبي Empirical Bias Amount (EBA)، **(٢)** مقدار التباين التجريبي Empirical Variance Amount (EVA)، **(٣)** الكفاءة النسبية التجريبية Empirical Relative Efficiency (ERE)، وأخيرًا **(٤)** تقديرات احتمالات البقاء التجريبية Empirical Survival Probabilities Estimates.

ويشمل هذا البحث "**ثمانية**" أجزاء، يتم عرض هيكل بيانات البقاء ثنائية المتغيرات المراقبة من اليمين في الجزء **الثاني**، ويتم عرض مقدر **الاستقلال** اللامعلمي الأصلي $\widehat{S}_{KM}(t_x, t_y)$ في الجزء **الثالث**، بينما يتم عرض المقدرات المقترحة بالتعديل على مقدر **الاستقلال** اللامعلمي الأصلي $\widehat{S}_{KM}(t_x, t_y)$ في الجزء **الرابع**، في حين يتم عرض الآلية التي اتبعت في دراسة المحاكاة في الجزء **الخامس**، ويتم عرض نتائج المحاكاة في الجزء **السادس**، بينما يتم عرض الدراسة التطبيقية ونتائجها في الجزء **السابع**، وأخيرا يتم عرض توصيات البحث في الجزء **الثامن**.

٢. هيكل بيانات البقاء ثنائية المتغيرات المراقبة من اليمين

يمكن وصف هيكل بيانات البقاء ثنائية المتغيرات المراقبة من اليمين من خلال وجود n من أزواج "**أوقات وقوع الحدث**" المستقلة والتي لها نفس التوزيع $(T_{x_i},\ T_{y_i})$، وكذلك وجود n من أزواج "**أوقات المراقبة**" Censored Times المستقلة والتي لها نفس التوزيع $(T_{C_{1i}}, T_{C_{2i}})$، وأيضًا "**استقلال**" المتجه $(T_{C_{1i}}, T_{C_{2i}})$ عن المتجه $(T_{x_i},\ T_{y_i})$. وتتكون "**البيانات المشاهدة**" the observed data من المتجهات العشوائية $(\widetilde{T}_{x_i},\ \widetilde{T}_{y_i}, \delta_i^x\ , \delta_i^y)(i = 1, 2, \dots, n)$، حيث:

$$(\widetilde{T}_{x_i}, \widetilde{T}_{y_i}) = \{min(T_{x_i}, T_{C_{1i}}),\ min(T_{y_i}, T_{C_{2i}})\}$$

$$(\delta_i^x, \delta_i^y) = \{I(T_{x_i} \le T_{C_{1i}}),\ I(T_{y_i} \le T_{C_{2i}})\}$$

$$I(T_{x_i} \le T_{C_{1i}}) = \begin{cases} 1, & T_{x_i} \le T_{C_{1i}} \\ 0, & T_{x_i} > T_{C_{1i}} \end{cases}$$

$$I(T_{y_i} \le T_{C_{2i}}) = \begin{cases} 1, & T_{y_i} \le T_{C_{2i}} \\ 0, & T_{y_i} > T_{C_{2i}} \end{cases}$$

وتمثل $S(t_x, t_y)$ "**دالة البقاء ثنائية المتغيرات**" Bivariate Survival Function وهي "احتمال أن يتعدى زمن وقوع الحدث **الأول** القيمة x والحدث **الثاني** القيمة y (**Rodríguez, 2005**)". حيث

$$S(t_x, t_y) = pr(T_x > t_x, T_y > t_y) \qquad (1)$$

تستخدم الصيغة **(1)**، بشكل أساسي في التطبيقات العملية من أجل التنبؤ باحتمالات البقاء المشتركة، وتقييم الاعتماد بين أوقات البقاء المزدوجة Paired Survival Times وتسهيل وتعزيز التحليل الهامشي Marginal Analysis. وقد اتضح أن التحليل في وجود "**البيانات المراقبة**"، يكون صعبًا حيث إن الأشكال المختلفة للمراقبة تشكل عائقًا عند تحليل البيانات. وللتغلب على هذه المشكلة، تم اقتراح العديد من مقدرات دالة البقاء ثنائية المتغيرات في الأدبيات الإحصائية .

٣. مقدر الاستقلال اللامعلمي ثنائي المتغيرات المعتمد على أسلوب "KM"

أشار كل من **Wang and Zafra (2009)** إلى أنه، في حالة استقلال وقت وقوع الحدثين T_x & T_y تصبح دالة البقاء ثنائية المتغيرات $S(t_x, t_y)$ كالآتي:

$$S(t_x, t_y) = P(T_x > t_x)P(T_y > t_y) = S(t_x)S(t_y) \qquad (2)$$

حيث إن $S(t_x)$ & $S(t_y)$ هي "**دوال البقاء الهامشية**" Marginal Survival Functions لـ T_x & T_y على التوالي. ويصبح مقدر **الاستقلال** اللامعلمي ثنائي المتغيرات المعتمد على أسلوب "**KM**" للدالة $S(t_x, t_y)$ كالتالي:

$$\widehat{S}_{KM}(t_x, t_y) = \prod_{i:\, \widehat{t}_{x(i)} \le t_x} \left(\frac{n - R_{1i}}{n - R_{1i} + 1}\right)^{\delta_i^x} \prod_{i:\, \widehat{t}_{y(i)} \le t_y} \left(\frac{n - R_{2i}}{n - R_{2i} + 1}\right)^{\delta_i^y} \qquad (A)$$

حيث إن:

- $\widehat{S}_{KM}(t_x, t_y)$: تشير إلى مقدر **الاستقلال** اللامعلمي ثنائي المتغيرات المعتمد على أسلوب "**KM**".
- $\widetilde{t}_{y_{(i)}}$ & $\widetilde{t}_{x_{(i)}}$: تشيران إلى وقت وقوع الحدث الأول والثاني المشاهد المرتب المناظر للمفردة i على التوالي.
- R_{2i} & R_{1i}: تشيران إلى رتبة $\widetilde{t}_{y(i)}$ & $\widetilde{t}_{x(i)}$ على التوالي.
- δ_i^y & δ_i^x: تشيران إلى مؤشر الحدث الأول والثاني على التوالي للمشاهدة i.
- n: تشير إلى عدد المشاهدات المعرضة للحدث أو المراقبة.

ويعاب على أسلوب "**KM**" لـ **Kaplan and Meier (1958)** المستخدم في تقدير "**دالتي البقاء الهامشيتين لوقت وقوع الحدثين**" بأنه ينتج عنه تقديرات "**متحيزة لأعلى**" لاحتمالات البقاء، ويكون "**غير مناسب**" في حالة "**عدم معلومية مؤشر الحدث δ**" لكل مفردة في الدراسة أو في ظل وجود "**الأخطار المتنافسة**"، كما أنه **يفشل** في التعامل مع أطراف Tails منحنى البقاء إذا كانت "**آخر**" مشاهدة "**مراقبة**"، وأيضًا **تتغير** احتمالات البقاء له عند كل "**وقت وقوع حدث فقط**" **(Shafiq *et al.*, 2007)**. ولكن، أشار **Wang and Zafra (2009)** إلى أنه، إذا كان حجم العينة "**كبيرًا**" بدرجة كافية، فإن مقدر **الاستقلال** اللامعلمي الأصلي $\widehat{S}_{KM}(t_x, t_y)$ "**يؤول**" إلى دالة البقاء الحقيقية، ويكون المقدر "**متسقًا**" في حالة "**المراقبة المستقلة**".

٤. المقدرات المقترحة على مقدر الاستقلال اللامعلمي ثنائي المتغيرات المعتمد على أسلوب "KM"

يقترح البحث "**خمسة**" مقدرات لامعلمية "**بالتعديل**" على مقدر **الاستقلال** اللامعلمي الأصلي ثنائي المتغيرات المعتمد على أسلوب "**KM**" $\widehat{S}_{KM}(t_x, t_y)$ **(المعادلة A)** والذي يعاني من العديد من العيوب، كما تم ذكره أعلاه، في محاولة للوصول إلى مقدر لامعلمي مناسب لتقدير دالة البقاء ثنائية المتغيرات للبيانات المراقبة من اليمين في ظل استقلال وقت وقوع الحدثين ووجود مراقبة مستقلة. ونعرض هذه المقدرات المقترحة في الآتي:

(أ) مقدر الاستقلال اللامعلمي ثنائي المتغيرات المعتمد على أسلوب "NA"

يتسم أسلوب "**NA**" بالمقارنة بأسلوب "**KM**" بسهولة حسابه، كما أنه يعتبر مقدرًا غير مباشرًا للدالة $S(t_x, t_y)$. وتعتمد فكرة هذا المقدر المقترح على تقدير "**دالتي البقاء الهامشيتين لوقت وقوع الحدثين**" باستخدام أسلوب "**NA**" لـ **Nelson–Aalen (1972, 1978)** بدلاً من أسلوب "**KM**" ليصبح مقدر **الاستقلال** اللامعلمي المقترح لدالة البقاء ثنائية المتغيرات كالتالي:

$$\widehat{S}_{NA}(t_x, t_y) = exp\left(-\sum_{i:\, \tilde{t}_{x(i)} \le t_x} \frac{d_{1i}}{n_{1i}}\right) exp\left(-\sum_{i:\, \tilde{t}_{y(i)} \le t_y} \frac{d_{2i}}{n_{2i}}\right) \quad \textbf{(B)}$$

حيث إن:

- $\widehat{S}_{NA}(t_x, t_y)$: تشير إلى مقدر **الاستقلال** اللامعلمي المقترح ثنائي المتغيرات المعتمد على أسلوب "**NA**".
- d_{1i}: تشير إلى عدد المشاهدات التي تعرضت للحدث الأول عند الوقت $\tilde{t}_{x(i)}$.
- d_{2i}: تشير إلى عدد المشاهدات التي تعرضت للحدث الثاني عند الوقت $\tilde{t}_{y(i)}$.
- n_{1i}: تشير إلى عدد المشاهدات المعرضة للحدث الأول أو للمراقبة عند الوقت $\tilde{t}_{x(i)}$.
- n_{2i}: تشير إلى عدد المشاهدات المعرضة للحدث الثاني أو للمراقبة عند الوقت $\tilde{t}_{y(i)}$.

وقد ترتب على استخدام أسلوب "**NA**" بدلاً من أسلوب "**KM**" لتقدير الدالة $S(t_x, t_y)$ الآتي: الحصول على تقديرات "**أعلى**" لاحتمال البقاء عند "**نهاية**" منحنى البقاء عن "**بدايته**" مقارنةً بأسلوب "**KM**" وذلك في حالة وجود "**مشاهدات مراقبة بكثافة**"، إن $\widehat{S}_{KM}(.) < \widehat{S}_{NA}(.)$، كما نجد أيضًا أن $\widehat{S}_{KM}(t_x, t_y) < \widehat{S}_{NA}(t_x, t_y)$.

(ب) مقدر الاستقلال اللامعلمي المقترح ثنائي المتغيرات المعتمد على أسلوب "S-VR"

يمتاز أسلوب "**S-VR**" بالمقارنة بأسلوب "**KM**" بتغير احتمالات البقاء عند كل من "**وقت وقوع الحدث والوقت المراقب من اليمين**"، كما أنه يتمكن من تقدير أطراف منحنى البقاء سواء كانت آخر مشاهدة مراقبة أو غير مراقبة **(Park and Na, 2005)**. وتعتمد فكرة هذا المقدر المقترح على تقدير "**دالتي البقاء الهامشيتين لوقت وقوع الحدثين**" باستخدام أسلوب "**S-VR**" لـ **Susarla and Van Ryzin (1980)** بدلاً من أسلوب "**KM**" ليصبح مقدر **الاستقلال** اللامعلمي المقترح لدالة البقاء ثنائية المتغيرات كالتالي:

$$\widehat{S}_{S-VR}(t_x, t_y) = \left(\begin{array}{l} \left\{ \frac{\sum_{i=1}^{n} I(\tilde{T}_{x(i)} > t_x)}{n} \prod_{i:\tilde{t}_{x(i)} \le t_x} \left\{ 1 + \frac{I(\delta_{(i)}^{x} = 0)}{n - i + 1} \right\} \right\} \\ \times \\ \left\{ \frac{\sum_{i=1}^{n} I(\tilde{T}_{y(i)} > t_y)}{n} \times \prod_{i:\tilde{t}_{y(i)} \le t_y} \left\{ 1 + \frac{I\left(\delta_{(i)}^{y} = 0\right)}{n - i + 1} \right\} \right\} \end{array} \right) \quad \textbf{(C)}$$

حيث تشير $\widehat{S}_{S-VR}(t_x, t_y)$ إلى مقدر **الاستقلال** اللامعلمي المقترح ثنائي المتغيرات المعتمد على أسلوب "**S-VR**".

وقد نتج عن استخدام أسلوب "**S-VR**" بدلاً من أسلوب "**KM**" لتقدير الدالة $S(t_x, t_y)$ الآتي: إن $\widehat{S}_{KM}(.) \geq \widehat{S}_{S-VR}(.)$، وأيضًا إن $\widehat{S}_{KM}(t_x, t_y) \geq \widehat{S}_{S-VR}(t_x, t_y)$.

(ج) مقدر الاستقلال اللامعلمي المقترح ثنائي المتغيرات المعتمد على أسلوب "WKM"

يتسم أسلوب "**WKM**" بالمقارنة بأسلوب "**KM**" بتغير احتمالات البقاء عند كل من "**وقت وقوع الحدث والوقت المراقب من اليمين**"، وأنه غير متحيز وموثوق في نتائجه، كما أنه

يتمكن من تقدير أطراف منحنى البقاء سواء كانت آخر مشاهدة مراقبة أو غير مراقبة (**Jan *et al.*, 2005**). وتعتمد فكرة هذا المقدر المقترح على تقدير "**دالتي البقاء الهامشيتين لوقت وقوع الحدثين**" باستخدام أسلوب "**WKM**" لـ **Jan *et al.* (2005)** بدلاً من أسلوب "**KM**" ليصبح مقدر **الاستقلال** اللامعلمي المقترح لدالة البقاء ثنائية المتغيرات كالتالي:

$$\widehat{S}_{WKM}(t_x, t_y) = \prod_{i:\, \tilde{t}_{x(i)} \le t_x} W_{1i}\left\{1 - \frac{d_{1i}}{n_{1i}}\right\} \prod_{i:\, \tilde{t}_{y(i)} \le t_y} W_{2i}\left\{1 - \frac{d_{2i}}{n_{2i}}\right\} \quad \text{(D)}$$

حيث إن:

- $\widehat{S}_{WKM}(t_x, t_y)$: تشير إلى مقدر **الاستقلال** اللامعلمي المقترح ثنائي المتغيرات المعتمد على أسلوب "**WKM**".
- W_{2i} **&** W_{1i}: تشيران إلى الوزن المرجح للمشاهدة i، ويأخذ هذا الوزن "**ثلاث**" قيم (القيمة "**0**" إذا كانت آخر مشاهدة "**مراقبة**"، القيمة "**1**" إذا كانت المشاهدة "**كاملة**"، وأخيرًا قيمة "**أقل من 1**" إذا كانت المشاهدة "**مراقبة**" عند وقت وقوع الحدث المشاهد المرتب الأول والثاني على التوالي)، حيث:

$$W_{1i} = \left\{\frac{n_{1i} - m_{1i}}{n_{1i}}\right\} \quad \& \quad W_{2i} = \left\{\frac{n_{2i} - m_{2i}}{n_{2i}}\right\}$$

- m_{2i} **&** m_{1i}: تشيران إلى عدد المشاهدات المراقبة عند وقت وقوع الحدث المشاهد المرتب الأول والثاني على التوالي.

وقد ترتب على استخدام أسلوب "**WKM**" بدلاً من أسلوب "**KM**" لتقدير الدالة $S(t_x, t_y)$ الآتي: إن $\widehat{S}_{KM}(.) \ge \widehat{S}_{WKM}(.)$، وأيضًا إن $\widehat{S}_{KM}(t_x, t_y) \ge \widehat{S}_{WKM}(t_x, t_y)$.

(د) مقدر الاستقلال اللامعلمي المقترح ثنائي المتغيرات المعتمد على أسلوب "MWKM"

يمتاز أسلوب "**MWKM**" بالمقارنة بأسلوب "**KM**" بتغير احتمالات البقاء عند كل من "**وقت وقوع الحدث والوقت المراقب من اليمين**"، وأنه غير متحيز وموثوق في نتائجه، كما أنه يتمكن من تقدير أطراف منحنى البقاء سواء كانت آخر مشاهدة مراقبة أو غير مراقبة، وأيضًا يعطي أسلوب "**MWKM**" وزنًا "**غير صفريًا**" لآخر مشاهدة إذا كانت "**مراقبة**"، بحيث يصبح "**احتمال البقاء**" لهذه المشاهدة "**صغيرًا جدًا**" (**Shafiq *et al.*, 2007**). وتعتمد فكرة هذا المقدر المقترح على تقدير "**دالتي البقاء الهامشيتين لوقت وقوع الحدثين**" باستخدام أسلوب "**MWKM**" لـ **Shafiq *et al.* (2007)** بدلاً من أسلوب "**KM**" ليصبح مقدر **الاستقلال** اللامعلمي المقترح لدالة البقاء ثنائية المتغيرات كالتالي:

$$\widehat{S}_{MWKM}(t_x, t_y) = \prod_{i:\, \tilde{t}_{x(i)} \le t_x} W_{1i}^{*}\left\{\frac{n_{1i} - d_{1i}}{n_{1i}}\right\} \quad \prod_{i:\, \tilde{t}_{y(i)} \le t_y} W_{2i}^{*}\left\{\frac{n_{2i} - d_{2i}}{n_{2i}}\right\} \qquad \textbf{(E)}$$

حيث إن:

- $\widehat{S}_{MWKM}(x, y)$: تشير إلى مقدر **الاستقلال** اللامعلمي المقترح ثنائي المتغيرات المعتمد على أسلوب "**MWKM**".
- W_{1i}^{*} **&** W_{2i}^{*}: تشيران إلى الوزن المرجح المعدل للمشاهدة i، ويأخذ هذا الوزن "**ثلاث**" قيم (القيمة "**0**" إذا كانت آخر مشاهدة "**مراقبة**"، القيمة "**1**" إذا كانت المشاهدة "**كاملة**"، وأخيرًا قيمة **أقل من** "**1**" إذا كانت المشاهدة "**مراقبة**" عند وقت وقوع الحدث المشاهد المرتب الأول والثاني على التوالي)، حيث:

$$W_{1i}^{*} = 1 - Sin\left\{\frac{m_{1i} \times P_{1i}^{*}}{n_{1i}}\right\} \quad \& \quad W_{2i}^{*} = 1 - Sin\left\{\frac{m_{2i} \times P_{2i}^{*}}{n_{2i}}\right\}$$

- P_{2i}^{*} **&** P_{1i}^{*}: تشيران إلى احتمال البقاء للمشاهدة i، حيث:

$$P_{1i}^{*}=\left\{\frac{n_{1i}-d_{1i}}{n_{1i}}\right\} \quad \& \quad P_{2i}^{*}=\left\{\frac{n_{2i}-d_{2i}}{n_{2i}}\right\}$$

وقد ترتب على استخدام أسلوب "**MWKM**" بدلاً من أسلوب "**KM**" لتقدير الدالة $S(t_x,t_y)$ الآتي: إن $\widehat{S}_{KM}(.)\geq\widehat{S}_{MWKM}(.)$، وأيضًا إن $\widehat{S}_{KM}(t_x,t_y)\geq\widehat{S}_{MWKM}(t_x,t_y)$.

(هـ) مقدر الاستقلال اللامعلمي المقترح ثنائي المتغيرات المعتمد على أسلوب "Huang"

يتسم أسلوب "**Huang**" بالمقارنة بأسلوب "**KM**" بتغير احتمالات البقاء عند كل من "**وقت وقوع الحدث والوقت المراقب من اليمين**"، وأنه أسلوب غير متحيز وموثوق في نتائجه، كما أنه يتمكن من تقدير أطراف منحنى البقاء سواء كانت آخر مشاهدة مراقبة أو غير مراقبة، وأيضًا يعتبر أكثر "**كفاءة**" مقارنةً بأسلوب "**KM**" وذلك عن طريق اختيار الأوزان التي يتم تعيينها للمشاهدات (**Huang, 2008**). وتعتمد فكرة هذا المقدر المقترح على تقدير "**دالتي البقاء الهامشية لوقت وقوع الحدث**" باستخدام أسلوب "**Huang**" لـ **Huang (2008)** بدلاً من أسلوب "**KM**" ليصبح مقدر **الاستقلال** اللامعلمي المقترح لدالة البقاء ثنائية المتغيرات كالتالي:

$$\widehat{S}_{H}(t_x,t_y)=\left\{\prod_{i:\,\hat{t}_{x(i)}\leq t_x}\left(1-\frac{d_{1i}}{n_{1i}}\right)(1-p_{n,i})\right\}\left\{\prod_{i:\,\hat{t}_{y(i)}\leq t_y}\left(1-\frac{d_{2i}}{n_{2i}}\right)(1-p_{n,i})\right\} \qquad (F)$$

حيث إن:

- $\widehat{S}_{H}(t_x,t_y)$: تشير إلى مقدر **الاستقلال** اللامعلمي المقترح ثنائي المتغيرات المعتمد على أسلوب "**Huang**".
- $p_{n,i}$: تشير إلى الوزن المقترح من **Huang**، ويعرف كالتالي:

$$p_{n,i}=\begin{cases}\dfrac{1}{\sqrt{n(n-1)}}>\dfrac{1}{n}, & if\ \ i=2,\ldots..,n-1,\\ \dfrac{1}{2}\left[1-\dfrac{n-2}{\sqrt{n(n-1)}}\right]<\dfrac{1}{n}, & if\ \ i=1,n.\end{cases}$$

وقد ترتب على استخدام أسلوب "**Huang**" بدلاً من أسلوب "**KM**" لتقدير الدالة $S(t_x,t_y)$ الآتي: إن $\widehat{S}_{KM}(.)>\widehat{S}_{H}(.)$، وأيضًا إن $\widehat{S}_{KM}(t_x,t_y)>\widehat{S}_{H}(t_x,t_y)$.

٥. دراسة المحاكاة Simulation Study

يعتمد البحث على أخذ عددٍ من العوامل في الاعتبار في دراسة المحاكاة، وهذه العوامل هي: **حجم العينة** Sample Size حيث يعتمد البحث على حجمي للعينة وهما **(n = 30 ، 60)**، **نسبة المراقبة** Censored Ratio (CR) حيث يستخدم البحث نسبتي للمراقبة **(CR = 30% ، 40%)**، وأخيرًا **طرق التقدير** Estimation Methods حيث تركز دراسة المحاكاة على تقدير دالة البقاء ثنائية المتغيرات "**لامعلمياً**" للبيانات المعرضة "**لحدثين مستقلين**" والخاضعة "**لمراقبة من اليمين المستقلة**" باستخدام مقدر **الاستقلال** اللامعلمي الأصلي $\widehat{S}_{KM}(t_x, t_y)$ **(معادلة A)**، وباستخدام المقدرات المقترحة بالتعديل عليه **(**مقدر **الاستقلال** اللامعلمي المقترح $\widehat{S}_{NA}(t_x, t_y)$ **(معادلة B)**، مقدر **الاستقلال** اللامعلمي المقترح $\widehat{S}_{S-VR}(t_x, t_y)$ **(معادلة C)**، مقدر **الاستقلال** اللامعلمي المقترح $\widehat{S}_{WKM}(t_x, t_y)$ **(معادلة D)**، مقدر **الاستقلال** اللامعلمي المقترح $\widehat{S}_{MWKM}(t_x, t_y)$ **(معادلة E)**، مقدر **الاستقلال** اللامعلمي المقترح $\widehat{S}_{H}(t_x, t_y)$ **(معادلة F)**.

ويتم المفاضلة بين المقدرات اللامعلمية محل البحث وفق:

أ. **مقدار التحيز التجريبي Empirical Bias Amount**: حيث نلاحظ أنه، إذا "**قل**" مقدار التحيز التجريبي لمقدر **الاستقلال** اللامعلمي الأصلي $\widehat{S}_{KM}(t_x, t_y)$ مقارنةً بمقدار التحيز التجريبي للمقدر المقترح في البحث كلما كان مقدر **الاستقلال** اللامعلمي الأصلي $\widehat{S}_{KM}(t_x, t_y)$ "**أفضل**" من المقدر المقترح **(Wang and Wells, 1997)**.

ب. **مقدار التباين التجريبي Empirical Variance Amount**: حيث نلاحظ أنه، إذا "**قل**" مقدار التباين التجريبي لمقدر **الاستقلال** اللامعلمي الأصلي $\widehat{S}_{KM}(t_x, t_y)$ مقارنةً بمقدار التباين التجريبي للمقدر المقترح في البحث كلما كان مقدر **الاستقلال** اللامعلمي الأصلي $\widehat{S}_{KM}(t_x, t_y)$ "**أفضل**" من المقدر المقترح في البحث **(Wang and Wells, 1997)**.

ج. **الكفاءة النسبية التجريبية Empirical Relative Efficiency (ERE)**: حيث تعرف "**الكفاءة النسبية التجريبية**" على أنها $\left(\frac{\text{متوسط مربع الخطأ للمقدر المقترح}}{\text{متوسط مربع الخطأ للمقدر الأصلي}}\right)$ **(Zhao, 2008)**. وكلما "**زادت**" قيمة الكفاءة النسبية التجريبية عن القيمة "**1**" كلما كان مقدر **الاستقلال** اللامعلمي الأصلي $\widehat{S}_{KM}(t_x, t_y)$ "**أكثر كفاءة**" مقارنةً بالمقدرات المقترحة الأخرى لمقدر **الاستقلال** اللامعلمي الأصلي $\widehat{S}_{KM}(t_x, t_y)$. وأيضًا كلما "**اقتربت**" قيمة الكفاءة النسبية التجريبية لإحدى المقدرات اللامعلمية المقترحة من القيمة "**1**" كلما كان هذا المقدر المقترح "**أكثر كفاءة**" مقارنةً بالمقدرات المقترحة الأخرى.

د. **تقديرات احتمالات البقاء ثنائية المتغيرات التجريبية Empirical Bivariate Survival Probabilities Estimates**: حيث نلاحظ أنه، إذا "**زادت**" تقديرات احتمالات البقاء أحادية المتغيرات التجريبية لمقدر "**KM**" اللامعلمي مقارنةً بتقديرات احتمالات البقاء أحادية المتغيرات التجريبية للمقدر اللامعلمي المقترح في البحث كلما كان المقدر المقترح "**أفضل**" من مقدر "**KM**" **(Shafiq *et al.*, 2006)**. ويقوم البحث بتعميم هذا المعيار في الحالة ثنائية المتغيرات كمحاولة للحصول على مقدر مناسب لدالة البقاء ثنائية المتغيرات.

ويتم تنفيذ "**1000**" معاودة، أي توليد "**1000**" مجموعة مختلفة من البيانات لتقدير دالة البقاء ثنائية المتغيرات عند بعض قيم أوقات وقوع الحدث المشاهدة المناظرة لمستويات البقاء **(25% ، 50% ، 70% ، 75% ، 99%)**، وذلك عند حجمين للعينة **(n = 30 ، 60)** ونسبتين للمراقبة **(CR = 30% ، 40%)**. كما يتم إجراء دراسة المحاكاة على عددٍ من الافتراضات وهي: تعرض كل مشاهدة **لحدثين مستقلين**، وأيضًا **استقلال** وقت وقوع الحدثين T_x & T_y، وكذلك **استقلال** وقتي المراقبة T_{C_1} & T_{C_2}، بالإضافة إلى **استقلال** متجه وقت وقوع الحدثين (T_x, T_y) عن متجه وقتي المراقبة (T_{C_1}, T_{C_2}). ويتبع البحث مجموعة من الخطوات لتوليد بيانات دراسة المحاكاة وهي:

- توليد المتغير العشوائي T_x (وقت وقوع الحدث **الأول**) من "**التوزيع الأسي**" Exponential Distribution بمعلمة[١] $(\lambda = 0.5)$، وأيضًا توليد المتغير العشوائي T_y (وقت وقوع الحدث **الثاني**) من "**التوزيع الأسي**" بمعلمة $(\lambda = 0.5)$.
- توليد المتغير العشوائي[٢] T_{C_1} (وقت المراقبة **الأول**) من "**التوزيع المنتظم**" Uniform Distribution ذي المعلمتين $(a = 0 \ ، \ b)$، بالإضافة إلى توليد المتغير العشوائي T_{C_2} (وقت المراقبة **الثاني**) من "**التوزيع المنتظم**" ذي المعلمتين $(a = 0 \ ، \ b)$. وتجدر الإشارة هنا إلى أن قيمة المعلمة b سوف تختلف تبعًا "**لنسبة المراقبة**".
- توليد المتغير العشوائي $\widetilde{T}_x$ (وقت وقوع الحدث **الأول المشاهد**) عن طريق مقارنة وقت وقوع الحدث **الأول** T_x بوقت المراقبة **الأول** T_{C_1} حيث $\widetilde{T}_x = (T_x \wedge T_{C_1})$، وأيضًا توليد المتغير العشوائي $\widetilde{T}_y$ (وقت وقوع الحدث **الثاني المشاهد**) عن طريق مقارنة وقت وقوع الحدث **الثاني** T_y بوقت المراقبة **الثاني** T_{C_2} حيث $\widetilde{T}_y = (T_y \wedge T_{C_2})$.
- تعريف المؤشرين $\delta^x = I(T_x \leq T_{C_1})$ & $\delta^y = I(T_y \leq T_{C_2})$.

---

[١] تشير λ إلى "**معلمة المعدل**" Rate Parameter وهي مقلوب "**معلمة المقياس**" Scale Parameter للتوزيع الأسي. واعتمد البحث على المعلمة $(\lambda = 0.5)$ بعد إجراء عدد من المحاولات التجريبية قبل إجراء دراسة المحاكاة الحالية وأيضًا ووفقًا لدراسة **Wang and Zafra (2009)**.

[٢] اعتمد البحث على التوزيع المنتظم ذي المعلمتين $(a = 0 \ ، \ b)$ للتبسيط، ويتم تحديد قيمة المعلمة b وفقًا لنسبة المراقبة باستخدام المعادلة $\left(b = \frac{\text{متوسط التوزيع الأسي}}{\text{نسبة المراقبة}}\right)$ **(Zhao, 2008)**.

٦. نتائج المحاكاة

أولاً: النتائج الخاصة

يتضح من الجدولين (١) & (٢) التاليين "**انخفاض**" مقدار التحيز التجريبي لمقدر **الاستقلال** اللامعلمي الأصلي $\widehat{S}_{KM}(t_x, t_y)$ والمقدرات المقترحة بالتعديل عليه "**بزيادة**" حجم العينة، كما يتبين أن مقدار التحيز التجريبي لمقدر **الاستقلال** اللامعلمي المقترح $\widehat{S}_{H}(t_x, t_y)$ "**أقل**" من مقدر **الاستقلال** اللامعلمي المقترح $\widehat{S}_{WKM}(t_x, t_y)$، يليه مقدر **الاستقلال** اللامعلمي المقترح $\widehat{S}_{MWKM}(t_x, t_y)$، يليه مقدر **الاستقلال** اللامعلمي المقترح $\widehat{S}_{S-VR}(t_x, t_y)$، ثم مقدر **الاستقلال** اللامعلمي الأصلي $\widehat{S}_{KM}(t_x, t_y)$، وأخيرًا مقدر **الاستقلال** اللامعلمي المقترح $\widehat{S}_{NA}(t_x, t_y)$ وذلك عند جميع مستويات البقاء ونسبتي المراقبة وحجمي العينة المبحوثتين.

جدول (١) مقدار التحيز التجريبي لمقدر الاستقلال والمقدرات المقترحة بالتعديل عليه من بيانات التوزيع الأسي($\lambda = 0.5$) عند نسبتي المراقبة 30% وحجمي العينة المبحوثتين باستخدام المعادلات F – A

n	(t_x, t_y)	True $S(t_x, t_y)$	B	A	C	E	D	F
30	(0.003, 0.003)	99%	-0.00036	-0.04156	-0.05798	-0.08252	-0.08769	-0.09268
	(0.132, 0.132)	75%	-0.00016	-0.01028	-0.03912	-0.04217	-0.04882	-0.05298
	(0.273, 0.273)	70%	-0.00064	-0.07124	-0.08468	-0.09216	-0.09698	-0.10361
	(0.599, 0.599)	50%	-0.00121	-0.06166	-0.07685	-0.09232	-0.09882	-0.10279
	(0.960, 0.960)	25%	-0.00101	-0.03113	-0.04235	-0.05211	-0.05789	-0.10236
60	(0.003, 0.003)	99%	-0.09124	-0.13206	-0.16924	-0.18368	-0.18797	-0.194
	(0.132, 0.132)	75%	-0.00026	-0.02046	-0.04669	-0.05262	-0.05968	-0.10302
	(0.273, 0.273)	70%	-0.00156	-0.08235	-0.09975	-0.10346	-0.10979	-0.15458
	(0.599, 0.599)	50%	-0.00133	-0.07232	-0.09679	-0.10325	-0.10618	-0.10626
	(0.960, 0.960)	25%	-0.00116	-0.04127	-0.05722	-0.0632	-0.06579	-0.11468

المصدر: الجدول من إعداد الباحثة في ضوء نتائج دراسة المحاكاة.

جدول (٢) مقدار التحيز التجريبي لمقدر الاستقلال والمقدرات المقترحة بالتعديل عليه من بيانات التوزيع الأسي($\lambda = 0.5$) عند نسبتي المراقبة 40% وحجمي العينة المبحوثتين باستخدام المعادلات F – A

n	(t_x, t_y)	True $S(t_x, t_y)$	B	A	C	E	D	F
30	(0.003, 0.003)	99%	-0.00059	-0.04078	-0.07798	-0.09252	-0.09789	-0.11768
	(0.132, 0.132)	75%	-0.00017	-0.03058	-0.04912	-0.05217	-0.05648	-0.07976
	(0.273, 0.273)	70%	-0.00146	-0.07362	-0.09468	-0.10217	-0.10547	-0.13362
	(0.599, 0.599)	50%	-0.00156	-0.06154	-0.09685	-0.10232	-0.10679	-0.16279
	(0.960, 0.960)	25%	-0.00027	-0.03034	-0.04235	-0.07211	-0.07979	-0.10327
60	(0.003, 0.003)	99%	-0.00121	-0.05133	-0.16924	-0.18368	-0.18887	-0.20679
	(0.132, 0.132)	75%	-0.0012	-0.04132	-0.05669	-0.05992	-0.05999	-0.11702
	(0.273, 0.273)	70%	-0.00324	-0.09416	-0.10975	-0.20346	-0.20679	-0.29457
	(0.599, 0.599)	50%	-0.0016	-0.08162	-0.10679	-0.17325	-0.17789	-0.19755
	(0.960, 0.960)	25%	-0.00062	-0.04075	-0.06722	-0.0932	-0.09768	-0.13491

المصدر: الجدول من إعداد الباحثة في ضوء نتائج دراسة المحاكاة.

ويتبين من الجدولين **(٣)** & **(٤)** التاليين "**انخفاض**" مقدار التباين التجريبي لمقدر **الاستقلال** اللامعلمي الأصلي $\widehat{S}_{KM}(t_x, t_y)$ والمقدرات المقترحة بالتعديل عليه "**بزيادة**" حجم العينة، كما يتضح أن مقدار التباين التجريبي لمقدر **الاستقلال** اللامعلمي المقترح $\widehat{S}_{H}(t_x, t_y)$ "**أقل**" من مقدر **الاستقلال** اللامعلمي المقترح $\widehat{S}_{WKM}(t_x, t_y)$، يليه مقدر **الاستقلال** اللامعلمي المقترح $\widehat{S}_{MWKM}(t_x, t_y)$، يليه مقدر **الاستقلال** اللامعلمي المقترح $\widehat{S}_{S-VR}(t_x, t_y)$، ثم مقدر **الاستقلال** اللامعلمي الأصلي $\widehat{S}_{KM}(t_x, t_y)$، وأخيرًا مقدر **الاستقلال** اللامعلمي المقترح $\widehat{S}_{NA}(t_x, t_y)$ وذلك عند جميع مستويات البقاء ونسبتي المراقبة وحجمي العينة المبحوثتين.

جدول (٣) مقدار التباين التجريبي لمقدر الاستقلال والمقدرات المقترحة بالتعديل عليه من بيانات التوزيع الأسي($\lambda = 0.5$) عند نسبة المراقبة 30% وحجمي العينة المبحوثتين باستخدام المعادلات F – A

n	(t_X, t_Y)	True $S(t_X, t_Y)$	B	A	C	E	D	F
30	(0.003, 0.003)	99%	0.027825	0.026209	0.025094	0.024356	0.023441	0.022529
	(0.132, 0.132)	75%	0.075352	0.063989	0.055472	0.040368	0.039849	0.030124
	(0.273, 0.273)	70%	0.085477	0.072453	0.064481	0.056354	0.046371	0.043620
	(0.599, 0.599)	50%	0.089838	0.072590	0.064784	0.043542	0.033097	0.024494
	(0.960, 0.960)	25%	0.096016	0.088269	0.084740	0.077063	0.067063	0.053191
60	(0.003, 0.003)	99%	0.026517	0.025778	0.024512	0.023533	0.019586	0.015065
	(0.132, 0.132)	75%	0.072424	0.061199	0.054172	0.039712	0.035214	0.020935
	(0.273, 0.273)	70%	0.075743	0.064207	0.052598	0.044942	0.039814	0.035434
	(0.599, 0.599)	50%	0.078679	0.064696	0.052968	0.040740	0.024306	0.023123
	(0.960, 0.960)	25%	0.087387	0.075859	0.06468	0.054494	0.052419	0.043242

المصدر: الجدول من إعداد الباحثة في ضوء نتائج دراسة المحاكاة.

جدول (٤) مقدار التباين التجريبي لمقدر الاستقلال والمقدرات المقترحة بالتعديل عليه من بيانات التوزيع الأسي($\lambda = 0.5$) عند نسبة المراقبة 40% وحجمي العينة المبحوثتين باستخدام المعادلات F – A

n	(t_X, t_Y)	True $S(t_X, t_Y)$	B	A	C	E	D	F
30	(0.003, 0.003)	99%	0.026812	0.025718	0.024973	0.023816	0.023584	0.022027
	(0.132, 0.132)	75%	0.075538	0.064318	0.063649	0.057355	0.042594	0.035460
	(0.273, 0.273)	70%	0.084753	0.079826	0.074231	0.068389	0.056442	0.048217
	(0.599, 0.599)	50%	0.075851	0.066731	0.065291	0.054871	0.048736	0.044579
	(0.960, 0.960)	25%	0.030025	0.027524	0.029655	0.028253	0.028653	0.027354
60	(0.003, 0.003)	99%	0.025802	0.024283	0.023162	0.022194	0.021914	0.021946
	(0.132, 0.132)	75%	0.064353	0.062351	0.053502	0.043488	0.028346	0.027345
	(0.273, 0.273)	70%	0.075728	6.82E-02	6.16E-02	5.44E-02	4.27E-01	0.039736
	(0.599, 0.599)	50%	0.071538	0.065279	0.064081	0.051468	0.041648	0.037842
	(0.960, 0.960)	25%	0.029961	0.027094	0.026962	0.025916	0.023044	0.022144

المصدر: الجدول من إعداد الباحثة في ضوء نتائج دراسة المحاكاة.

بينما نستنتج من الجدولين **(٥) & (٦)** التاليين أن التوفيقات اللامعلمية المرجحة المقترحة المعتمدة على الأساليب "**Huang**" & "**WKM**" & "**MWKM**" "**أكثر كفاءة**" من التوفيق اللامعلمي الأصلي غير المرجح المعتمد على أسلوب "**KM**"، بينما التوفيقات اللامعلمية غير المرجحة المقترحة المعتمدة على أسلوبي "**S–VR**" & "**NA**" "**أقل كفاءة**" من التوفيق اللامعلمي الأصلي غير المرجح المعتمد على أسلوب "**KM**"، وأيضًا نجد أن التوفيق

اللامعلمي المقترح المعتمد على أسلوب "**MWKM**" "**أكثر كفاءة**" من المعتمد على أسلوب "**WKM**"، يليه المعتمد على أسلوب "**Huang**" وذلك مقارنةً بالتوفيق اللامعلمي الأصلي المعتمد على أسلوب "**KM**" وذلك عند نسبتي المراقبة وحجمي العينة المبحوثتين وجميع مستويات البقاء.

كما نلاحظ أيضًا أنه "**بزيادة**" حجم العينة عند نسبتي المراقبة المبحوثتين، "**تتحسن**" الكفاءة النسبية التجريبية لجميع المقدرات اللامعلمية المقترحة لمقدر **الاستقلال** اللامعلمي الأصلي $\widehat{S}_{KM}(t_x, t_y)$ وذلك عند مستويي البقاء **25%** & **99%**. بينما "**تسوء**" الكفاءة النسبية التجريبية لهذه المقدرات عند مستويات البقاء **50%** & **70%** & **75%**. ويرجع "**السبب**" في ذلك إلى أنه "**بزيادة**" حجم العينة "**تقل**" قيمة متوسط مربعات الخطأ "**MSE**" للمقدرات اللامعلمية المقترحة بدرجة "**أعلى (أقل)**" مما "**تقل**" بها قيمة متوسط مربعات الخطأ "**MSE**" بالنسبة لمقدر **الاستقلال** اللامعلمي الأصلي $\widehat{S}_{KM}(t_x, t_y)$ وبالتالي "**تتحسن (تسوء)**" كفاءة هذه المقدرات.

جدول (٥) الكفاءة النسبية التجريبية لمقدر الاستقلال والمقدرات المقترحة بالتعديل عليه من بيانات التوزيع الأسي($\lambda = 0.5$) عند نسبة المراقبة 30% وحجمي العينة المبحوثتين باستخدام المعادلات F – A

n	(t_x, t_y)	True $S(t_x, t_y)$	B	C	E	D	F
30	(0.003, 0.003)	99%	2.955636	2.361808	0.964121	0.937289	0.876204
	(0.132, 0.132)	75%	2.988746	2.967786	0.952615	0.945371	0.857544
	(0.273, 0.273)	70%	2.76373	2.046554	0.891744	0.853168	0.841809
	(0.599, 0.599)	50%	2.12186	1.94385	0.954075	0.952637	0.91374
	(0.960, 0.960)	25%	2.847924	1.98647	0.987592	0.957345	0.817299
60	(0.003, 0.003)	99%	2.854685	1.364618	0.952533	0.886935	0.862409
	(0.132, 0.132)	75%	3.224566	2.971761	0.988632	0.964465	0.864386
	(0.273, 0.273)	70%	2.966143	2.948842	0.904512	0.869278	0.865865
	(0.599, 0.599)	50%	2.648394	1.965836	0.981884	0.974994	0.933422
	(0.960, 0.960)	25%	2.626409	1.887502	0.940833	0.935373	0.803195

المصدر: الجدول من إعداد الباحثة في ضوء نتائج دراسة المحاكاة.

جدول (٦) الكفاءة النسبية التجريبية لمقدر الاستقلال والمقدرات المقترحة بالتعديل عليه من بيانات التوزيع الأسي($\lambda = 0.5$) عند نسبة المراقبة 40% وحجمي العينة المبحوثتين باستخدام المعادلات A – F

n	(t_x, t_y)	True $S(t_x, t_y)$	B	C	E	D	F
30	(0.003, 0.003)	99%	3.042334	2.97803	0.98376	0.917021	0.88505
	(0.132, 0.132)	75%	2.91354	1.97924	0.892427	0.876405	0.86588
	(0.273, 0.273)	70%	2.918478	2.07974	0.88453	0.864234	0.857675
	(0.599, 0.599)	50%	2.73513	1.942885	0.954306	0.843604	0.826045
	(0.960, 0.960)	25%	2.036148	1.014173	0.926877	0.804103	0.784322
60	(0.003, 0.003)	99%	2.850948	1.989414	0.943867	0.902834	0.874095
	(0.132, 0.132)	75%	3.146602	2.035495	0.953075	0.928616	0.892718
	(0.273, 0.273)	70%	3.531178	2.377594	0.974372	0.880241	0.860582
	(0.599, 0.599)	50%	2.936164	2.203287	0.961828	0.928451	0.890648
	(0.960, 0.960)	25%	2.014761	1.345157	0.901317	0.895605	0.871139

المصدر: الجدول من إعداد الباحثة في ضوء نتائج دراسة المحاكاة.

ويتضح من الجدولين (٧) & (٨) التاليين "**انخفاض**" تقديرات احتمالات البقاء ثنائية المتغيرات التجريبية لمقدر **الاستقلال** اللامعلمي الأصلي $\widehat{S}_{KM}(t_x, t_y)$ والمقدرات المقترحة بالتعديل عليه "**بزيادة**" حجم العينة، كما يتبين أن تقديرات احتمالات البقاء ثنائية المتغيرات التجريبية لمقدر **الاستقلال** اللامعلمي المقترح $\widehat{S}_{H}(t_x, t_y)$ "**أقل**" من مقدر **الاستقلال** اللامعلمي المقترح $\widehat{S}_{WKM}(t_x, t_y)$، يليه مقدر **الاستقلال** اللامعلمي المقترح $\widehat{S}_{MWKM}(t_x, t_y)$، يليه مقدر **الاستقلال** اللامعلمي المقترح $\widehat{S}_{S-VR}(t_x, t_y)$، ثم مقدر **الاستقلال** اللامعلمي الأصلي $\widehat{S}_{KM}(t_x, t_y)$، وأخيرًا مقدر **الاستقلال** اللامعلمي المقترح $\widehat{S}_{NA}(t_x, t_y)$ وذلك عند جميع مستويات البقاء ونسبتي المراقبة وحجمي العينة المبحوثتين

جدول (٧) تقديرات احتمالات البقاء ثنائية المتغيرات التجريبية لمقدر الاستقلال والمقدرات المقترحة بالتعديل عليه من بيانات التوزيع الأسي($\lambda = 0.5$) عند نسبة المراقبة 30% وحجمي العينة المبحوثتين باستخدام المعادلات F – A

n	(t_x, t_y)	True $S(t_x, t_y)$	B	A	C	E	D	F
30	(0.003, 0.003)	99%	0.989636	0.948443	0.932021	0.907476	0.90231	0.897321
	(0.132, 0.132)	75%	0.749835	0.739716	0.710881	0.707835	0.70118	0.697023
	(0.273, 0.273)	70%	0.699357	0.628761	0.61532	0.607835	0.60302	0.596387
	(0.599, 0.599)	50%	0.498794	0.43834	0.423152	0.407681	0.40118	0.397207
	(0.960, 0.960)	25%	0.248991	0.21887	0.207651	0.197893	0.19211	0.147637
60	(0.003, 0.003)	99%	0.898756	0.857944	0.820761	0.806324	0.80203	0.796001
	(0.132, 0.132)	75%	0.749736	0.729543	0.703312	0.697384	0.69032	0.646981
	(0.273, 0.273)	70%	0.698435	0.617652	0.600251	0.596538	0.59021	0.545421
	(0.599, 0.599)	50%	0.498673	0.427683	0.40321	0.396746	0.39312	0.395739
	(0.960, 0.960)	25%	0.248839	0.208731	0.192778	0.186798	0.18421	0.135321

المصدر: الجدول من إعداد الباحثة في ضوء نتائج دراسة المحاكاة.

جدول (٨) تقديرات احتمالات البقاء ثنائية المتغيرات التجريبية لمقدر الاستقلال والمقدرات المقترحة بالتعديل عليه من بيانات التوزيع الأسي($\lambda = 0.5$) عند نسبة المراقبة 40% وحجمي العينة المبحوثتين باستخدام المعادلات F – A

n	(t_x, t_y)	True $S(t_x, t_y)$	B	A	C	E	D	F
30	(0.003, 0.003)	99%	0.989413	0.949221	0.912021	0.897476	0.89211	0.872321
	(0.132, 0.132)	75%	0.749834	0.719423	0.700881	0.697835	0.69352	0.670243
	(0.273, 0.273)	70%	0.698543	0.626378	0.60532	0.597835	0.59453	0.566382
	(0.599, 0.599)	50%	0.498438	0.438461	0.403152	0.397681	0.39321	0.337206
	(0.960, 0.960)	25%	0.249734	0.219656	0.207651	0.177893	0.17021	0.146732
60	(0.003, 0.003)	99%	0.988789	0.938673	0.820761	0.806324	0.80113	0.78321
	(0.132, 0.132)	75%	0.748798	0.708684	0.693312	0.690084	0.69001	0.632981
	(0.273, 0.273)	70%	0.696764	0.605837	0.590251	0.496538	0.49321	0.405432
	(0.599, 0.599)	50%	0.498401	0.418377	0.39321	0.326746	0.32211	0.302451
	(0.960, 0.960)	25%	0.249379	0.209248	0.182778	0.156798	0.15231	0.115087

المصدر: الجدول من إعداد الباحثة في ضوء نتائج دراسة المحاكاة.

وبناءً على ما سبق، نستنتج "**أفضلية**" مقدر **الاستقلال** اللامعلمي المقترح $\widehat{S}_H(t_x, t_y)$ على جميع المقدرات اللامعلمية الأخرى المبحوثة (المستطيلات الداكنة في مجموعة الجداول من (١) إلى (٨)). ويرجع "**السبب**" في ذلك إلى أنه، عند كل نسبة من نسبتي المراقبة وحجمي العينة المبحوثتان وجميع مستويات البقاء، نلاحظ أن مقدار كل من التحيز التجريبي والتباين التجريبي والكفاءة النسبية التجريبية وتقديرات احتمالات البقاء ثنائية المتغيرات التجريبية لمقدر **الاستقلال** اللامعلمي المقترح $\widehat{S}_H(t_x, t_y)$ "**أقل**" مقارنةً بالمقدرات اللامعلمية الأخرى المبحوثة.

ثانيًا: النتائج العامة

تشير النتائج العامة لدراسة المحاكاة إلى "**أفضلية**" التوفيق اللامعلمي المرجح المقترح المعتمد على أسلوب "**Huang**" على جميع التوفيقات اللامعلمية المقترحة الأخرى لمقدر **الاستقلال** اللامعلمي الأصلي $\widehat{S}_{KM}(t_x, t_y)$ لأنه "**الأقل تحيزًا والأقل تباينًا والأقل تقديرًا لاحتمالات البقاء ثنائية المتغيرات التجريبية والأكثر كفاءة**".

كما خلصت نتائج المحاكاة إلى أنه "**بزيادة**" حجم العينة، "**تزداد**" قيمة الكفاءة النسبية التجريبية للتوفيقات اللامعلمية غير المرجحة المقترحة المعتمدة على أسلوبي "**S-VR**" & "**NA**" وبالتالي "**تسوء**" كفاءة هذه التوفيقات. ويرجع "**السبب**" في ذلك إلى أنه "**بزيادة**" حجم العينة "**تقل**" قيمة متوسط مربعات الخطأ "**MSE**" لهذه المقدرات اللامعلمية بدرجة "**أقل**" مما "**تقل**" بها قيمة متوسط مربعات الخطأ "**MSE**" بالنسبة لمقدر **الاستقلال** اللامعلمي الأصلي $\widehat{S}_{KM}(t_x, t_y)$.

وكذلك أوضحت نتائج المحاكاة إلى أنه "**بزيادة**" حجم العينة، "**لا تتغير**" قيمة الكفاءة النسبية التجريبية للتوفيقات اللامعلمية المقترحة المعتمدة على "**Huang**" & "**WKM**" & "**MWKM**". ويرجع "**السبب**" في ذلك إلى أنه "**بزيادة**" حجم العينة "**تقل**" قيمة متوسط مربعات الخطأ "**MSE**" لهذه المقدرات اللامعلمية تقريبًا "**بنفس**" الدرجة التي "**تقل**" بها قيمة متوسط مربعات الخطأ "**MSE**" بالنسبة لمقدر **الاستقلال** اللامعلمي الأصلي $\widehat{S}_{KM}(t_x, t_y)$.

٧- الدراسة التطبيقية

بتطبيق مقدر **الاستقلال** اللامعلمي الأصلي $\widehat{S}_{KM}(t_x, t_y)$ والمقدرات المقترحة بالتعديل عليه على عينة تمت متابعتها خلال الفترة **(2004/7-2023/8)** حجمها "**15**" مشاهدة من الهيئة المعاونة **(فئة معيد)** بكلية تجارة جامعة دمنهور تعيين خلال الفترة **(2004/7-2007/1)** وذلك لتتبع تجاوزهم سن "**التاسعة والثلاثون وثمانية أشهر**" كحدث "**أول**" والحصول على "**درجة الدكتوراه قبل بلوغ عام 2020**" كحدث "**ثاني**"، حصلنا على "**15**" زوجًا لوقت وقوع الحدثين المستقلين (T_x, T_y) وكانت نسبة المراقبة "**المستقلة**" "**33.33%**" **(جدولي "٩" & "١٠")**.

جدول (٩) بيانات عينة من الهيئة المعاونة (فئة معيد) بكلية تجارة جامعة دمنهور تعيين خلال الفترة (2005–2007) حجمها $(15 = n)$

z_2	d_2	T_y	z_1	d_1	T_x	تاريخ التعيين كمعيد	تاريخ الحصول على الدكتوراه	تاريخ الميلاد	المشاهدة
0.0690	0	16.58+	0.0515	0	16.58+	1/2007	-	7/1984	1
0.0690	0	16.58+	0.0690	0	16.58+	1/2007	-	9/1984	2
0.0690	0	16.58+	0.0690	0	16.58+	1/2007	-	4/1984	3
0.0690	0	16.58+	0.0690	0	16.58+	1/2007	-	9/1984	4
0.0690	1	13.17	0.0690	1	17	10/2005	12/2018	1/1983	5
0.0690	1	16.75	0.0690	1	17	7/2004	4/2021	10/1981	6
0.0515	1	10.75	0.0690	1	17.17	10/2005	7/2016	3/1983	7
0.0690	1	11.5	0.0690	1	17.33	10/2005	4/2017	5/1983	8
0.0690	1	13.33	0.0690	1	17.42	7/2004	11/2017	3/1982	9
0.0690	1	13.67	0.0690	1	17.42	10/2005	6/2019	6/1983	10
0.0690	1	13.92	0.0690	1	17.42	4/2006	3/2020	12/1983	11
0.0515	0	17.83+	0.0690	0	17.83+	10/2005	-	1/1984	12
0.0690	1	16.03	0.0690	1	17.92	10/2005	11/2021	12/1983	13
0.0690	1	16.5	0.0690	1	18.17	7/2004	1/2021	12/1982	14
0.0690	1	13.17	0.0515	1	18.25	7/2004	9/2017	1/1983	15

المصدر: من سجلات شئون هيئة أعضاء التدريس ومعاونيهم بكلية تجارة جامعة دمنهور.

- T_x: الوقت المستغرق حتى تخطي المشاهدة سن ٣٩ سنة & ٨ شهور((عمر ٣٩ و٩شهور)– (العمر ما بين تاريخ الميلاد وتاريخ التعيين)).
- T_y: الوقت المستغرق حتى الحصول على الدكتوراه.
- z_1 & z_2 : وزني المقدر **Huang** .
- d_1 : عدد المشاهدات التي تخطيت سن ٣٩ سنة & ٨ شهور وتأخذ القيمة **1**.
- d_2 : عدد المشاهدات التي حصلت على درجة الدكتوراه وتأخذ القيمة **1**.

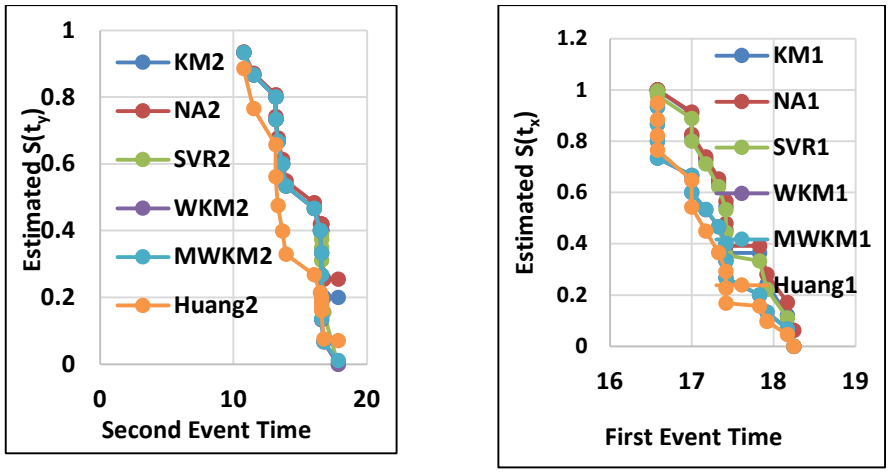
جدول (١٠) تقديرات احتمالات البقاء ثنائية المتغيرات لمقدر الاستقلال والمقدرات المقترحة بالتعديل عليه لعينة من أعضاء هيئة التدريس (فئة مدرس) بكلية تجارة جامعة دمنهور تعيين خلال الفترة (2005–2007) حجمها $(15 = n)$ باستخدام المعادلات F – A

i	(t_x, t_y)	B	A	C	E	D	F
1	(16.58+,16.58+)	0.419694	0.4	0.38716	0.311415	0.311111	0.189096
2	(16.58+,16.58+)	0.419694	0.4	0.369778	0.231737	0.231111	0.163901
3	(16.58+,16.58+)	0.419694	0.4	0.344615	0.161002	0.16	0.142063
4	(16.58+,16.58+)	0.419694	0.4	0.304198	0.099306	0.097778	0.123135
5	(17,13.17)	0.73644	0.727273	0.711111	0.533496	0.533333	0.426059
6	(17,16.75)	0.210317	0.163636	0.124444	0.040625	0.04	0.040649
7	(17.17,10.75)	0.691642	0.678788	0.663704	0.49793	0.497778	0.397655
8	(17.33,11.5)	0.568295	0.551515	0.539259	0.404568	0.404444	0.280046
9	(17.42,13.33)	0.383223	0.363636	0.355556	0.266748	0.266667	0.13872
10	(17.42,13.67)	0.293521	0.272727	0.266667	0.200061	0.2	0.090178
11	(17.42,13.92)	0.215043	0.193939	0.18963	0.142266	0.142222	0.055582
12	(17.83+,17.83+)	0.099781	0.072727	0	0.002154	0	0.011158
13	(17.92,16.03)	0.13598	0.113131	0.103704	0.062457	0.062222	0.026164
14	(18.17,16.5)	0.071497	0.048485	0.044444	0.026767	0.026667	0.009719
15	(18.25,13.17)	0.046503	0	0	0	0	0

المصدر: نتائج الدراسة التطبيقية الأولى باستخدام برنامج "**R**".

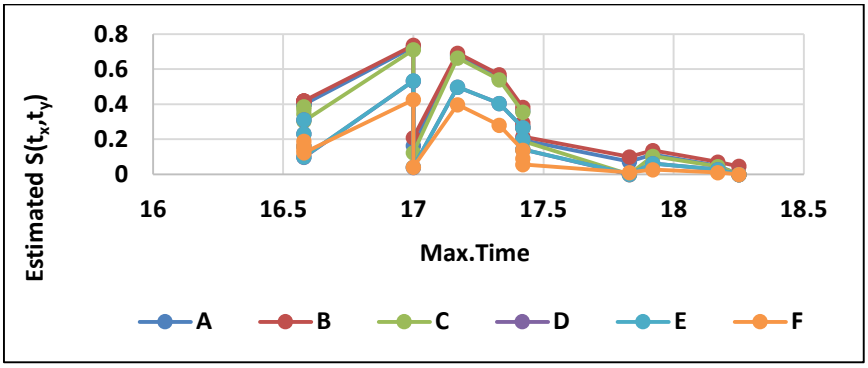
ويتضح من جدول (١٠) السابق أن متوسط تقديرات احتمالات البقاء ثنائية المتغيرات لمقدر **الاستقلال** اللامعلمي المقترح $\widehat{S}_H(t_x,t_y)$ "**أقل**" من مقدر **الاستقلال** اللامعلمي المقترح $\widehat{S}_{WKM}(t_x,t_y)$، يليه مقدر **الاستقلال** اللامعلمي المقترح $\widehat{S}_{MWKM}(t_x,t_y)$، يليه مقدر **الاستقلال** اللامعلمي المقترح $\widehat{S}_{S-VR}(t_x,t_y)$، ثم مقدر **الاستقلال** اللامعلمي الأصلي $\widehat{S}_{KM}(t_x,t_y)$، وأخيرًا مقدر **الاستقلال** اللامعلمي المقترح $\widehat{S}_{NA}(t_x,t_y)$.

ويتبين من شكل (١) التالي أن تقديرات احتمالات البقاء الهامشية للحدثين الأول والثاني لأسلوب "**Huang**" "**أقل**" من جميع الأساليب الأخرى وهي: "**KM**" ، "**NA**" ، "**SVR**" ، "**WKM**" ، "**MWKM**"، كما يتضح أن أسلوب "**Huang**" "**أقل تحيزًا**" مقارنةً بالأساليب الأخرى وهي: "**KM**" ، "**NA**" ، "**SVR**" ، "**WKM**" ، "**MWKM**".



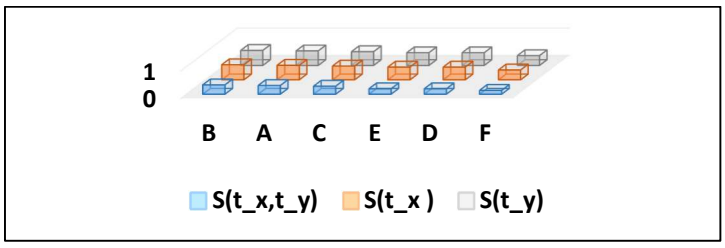
شكل (١) تقديرات احتمالات البقاء الهامشية للحدثين الأول والثاني عند حجم العينة $(n = 15)$

ويوضح شكل (٢) التالي أن تقديرات احتمالات البقاء ثنائية المتغيرات لمقدر **الاستقلال** اللامعلمي المقترح $\widehat{S}_H(t_x,t_y)$ "**أقل**" من جميع المقدرات اللامعلمية الأخرى لمقدر **الاستقلال** اللامعلمي الأصلي $\widehat{S}_{KM}(t_x,t_y)$، كما تبين أن مقدر **الاستقلال** اللامعلمي المقترح $\widehat{S}_H(t_x,t_y)$ "**أقل تحيزًا**" مقارنةً بالمقدرات اللامعلمية الأخرى لمقدر **الاستقلال** اللامعلمي الأصلي $\widehat{S}_{KM}(t_x,t_y)$.



شكل (٢) تقديرات احتمالات البقاء ثنائية المتغيرات لمقدر الاستقلال والمقدرات المقترحة بالتعديل عليه عند حجم العينة (n = 15)

بينما يعكس شكل **(٣)** التالي أن متوسط تقديرات احتمالات البقاء ثنائية المتغيرات لمقدر **الاستقلال** اللامعلمي المقترح $\widehat{S}_H(t_x,t_y)$ "**أقل**" من مقدر **الاستقلال** اللامعلمي المقترح $\widehat{S}_{WKM}(t_x,t_y)$، يليه مقدر **الاستقلال** اللامعلمي المقترح $\widehat{S}_{MWKM}(t_x,t_y)$، يليه مقدر **الاستقلال** اللامعلمي المقترح $\widehat{S}_{S-VR}(t_x,t_y)$، ثم مقدر **الاستقلال** اللامعلمي الأصلي $\widehat{S}_{KM}(t_x,t_y)$، وأخيرًا مقدر **الاستقلال** اللامعلمي المقترح $\widehat{S}_{NA}(t_x,t_y)$. كما يشير شكل **(٣)** إلى أن متوسط تقديرات احتمالات البقاء "**ثنائية**" المتغيرات $\widehat{S}(t_x,t_y)$ لمقدر **الاستقلال** اللامعلمي الأصلي $\widehat{S}_{KM}(t_x,t_y)$ والمقدرات المقترحة بالتعديل عليه "**أقل**" من متوسط تقديرات احتمالات البقاء "**أحادية**" المتغيرات للحدث "**الأول**" $\widehat{S}(t_x)$، يليه متوسط تقديرات احتمالات البقاء "**أحادية**" المتغيرات للحدث "**الثاني**" $\widehat{S}(t_y)$.



شكل (٣) متوسط تقديرات احتمالات البقاء أحادية وثنائية المتغيرات لمقدر الاستقلال والمقدرات المقترحة بالتعديل عليه عند حجم العينة (n = 15)

وأشار البحث[٣] إلى "**إمكانية**" تطبيق مقدر **الاستقلال** اللامعلمي الأصلي $\widehat{S}_{KM}(t_x, t_y)$ والمقدرات اللامعلمية المقترحة بالتعديل عليه لجميع أحجام العينات **(الصغيرة ، المتوسطة ، الكبيرة)** ولجميع نسب المراقبة **(الخفيفة ، المتوسطة ، الكثيفة)**.

٨- التوصيات Recommendations

في ضوء نتائج المحاكاة والتي تمثلت أبرزها في "**أفضلية**" التوفيق اللامعلمي لمقدر **الاستقلال** اللامعلمي المقترح $\widehat{S}_{H}(t_x, t_y)$ على جميع التوفيقات اللامعلمية الأخرى المبحوثة لأنه "**الأقل تحيزًا والأقل تباينًا والأقل تقديرًا لاحتمالات البقاء ثنائية المتغيرات التجريبية والأكثر كفاءة**"، بالإضافة إلى "**تفوق**" مقدر **الاستقلال** اللامعلمي المقترح $\widehat{S}_{H}(t_x, t_y)$ على جميع المقدرات اللامعلمية المقترحة الأخرى بزيادة حجم العينة، يوصي البحث بالآتي:

أ. استخدام مقدر **الاستقلال** اللامعلمي المقترح $\widehat{S}_{WKM}(t_x, t_y)$ عن مقدر **الاستقلال** اللامعلمي الأصلي $\widehat{S}_{KM}(t_x, t_y)$ حيث إنه يمتاز بتغير احتمالات البقاء عند كل من "**وقت وقوع الحدث والوقت المراقب من اليمين**"، وأنه غير متحيز وموثوق في نتائجه، كما أنه يتمكن من تقدير أطراف منحنى البقاء سواء كانت آخر مشاهدة مراقبة أو غير مراقبة.

ب. استخدام مقدر **الاستقلال** اللامعلمي المقترح $\widehat{S}_{MWKM}(t_x, t_y)$ عن مقدر **الاستقلال** اللامعلمي المقترح $\widehat{S}_{WKM}(t_x, t_y)$ حيث إنه يعطي وزنًا "**غير صفريًا**" لآخر مشاهدة إذا كانت "**مراقبة**"، بحيث يصبح "**احتمال البقاء**" لهذه المشاهدة "**صغيرًا جدًا**".

---

[٣] قام البحث قبل إجراء دراسة المحاكاة بعدد من المحاولات التجريبية، حيث استخدم أحجام العينات $(n = 10, 20, 30, 60, 100, 200)$، كما استخدم نسب المراقبة $(CR = 5\%, 10\%, 20\%, 30\%, 40\%)$.

ج. استخدام مقدر **الاستقلال** اللامعلمي المقترح $\widehat{S}_H(t_x, t_y)$ عن مقدر **الاستقلال** اللامعلمي المقترح $\widehat{S}_{MWKM}(t_x, t_y)$ لأنه "**الأقل تحيزًا والأقل تباينًا والأقل تقديرًا لاحتمالات البقاء ثنائية المتغيرات التجريبية والأكثر كفاءة**".

د. تفضيل استخدام طرق التوفيق اللامعلمية المرجحة المعتمدة على أساليب "**Huang**" & "**WKM**" & "**MWKM**" عن طرق التوفيق اللامعلمية غير المرجحة المعتمدة على أساليب "**S−VR**" & "**NA**" & "**KM**" في حالة استقلال وقت وقوع الحدثين لنفس المشاهدة لأنها تؤدي إلى نتائج كفء.

أما بالنسبة للبحوث المستقبلية في هذا المجال، فيقترح البحث الآتي:

١. مقارنة بعض المقدرات المعلمية واللامعلمية لدالة البقاء ثنائية المتغيرات في حالة وجود "**أخطار متنافسة وبيانات مراقبة من جهة اليمين**" من خلال إجراء دراسة محاكاة عند مستويات مختلفة من أحجام العينة ونسب المراقبة ووجود سببين لوقوع الحدث.

٢. مقارنة بعض المقدرات المعلمية واللامعلمية لدالة البقاء ثنائية المتغيرات في حالة وجود "**بيانات مراقبة من جهة اليمين ومعرضة لمراقبة أحادية المتغيرات**" من خلال إجراء دراسة محاكاة عند مستويات مختلفة من أحجام العينة ونسب المراقبة.

٣. مقارنة بعض المقدرات المعلمية واللامعلمية لدالة البقاء ثنائية المتغيرات في حالة وجود "**بيانات مراقبة من جهة اليمين ومعرضة لمراقبة مستقلة**" من خلال إجراء دراسة محاكاة عند مستويات مختلفة من أحجام العينة ونسب المراقبة.

٤. تطبيق مقدر **الاستقلال** اللامعلمي الأصلي $\widehat{S}_{KM}(t_x, t_y)$ والمقدرات المقترحة بالتعديل عليه في حالة المتغيرات المتعددة.

٩- المراجع References

1. Aalen, O. (1978), "Nonparametric Estimation of Partial Transition Probabilities in Multiple Decrement Models," ***Annals of Statistics***, Vol. **6**, PP. 534-545.
2. Dabrowska, D. M. (1988), "Kaplan–Meier Estimate on the Plane," ***Annals of Statistics***, Vol. **16**, PP. 1475–1489.
3. Dai, H., Restaino, M., & Wang, H. (2016), "A Class of Nonparametric Bivariate Survival Function Estimators for Randomly Censored and Truncated Data," ***Journal of Nonparametric Statistics***, Vol. **28**, No. 4, PP. 736-751.
4. Emura, T. and Michimae, H. (2022), "Left-Truncated and Right-Censored Field Failure Data: Review of Parametric Analysis for Reliability," ***Qual Reliab Engng Int***, Vol. **38**, PP. 3919-3934.
5. _____, Matsui, S. and Rondeau, V. (2019). Survival Analysis with Correlated Endpoints: Joint Frailty-Copula Models. ***Springer Singapore***.
6. Hougaard, P. (2000), Analysis of Multivariate Survival Data. ***Springer Science and Business Media***.
7. Huang, M. L. (2008), "A Weighted Estimation Method for Survival Function," ***Applied Mathematical Sciences***, Vol. **2**, No. 16, PP. 753-762.
8. Jan, B., Shah, S. W. A., Shah, S., and Qadir, M. F. (2005), "Weighted Kaplan Meier Estimation of Survival Function in Heavy Censoring," ***Pakistan Journal of Statistics***, Vol. **21**, No. 1, PP. 55-63.
9. Janssen, P. and Veraverbeke, N. (2023), "Nonparametric Estimation of Univariate and Bivariate Survival Functions under Right Censoring: A Survey," ***Metrika***.
10. Kaplan, E. L. and Meier, P. (1958), "Nonparametric Estimation from Incomplete Observations," ***Journal of the American Statistical Association***, Vol. **53**, No. 282, PP. 457- 481.
11. Lin, D. Y. and Ying, Z. (1993), "A Simple Nonparametric of the Bivariate Survival Function under Univariate Censoring," ***Biometrika***, Vol. **80**, No. 3, PP. 573-581.
12. Miller, R., Gong, G. and Munoz, A. (1980). Survival Analysis. Stanford University. Wiley Classics Library Edition.
13. Nelson, W. (1972), "Theory and Applications of Hazard Plotting for Censored Failure Data," ***Technometrics***, Vol. **14**, PP. 945-966.
14. Park, H. I. & Na, J. H. (2005), "Estimation of Bivariate Survival Function for Possibly Censored Data," ***The Korean Communications in Statistics***, Vol. 12, No. 3, PP. 783-795.
15. Prentice, R. L. and Cai, J. (1992), "Covariance and Survival Function Estimation Using Censored Multivariate Failure Time Data," ***Biometrika,*** Vol. **79**, PP. 495-512.

16. Pruitt, R. C. (1991), "Strong Consistency of Self-Consistent Estimators; General Theory and an Application to Bivariate Survival Analysis," ***Technical report*****,** Vol. ***543, University of Minnesota, Minneapolis***.

17. Qi, X., Zhou, S. and Plummer, M. (2022), "On Bayesian Modeling of Censored Data in JAGS," ***BMC Bioinformatics***, Vol. **13**, No. 1, PP. 1-13.

18. Ramos, P. L., Guzman, D. C. F., Mota, A. L., Rodrigues, F. A. and Louzada, F. (2020), "Sampling with Censored Data: A Practical Guide," Available @ ***https://webcache. Googleuser content. com/ search? q=cache:-z-HnzhN1Z8J: https:// arx iv. org/ pdf /2011 .084 17+ &cd =33 & hl=ar&ct=clnk&gl=eg***.

19. Rodriguez, G. (2005), "Multivariate Survival Models," Available @ ***https: // data. Princeton.edu/ pop509 /Multivariate Survival. pdf.***

20. Shafiq, M., Alamgir, Shah, S., and Khan, S. A. (2006), "Life Expectancy for Hemophilic Patients: A non-parametric approach," ***Journal of Pakistan Academy of Sciences PAS***, Vol. **2**, No. 3, PP. 15-19.

21. _____, Shah, S., and Alamgir. (2007), "Modified Weighted Kaplan-Meier Estimator," ***Pakistan Journal of Statistics Operation Research***, Vol. **3**, No. 1, PP. 39-44.

22. Susarla, V. and Van Ryzin, J. (1980), "Large Sample Theory for an Estimator of the Mean Survival Time from Censored Samples," ***Annals of Statistics***, Vol. **8**, PP. 1002-1016.

23. van der Laan, M. J. (1996), "Efficient Estimation in the Bivariate Censoring Model and Repairing NPMLE," ***The Annals of Statistics***, Vol. **24**, No. 2, PP. 596-627.

24. Wang, W. and Wells, M. T. (1997), "Nonparametric Estimators of the Bivariate Survival Function under Simplified Censoring Conditions," ***Biometrika,*** Vol. **84**, No. 3, PP. 863-880.

25. Wang, J. and Zafra, P. (2009), "Estimating Bivariate Survival Function by Volterra Estimator Using Dynamic Programming Techniques," ***Journal of Data Science***, Vol. **7**, PP. 365-380.

26. Zhao, G. M. A. (2008), "Nonparametric and Parametric Survival Analysis of Censored Data with Possible Violation of Method Assumptions," ***Master of Arts, Greensboro, North Carolina***.