Use Generalized Pareto Survival Models to Estimation Optimal Survival Time for Myocardial Infarction Patients

أستخدام أنموذج بقاء باريتو العام لتقدير فترة البقاء المثلى لمرضى احتشاء عضلة القلب

أ.د.افتخار عبد الحميد النقاش/ الجامعة المستنصرية
Iftikar.alnaqash@gmail.com
م.م.مروة علي مكلف/ الجامعة المستنصرية
marwahalsodani1982@gmail .com

OPEN ACCESS
P - ISSN 2518 - 5764
E - ISSN 2227 - 703X

Received:4/8/2019
Accepted: 5/9/2019

مستخلص البحث

يعد تحليل البقاء من طرائق التحليل الحديثة التي بنيت على اساس إن المتغير المعتمد يمثل الوقت حتى حدوث الحدث المهتمين به في الدراسة، وقد تعددت نماذج البقاء التي عرضت دراسة تأثير العوامل التوضيحية على أحتمال البقاء منها النماذج المقترحة من قبل العالم David Cox وهي أحد أهم نماذج البقاء وأكثرها شيوعاً، ويتكون من دالتين أحداهما دالة معلمية لا تعتمد على أوقات البقاء والاخرى دالة لامعلمية تعتمد على أوقات البقاء لذلك يعرف أنموذج Cox بأنه أنموذج شبه معلمي، ومجموعة النماذج المعلمية التي تعتمد على معلمات توزيع اوقات البقاء مثل Exponential Model, Weibull Model, Log-logistic Model. يهدف البحث الحالي في اعتماد بعض معايير الامثلية لبيز في الوصول الى التصميم الامثل لتقدير فترة البقاء المثلى لمرضى احتشاء عضلة القلب من خلال بناء أنموذج بقاء معلمي اعتماداً على التوزيع الاحتمالي لعينة اوقات بقاء مرضى احتشاء عضلة القلب الذي يعد من بين الامراض الأكثر خطورة التي تهدد حياة الانسان والمسبب الرئيسي للوفيات في جميع انحاء العالم، اذ تتباين فترات بقاء مرضى احتشاء عضلة القلب بتباين العامل او العوامل المسببة للاصابة، أذ هناك العديد من العوامل التي تؤدي الى الاصابة بهذا المرض مثل داء السكري وارتفاع ضغط الدم وارتفاع نسبة الكوليسترول الضار والضغوط النفسية والسمنة، لذا ظهرت الحاجة الى تقدير فترة البقاء المثلى من خلال بناء أنموذج للعلاقة بين العوامل التي تؤدي الى حدوث المرض وبين أوقات بقاء المرضى، وتوصلنا الى ان امثل معدل لفترة البقاء هو 18 يوم .

المصطلحات الرئيسة للبحث/ نماذج البقاء, أنموذج باريتو العام للمخاطرة النسبية, معيار امثل- C_B, معيار امثل- A_B, التصميم الامثل, مصفوفة المعلومات, طريقة الامكان الاعظم الكاملة.

*البحث مستل من اطروحة دكتوراه

Journal of Economics and Administrative Sciences
2019; Vol. 25, No.115
Pages: 452- 463

1- المقدمة

من خلال التقصي والبحث في أدبيات نظرية الامثلية وأستعمال معايير امثلية بيز لتفوقها على معايير الامثلية الكلاسكية في تحديد التصميم الامثل الذي يعطي مقدرات معلمات نماذج البقاء اللاخطية بأقل تباين يتم اعتماده في تقدير فترة البقاء المثلى على اساس قيم المتغيرات التوضيحية للأنموذج، اتضح انه لم يتم دراسة واستعمال هذه النظرية والنماذج في العراق وفي أي من البلدان العربية الاخرى لحد الان على حد علمنا، ولا يخفى على احد اهمية اعتماد اسلوب علمي دقيق لا يتطلب تحقق افتراضات صعبة ومعقدة للقيام بعملية التقدير والحصول على مقدرات بأقل تباين يمكن تفسيرها وتطبيقها بسهولة في الواقع، خاصة مع التقدم الكبير في مجال البرامجيات الجاهزة للتطبيقات الاحصائية.

يعد تحليل البقاء من طرائق التحليل الحديثة التي بنيت على اساس إن المتغير المعتمد يمثل الوقت حتى حدوث الحدث المهتمين به في الدراسة, تم تطبيق نماذج وأختبارات تحليل البقاء في العديد من المجالات الصناعية والطبية التي تعتبر الوقت عاملاً رئيسياً في دراسة وتحليل بيانات الظواهر المدروسة وقد تعددت نماذج البقاء التي تتضمن دراسة تأثير العوامل التوضيحية على أحتمال البقاء منها الأنموذج المقترح من العالم **David Cox** في عام **1972** وهو أحد أهم نماذج البقاء وأكثرها شيوعاً، ويتكون من دالتين أحداهما دالة معلمية متمثلة في دراسة تأثير المتغيرات التوضيحية وهذه الدالة لا تعتمد على أوقات البقاء والاخرى دالة لامعلمية تعتمد على أوقات البقاء لذلك يعرف أنموذج **Cox** بأنه أنموذج شبه معلمي، والنماذج الاخرى المعلمية التي تعتمد على معلمات توزيع اوقات البقاء مثل **Exponential Model, Weibull Model, Log-logistic Model** ونظراً لأهمية هذه النماذج في الحياة العملية قام العديد من الباحثين بدراستها من وجهات نظر عديدة ومختلفة ومازال اهتمام الباحثين بدراسة هذه النماذج قائماً حتى الآن. ومن هنا جاءت فكرة البحث في دراسة تأثير مجموعة من العوامل المتمثلة بالمتغيرات التوضيحية على اوقات البقاء المتمثلة بالمتغير المعتمد لمرضى احتشاء عضلة القلب وهي حالة مرضية شائعة الحدوث وتشكل خطراً على حياة الاشخاص وتعد وفق العديد من الدراسات المسبب الرئيسي للوفاة في العالم و يحصل احتشاء عضلة القلب عند قطع امداد الدم للقلب بشكل مفاجئ.

2- مشكلة البحث Problem Of Research

تعد امراض القلب والاوعية الدموية من بين الامراض الأكثر خطورة التي تهدد حياة الانسان ومنها مرض احتشاء عضلة القلب المسبب الرئيسي للوفيات في جميع انحاء العالم. تتباين فترات بقاء مرضى احتشاء عضلة القلب بتباين العامل او العوامل المسببة للاصابة بهذا المرض حسب اراء الاطباء المختصين بأمراض القلب والاوعية الدموية، أذ هناك العديد من العوامل التي تؤدي الى الاصابة بهذا المرض مثل داء السكري وارتفاع ضغط الدم وارتفاع نسبة الكوليسترول الضار والضغوط النفسية والسمنة، لذا ظهرت الحاجة الى تقدير فترة البقاء المثلى من خلال بناء أنموذج للعلاقة بين العوامل التي تؤدي الى حدوث المرض وبين أوقات بقاء مرضى احتشاء عضلة القلب.

3- هدف البحث Objective Of Research

تأسيساً على المشكلة المذكورة في الفقرة السابقة وبعد تسجيل فترات بقاء مرضى احتشاء عضلة القلب والعوامل المؤدية الى الاصابة بهذا المرض في مجموعة من المستشفيات المتخصصة، تم تحديد هدف البحث في اعتماد بعض معايير الامثلية لبيز في الوصول الى التصميم الامثل لتقدير فترة البقاء المثلى لمرضى احتشاء عضلة القلب من خلال بناء أنموذج بقاء معلمي اعتماداً على التوزيع الاحتمالي لعينة اوقات بقاء مرضى احتشاء عضلة القلب التي سنعتمدها في بحثنا والعوامل المؤدية للاصابة بهذا المرض.

4- نماذج البقاء Survival Models (2,3,4,7)

الميزة الرئيسية لبيانات البقاء (بيانات المراقبة) عدم ملائمة الطرق القياسية للنمذجة والتحليل في تقدير وتفسير معلماتها، نماذج البقاء عادةً تكون نماذج لاخطية و يمثل متغير الاستجابة قيم اوقات البقاء (t_i) أما المتغيرات التوضيحية فتمثل مجموعة العوامل التي تؤثر على فترة البقاء على قيد الحياة للمفردة تحت البحث. ففي عام 1984 افترض الباحثان Cox and Oakes (صفحة 62) أن مقابل كل وقت بقاء متجه من المتغيرات التوضيحية عناصره تمثل العوامل المختلفة التي تؤثر على وقت البقاء تتمثل في

1. المعالجة تحت البحث (متغير الحالة): متغير مقياس نوعي ذو نطاق ثنائي يأخذ القيمة (1) عند أوقات البقاء المشاهدة وتكون قيمته (0) عند أوقات البقاء المراقبة.
2. الصفات الجوهرية (الاساسية) للمشاهدات: يعبر عنها غالباً بمتغيرات نوعية كالجنس والعمر والعمل والتحصيل الدراسي وغيرها.
3. المتغيرات الخارجية: تمثل مجموعة القيم للمتغيرات المؤثرة على الظاهرة المبحوثة.

أن تسجيل قيم المتغير العشوائي لوقت البقاء يحتاج الى تحديد الاتي:

- اصل الوقت (time origin)او وقت الصفر (Time 0) يمثل وقت بداية التجربة وفي الدراسات الطبية وقت الصفر يمثل تاريخ بداية متابعة المرضى ويعرف بوقت البدء.
- مقياس الوقت (time scale) عبارة عن مقياس ثابت لقياس مرور الوقت خلال التجربة بأحدى وحدات الزمن وهي الدقائق, الساعات, ... السنوات.
- الحدث (event) الذي نهتم به في التجربة يجب ان يحدد بشكل واضح مثلاً في الدراسات الطبية الحدث يعني حدوث الوفاة او الشفاء.

وفي عام 2010 صنف الباحث German Rodriguez (صفحة 7) نماذج البقاء عندما يكون التوزيع الاحتمالي لاوقات البقاء معروف (نماذج البقاء المعلمية) على وفق اصناف النماذج التي اقترحها Cox and Oakes عام 1984 كما مبين ادناه:

1. العائلة المعلمية Parametric Families
2. الحياة المسرعة Accelerated Life
3. الخطورة النسبية Proportional Hazards
4. الاحتمالات النسبية Proportional Odds

ويعد صنف نماذج المخاطرة النسبية المعلمية Parametric Proportional Hazards Models احد اكثر نماذج البقاء استعمال وشيوعاً عندما يكون التوزيع الاحتمالي لاوقات البقاء معروف، حيث يتم في هذا النوع من النماذج تحديد شكل الدالتين المكونة لأنموذج المخاطرة النسبية وهما دالة المخاطرة النسبية $\psi(X_i)$ ودالة المخاطرة الاساسية $h_0(t)$، بالاعتماد على التوزيعات الاحتمالية المستعملة في وصف بيانات اوقات البقاء (الوقت حتى الحدث time to event) ليتم من خلالها الحصول على شكل معلمي محدد لدالة المخاطرة الاساسية، لذلك سوف نعتمد احد التوزيعات الاحتمالية الموجبة من مجموعة التوزيعات الاحتمالية التي مجالها $(0, \infty)$، الذي يثبت أختبار حسن المطابقة ملائمته لتوزيع بيانات اوقات البقاء وحيث أن بيانات اوقات البقاء في بحثنا تتبع توزيع Generalized Pareto بثلاث معلمات، لذا سنوضح في الفقرة اللاحقة الدالة الاحتمالية للتوزيع وشكل دوال وطريقة تقدير معلمات أنموذج باريتو العام للمخاطرة النسبية.

5- أنموذج باريتو العام للمخاطرة النسبية(2)

The Generalized Pareto Proportional Hazards Model

بالاعتماد على دالة الكثافة الاحتمالية لتوزيع باريتو العام بثلاث معلمات المدرجة ادناه

$$f(t) = \frac{1}{\sigma}\{1 + k\left(\frac{t-\mu}{\sigma}\right)\}^{-1-\frac{1}{k}} \quad \dots(1)$$

أذ أن

t : يمثل متغير اوقات البقاء

σ : تمثل معلمة القياس (Scale Parameter) $\sigma \in (0, \infty)$

k : تمثل معلمة الشكل Shape Parameter $k \in (-\infty, \infty)$

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

أما دالة البقاء لهذا التوزيع تكون حسب الصيغة التالية:

$$S(t) = (1 + k\left(\frac{t-\mu}{\sigma}\right))^{-\frac{1}{k}} \quad \dots(2)$$

أما دالة المخاطرة للتوزيع فيمكن الحصول عليها من خلال العلاقة بين دالة الكثافة الاحتمالية ودالة البقاء للتوزيع وكما موضح ادناه:

$$h(t) = \frac{f(t)}{S(t)} \quad \dots(3)$$

$$h(t) = \frac{\frac{1}{\sigma}(1 + k\frac{(t-\mu)}{\sigma})^{-1}(1 + k\frac{(t-\mu)}{\sigma})^{-\frac{1}{k}}}{(1 + k\frac{(t-\mu)}{\sigma})^{-\frac{1}{k}}} \quad \dots(4)$$

$$h(t) = \frac{1}{\sigma\{1 + k\left(\frac{t-\mu}{\sigma}\right)\}} \quad \dots(5)$$

$$H(t) = \int_{-\infty}^{x} h(t)dt \quad \dots(6)$$

وبذلك سوف يتم تعويض دالة المخاطرة الخاصة بتوزيع باريتو العام في أنموذج المخاطرة النسبية ويصبح أنموذج باريتو العام للخطورة النسبية حسب الصيغة:

$$h_i(t) = \exp(\beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_p x_{pi}) h_0(t) \quad \dots(7)$$

$$h_i(t) = \exp(\beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_p x_{pi}) \frac{1}{\sigma\{1 + k\left(\frac{t-\mu}{\sigma}\right)\}} \quad \dots(8)$$

$$h_i(t) = \exp(\acute{\beta} X_i) \frac{1}{\sigma\{1 + k\left(\frac{t-\mu}{\sigma}\right)\}} \quad \dots(9)$$

يتم تقدير معلمات أنموذج باريتو العام للخطورة النسبية بأستعمال طريقة الامكان الاعظم التي تجعل دالة الامكان للتوزيع فيها اعظم ما يمكن وفق الصيغة:

$$L = \prod_{i=1}^{n} \{f(t_i)\}^{\delta_i} \{S(t_i)\}^{1-\delta_i} \qquad \dots(10)$$

أذ أن

n : حجم العينة.

δ_i: يمثل مؤشر اوقات البقاء الذي يأخذ قيمة صفر عندما يكون وقت البقاء المراقب t_i ويأخذ القيمة واحد عندما يكون وقت البقاء المشاهد t_i.

$f(t_i)$: تمثل دالة التوزيع الاحتمالية لأنموذج باريتو العام للخطورة النسبية.

$S(t_i)$: تمثل دالة البقاء لأنموذج باريتو العام للخطورة النسبية.

اما مصفوفة المعلومات لأنموذج باريتو العام للمخاطرة النسبية فتعتمد على لوغارتم دالة الامكان الاعظم للانموذج و يتم حساب المشتقة الثانية للوغارتم دالة الامكان بالنسبة لكل معلمة من معلمات الانموذج غير المعروفة للحصول على مصفوفة Hessian ومن خلال مصفوفة Hessian نحصل على مصفوفة المعلومات وحسب الصيغة:

$$I(\beta) = -H(\beta) \qquad \dots(11)$$

$$I(\beta) = -\left[\frac{\partial^2 \log L(\hat{\beta})}{\partial \beta_j \, \partial \, \beta_k}\right] \qquad j,k = 1,2,\dots.p \qquad \dots(12)$$

6- معايير بيز المثلى(1,5,8) The Bayesian Optimal Criteria

اول من استعمال اسلوب تصميم بيز في ادبيات اختيار التصميم الامثل هما الباحثان **Sandor and Wedel** في عام **2001** وقاموا بأنشاء تصاميم بيز بأستعمال معيار امثل **D** – الخاص ابنموذج **Logit** متعدد الحدود ويشار له بمعيار امثل **D_B** – **D_B** واستنتجوا ان تصاميم امثل **D** – تتفوق على تصاميم امثل **D** – المحلية. اكتسب التصميم التجريبي لبيز انتشاراً واسعاً في العقدين الماضيين حيث يعتمد في بناء التصاميم المثلى للنماذج اللاخطية على تحسين المعدل لدالة من مصفوفة المعلومات من خلال توزيع سابق مفترض للمعلمات غير المعروفة لذلك يكون معيار الامثلية لبيز عبارة عن التوقع الرياضي لتلك الدالة من مصفوفة المعلومات, أيجاد تصاميم بيز المثلى رياضياً يكون معقد ويتطلب استعمال الامثلية العددية **(Numerical Optimization)** .

أن مشكلة التصميم يمكن صياغتها من خلال تعظيم القيمة المتوقعة التي غالباً ماتكون عبارة عن تكامل معقد لذا يتطلب الحل استعمال التقريب الرياضي. معظم التقريبات الى القيمة المتوقعة تتضمن استخدام التقريب الطبيعي الى التوزيع اللاحق وتأسيساً على ذلك يمكن الافتراض أن مقدرات الامكان الاعظم لمتجه المعلمات $\hat{\beta}$ تتبع التوزيع الطبيعي

$$\beta/Y,\xi \sim N\left(\hat{\beta},[M(\hat{\beta},\xi)]^{-1}\right) \qquad \dots(13)$$

وفي بحثنا سنسلط الضوء على دراسة معيارين من معايير بيز المثلى هما

1. معيار امثل – C_B Bayesian C- Optimal Criteria

دالة معيار الهدف وفق هذا المعيار تتمثل في تقليل تباين العلاقة الخطية بين المعلمات المقدرة من خلال ايجاد القيمة الصغرى لتباينات المعلمات اي ان المعيار عبارة عن تصغير القيمة المتوقعة لحاصل ضرب مصفوفة التباين والتباين المشترك في دالة خطية بالمعلمات المقدرة وفق الصيغة ادناه:

$$\Psi_C(\xi) = E_\beta[C(\beta)^T M^{-1}(\xi,\beta)C(\beta)] = \int_\beta [C(\beta)^T M^{-1}(\xi,\beta)C(\beta)]\, g(\beta)d\beta \dots(14)$$

$\Psi_C(\xi)$:معيار امثل- C العام

$M^{-1}(\xi,\beta)$: مصفوفة التباين والتباين المشترك للمعلمات المقدرة للتصميم ξ

$C(\beta)$: عبارة عن متجه الانحدار لدالة المعلمات $C_i(\beta) = \frac{\partial P(\beta)}{\partial \beta_i}$

$P(\beta)$:عبارة عن دالة خطية بالمعلمات يتم اعطائها اوزان متساوية او غير متساوية حسب طبيعة واهمية المتغيرات التوضيحية الداخلة في أنموذج الانحدار.

حيث تكون $P(\beta)$ في حالة أنموذج المخاطرة النسبية لباريتو العام وفق الصيغة:

$$P(\beta) = \alpha_1 \beta_1 + \alpha_2 \beta_2 + \alpha_3 \beta_3 + \alpha_4 \sigma + \alpha_5 \mu + \alpha_6 k \quad \dots(15)$$

وقد تم أعطاء وزن متساوي لجميع المعلمات في الدالة الخطية مساوي الى $\left(\frac{1}{\text{عدد المعلمات}}\right)$ وعند اشتقاق دالة المعلمات الخطية اعلاه مشتقة اولى بالنسبة لكل معلمة يكون المتجه $C(\beta)$ متجه عمودي يمثل الاوزان المتساوية للمعلمات.

$g(\beta)$: تمثل دالة بمتجه مقدرات الامكان الاعظم لمعلمات الأنموذج وعلى وفق التقريب الطبيعي كما في المعادلة ادناه

$$\beta/y,\xi \sim N\left(\hat{\beta},[M(\hat{\beta},\xi)]^{-1}\right) \quad \dots\dots(16)$$

$$g(\beta) = \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-\frac{(\beta-\mu)^2}{2\sigma^2}} \quad \dots\dots(17)$$

2. معيار امثل – A_B **Bayesian A- Optimal Criteria**

دالة معيار الهدف لهذا المعيار تتمثل في تقليل مجموع او متوسط تباينات المعلمات المقدرة وبتطبيق اسلوب بيز تكون صيغة المعيار عبارة عن تصغير القيمة المتوقعة لمجموع تباينات المعلمات المقدرة وفق الصيغة ادناه:

$$\Psi_A(\xi) = E_\beta\left(tr\ A(\beta)\ M^{-1}(\xi,\beta)\right) = \int_\beta tr\ A(\beta)\ M^{-1}(\xi,\beta)\ g(\beta)d\beta \quad \dots(18)$$

$$A(\beta) = C(\beta)C(\beta)^T \quad \dots(19)$$

$\Psi_A(\xi)$: معيار امثل- C العام

$\{C(\beta), g(\beta), M^{-1}(\xi,\beta)\}$كما معرفة سابقاً في معيار امثل - C_B الا انه $C(\beta)$ في هذا المعيار يتم أعطاءها اوزان غير متساوية للمعلمات في الدالة الخطية.

7. الجانب التطبيقي

7-1 بيانات البحث

لغرض تحقيق اهداف البحث المذكورة انفاً قام الباحث بزيارات ميدانية متعددة الى المستشفيات الثلاثة التي تم أختيارها عشوائياً من بين المستشفيات المختصة بأمراض القلب في محافظة بغداد وهي (مستشفى ابن النفيس, مركز ابن البيطار التخصصي لجراحة القلب, المركز العراقي لامراض القلب) بغية الحصول على بيانات عن اوقات البقاء والعوامل المؤثرة على اوقات البقاء والتي تم تحديدها من قبل مجموعة من الاطباء المختصين بأمراض القلب للمرضى المصابين بمرض احتشاء عضلة القلب والتي يفترض ان تكون مسجلة في الملف الخاص بكل مريض من المرضى الراقدين في المراكز التخصصية والمستشفيات المذكورة اعلاه خلال شهر نيسان للعام 2017 والبالغ عددهم (722) مريض، ونظراً لعدم الوعي بأهمية تسجيل وحفظ بيانات الفحوصات الطبية التي يتم اجرائها للمريض في الملفات الخاصة لهم تم الحصول على بيانات (12)

عامل من العوامل التي تؤثر على فترة بقاء المريض الى (24) مريض يمثلون عينة عشوائية بسيطة للبحث، نظراً لصعوبة دراسة جميع هذه المتغيرات ضمن نماذج البقاء لذا تم أستعمال أسلوب التحليل العاملي اكثر من مرة لتقليص عدد المتغيرات المشمولة بالبحث وبناء أنموذج البقاء على اساس أهم ثلاثة متغيرات وهي متغير الدهون الثلاثية Trig ومتغير البروتين الشحمي منخفض الكثافة VLDL ومتغير ضغط الدم الواطي وبأستعمال التطبيق الجاهز (Easy Fit 5.6 Professional) قام البرنامج بتطبيق ثلاث اختبارات احصائية للتأكد من توزيع البيانات النتائج التي حصلنا عليها موضحة في الجدول ادناه

جدول (1-7)

جدول يوضح توزيع البيانات حسب ثلاث اختبارات احصائية

Gen. Pareto					
Kolmogorov-Smirnov					
Sample Size **Statistic** **P-Value** **Rank**	**65** **0.05124** **0.99231** **1**				
□	**0.2**	**0.1**	**0.05**	**0.02**	**0.01**
Critical Value	**0.13052**	**0.14913**	**0.16567**	**0.18525**	**0.19877**
Reject?	**No**	**No**	**No**	**No**	**No**
Anderson-Darling					
Sample Size **Statistic** **Rank**	**65** **0.18582** **1**				
□	**0.2**	**0.1**	**0.05**	**0.02**	**0.01**
Critical Value	**1.3749**	**1.9286**	**2.5018**	**3.2892**	**3.9074**
Reject?	**No**	**No**	**No**	**No**	**No**
Chi-Squared					
Deg. of freedom **Statistic** **P-Value** **Rank**	**6** **0.84228** **0.99089** **1**				
□	**0.2**	**0.1**	**0.05**	**0.02**	**0.01**
Critical Value	**8.5581**	**10.645**	**12.592**	**15.033**	**16.812**
Reject?	**No**	**No**	**No**	**No**	**No**

نلاحظ من الجدول اعلاه ان قيم الاختبارات الاحصائية الثلاثة (Anderson- Darling -Kolmogorov-Smirnov- Chi-Squared) تكون اصغر من القيم الحرجة لقيم (α) المختلفة لذلك سوف يتم قبول فرضية العدم أي أن بيانات اوقات البقاء للعينة المدروسة تتبع توزيع باريتو العام

ثم تم تشكيل التصاميم بالاعتماد على فضاء التصميم الذي يتكون من (8) نقاط وبأستعمال الاسلوب العشوائي تم أختيار ثلاث أوقات بقاء لكل نقطة من نقاط فضاء التصميم وكانت البيانات حسب فضاء التصميم كما موضحة في الجدول ادناه:

جدول (7-2) البيانات حسب فضاء التصميم

تسلسل	الدفع	فترة البقاء	Status	Trig	VLDL	الضغط الواطي
A	A_1	1	0	93	19	8.7
	A_2	1	0	226	45	6.9
	A_3	2	1	86	17	8.5
B	B_1	4	1	113	23	7
	B_2	4	0	108	22	7
	B_3	5	0	246	49	9.5
C	C_1	9	0	127	25	8
	C_2	10	1	116	23	8
	C_3	10	0	159	32	8
D	D_1	11	0	431	86	10
	D_2	12	0	52	10	8
	D_3	12	1	171	34	4.2
E	E_1	16	0	270	54	7
	E_2	16	0	86	17	8
	E_3	16	1	237	47	9
F	F_1	19	1	98	20	8
	F_2	20	1	164	33	8
	F_3	21	0	99	25	8
G	G_1	22	1	125	25	9
	G_2	25	0	222	44	9
	G_3	25	0	161	32	11
H	H_1	26	1	52	10	6.8
	H_2	28	0	405	81	9
	H_3	29	0	71	14	8

2-7 تحليل البيانات

من خلال هذه البيانات تم الحصول على (44469) تصميم ويأحجام تصاميم مختلفة تبدأ من حجم تصميم (6) الى حجم تصميم (8) بتطبيق طريقة البحث العشوائي التي تمثل أحدى الخوارزميات العشوائية للامثلية العددية على نقاط فضاء التصميم التي تمثل اوقات البقاء لمرضى احتشاء عضلة القلب (بالأيام) أذ تقوم الخوارزمية على أجراء استكشاف على نقاط فضاء التصميم بأستعمال الاسلوب العشوائي لايجاد النقاط التي تقلل دالة معيار الهدف الخاصة بمعايير امثلية بيز المستخدمة وذلك من خلال اختيار القيمة الصغرى لمعيار امثل – C_B ومعيار امثل – A_B أذ كانت عدد نقاط فضاء التصميم هي ثمانية (A,B,C,D,E,F,G,H) كل نقطة من فضاء التصميم تتضمن ثلاث حالات على اساسها يتم حساب القيم التقديرية لمعلمات كل أنموذج بالاضافة الى حساب مصفوفة المعلومات ومصفوفة التباين والتباين المشترك لكل تصميم وعند كل حجم من احجام التصاميم ثم الحصول على القيمة المثلى للمعيار للتوصل الى التصميم الامثل الذي يؤدي الى امثل فترة بقاء لمرضى احتشاء عضلة القلب بأقل تباين ممكن تم الحصول على النتائج من خلال كتابة برنامج بأستعمال برنامج MATLAB وكانت النتائج موضحة في الجدول ادناه:

جدول (7-3)

التصاميم المثلى بأستخدام معيار امثل - C_B ومعيار امثل - A_B لأنموذج باريتو العام للمخاطرة النسبية

المعايير	حجم التصميم	التصميم الرئيسي	التصميم الفرعي	القيمة التقديرية	مصفوفة التباين والتباين المشترك						القيمة المثلى للمعيار	معدل فترة البقاء
معيار امثل - C_B	6	BDEFGH	$B_2D_1E_2F_2G_2H_2$	-0.1694	1.5E-08	6.0E-15	7.1E-12	2.5E-05	8.8E-08	1.2E-09	1.16E-08	18
				0.8369	6.0E-15	0.0E+00	2.8E-14	3.2E-10	1.1E-13	4.3E-11		
				1.4299	7.1E-12	2.8E-14	2.4E-10	1.3E-06	2.7E-06	5.9E-13		
				13.6489	2.5E-05	3.2E-10	1.3E-06	1.2E-01	9.3E-05	5.8E-06		
				0.3512	8.8E-08	1.1E-13	2.7E-06	9.3E-05	3.4E-07	4.5E-09		
				0.4707	1.2E-09	4.3E-11	2.5E-13	5.8E-06	4.5E-09	2.8E-10		
معيار امثل - A_B	6	BDEFGH	$B_2D_1E_2F_2G_2H_2$	-0.1694	1.5E-08	6.0E-15	7.1E-12	2.5E-05	8.8E-08	1.2E-09	2.89E-09	18
				0.8369	6.0E-15	0.0E+00	2.8E-14	3.2E-10	1.1E-13	4.3E-11		
				1.4299	7.1E-12	2.8E-14	2.4E-10	1.3E-06	2.7E-06	5.9E-13		
				13.6489	2.5E-05	3.2E-10	1.3E-06	1.2E-01	9.3E-05	5.8E-06		
				0.3512	8.8E-08	1.1E-13	2.7E-06	9.3E-05	3.4E-07	4.5E-09		
				0.4707	1.2E-09	4.3E-11	2.5E-13	5.8E-06	4.5E-09	2.8E-10		
معيار امثل - C_B	7	BCDEFGH	$B_2C_2D_1E_2F_2G_2H_2$	-0.1694	1.5E-08	7.0E-15	7.0E-12	2.5E-04	8.8E-08	1.2E-09	1.15E-08	17
				0.8369	7.0E-15	0.0E+00	2.9E-14	3.3E-10	1.0E-13	4.4E-11		
				1.4299	7.0E-12	2.9E-14	2.5E-10	1.3E-06	2.7E-06	6.0E-13		
				13.6097	2.5E-05	3.3E-10	1.3E-06	1.2E-01	9.5E-05	5.8E-06		
				0.3511	8.8E-08	1.0E-13	2.7E-06	9.5E-05	3.5E-07	4.6E-09		
				0.4708	1.2E-09	4.4E-11	6.0E-13	5.8E-06	4.6E-09	2.8E-10		

معيار امثل - A_B	7	BCDEFGH	$B_2C_2D_1E_2F_2G_2H_2$	-0.1694	1.5E-08	7.0E-15	7.0E-12	2.5E-04	8.8E-08	1.2E-09	2.88E-09	17
				0.8369	7.0E-15	0.0E+00	2.9E-14	3.3E-10	1.0E-13	4.4E-11		
				1.4299	7.0E-12	2.9E-14	2.5E-10	1.3E-06	2.7E-06	6.0E-13		
				13.6097	2.5E-05	3.3E-10	1.3E-06	1.2E-01	9.5E-05	5.8E-06		
				0.3511	8.8E-08	1.0E-13	2.7E-06	9.5E-05	3.5E-07	4.6E-09		
				0.4708	1.2E-09	4.4E-11	6.0E-13	5.8E-06	4.6E-09	2.8E-10		
معيار امثل - C_B	8	ABCDEFGH	$A_1B_2C_2D_2E_2F_2G_2H_2$	-0.1721	1.6E-07	2.3E-12	6.7E-09	2.0E-04	2.0E-06	9.1E-09	4.55E-09	15
				0.8370	2.3E-12	0.0E+00	3.2E-13	1.5E-09	9.9E-13	5.8E-10		
				1.4380	6.7E-09	3.2E-13	2.9E-09	5.0E-06	2.9E-05	8.4E-10		
				14.9922	2.0E-04	1.5E-09	5.0E-06	5.7E-01	6.6E-05	2.6E-05		
				0.3661	2.0E-06	9.9E-13	2.9E-05	6.6E-05	5.3E-07	3.0E-09		
				0.4708	9.1E-09	5.8E-10	8.4E-10	2.6E-05	3.0E-09	1.2E-09		
معيار امثل - A_B	8	ABCDEFGH	$A_1B_2C_2D_2E_2F_2G_2H_2$	-0.1721	1.6E-07	2.3E-12	6.7E-09	2.0E-04	2.0E-06	9.1E-09	1.14E-09	15
				0.8370	2.3E-12	0.0E+00	3.2E-13	1.5E-09	9.9E-13	5.8E-10		
				1.4380	6.7E-09	3.2E-13	2.9E-09	5.0E-06	2.9E-05	8.4E-10		
				14.9922	2.0E-04	1.5E-09	5.0E-06	5.7E-01	6.6E-05	2.6E-05		
				0.3661	2.0E-06	9.9E-13	2.9E-05	6.6E-05	5.3E-07	3.0E-09		
				0.4708	9.1E-09	5.8E-10	8.4E-10	2.6E-05	3.0E-09	1.2E-09		

8. مناقشة النتائج

توصلنا من تطبيق معياري A_B,C_B للامثلية في تقدير معلمات أنموذج بقاء باريتو العام بأقل تباين أن قيم معيار امثلية – A_B أصغر من قيم معيار امثلية -C_B عند مختلف احجام التصاميم المدروسة من 6-8 نقطة السبب في ذلك يعود الى اختلاف متجه المعلمات $P(\beta)$ لكل معيار، وأن حجم التصميم 8 نقاط كان افضل حجم تصميم لانه يمتلك اصغر قيمة لمعيار الامثلية من بين احجام التصاميم المدروسة سواء تم حساب معيار امثل – A_B او معيار امثل – C_B, كما أن التصاميم المثلى التي حصلنا عليها من المعيارين هي نفس التصاميم الرئيسية والفرعية عند جميع احجام التصاميم المدروسة.

توصلنا الى أن تطبيق طريقة الامكان الاعظم في تقدير معلمات الأنموذج ومعلمات دالة الخطورة الاساسية اعطت نفس المقدرات عند احجام التصاميم 6،7 وهي قريبة من القيم مقدرات الامكان الاعظم لجميع نقاط فضاء التصميم، مما يؤكد أن اختلاف قيمة معياري الامثلية يرجع الى صيغة المعيار نفسه أذ أن تساوي المقدرات يؤدي الى تساوي مصفوفة التباين والتباين مشترك وبالتالي تساوي محدد ومعكوس مصفوفة المعلومات عند احجام التصاميم المدروسة.

توصلنا الى أن فترة البقاء لمرضى احتشاء عضلة القلب تكون كبيرة (18) يوم عند حجم تصميم (6) مشاهدة للمعيارين ثم تبدأ بالانخفاض تدريجياً الى أن تصل الى (15) يوم عند حجم تصميم (8) مشاهدة.

11- المصادر

1.Chaloner K. & Larntz K., 1989 "Optimal Bayesian Design Applied to Logistic Regression Experiments", Department of Applied Statistics, University of Minnesota, pp1-18.

2.Collett D., 1994 "Modelling Survival Data in Medical Research", Texts in Statistical Science.

3.Cox D.R. & Oakes D. 1983 " Analysis of Survival Data", Department of Mathematics, Imperial College, University of London, UK & Department of Statistics, University of Rochester, USA.

4.Ibrahim J., 2005 "Survival Analysis: Introduction".

5.Kessels R., Jones B., Goos P. & Vandebroek M., "Recommendations on the Use of Bayesian Optimal Designs for Choice Experiments". Department of Decision Sciences and Information Management, Faculty of Economics and Applied Economics, Katholieke Universiteit Leuven, Naamsestraat 69, 3000 Leuven, Belgium, Department of Mathematics, Statistics and Actuarial Sciences, Faculty of Applied Economics, Universtiteit Antwerpen, Prinsstraat 13,2000 Antwerpen, Belgium, Department of Decision Sciences and Information Management & University Center for Statistics, Katholieke Universiteit, 3000 Leuven, Belgium, pp1-19.

6.Konstantinou M., 2013 "Locally Optimal and Robust Designs for Two-Parameter Nonlinear Models with Application to Survival Models", University of Southampton, Faculty of Human and Social Sciences, Mathematical Sciences.

7.Rodriguez G., 2010 "Parametric Survival Models", pp1-14.

8.Wu X., 2007 "Optimal Designs for Segmented Polynomial Regression Models and Web-based Implementation of Optimal Design Software", Stony Brook University, pp1-82.

Use Generalized Pareto Survival Models to Estimation Optimal Survival Time for Myocardial Infarction Patients

Marwah Ali Maklef Al Mustansiriyah University
Prof Dr. Iftikhar Abdulhameed Al Naqash Al Mustansiriyah University

Abstract

The survival analysis is one of the modern methods of analysis that is based on the fact that the dependent variable represents time until the event concerned in the study. There are many survival models that deal with the impact of explanatory factors on the likelihood of survival, including the models proposed by the world, David Cox, one of the most important and common models of survival, where it consists of two functions, one of which is a parametric function that does not depend on the survival time and the other a nonparametric function that depends on times of survival, which the Cox model is defined as a semi parametric model, The set of parametric models that depend on the time-to-event distribution parameters such as Exponential Model, Weibull Model, Log-logistic Model. Our research aims to adopt some of the Bayesian Optimal Criteria in achieving optimal design to estimate the optimal survival time for patients with myocardial infarction by constructing a parametric survival model based on the probability distribution of the survival times of myocardial infarction patients, which is among the most serious diseases that threaten human life and the main cause of death all over the world, as the duration of survival of patients with myocardial infarction varies with the factor or factors causing the injury, there are many factors that lead to the disease such as diabetes, high blood pressure, high cholesterol, psychological pressure and obesity. Therefore, the need to estimate the optimal survival time was expressed by constructing a model of the relationship between the factors leading to the disease and the patient survival time, and we found that the optimal rate of survival time is 18 days.

Keywords: **Survival Models, The Generalized Pareto Proportional Hazards Model, C_B- Optimal Criterion, A_B- Optimal Criterion, Optimal Design, Information Matrix, Complete Maximum Likelihood Method.**