

مقارنة بين أسلوب بيز التجريبي مع طريقة العزوم لتقدير معلمات الانتماء في المحاولات السريرية بأستعمال المحاكاة

أ. صباح هادي الجاسم الباحث محمد صباح خالد
كلية الادارة والاقتصاد جامعة بغداد

1. المستخلص:-

في هذا البحث تتم مقارنة تقديرات بيز التجريبي لمعلمات الانتماء في التجارب السريرية مع تقديرات معلمات الانتماء بطريقة العزوم، وذلك بتوظيف أسلوب المحاكاة بطريقة مونت كارلو حيث تم استعمال المقياس الاحصائي متوسط مربعات الخطأ (MSE) من اجل المقارنة بين كفاءة التقديرات ولحجوم عينات مختلفة علماً انه تم افتراض نسب فقدان مختلفة للبيانات وعلى فرض ان توزيع المشاهدات هو توزيع ثنائي الحدين في حين كان توزيع المعلمات العشوائية المجهولة هو توزيع بيتا وتم التوصل في هذا البحث الى ان تقديرات بيز التجريبي لمعلمات الانتماء العشوائية افضل من تقديرات العزوم لهذه المعلمات ولحجوم العينات المفترضة في الجانب التجريبي.

الكلمات المفتاحية :- البيانات المفقودة ، التجارب السريرية ، معلمات الانتماء

Comparison between the empirical bayes method with moments method to estimate the affiliation parameter in the clinical trials using simulation

1. Abstract:-

In this research the Empirical Bayes method is used to Estimate the affiliation parameter in the clinical trials and then we compare this with the Moment Estimates for this parameter using Monte Carlo stimulation , we assumed that the distribution of the observation is binomial distribution while the distribution with the unknown random parameters is beta distribution ,finally we conclude that the Empirical bayes method for the random affiliation parameter is efficient using Mean Squares Error (MSE) and for different Sample size .

Key word :- missing data , clinical trials , affiliation parameter

ملاحظة / بحث مستل من رسالة ماجستير لم تناقش



مجلة العلوم

الاقتصادية والإدارية

المجلد 18

العدد 69

الصفحات 226 - 236



مقارنة بين أسلوب بيز التجريبي مع طريقة العزوم لتقدير معلمات

الأنتماء في المحاولات السريرية باستعمال المحاكاة

2. المقدمة [1][3][4][5][7][9] :-

تُعرف التجارب السريرية (clinical trials) على أنها الدراسات التي تهتم بتقييم التداخلات العلاجية أو الدوائية أو الجراحية ، وذلك عن طريق تقسيم المرضى أو الأشخاص الذين ستجرى عليهم التجربة إلى مجموعتين بشكل عشوائي، حيث يطلق على المجموعة الأولى : مجموعة التجربة (experiment group) و الأخرى تكون مجموعة المراقبة (control group) .

في اغلب التجارب السريرية (Clinical trials) غالباً ما تظهر بيانات مفقودة (Missing data) ويرجع ذلك لأسباب مختلفة منها كون المريض يرفض مواصلة العلاج وذلك بعدم تناوله للدواء بسبب أعراض جانبية قد تزعجه أو كونه غير ملتزم بالعلاج بسبب إهماله لمواعيد الدواء الى غير ذلك من الأسباب. ان وجود بيانات مفقودة في التجارب السريرية تؤثر على معنوية ونتائج الاختبار في البحوث المتعلقة بهذا الموضوع لذلك نستعمل طرائق تسمى طرائق التعويض (Imputation method) وهي ادوات مهمة والتي يمكن استعمالها لمعالجة مشكلة البيانات المفقودة ويكون ذلك عن طريق تقسيم الحالات الى مجاميع عشوائية بغض النظر عن اكمالهم العلاج من عدمه لكي يكون من السهل تطبيق طرائق الاسناد ومن هذه الطرائق هي الآتية :-

1. استبعاد البيانات المفقودة (Excluding non-complete data).
2. افضل او اسوء حالة تعويض (best or worst case Imputation).
3. اخر مشاهدة مختارة الى الامام (Last observation carried forward).
4. طريقة قيمة الوسط الحسابي (Mean value method).

وما يلي ملخص لكل طريقة [3]

1. طريقة استبعاد البيانات المفقودة (Excluding non-complete data) :-

تستعمل هذه الطريقة عندما يكون عدد البيانات المفقودة بنسب او مستويات ليست عالية وبالشكل الذي لا يؤثر على حجم العينة وبالتالي معنوية نتائج البحوث المتعلقة بموضوع التجارب السريرية عندئذ يتم استبعاد هذه البيانات.

2. طريقة افضل او اسوأ حالة (best or worst case Imputation) :-

يتم بموجب هذه الطريقة الاستناد الى افضل او اسوأ مشاهدة ، وليس كليهما بعد ذلك يتم الاستعانة بهذه المشاهدات عوضاً عن البيانات المفقودة وخاصة عندما تكون مستويات فقدان البيانات عالية مما يجعل حجم العينة صغيرة وبالتالي سوف يؤثر على معنوية النتائج المتعلقة بالبحث.

3. اخر مشاهدة مختارة الى الامام (Last observation carried forward) :-

تستعمل اخر مشاهدة قبل اول مشاهدة مفقودة ، ويمكن استعمال هذه الطريقة عندما تكون نسب البيانات المفقودة قليلة ، ولكن عندما يكون هنالك زيادة او نقصان لنسب البيانات المفقودة مع الزمن فان هذه الطريقة تصبح غير مفضلة وذلك بسبب تغير قيم البيانات المفقودة مما يجعل هنالك زيادة في الفروقات بينها .

4. طريقة قيمة الوسط الحسابي (Mean value method) :-

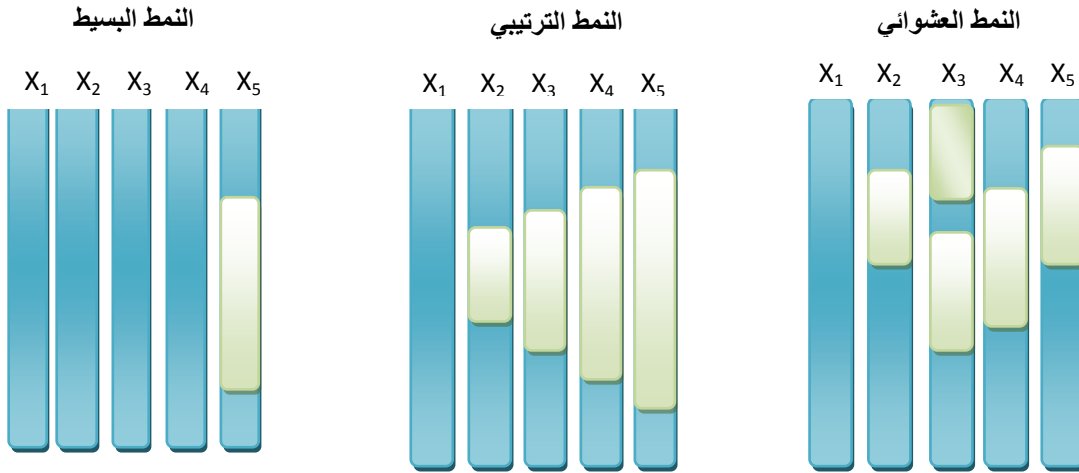
في هذه الطريقة يتم استعمال الوسط الحسابي للبيانات المسجلة للحصول على اصغر تباين للمشاهدات المتعلقة بموضوع البحث ومن الممكن في هذه الحالة ان تظهر قيم متطرفة بسبب عدم مشاهدة الحالات المريضة جداً او الجيدة جداً ، بعبارة اخرى انه باستعمال هذه الطريقة لا تكون البيانات واضحة دائماً لانها تقدر قيمة مركزية وتهمل الغموض وهنالك اسلوب للتعامل مع هذه المشكلة وهي استعمال اساليب تعويض متعددة وذلك بتوليد نسخ متعددة من البيانات ويتم ابدال هذه القيم المتولدة عشوائياً بدل القيم المفقودة.

وكما هو معلوم هناك بعض الانماط المختلفة من البيانات المفقودة وسوف أتطرق الى الانماط المتعلقة بالبحث وهي :- [9]

مقارنة بين أسلوب بيز التجريبي مع طريقة العزوم لتقدير معلمات

الانتماء في المحاولات السريرية باستعمال المحاكاة

النوع البسيط حيث تكون جميع المتغيرات مشاهدة عدا متغير واحد بعبارة أخرى ان هناك قيمة مفقودة في قسم من مشاهدات هذا المتغير، وفي حالة النمط الترتيبي للقيم المفقودة فيتم ترتيب بيانات المتغيرات بحسب عدد القيم المفقودة بشكل تصاعدي او تنازلي وبالشكل الذي يكون ترتيب المتغير تام المشاهدات اولاً يليه المتغير الذي يضم عدد اقل من القيم المفقودة وهكذا ،واخيراً يمكن ان يكون فقدان عشوائي لاي قيمة من قيم المتغيرات قيد البحث.



اما بالنسبة لالية فقدان البيانات فاذا كان سبب فقدان مستقل عن القيمة المفقودة وعن قيم المتغيرات الاخرى في العينة فان البيانات تفقد تماماً بشكل عشوائي وفي حالة كون سبب فقدان له علاقة بقيم المتغيرات الاخرى ومستقل عن القيمة المفقودة يكون فقدان عشوائي واخيراً اذا كان فقدان ناتج من القيمة المفقودة يكون سبب فقدان غير عشوائي وكما في النمط الاول.

3. هدف البحث :-

يهدف هذا البحث الى مقارنة اسلوب بيز التجريبي (Empirical bayes method) مع طريقة العزوم (Moments method) لتقدير معلمات الانتماء (اي نسبة العلاج مع الزمن) ولغرض تحقيق هذا الهدف يتم توظيف المحاكاة بطريقة مونت كارلو من اجل المقارنة بين كفاءة المقدرين ولحجوم عينات مختلفة ونسب فقدان مختلفة وذلك باستعمال المقياس الاحصائي (MSE) .

4 . الفرضيات المتعلقة بموضوع البحث:-

سنفرض ان هناك عدد N من المرضى ، وان كل مريض مستقل عن الاخر كما ان كل مريض يمتلك مشاهدات (t) حيث ان $t=1, \dots, n_i$ وان $i=1, \dots, N$ والتي يفترض بانها مستقلة ايضاً وسوف يتم تخصيص $Z_{it} \in [-3, \dots, 3]$ كمتغير لملاحظة واحدة حيث ان Z_{it} اما ان يكون جيد وفي هذه الحالة $Y_{it}=1$ اذا كان $Z_{it} \in [-1, 0, 1]$ او ان يكون غير جيد في هذه الحالة $Y_{it}=0$ اذا كان $Z_{it} \in [-3, -2, 2, 3]$ حيث ان كل من الارقام ضم الفترات اعلاه يستعمل لوصف حالة معينة من حالات المريض فمثلاً



مقارنة بين أسلوب بيز التجريبي مع طريقة العزوم لتقدير معلمات

الانتماء في المحاولات السريرية بأستعمال المحاكاة

- 3-:- عدم وجود تحسن في حالة المريض والحالة سيئة جداً
 - 3-:- وجود تحسن في حالة المريض لكن الحالة سيئة جداً
 - 2-:- عدم وجود تحسن في حالة المريض وحالة المريض سيئة
 - 2-:- وجود تحسن في حالة المريض وحالة المريض سيئة
 - 1-:- عدم وجود تحسن في حالة المريض وحالة المريض سيئة
 - 1-:- وجود تحسن في حالة المريض وحالة المريض سيئة
 - 0-:- هي افضل حالة للمريض
- ان كل رقم من الارقام اعلاه يعبر به عن المريض من خلال ملاحظة الفرق بين مشاهدت المريض السابقة واللاحقة.

ان المتغيرات العشوائية لملاحظة واحدة حتى تكون جيدة تتوزع توزيع برنولي π_i iid Y ولملاحظات عددها n_i سوف نخصص المتغير العشوائي $X_i = \sum_{t=1}^{n_i} Y_{it}$ الذي يتوزع توزيعاً ثنائياً اي ان $X_i/\pi_i \stackrel{ind}{\sim} Bi(n_i, \pi_i)$ وذلك لانه المرضى احدهم مستقل عن الآخر ولكن بسبب اختلاف n_i , π_i فان X_i/π_i سوف لا يكون متطابقة التوزيع وان π_i معلمة الانتماء لكل مريض خلال فترة المعالجة .

5. الجانب النظري [2][6]-:-

1.5 اسلوب بيز التجريبي (Emprirical Bayes Precedure) :-

بفرض توفر دالة خسارة تربيعية فان مقدر بيز القياسي للمعلمة π_i هو متوسط التوزيع اللاحق

(Posterior mean)

حيث ان التوزيع اللاحق للمعلمة العشوائية الرئيسية π_i هو :-

$$h(\theta/\underline{x}, \underline{\eta}) = \frac{f(\underline{x}/\theta) g(\theta/\eta)}{m(\underline{x}/\eta)} \quad \dots (1)$$

حيث ان η هي معلمة فوقية (hyper parameter) وان $\underline{x} = (x_1, x_2, \dots, x_n)$ هو متجه المشاهدات المتاحة و $m(\underline{x}/\eta)$ هو التوزيع الحدي (Marginal distribution) لمتجه المشاهدات $\underline{X}$ والذي هو بالصيغة الاتية :-

$$m(\underline{x}/\eta) = \int_{\theta} f(\underline{x}/\theta) g(\theta/\eta) d\eta \quad \dots (2)$$

وان توقع التوزيع اللاحق (posterior expectation) المتعلق بدالة الكثافة الاحتمالية اللاحقة في المعادلة (1)

$$E(\theta/\underline{x}) = \frac{\int_{\theta} \theta f(\underline{x}/\theta) g(\theta/\eta) d\theta}{\int_{\theta} f(\underline{x}/\theta) g(\theta/\eta) d\theta} \quad \dots (3)$$



مقارنة بين أسلوب بيز التجريبي مع طريقة العزوم لتقدير معالم

الأنتماء في المحاولات السريية بأستعمال المحاكاة

في أسلوب بيز في تقدير المعادلة (1) يمكن استعمالها لتقدير θ في حالة كون المعلمة الفوقية (η) تقديرها معلومة كما في (محاولات باركسون السريية) عندئذ فإن أسلوب بيز التجريبي سوف يستعمل التوزيع الفوقي ($hyper distribution$) للمعلمة الفوقية والمتمثل بدالة الكثافة الاحتمالي $\Phi(\eta)$ وعليه فإن التوزيع اللاحق للمعلمة العشوائية π في هذه الحالة هو :-

$$h(\theta/\underline{x}) = \frac{\int_{\forall \theta} f(\underline{x}/\theta) g(\theta/\eta) \Phi(\eta) d\eta}{\int_{\forall \theta} \int_{\forall \eta} f(\underline{x}/\theta) g(\theta/\eta) h(\eta) d\theta d\eta} \quad \dots \quad (4)$$

في المعادلة (4) يمكن استعمال مقدر الامكان الاعظم الحدي ($MMLE$) للمعلمة الفوقية η وليكن $\hat{\eta}$ وذلك باستعمال البيانات المتعلقة بالموضوع ، عندئذ يمكن عمل استدلال احصائي حول دالة الكثافة الاحتمالية اللاحقة ($h(\underline{x}/\hat{\eta})$) بواسطة تعويض $\hat{\eta}$ في المعادلة (1) وان هذا الأسلوب وبسبب اعتماده على المشاهدة الحديثة والملاحظات السابقة هو أسلوب بيز التجريبي (EB) في التقدير.

ان اختيار توزيع بيتا $\pi_i \sim Beta(r, s)$ توزيع اولي (Prior distribution) للمعلمة العشوائية π_i وهي مرافقة اولية (Prior conjugate) لتوزيع ثنائي الحدين الشرطي $X_i/\pi_i \sim Bi(n_i, \pi_i)$ ، وذلك لان التوزيع اللاحق للمعلمة العشوائية π_i سوف ينتمي لنفس العائلة المعلمية ($Parametric family$) ، حيث ان دالة الكتلة الاحتمالية للمتغير العشوائي (X_i/π_i) هي :-

$$P(X_i/\pi_i) = \binom{n_i}{x_i} \pi_i^{x_i} (1 - \pi_i)^{n_i - x_i} \quad \dots \quad (5)$$

ولغرض ايجاد التوقع والتباين للمتغير العشوائي X_i/π_i حيث ان $i=1,2,\dots,N$ نتبع الخطوات التالية :-

$$\begin{aligned} E(X_i/\pi_i) &= \sum_{x_i=0}^{n_i} x_i P(X_i/\pi_i) \\ E(X_i/\pi_i) &= \sum_{x_i=0}^{n_i} x_i \binom{n_i}{x_i} \pi_i^{x_i} (1 - \pi_i)^{n_i - x_i} \\ E(X_i/\pi_i) &= \sum_{x_i=0}^{n_i} x_i \frac{n_i!}{x_i!(n_i - x_i)!} \pi_i^{x_i} (1 - \pi_i)^{n_i - x_i} \\ E(X_i/\pi_i) &= \sum_{x_i=0}^{n_i} x_i \frac{n_i(n_i - 1)!}{x_i(x_i - 1)!(n_i - x_i)!} \pi_i \pi_i^{x_i - 1} (1 - \pi_i)^{n_i - x_i} \\ E(X_i/\pi_i) &= n_i \pi_i \sum_{x_i=0}^{n_i} \frac{(n_i - 1)!}{(x_i - 1)!(n_i - x_i)!} \pi_i^{x_i - 1} (1 - \pi_i)^{n_i - x_i} \\ E(X_i/\pi_i) &= n_i \pi_i \quad \dots \quad (6) \end{aligned}$$

$$E\left(\frac{X_i}{n_i} / \pi_i\right) = \frac{E(X_i = x_i / \pi_i)}{n_i} = \frac{n_i \pi_i}{n_i} = \pi_i \quad \dots \quad (7)$$



مقارنة بين أسلوب بيز التجريبي مع طريقة العزوم لتقدير معالم الأنتماء في المحاولات السريية بأستعمال المحاكاة

وبنفس الأسلوب يمكن إيجاد $[E(X_i(X_i - 1)/\pi_i)]$ حيث ان

$$\begin{aligned} [E(X_i(X_i - 1)/\pi_i)] &= \sum_{x_i=0}^{n_i} x_i(x_i - 1) P(X_i/\pi_i) \\ [E(X_i(X_i - 1)/\pi_i)] &= \sum_{x_i=0}^{n_i} x_i(x_i - 1) \binom{n_i}{x_i} \pi_i^{x_i} (1 - \pi_i)^{n_i - x_i} \\ [E(X_i(X_i - 1)/\pi_i)] &= \sum_{x_i=0}^{n_i} x_i(x_i - 1) \frac{n_i!}{x_i!(n_i - x_i)!} \pi_i^{x_i} (1 - \pi_i)^{n_i - x_i} \\ [E(X_i(X_i - 1)/\pi_i)] &= \sum_{x_i=0}^{n_i} x_i(x_i - 1) \frac{n_i(n_i - 1)(n_i - 2)!}{x_i(x_i - 1)(x_i - 2)!(n_i - x_i)!} \pi_i^2 \pi_i^{x_i - 2} (1 - \pi_i)^{n_i - x_i} \\ [E(X_i(X_i - 1)/\pi_i)] &= n_i(n_i - 1) \pi_i^2 \sum_{x_i=0}^{n_i} \frac{(n_i - 2)!}{(x_i - 2)!(n_i - x_i)!} \pi_i^{x_i - 2} (1 - \pi_i)^{n_i - x_i} \\ [E(X_i(X_i - 1)/\pi_i)] &= n_i(n_i - 1) \pi_i^2 \quad \dots \quad (8) \end{aligned}$$

من المعادلة (6) و المعادلة (8) اعلاه يمكن إيجاد التباين وكما يلي :-

$$\begin{aligned} var(X_i/\pi_i) &= E(X_i^2/\pi_i) - [E(X_i/\pi_i)]^2 \\ \text{بإضافة وطرح } E(X_i/X_i) \text{ ينتج} \\ var(X_i/\pi_i) &= E(X_i^2/\pi_i) - [E(X_i/\pi_i)]^2 - E(X_i/\pi_i) + E(X_i/\pi_i) \\ var(X_i/\pi_i) &= E(X_i^2/\pi_i) - E(X_i/\pi_i) - [E(X_i/\pi_i)]^2 + E(X_i/\pi_i) \\ var(X_i/\pi_i) &= E((X_i^2 - X_i)/\pi_i) - [E(X_i/\pi_i)]^2 + E(X_i/\pi_i) \\ var(X_i/\pi_i) &= n_i(n_i - 1) \pi_i^2 - n_i^2 \pi_i^2 + n_i \pi_i \\ var(X_i/\pi_i) &= n_i^2 \pi_i^2 - n_i \pi_i^2 - n_i^2 \pi_i^2 + n_i \pi_i \\ var(X_i/\pi_i) &= n_i \pi_i - n_i \pi_i^2 \\ var(X_i/\pi_i) &= n_i \pi_i (1 - \pi_i) \quad \dots \quad (9) \\ var\left(\frac{X_i}{n_i}/\pi_i\right) &= \frac{n_i \pi_i (1 - \pi_i)}{n_i^2} = \frac{\pi_i (1 - \pi_i)}{n_i} \quad \dots \quad (10) \end{aligned}$$

ان دالة الكثافة الاحتمالية للتوزيع الاولي للمعلمة العشوائية $\pi \sim Beta(a, b)$ هي :-

$$p(\Pi_i = \pi_i/a, b) = \frac{\Gamma(r+s)}{\Gamma(r)\Gamma(s)} \pi_i^{r-1} (1 - \pi_i)^{s-1} \quad ; r, s > 0 \quad \dots \quad (11)$$

ولجعل توزيع ثنائي الحدين و بيتا اكثر مرونة في الاستعمال ، نقوم بابدال معالم بيتا (a, b) الى (M, μ) حيث ان :-

بفرض ان [2]

$$\mu = \frac{s}{(r+s)} \quad \dots \quad (12) \quad M = r + s$$

ينتج

$$\mu = \frac{s}{M}$$

$$\therefore s = M\mu$$

$$r = M - M\mu$$

$$\therefore r = M(1 - \mu)$$



مقارنة بين أسلوب بيز التجريبي مع طريقة العزوم لتقدير معلمات الأنتماء في المحاولات السريية بأستعمال المحاكاة

وتصبح المعادلة (11) بالصيغة التالية :-

$$g(\Pi_i = \pi_i / \mu, M) = \frac{\Gamma(M)}{\Gamma(M\mu)\Gamma(M(1-\mu))} \pi_i^{M\mu-1} (1 - \pi_i)^{M(1-\mu)-1} \quad \dots (13)$$

حيث ان M, μ هي معلمات فوقية (hyperparameters).

ان التوقع والتباين للمتغير العشوائي π_i هي كالآتي:-

التوقع او العزم الاول هو

$$E(\pi_i / \mu, M) = \mu \quad \dots (14)$$

اما العزم الثاني فهو

$$E(\pi_i^2 / \mu, M) = \frac{\mu(M\mu+1)}{(M+1)} \quad \dots (15)$$

اما التباين فيكون

$$var(\pi_i / \mu, M) = \frac{\mu - \mu^2}{(M+1)} \quad \dots (16)$$

وتكون دالة الكثافة الاحتمالية الحدية هي :-

$$m(X_i = x_i / \mu, M) = \int_0^1 p(X_i / \pi_i) g(\Pi_i / \mu, M) d\pi_i$$

$$m(X_i = x_i / \mu, M) = \binom{n_i}{x_i} \frac{\Gamma(M)}{\Gamma(M\mu)\Gamma(M(1-\mu))} * \frac{\Gamma(x_i + M\mu)\Gamma(n_i - x_i + M(1-\mu))}{\Gamma(n_i + M)} \quad \dots (17)$$

اما التوقع للمتغير العشوائي الحدي X_i فهو:-

$$E\left(\frac{X_i}{n_i}\right) = E\left[E\left(\frac{X_i}{n_i} / \pi_i\right)\right] = E[\pi_i] = \mu \quad \dots (18)$$

$$E(X_i) = n_i \mu$$

ان التباين يمكن ايجاده كما ياتي:-

$$var\left(\frac{X_i}{n_i}\right) = E\left[var\left(\frac{X_i}{n_i} / \pi_i\right)\right] + var\left[E\left(\frac{X_i}{n_i} / \pi_i\right)\right]$$

من المعادلة (7) لدينا

$$But \quad E\left(\frac{X_i}{n_i} / \pi_i\right) = \pi_i$$

$$And \text{ also } var\left(\frac{X_i}{n_i} / \pi_i\right) = \frac{\pi_i(1-\pi_i)}{n_i} \quad \text{من المعادلة (10)}$$

$$\therefore var\left(\frac{X_i}{n_i}\right) = E\left[\frac{\pi_i(1-\pi_i)}{n_i}\right] + var(\pi_i)$$

$$var\left(\frac{X_i}{n_i}\right) = \frac{1}{n} E[\pi_i(1 - \pi_i)] + var(\pi_i)$$

$$But \quad E[\pi_i(1 - \pi_i)] = E(\pi_i) - [E(\pi_i^2)]$$

$$E[\pi_i(1 - \pi_i)] = \mu - \mu^2 - var(\pi_i)$$

$$\therefore var\left(\frac{X_i}{n_i}\right) = \frac{1}{n_i} [\mu(1 - \mu) - var(\pi_i)] + var(\pi_i)$$

$$var\left(\frac{X_i}{n_i}\right) = \frac{1}{n_i} [\mu(1 - \mu)] - \frac{1}{n_i} \left(\frac{\mu(1-\mu)}{M+1}\right) + \left(\frac{\mu(1-\mu)}{M+1}\right)$$

$$var\left(\frac{X_i}{n_i}\right) = \frac{1}{n_i} [\mu(1 - \mu)] + \frac{1}{n_i} \left(\frac{\mu(1-\mu)}{M+1}\right) (n_i - 1)$$

$$var\left(\frac{X_i}{n_i}\right) = \frac{1}{n_i} [\mu(1 - \mu)] \left(1 + \frac{n_i - 1}{M+1}\right) \quad \dots (19)$$



مقارنة بين أسلوب بيز التجريبي مع طريقة العزوم لتقدير معلمات الانتماء في المحاولات السريية بأستعمال المحاكاة

ان التوزيع اللاحق (*posterior distribution*) للمعلمة العشوائية π_i هو بالصيغة الآتية :-

$$h(\pi_i/X_i) \propto P(X_i/\pi_i)p(\pi_i)$$

وباستعمال ثابت التناسب $m(X_i = x_i/\mu, M)$ تكون دالة الكثافة الاحتمالية اللاحقة للمعلمة العشوائية π_i هي :-

$$P(\pi_i/X_i) = \binom{n_i}{X_i} \frac{\Gamma(M)}{\Gamma(M\mu)\Gamma(M(1-\mu))} \pi_i^{X_i+M\mu-1} (1-\pi_i)^{n_i-X_i+M(1-\mu)-1} \dots (20)$$

وهذا أيضاً توزيع بيتا ولكن بالعلمات $[X_i+M\mu, n_i-X_i+M(1-\mu)]$.

ان مقدار بيز للمعلمة العشوائية π_i هي التوقع اللاحق (*Posterior expectation*) :-

$$\begin{aligned} \hat{\pi}_{EBi} &= \frac{X_i+M\mu}{X_i+M\mu+n_i-X_i+M(1-\mu)} \\ &= \frac{X_i+M\mu}{n_i+M} \dots (21) \end{aligned}$$

ولغرض ايجاد المقدّر $\hat{\pi}_{BEi}$ نقوم بتوضيف مقدرات الامكان الاعظم الحدية لكل من M و μ حيث ان X هي المشاهدات الحديثة وان n هو حجم العينة وهي كميات معلومة.

2.5 طريقة العزوم (*Moments method*) :-

اما مقدار العزوم (*Moments Estimates*) لمعلمة الانتماء π_i (نسبة العلاج مع الزمن) فيمكن الحصول عليه كما يأتي :-

$$\hat{\pi}_i = \frac{x_i}{n_i} \dots (22)$$

اما المقدّر الموزون للوسط (μ) فهو الآتي

$$\hat{\mu} = \frac{\sum_{i=1}^N n_i \hat{\pi}_i}{\sum_{i=1}^N n_i} = \frac{\sum_{i=1}^N x_i}{\sum_{i=1}^N n_i} \dots (23)$$

ولايجاد عزم المقدّر لـ (M) نستخدم المعادلة (19) حيث ان عزم المقدّر الموزون لـ (S^2) هو

$$S^2 = \frac{N \sum_{i=1}^N n_i (\hat{\pi}_i - \hat{\mu})^2}{(N-1) \sum_{i=1}^N n_i} \dots (24)$$

كذلك

$$S^2 = \frac{1}{N} \sum_{i=1}^N \text{var} \left(\frac{x_i}{n_i} \right) = \frac{1}{N} \sum_{i=1}^N \frac{\hat{\mu}(1-\hat{\mu})}{n_i} \left[1 + \frac{n_i-1}{\hat{M}+1} \right] \dots (25)$$

وبحل المعادلات (24) و (25) اعلاه نحصل على

$$\hat{M} = \frac{\hat{\mu}(1-\hat{\mu})-S^2}{S^2 - \frac{\hat{\mu}(1-\hat{\mu})}{N} \sum_{i=1}^N \frac{1}{n_i}} \dots (26)$$

اما بالنسبة لتوقع التوزيع اللاحق فهو

$$\hat{\pi}_i = E(\pi/\hat{\mu}, \hat{M}) = \frac{\pi_{EB}}{\pi_{EB} + s_{EB}} = \frac{\frac{x_i}{n_i} + \hat{M}\hat{\mu}}{n_i + \hat{M}} = \frac{\hat{M}}{n_i + \hat{M}} \hat{\mu} + \frac{n_i}{n_i + \hat{M}} * \frac{x_i}{n_i} \dots (27)$$



مقارنة بين أسلوب بيز التجريبي مع طريقة العزوم لتقدير معلمات الانتماء في المحاولات السريرية بأستعمال المحاكاة

وتباين المقدّر اللاحق هو

$$\tilde{\sigma}^2 = \frac{r_{EB} s_{EB}}{(r_{EB} + s_{EB})^2 (r_{EB} + s_{EB} + 1)} = \frac{\tilde{\pi}_i (1 - \tilde{\pi}_i)}{n_i + \tilde{M} + 1} \quad \dots (28)$$

حيث ان

r_{EB}, s_{EB} - هي المعلمات للتوزيع اللاحق (posterior distribution)

$$\tilde{\pi}_i = \tilde{B} \hat{\mu} + (1 - \tilde{B}) \hat{\pi}_i \quad \dots (29)$$

حيث ان

$$\tilde{B} = \frac{\tilde{M}}{\tilde{M} + n_i}$$

6. الجانب التجريبي :-

في هذا الجانب سوف تتم المقارنه بين مقدري العزوم ومقدر بيز التجريبي لمعلمات الانتماء في المحاولات السريرية تجريبياً وفق ما ورد من تعاريف ومفاهيم لهذه التجارب في مقدمة البحث ، حيث ليس هنالك بيانات حقيقية وانما بيانات تحاكي الواقع التطبيقي وذلك بتوظيف اسلوب المحاكاة بطريقة مونت كارلو (Monte Carlo) وكانت الفرضيات المتعلقة بتجربة المحاكاة هي كالآتي :-

1. القيم الافتراضية لحجوم العينات :-

تم اختيار حجوم عينات مختلفة بشكل يتناسب مع مدى تأثير حجم العينة على دقة وكفاءة النتائج المستحصلة من طرائق التقدير المستخدمة ضمن هذا البحث ، حيث تم اخذ حجم عينة صغيرة وهو (20) وحجم عينة متوسطة (50) واخيراً حجم عينة كبير هو (100).

2. القيم الافتراضية للمعلمات :-

القيم الافتراضية للمعلمات هي $\mu = 0.65$ و $M = 10$.

3. طريقة التوليد :-

تم في هذه المرحلة توليد بيانات عشوائية تتبع توزيع ثنائي الحدين وتوزيع بيتا وتم التوليد باستخدام طريقة الرفض والقبول وذلك لسهولة وكفاءتها .

4. مقياس الكفاءة للمقدرين (MSE) :-

متوسطات مربعات الخطأ للمقدرات $\hat{\pi}_i$

$$MSE(\hat{\pi}_i) = \frac{1}{L} \sum_{j=1}^L (\hat{\pi}_{ij} - \hat{\pi}_i)^2$$

L : عدد تكرارات التجربة حيث ان (L = 1000)

$\hat{\pi}_{ij}$: مقدر معلمة الانتماء $\hat{\pi}_i$ بحسب الطريقة ، حيث ان $i = 1, 2, \dots$



مقارنة بين أسلوب بيز التجريبي مع طريقة العزوم لتقدير معلمات
الأنتماء في المحاولات السريرية بأستعمال المحاكاة

حجم العينة	نسبة الفقدان	$\pi \sim \text{Beta}(M\mu, M(1-\mu))$	$\hat{\pi}_{EB}$	$\text{MSE}(\hat{\pi}_{EB})$	$\hat{\pi}_{ME}$	$\text{MSE}(\hat{\pi}_{ME})$
20	0%	0.699353	0.8833	0.1856	0.952381	0.24983
	10%	0.644192	0.875	0.20398	0.947368	0.27459
	20%	0.641741	0.86538	0.2221	0.941176	0.29922
	25%	0.619721	0.85416	0.23931	0.933333	0.32304
50	0%	0.662107	0.94166	0.00262	0.98039	0.00807
	10%	0.704686	0.93636	0.01424	0.978260	0.02583
	20%	0.725791	0.93	0.03440	0.97153	0.05339
	25%	0.78532	0.92222	0.06284	0.97222	0.09041
100	0%	0.63260	0.96818	0.00603	0.99009	0.009917
	10%	0.62325	0.965	0.02175	0.98909	0.029409
	20%	0.55724	0.96111	0.04690	0.987654	0.05911
	25%	0.55689	0.95880	0.06289	0.986842	0.077736

جدول بالنتائج المستحصل عليها من تجربة المحاكاة



مقارنة بين أسلوب بيز التجريبي مع طريقة العزوم لتقدير معلمات الأنتماء في المحاولات السريرية بأستعمال المحاكاة

7. الاستنتاجات:-

1. في ضوء تحليل تجربة المحاكاة تم التوصل الى الاستنتاجات التالية :-
1. طريقة بيز التجريبي نافعة لكميات مختلفة من البيانات المفقودة ولقيم مختلفة من معلمات البيانات لعدد من المشاهدات .
2. أسلوب بيز التجريبي في التقدير اكثر كفاءة من أسلوب العزوم في التقدير لجميع حجوم العينات ولجميع نسب فقدان المختلفة .
3. في حالة ان البيانات تامة ولا تعاني اي نسبة فقدان فان افضل مقدر لاسلوب بيز يكون عند الحجم (50) .
4. كلما زادت نسبة البيانات المفقودة بالنسبة للبيانات الاصلية زادت قيم متوسط مربعات الخطأ (MSE) للمقدرات وبالتالي تقل كفاءة المقدرات .

8. المصادر:-

- [1] Basu,s.& Heitjan ,D.F.(1996) "Distinguishing' Missing at Random' &'Missing completely at Random' The JASA,vol.50,No.30,p.207-231.
- [2] Carlin ,B.P. and Louis,T.A. (2000) Bayes and Empirical Bayes Method for Data Analysis , New York: Chapman & Hall.
- [3] Committee for Proprietary Medicinal Products (2001) Points to consider on Missing Data, London :The European Agency for the Evaluation of Medicinal Products.
- [4] Encyclopedia of Biostatistics (1998),Wiley.
- [5] EMEA (European Medicines Agency) (1998).
- [6] Gut A. (1995) An Intermediate Course in Probability, New York: Springer-Verlag.
- [7] Kazemi ,I,(2005) "method for Missing data"
http://www.cas.lancs.ac.uk/short_courses/notes/missingdata/session6.pdf
- [8] Little,R.J.A& Rubin ,D.B(2003)"statistical analysis with Missing Data" 2thed, John Wiley & Sons, New York.
- [9] Rapur ,N,(2003) " TREATMENT OF DATA WITH MISSING ELEMENTS IN PROCESS MODELLING" the University of Cincinnati, Department of Industrial Engineering, M.Sc Thesis.