مجلة العلوم الرياضية والمرتبطة
المجلد الرابع / العدد الحادي والثلاثون – ديسمبر 2021 ISSN 2518-5470

تطبيق طريقة البوتستراب لحساب فترات ثقة لإحصاءات مختلفة باستخدام ماتلاب

Applying Bootstrap Method to Compute Confidence Intervals for Various Statistics Using MATLAB

د. البهلول عمر علي شلابي
الايميل : b.shalabi@uot.edu.ly

ا. سعاد محمد أحمد البرقاوي
الايميل :s.elbargawi@uot.edu.ly

تاريخ الاستلام / 2021/7/12
تاريخ القبول /2021/10/28

الكلمات المفتاحية : طريقة البوتستراب – طريقة ماتلاب

الملخص البحث

أتاحت طريقة البوتستراب (Bootstrap Method) التي قدمها العالم إيفرون (Efron, 1979) إجراء العديد من الأبحاث في مجال الإحصاء وغيرها من المجالات، والتي كانت لولا ذلك ستكون صعبة إن لم تكن مستحيلة. لقد أصبحت طريقة البوتستراب تأخذ جانباً مهماً في أغلب البحوث الإحصائية خاصة فيما يتعلق بالبحوث الطبية.

تم في هذه الورقة تزويد الباحث بمقدمة مختصرة عن طريقة البوتستراب، كما تم تقديم شرح مبسط لكيفية استخدام طريقة البوتستراب لحساب الخطأ المعياري لمتوسط العينة، والخطأ المعياري لوسيط العينة، بالإضافة إلى شرح كيفية استخدام طريقة البوتستراب في تكوين فترات ثقة لكل من المتوسط الحسابي والوسيط بمستوى ثقة معين باستخدام طريقة فترة بوتستراب المئوية (Percentile Bootstrap Interval). لتمكين الباحث من تطبيق طريقة البوتستراب بكل سهولة ويسر فقد تم في هذه الورقة تقديم شرح وافي وواضح لكيفية تطبيق طريقة البوتستراب باستخدام برنامج ماتلاب (MATLAB Software)، وفي نهاية الورقة تم تقديم دالة ماتلاب، تحمل اسم bootstrap_ci، وظيفتها الأساسية هي ايجاد فترات ثقة لمتوسط ووسيط المجتمع باستخدام طريقة فترة البوتستراب المئوية.

المقدمة

تعتبرعملية تحديد قيمة الخطأ المعياري (Standard Error) لبعض الإحصاءات الهامة من المسائل الهامة في الإحصاء الإستدلالي (Inferential statistics).

عملية سحب عينات مختلفة ذات نفس الحجم من مجتمع إحصائي محدود أو غير محدود، ثم القيام بحساب قيمة إحصاءة معينة لكل عينة لتقدير معلمة المجتمع المجهولة (مثلاً: حساب المتوسط الحسابي لكل عينة لتقدير معلمة المتوسط الحسابي للمجتمع) ثم وضع قيم الإحصاءات المحسوبة لجميع العينات في جدول تكرار، يعطي ما يعرف بتوزيع المعاينة للإحصاءة (The Sampling Distribution of the Statistic) .

قبل الانتشار الواسع لأجهزة الحاسوب السريعة، كانت عملية إيجاد توزيعات المعاينة للإحصاءات المختلفة يتم إشتقاقها رياضياً، ولكن الاشتقاقات الرياضية لتوزيع المعاينة لإحصاءة ما غالباً ما تتطلب وضع افتراضات حول توزيع الاحتمال للمجتمع الذي جاءت منه تلك الإحصاءة، وتنطبق هذه الافتراضات فقط على إحصاءات معينة مثل المتوسط الحسابي للعينة.

تتم عملية سحب عينات بشكل متكرر عدة مرات من المجتمع الإحصائي محل الدراسة بمساعدة أجهزة الحاسوب ذات المواصفات العالية والمتوفرة بسهولة في الوقت الحالي.

طرق إعادة المعاينة (Resampling Methods) هي التي يتم فيها أخذ عينات من مشاهدات عينة لاستخلاص استنتاجات معينة حول المجتمع الذي سحبت منه هذه العينة. هذا وتوجد طريقتان من أكثر طرق إعادة المعاينة شيوعًا هما: طريقة جاك نايف (Jackknife Method) وطريقة البوتستراب (Bootstrap Method)، كلتا الطريقتين تعتبر أمثلة على الأساليب الإحصائية اللامعلمية (Efron, 1982) .

مثل هذه الطرق تتميز بأنها لا تستلزم أي افتراض حول توزيع المجتمع الذي جاءت منه مشاهدات العينة كما أن الرياضيات المستخدمة فيها بسيطة نسبيًا ولا تتطلب سوى فهم الجبرعلى مستوى المدرسة الثانوية.

طريقة البوتستراب (Bootstrap Method) هي تقنية تعتمد بشكل كبيرعلى الحاسوب بحيث يمكن استخدامها لاستنتاج توزيع المعاينة لأي إحصاءة تقريبًا عن طريق أخذ عينات متكررة (مع الإرجاع) من العينة نفسها.

تم تقديم طريقة البوتستراب بواسطة إيفرون سنة 1979 (Efron, 1979) كطريقة عامة لحساب الخطأ المعياري لأي مقدر، وإيجاد فترة الثقة لمعلمة مجهولة عندما يكون توزيع المجتمع مجهول ولا يمكن اشتقاق صيغة رياضية لحساب الخطأ المعياري وحدود فترة الثقة للمعلمة المجهولة.

عندما لا توجد صيغة بسيطة ومعروفة لحساب الخطأ المعياري لإحصاءة ما، *SE*، فإن طريقة البوتستراب تكون مفيدة جداً.

الهدف من هذه الورقة

تهدف هذه الورقة إلى تقديم مقدمة قصيرة مختصرة عن طريقة البوتستراب، كما تهدف أيضاً إلى عرض موجز لكيفية استخدام طريقة البوتستراب لحساب الخطأ المعياري لمتوسط العينة، والخطأ المعياري لوسيط العينة، بالإضافة إلى توضيح كيفية استخدام طريقة البوتستراب في تكوين فترات ثقة لكل من المتوسط الحسابي والوسيط بمستوى ثقة معين وذلك بالاستعانة ببرنامج ماتلاب (MATLAB).

المجتمع الأفتراضي والعينة التي تم استخدامها في هذه الورقة

لتوضيح طريقة البوتستراب وكيفية الحصول على فترات ثقة حول المتوسط الحسابي والوسيط لمجتمع ما سنفترض أن لدينا مجتمع أوزان مرضى الغدة الدرقية (THS) مكون من 2000 مريض. تم قياس وزن كل مريض بالكيلو جرام (كجم) وحسبت قيمة المتوسط الحسابي لأوزانهم فكانت تساوي 54.097 كجم وبانحراف معياري قدره 8.9258 كجم، كما حسبت قيمة الوسيط لأوزانهم فكانت تساوي 54.367 كجم.

أخذت عينة عشوائية من مجتمع مرضى الغدة الدرقية المفترض حجمها 15 مريض فكانت أوزانهم كما هو مبين في جدول 1 المبين أدناه.

جدول 1: أوزان عينة عشوائية حجمها 15 مريض مقاسة بالكيلوجرام مسحوبة من مجتمع مرضى الغدة الدرقية المفترض.

61.415	55.994	46.304	55.249	74.575
60.318	46.448	49.869	58.373	46.733
75.694	56.9	45.322	35.188	57.411

حساب الخطأ المعياري للمتوسط الحسابي للعينة وفقاً لنظرية النهاية المركزية

لحساب الخطأ المعياري للمتوسط الحسابي للعينة المبينة مشاهداتها في جدول 1، نقوم أولاً بحساب قيمة المتوسط الحسابي للعينة كالتالي:

$$\bar{y} = \sum_{i=1}^{n} y_i / n = \sum_{i=1}^{15} y_i / 15 = 55.053$$

ثم نقوم بحساب الانحراف المعياري لمشاهدات العينة كالتالي:

$$s_y = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2} = \sqrt{\frac{1}{15-1}\sum_{i=1}^{15}(y_i - 55.053)^2} = 10.798$$

وفقاً للنظرية الإحصائية، فإن قيمة الخطأ المعياري لمتوسط العينة، $SE_{\bar{y}}$ ، يحسب كالتالي:

$$SE_{\bar{Y}} = \frac{s_y}{\sqrt{n}} = \frac{10.798}{\sqrt{15}} = 2.7879$$

الصيغة الأخيرة لحساب قيمة $SE_{\bar{y}}$ جاءت وفقاً لنظرية النهاية المركزية (Central Limit Theorem) التي تنص على أن إحصاءة المتوسط الحسابي للعينة، $\bar{Y}$ ، تتبع تقريباً التوزيع الطبيعي.

حساب الخطأ المعياري للمتوسط الحسابي للعينة وفقاً لطريقة البوتستراب

الآن سنتطرق إلى شرح مبسط لكيفية استخدام طريقة البوتستراب كطريقة بديلة لحساب الخطأ المعياري للمتوسط الحسابي للعينة.

تُعرف طريقة البوتستراب بأنها إجراء إعادة أخذ عينات من مشاهدات عينة واحدة للحصول على توزيع معاينة لإحصاءة معينة، وبالتالي حساب الخطأ المعياري للإحصاءة، وذلك عن طريق أخذ عينات عشوائية (مع الإحلال) بشكل متكرر من مشاهدات العينة التي لدينا، ذات الحجم n، والتي تم سحبها من المجتمع الإحصائي المستهدف بالدراسة (Efron and Tibshirani1993).

تتلخص خطوات إنشاء عينات البوتستراب (Bootstrap Samples) للحصول على الخطأ المعياري للمتوسط الحسابي للعينة (أو لأي إحصاءة أخرى) في التالي:

1- إحلال مشاهدات العينة، ذات الحجم n ، التي لدينا محل مشاهدات المجتمع.
2- أخذ عدد B من العينات العشوائية، ذات الحجم n ، (مع الإحلال) من بين مشاهدات العينة التي لدينا في الخطوة 1. عدد العينات B يجب أن يكون كبيراً، لنقل مثلاً B تساوي 1000. سنشير إلى عينات البوتستراب (Bootstrap Samples) التي عددها B بكتابة:

$$B^{(1)}, B^{(2)}, \cdots, B^{(B)}$$

3- حساب قيمة المتوسط الحسابي لكل عينة من العينات التي عددها B . سنشير إلى قيم المتوسطات بكتابة:

$$\bar{y}_{B^{(1)}}, \bar{y}_{B^{(2)}}, \cdots, \bar{y}_{B^{(B)}}$$

4- حساب المتوسط الحسابي الكلي لجميع المتوسطات الحسابية في الخطوة 3، وذلك وفقاً للصيغة التالي:

$$\bar{\bar{y}}_B = \frac{1}{B}\sum_{i=1}^{B} \bar{y}_{B^{(i)}} = \frac{1}{B}\left(\bar{y}_{B^{(1)}} + \bar{y}_{B^{(2)}} + \cdots \bar{y}_{B^{(B)}}\right)$$

5- تكوين توزيع المعاينة للمتوسط الحسابي للعينة وذلك بإنشاء جدول التكرار (أو برسم مدرج التكرار) لقيم المتوسطات الحسابية التي تم الحصول عليها في الخطوة 3 .
6- حساب قيمة الخطأ المعياري للمتوسط الحسابي للعينة وذلك بإيجاد قيمة الانحراف المعياري لتوزيع المعاينة للمتوسط الحسابي للعينة الذي تم الحصول عليه في الخطوة 4 وذلك باستخدام الصيغة التالية:

$$SE_{\bar{y}_B} = \sqrt{\frac{1}{B-1}\sum_{i=1}^{B}\left(\bar{y}_{B^{(i)}} - \bar{\bar{y}}_B\right)^2}$$

باستخدام مشاهدات العينة التي حجمها 15 مريض، والمبينة في جدول 1، تم أخذ عدد 1000 عينة بوتستراب، ذات الحجم 15، (مع الإحلال) من بين مشاهدات العينة التي لدينا في جدول 1. وتم حساب المتوسط الحسابي والوسيط لكل عينة كما هو مبين في جدول 2 الموضح أدناه.

قيمة الخطأ المعياري للمتوسط الحسابي للعينة باستخدام طريقة البوتستراب نتحصل عليها كالتالي:

$$SE_{\bar{y}_s} = \sqrt{\frac{1}{B-1}\sum_{i=1}^{B}\left(\bar{y}_{B^{(i)}} - \bar{\bar{y}}_B\right)^2} = \sqrt{\frac{1}{1000-1}\sum_{i=1}^{1000}\left(\bar{y}_{B^{(i)}} - 54.709\right)^2} = 2.8396$$

كما هو واضح، فإن الخطأ المعياري للمتوسط الحسابي للعينة بطريقة البوتستراب يساوي 2.8396 وهي قيمة قريبة جدًا من قيمة الخطأ المعياري للمتوسط الحسابي للعينة، 2.7879، والتي تم الحصول عليها مسبقًا، باستخدام الصيغة النظرية.

حساب الخطأ المعياري لوسيط العينة وفقاً لطريقة البوتستراب

على عكس المتوسط الحسابي للعينة، لا توجد صيغة بسيطة لحساب الخطأ المعياري لوسيط العينة. طريقة البوتستراب توفر طريقة بسيطة وسهلة لحساب الخطأ المعياري لوسيط العينة.

تمامًا كما هو الحال بالنسبة للمتوسط الحسابي للعينة، يتم حساب وسيط كل عينة من عينات البوتستراب التي تم الحصول عليها كما هو مبين في جدول 2.

الخطأ المعياري لوسيط العينة هو في الواقع الانحراف المعياري لعدد 1000 وسيط لعينات البوتستراب التي تم حسابها في جدول 2، ويتم الحصول عليه باستخدام الصيغة التالية:

$$SE_{\bar{y}_B^m} = \sqrt{\frac{1}{B-1}\sum_{i=1}^{B}\left(y^m_{B^{(i)}} - \bar{y}^m_B\right)^2} = \sqrt{\frac{1}{1000-1}\sum_{i=1}^{1000}\left(y^m_{B^{(i)}} - 54.048\right)^2} = 3.9491$$

جدول 2: مشاهدات عينات البوتستراب (التي عددها 1000) من $B^{(1)}$ إلى $B^{(1000)}$ والتي تم الحصول عليها من مشاهدات العينة المبينة في جدول 1.

رقم المشاهدة i	مشاهدات العينة	$B^{(1)}$	$B^{(2)}$	...	$B^{(1000)}$
1	61.415	46.304	60.318	...	55.249
2	60.318	75.694	46.733	...	58.373
3	75.694	49.869	46.733	...	46.304
4	55.994	56.9	55.249	...	46.733
5	46.448	61.415	55.249	...	46.733
6	56.9	46.733	35.188	...	58.373
7	46.304	46.304	46.733	...	46.448
8	49.869	74.575	35.188	...	58.373
9	45.322	55.994	75.694	...	45.322
10	55.249	35.188	60.318	...	55.249

11	58.373	74.575	35.188	...	46.448
12	35.188	35.188	46.448	...	45.322
13	74.575	49.869	46.733	...	75.694
14	46.733	46.304	74.575	...	58.373
15	57.411	55.994	35.188	...	55.994
$\bar{y}_{B^{(i)}}$ المتوسط الحسابي لعينة البوتستراب =		54.06	50.369	...	53.266
$\bar{\bar{y}}_B$ المتوسط الحسابي الكلي لعينات البوتستراب =		54.709			
$y^m_{B^{(i)}}$ الوسيط لعينة البوتستراب =		49.869	46.733	...	55.249
$\bar{y}^m_B$ المتوسط الحسابي الكلي لعينات البوتستراب =		54.048			

حساب الخطأ المعياري لوسيط العينة وفقاً لطريقة البوتستراب

على عكس المتوسط الحسابي للعينة، لا توجد صيغة بسيطة لحساب الخطأ المعياري لوسيط العينة. طريقة البوتستراب توفر طريقة بسيطة وسهلة لحساب الخطأ المعياري لوسيط العينة.

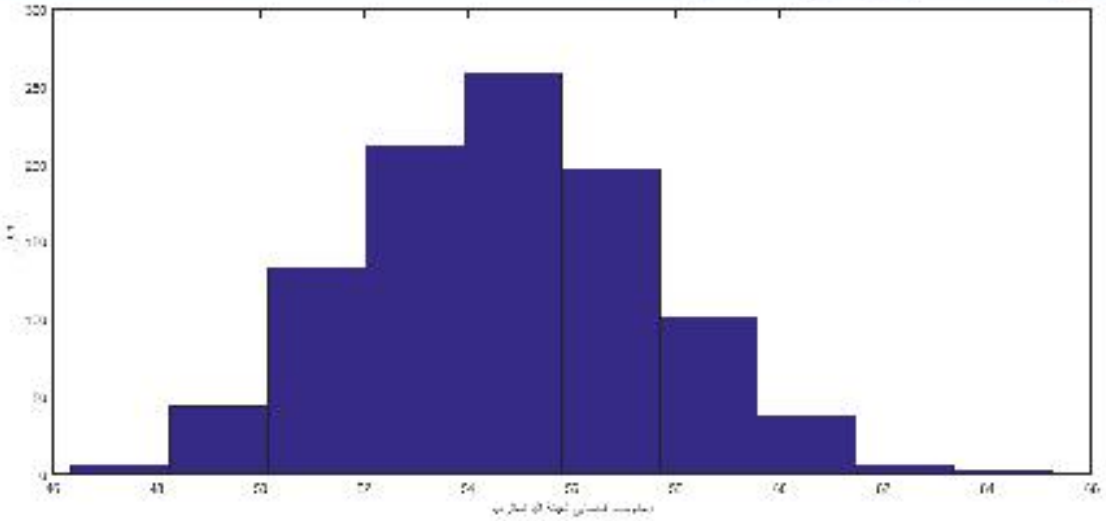
تمامًا كما هو الحال بالنسبة للمتوسط الحسابي للعينة، يتم حساب وسيط كل عينة من عينات البوتستراب التي تم الحصول عليها كما هو مبين في جدول 2.

الخطأ المعياري لوسيط العينة هو في الواقع الانحراف المعياري لعدد 1000 وسيط لعينات البوتستراب التي تم حسابها في جدول 2، ويتم الحصول عليه باستخدام الصيغة التالية:

$$SE_{\bar{y}^m_B} = \sqrt{\frac{1}{B-1}\sum_{i=1}^{B}\left(y^m_{B^{(i)}} - \bar{y}^m_B\right)^2} = \sqrt{\frac{1}{1000-1}\sum_{i=1}^{1000}\left(y^m_{B^{(i)}} - 54.048\right)^2} = 3.9491$$

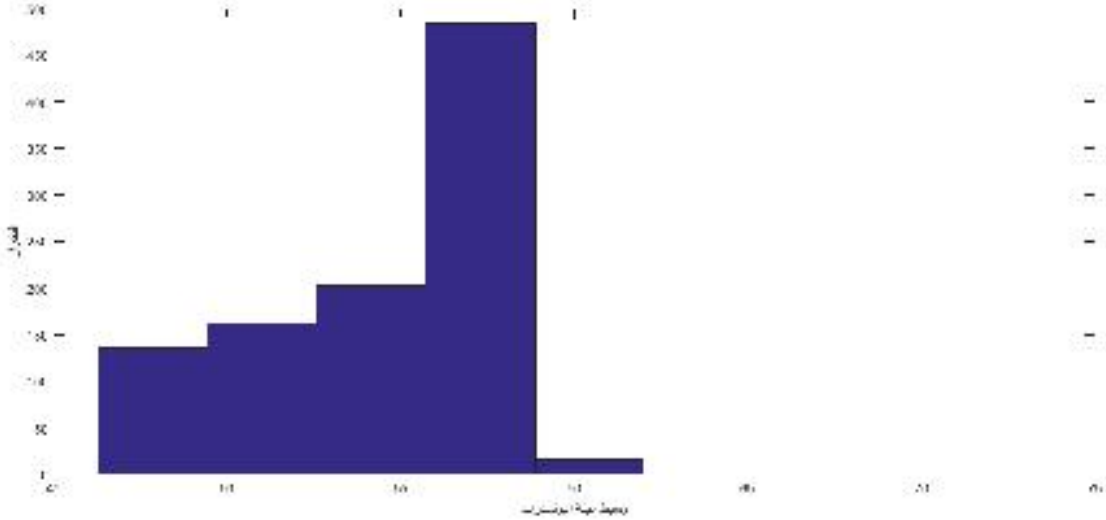
توزيع المعاينة للمتوسط الحسابي للعينة ولوسيط العينة

توزيع المعاينة للمتوسط الحسابي للعينة الذي تم الحصول عليه باستخدام طريقة البوتستراب من خلال استخدام بيانات جدول 2، والمتعلقة بقيم المتوسطات الحسابية لعينات البوتستراب، يمكننا التعبير عنه برسم شكل مدرج التكرار كما هو مبين في شكل 1 المبين أدناه.



شكل 1: مدرج التكرار لتوزيع المعاينة للمتوسط الحسابي للعينة باستخدام 1000 عينة بوتستراب تم توليدها من بيانات العينة المبينة في جدول 1 بعد حساب المتوسط الحسابي لكل عينة بوتستراب.

بالمثل توزيع المعاينة لوسيط العينة الذي تم الحصول عليه باستخدام طريقة البوتستراب من خلال استخدام بيانات جدول 2، والمتعلقة بقيم وسيطات عينات البوتستراب، يمكننا التعبير عنه برسم شكل مدرج التكرار كما هو مبين في شكل 2 المبين أدناه.



شكل 2: مدرج التكرار لتوزيع المعاينة لوسيط العينة باستخدام 1000 عينة بوتستراب تم توليدها من بيانات العينة المبينة في جدول 1 بعد حساب الوسيط لكل عينة بوتستراب.

استخدام ماتلاب (MATLAB) للحصول على عينات البوتستراب

تعتبر ماتلاب لغةً من لغات البرمجة عالية المستوى، كما أنها أيضاً عبارة عن بيئة تفاعلية يُعتمد عليها في تطوير الخوارزميات والقيام بتحليل البيانات، وتوفر للمستخدم مجموعةً من الأدوات والوظائف الرياضية التي تساعد في إيجاد حلولٍ سريعةٍ للغاية

بالاعتماد على جداول البيانات أو حتى لغات البرمجة التقليدية، ومن أبرزها جافا(JAVA, C++, C) ، وتزداد رقعة استخدامها بين أوساط مبرمجي أنظمة التحكم والبيولوجيا الحاسوبية وغيرها من المجالات. ويتميز برنامج ماتلاب بسهولة الاستخدام، حيث يتيح لمستخدميه الفرصة في الوصول إلى الحلول بالطرق الرياضية المألوفة. (للمزيد من الإطلاع حول برنامج ماتلاب أنظر: Jim and John, 2021).

للحصول على عينات البوتستراب باستخدام ماتلاب يجب أولاً إدخال مشاهدات العينة الأصلية، المبينة في جدول 1، لبرنامج ماتلاب باستخدام متجه يحمل الاسم data كما هو مبين في الشكل التالي:

```
>>
>> data=[61.413;60.318;75.694;55.994;46.418;36.8;46.304;48.869;47.322;55.249;58.273;35.188;74.573;46.733;37.411];
```

بعد ذلك نحدد عدد عينات البوتستراب، B، التي يجب أخذها ولتكن $B = 1000$ كما هو مبين في الشكل التالي:

```
>>
>> data=[61.413;60.318;75.694;55.994;46.418;36.8;46.304;48.869;47.322;55.249;58.273;35.188;74.573;46.733;37.411];
>> B=1000
fx >>
```

نقوم الآن بإنشاء دالة ماتلاب تحمل اسم bootsamples بواسطة ملف اسمه bootsamples.m بعد حفظه بداخل مجلد ماتلاب الفرعي الذي يحمل اسم bin.

الدالة bootsamples وظيفتها تكوين عدد B عينة بوتستراب يتم أخذها مع الإحلال من مشاهدات العينة الأصلية، المخزنة في المتجه data، بحيث يتم تخزين هذه العينات في مصفوفة تحتوي على n صف و B عمود، في مثالنا هذا $n = 15$ و $B = 1000$، هذه المصفوفة ستحمل اسم Bsamples.

الشكل التالي يبين محتويات الملف bootsamples.m:

```
1   function [Bsamples]=bootsamples(data,B)
2   rand('state', sum(B*clock));
3   n=length(data);
4   Bsamples=[];
5   sample=zeros(1,n);
6   for i=1:B
7   choose=round((n-1)*rand(1,n))+1;
8   for j=1:n
9   sample(j)=data(choose(j));
10  end;
11  Bsamples=[Bsamples;sample'];
12  end;
```

للحصول على عينات البوتستراب التي عددها 1000 باستخدام الدالة bootsamples نقوم داخل نافذة أوامر ماتلاب (Command Window) بكتابة [Bsamples]=bootsamples(data,B) وذلك بعد إدخال المتجه data والقيمة B كما هو موضح في الشكل التالي:

```
>> data=[61.415;60.318;75.694;55.994;46.448;56.9;46.304;40.869;45.322;55.249;58.373;35.188;74.575;46.733;57.411];
>> B=1000;
>> [Bsamples]=bootsamples(data,B);
>>
```

استخدام ماتلاب للحصول على الخطأ المعياري للمتوسط الحسابي والوسيط للعينة

للحصول على الخطأ المعياري للمتوسط الحسابي والوسيط للعينة باستخدام عينات البوتستراب التي تم الحصول عليها سابقاً من مشاهدات العينة الأصلية المخزنة في المتجه data نتبع الخطوات المبينة في الشكل التالي:

```
>> data=[61.415;60.318;75.694;55.994;46.448;56.9;46.304;40.869;45.322;55.249;58.373;35.188;74.575;46.733;57.411];
>> B=1000;
>> Bsamples=bootsamples(data,B);
>> Bmeans=mean(Bsamples);
>> SE_Bmean=std(Bmeans); % الخطأ المعياري لمتوسط العينة
SE_Bmean =

    2.8396

>> Bmedians=median(Bsamples);
>> SE_Bmedian=std(Bmedians); % الخطأ المعياري لوسيط العينة

SE_Bmedian =

    3.0491

fx >>
```

يمكننا الحصول على مدرج التكرار لتوزيع المعاينة لمتوسط العينة وتوزيع التكرار لوسيط العينة بطريقة البوتستراب من خلال نافذة الأوامر للماتلاب وذلك باستخدام الأمر hist كما هو مبين في الشكل التالي:

```
Command Window
>> hist(Bmeans,10); % شكل مدرج التكرار لتوزيع المعاينة للمتوسط الحسابي للعينة يتكون من 10 فئات
>> hist(Bmedians,8); % شكل مدرج التكرار لتوزيع المعاينة لوسيط العينة يتكون من 8 فئات
>>
```

فترات الثقة لمتوسط ووسيط المجتمع بطريقة البوتستراب

إذا كانت مشاهدات العينة التي حجمها n ومتوسطها الحسابي $\bar{x}$ قد سحبت عشوائياً من مجتمع إحصائي توزيعه معروف، التوزيع الطبيعي مثلاً، بمتوسط μ غير معلوم وانحراف معياري σ معلوم القيمة فإن فترة الثقة حول μ بمستوى ثقة $\gamma = 1-\alpha$ نتحصل عليها بالصيغة المعروفة التالية:

$$\left(\bar{x} - z_{1-\alpha/2}\sigma/\sqrt{n}, \bar{x} + z_{1-\alpha/2}\sigma/\sqrt{n}\right)$$

حيث $z_{1-\alpha/2}$ ترمز إلى قيمة المتغير العشوائي الطبيعي المعياري التي تترك إلى يسارها مساحة تحت المنحنى الطبيعي المعياري قدرها $(1-\alpha/2)$.

عندما تكون مشاهدات العينة مأخوذة من توزيع غير معروف لدينا ونود الحصول على فترة ثقة حول μ (أو أي معلمة مجتمع أخرى مجهولة مثل وسيط المجتمع) بمستوى ثقة $\gamma = 1-\alpha$ فإنه في الغالب لا توجد صيغة معروفة لإيجاد فترة الثقة المطلوبة. يمكننا في هذه الحالة استخدام طريقة البوتستراب لإيجاد فترة الثقة المطلوبة.

توجد عدة طرق لايجادة فترة الثقة باستخدام تقنية البوتستراب (Efron, 1982) أبسطها ما يعرف بطريقة فترة البوتستراب المئوية (Percentile Bootstrap Interval). تتلخص هذه الطريقة في الخطوات التالية:

1- توليد عدد B من عينات البوتستراب من مشاهدات العينة التي لدينا ذات الحجم n.
2- نقوم بحساب المتوسط الحسابي لكل عينة من عينات البوتستراب (في حالة رغبتنا في ايجاد فترة الثقة حول متوسط المجتمع μ).
3- نقوم باستخدام متوسطات عينات البوتستراب التي تم الحصول عليها في الخطوة الثانية في ايجاد الحد الأدنى لفترة الثقة المطلوبة وذلك بحساب المئين الذي رتبته $100\times(\alpha/2)$ ثم بعد ذلك نقوم بحساب المئين الذي رتبته $100\times(1-\alpha/2)$ للحصول على الحد الأعلى لفترة الثقة المطلوبة.

استخدام ماتلاب للحصول على فترات الثقة بطريقة البوتستراب

يمكننا من خلال نافذة أوامر ماتلاب الحصول على فترات ثقة لمتوسط ووسيط المجتمع باستخدام طريقة فترة البوتستراب المئوية (Percentile Bootstrap Interval) التي سبق شرحها في البند السابق من هذه الورقة وذلك باتباع الخطوات الموضحة في الشكل التالي بعد إنشاء ملف لدالة ماتلاب داخل مجلد ماتلاب الفرعي bin والذي يحمل اسم bootsamples.m.

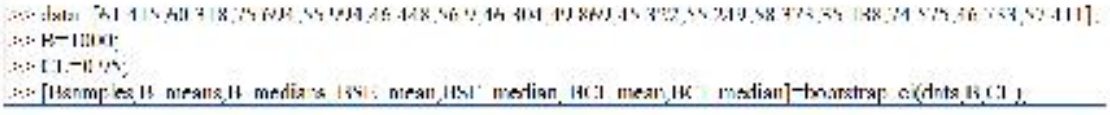
```
>> data=[61.415,60.518,75.694,55.991,46.446,56.9,49.304,49.869,45.322,55.219,58.375,35.168,74.575,46.733,57.411];
>> B=1000;
>> [Bsamples]=bootsample(data,B);
>> B_means=mean(Bsamples);
>> B_medians=median(Bsamples);
>> alpha=0.05;
>> BCI_mean=prctile(B_means,[alpha*100, 100-alpha*100]);

BCI_mean

   49.438   60.238

>> BCI_median=prctile(B_medians,[alpha*100, 100-alpha*100]);

BCI_median

   46.446   58.375
```

يمكننا إنشاء دالة ماتلاب تلخص الخطوات السابقة تحمل اسم bootstrap_ci وحفظها في ملف داخل مجلد ماتلاب الفرعي bin يحمل اسم bootstrap_ci.m. مدخلات هذه الدالة ستكون مشاهدات العينة الأصلية المسحوبة من المجتمع الإحصائي، data، عدد عينات البوتستراب،B، مستوى الثقة المرغوب، CL. مخرجات الدالة bootstrap_ci ستكون عينات البوتستراب، Bsamples، المتوسطات الحسابية لعينات البوتستراب، B_means، وسيطات عينات البوتستراب،B_medians، الخطأ المعياري للمتوسط الحسابي للعينة باستخدم طريقة البوتستراب، BSE_mean، الخطأ المعياري لوسيط العينة باستخدم طريقة

البوتستراب، BSE_median، فترة ثقة البوتستراب المئوية لمتوسط المجتمع، BCI_mean، وفترة ثقة البوتستراب المئوية لوسيط المجتمع، BCI_median.

الدالة boostrap_ci تم عرضها في ملحق هذه الورقة كما تم توضيحها في الشكل التالي:

لتنفيذ دالة ماتلاب boostrap_ci نتبع داخل نافذة أوامر ماتلاب الخطوات المبينة في الشكل التالي:



بعد إجراء الخطوات السابقة والضغط على زر Enter تتحصل داخل نافذة أوامر ماتلاب على محتويات المصفوفات التالية:

Bsamples = مصفوفة عينات البوتستراب بحيث كل عمود في المصفوفة يمثل مشاهدات عينة بوتستراب.

B_means = متجه المتوسطات الحسابية لعينات البوتستراب الظاهرة في المصفوفة Bsamples.

B_medians = متجه وسيطات عينات البوتستراب.

BSE_mean = قيمة الخطأ المعياري للمتوسط الحسابي للعينة باستخدم طريقة البوتستراب.

BSE_median = قيمة الخطأ المعياري لوسيط العينة باستخدم طريقة البوتستراب.

BCI_mean = فترة ثقة البوتستراب المئوية لمتوسط المجتمع.

BCI_median = فترة ثقة البوتستراب المئوية لوسيط المجتمع.

الاستنتاج

تعد طريقة البوتستراب مفيدة جداً وبشكل خاص لتقدير دقة التقديرعندما لا توجد صيغة معروفة مرتبطة بالخطأ المعياري للإحصاءة. يمكن تطبيق طريقة البوتستراب لمعرفة دقة التقدير لأي مقدر. عندما يتوفر لدينا حاسوب بمواصفات عالية فإنه يمكننا استخدامه في توليد عينات البوتستراب بسهولة، وبالتالي الحصول على توزيع المعاينة لأي إحصاءة. يمكننا أيضاً الحصول على فترة ثقة لمعلمة المجتمع الإحصائي المجهولة بمستوى ثقة $\gamma = 1-\alpha$ وذلك باستخدام طريقة فترة بوتستراب المئوية (Percentile Bootstrap Interval) التي سبق شرحها.

تقدم هذه الورقة لمحة مختصرة من العديد من تطبيقات طريقة البوتستراب. يمكن للمهتمين من البحاث في مجال الإحصاء الحيوي وغيرها من المجالات الراغبين في الحصول على فهم أوسع لموضوع البوتستراب الرجوع إلى كتاب إيفرون وتبشيراني (Efron and Tibshirani, 1993).

ملحق

دالة ماتلاب التي تحمل اسم bootstrap_ci، والتي يجب حفظها في ملف داخل مجلد ماتلاب الفرعي bin يحمل اسم bootstrap_ci.m، والتي وظيفتها الأساسية هي ايجاد فترات ثقة لمتوسط ووسيط المجتمع باستخدام طريقة فترة البوتستراب المئوية (Percentile Bootstrap Interval).

```
function [Bsamples,B_means,B_medians, BSE_mean,BSE_median, …
BCI_mean,BCI_median]=bootstrap_ci(data,B,CL);

rand('state', sum(B*clock));
n=length(data);
```

```
Bsamples=[];
sample=zeros(1,n);

for i=1:B
choose=round(((n-1)*rand(1,n))+1);
for j=1:n
sample(j)=data(choose(j));
end;
Bsamples=[Bsamples,sample'];
end;

B_means=mean(Bsamples);
B_medians=median(Bsamples);

BSE_mean=std(B_means);
BSE_median=std(B_medians);

alpha=1-CL;
BCI_mean=prctile(B_means,[alpha*100/2  100-alpha*100/2]);
BCI_median=prctile(B_medians,[alpha*100/2  100-alpha*100/2]);
```

المراجع

Efron, B. (1979). Bootstrap methods: another look at the jackknife. *Annals of Statistics* 7, 1-26.

Efron, B. (1982). "The Jackknife, the bootstrap, and other resampling plans," in *CBMS-NSF Regional Conference Series in Applied Mathematics, Monograph 38* (Philadelphia, PA: SIAM).

Efron, B. (1987). Better bootstrap confidence intervals. *J. Am. Stat. Assoc.* 82, 171–185.

Efron, B., and Tibshirani, R. J. (1993). *An Introduction to the Bootstrap*. New York, NY: Chapman and Hall.

Jim S., and John P. M. (2021). *MATLAB For Dummies*. New Jersey, John Wiley & Sons, Inc, Hoboken,

1- معاقلت ،القاهرة :2000