

استعمال طريقتي التقدير الإمكان الأعظم المعدلة والامكان التجريبية  
الحصينة في تقدير علاقة الانحدار الخطي العشوائية بين حجم كريات الدم  
الحمراء وسكر الدم

أ.د احمد شاكر محمد طاهر المتولي  
قسم الاحصاء / جامعة المستنصرية

الباحثة امانى عماد لعبي  
قسم الاحصاء / جامعة المستنصرية

تاريخ استلام البحث: 2023/07/29

تاريخ قبول البحث: 2023/09/01

نشر البحث في العدد الثالث والعشرين: حزيران / يونيو 2024

رمز التصنيف ديوي / النسخة الالكترونية (Online): 2522-64X/512.943

رمز التصنيف ديوي / النسخة الورقية (Print): 2519-948X/512.943

## استعمال طريقتي التقدير الإمكان الأعظم المعدلة والامكان التجريبية الحصينة في تقدير علاقة الانحدار الخطي العشوائية بين حجم كريات الدم الحمراء وسكر الدم

أ.د احمد شاكر محمد طاهر المتولي  
قسم الاحصاء / جامعة المستنصرية

الباحثة اماني عماد لعبي  
قسم الاحصاء / جامعة المستنصرية

### المستخلص

افتراضات نموذج الانحدار الخطي الخاصة بالمتغير التفسيري تنص على ان ذلك المتغير يكون ذو قيمة ثابتة للعينات المتكررة ومستقل عن حدود الخطأ العشوائي للنموذج. هذه الافتراضات من الناحية العملية قد تكون غير متحققة، اذ يكون المتغير التفسيري عشوائيا بمعنى أن قيمه تحدث باحتمالية معينة، في هذه الحالة يطلق على هذا النموذج بأنموذج الانحدار الخطي العشوائي الذي له استعمالات واسعة في تمثيل علاقة الانحدار بين المتغيرات في مختلف الظواهر الاقتصادية او الحياتية، ونتيجة لعدم تحقق الافتراضات المذكورة فإن مقدرات المربعات الصغرى تفقد بعض او جميع خواصها المثلى وذلك حسب طبيعة العلاقة بين المتغير التفسيري العشوائي وحدود الخطأ العشوائي. لذا يتم اللجوء الى استعمال طرائق بديلة تمكننا من الحصول على مقدرات تتصف بالخصائص المرغوبة، تعد طرائق التقدير الحصينة بديل مناسب لطريقة المربعات الصغرى الاعتيادية، في هذا البحث تم استعمال طريقتين حصينتين هما طريقة الإمكان الأعظم المعدلة (MMLE) و طريقة الإمكان التجريبية الحصينة (REL). الطريقة الأولى تم استعمالها سابقا من قبل العديد من الباحثين في تقدير معاملات انموذج الانحدار العشوائي، أما الطريقة الثانية فقد تم توظيفها من قبلنا اذ لم تستعمل سابقا في تقدير معاملات الانموذج موضوع البحث وانما استعملت في حالة معالجة المشاكل القياسية أو عندما تحتوي البيانات على قيمة شاذة أو متطرفة. يهدف البحث الى استعمال طريقتي التقدير المذكورتين في تقدير علاقة الانحدار الخطي العشوائية بين حجم كريات الدم الحمراء (PCV) كمتغير استجابة وسكر الدم (RBS) كمتغير توضيحي عشوائي بالاعتماد على بيانات عينة عشوائية مؤلفة من 30 مريض من المصابين بأمراض القلب.

**الكلمات المفتاحية:** المتغير التفسيري العشوائي، انموذج الانحدار الخطي العشوائي، طريقة الامكان الاعظم المعدلة (MMLE)، طريقة الامكان التجريبي الحصينة (REL).

### Abstract:

The assumptions of the linear regression model regarding the explanatory variable states that this variable has fixed values for repeated samples and is independent of the random error terms of the model. These assumptions in practice may not be realized, as the explanatory variable is random in the sense that its values occur with a certain probability. In this case, this model is called the random linear regression model, which has wide used in representing the regression relationship between variables in various economic or life phenomena. As a result of the non-fulfillment of the mentioned assumptions, the least

squares estimators lose some or all of their optimal properties, depending on the nature of the relationship between the random explanatory variable and the random error terms. Therefore, alternative methods are resorted to that enable us to obtain estimators that have the desired characteristics. Robust estimation methods are a suitable alternative to the ordinary least squares method. In this research, two robust methods were used, namely the Modified Maximum likelihood (MML) and Robust Empirical Likelihood (REL). The first method was previously used by many researchers to estimate the coefficients of the random regression model, as for the second method, it was employed by us, as it was not used previously in estimating the coefficients of the model under research, but rather it was used in the case of addressing standard problems or when the data contains outliers or extreme values. The research aims to use the two mentioned estimation methods to estimate the random linear regression relationship between red blood cell volume (PCV) as a response variable and blood sugar (RBS) as a random explanatory variable based on the data of a random sample of 30 patients with heart disease.

**Keywords:** Random Explanatory Variable, Random Linear Regression model, Modified Maximum likelihood Method, Robust Empirical Likelihood Method.

## 1- المقدمة:

من الفرضيات الأساسية لأنموذج الانحدار الخطي هي ان المتغيرات التفسيرية تكون ثابتة في العينات المتكررة، غير ان هذا الافتراض قد يكون من الناحية العملية غير متحقق، كما هو الحال في اغلب الظواهر الاقتصادية والحياتية، اذ تكون المتغيرات التفسيرية عشوائية، بمعنى قيم تلك المتغيرات ليست ثابتة، لذلك فإن القيم المختلفة للمتغيرات التفسيرية تحدث باحتمالية معينة. في مثل هذه الحالات يتم تحديد قيم المتغير التفسيري جنباً إلى جنب مع قيم متغير الاستجابة كنتيجة لبعض آليات الاحتمالات، بدلاً من التحكم بقيمها من قبل القائمين بعمل التجربة.

في حال اعتماد طريقة المربعات الصغرى لتقدير معاملات انموذج الانحدار بمتغيرات تفسيرية عشوائية فان تقديرات تلك الطريقة لمعاملات الانموذج تفقد بعض او جميع خواصها المثلى، وذلك يعتمد على نوع العلاقة بين المتغير أو المتغيرات التفسيري العشوائي وحدود الخطأ العشوائي. في هذه الحالة يجب البحث عن طرائق بديلة عن طريقة المربعات الصغرى تتميز بالكفاءة ولا تتأثر كثيراً بالانحراف عن افتراضات انموذج الانحدار الخطي ومن تلك الطرائق هي الطرائق الحصينة، ومن تلك الطرائق، طريقة الامكان الاعظم المعدلة ( MMLE ) وطريقة الاحتمال التجريبي الحصينة ( REL ).

أول من تعامل مع انموذج الانحدار بمتغيرات تفسيرية عشوائية هو الباحث Kerridge في عام 1967 ، [4] ، اذ افترض ان قيم المتغيرات التفسيرية المناظرة لقيم متغير الاستجابة مسحوبة



لعينة عشوائية بحجم  $n$  فإن  $Y_i$  هو متغير عشوائي ذو قيمة مستقلة يمثل متغير الاستجابة والذي يفترض ان يتوزع وفق التوزيع الطبيعي بمتوسط يساوي  $E(Y_i)$  وتباين  $\sigma^2$ ،  $X_{ij}$ ، اذ أن  $j = 1, 2, \dots, p$ ،  $i = 1, 2, \dots, n$ ، تمثل  $p$  من المتغيرات التفسيرية والذي يفترض ان تكون ثابتة للعينات المتكررة،  $u_i$  هي الأخطاء العشوائية للأنموذج والتي يفترض ان تكون قيمها مستقلة وتتوزع وفق التوزيع الطبيعي بمتوسط يساوي 0 وتباين ثابت يساوي  $\sigma^2$ . في الكثير من التطبيقات العملية المتغير التفسيري لا يكون ثابت وانما متغير عشوائي وفي هذه الحالة فإن انموذج الانحدار وفق المعادلة (1) يسمى بأنموذج الانحدار العشوائي، [2].

لغرض التبسيط سيتم التطرق الى حالة وجود متغير تفسيري واحد أي انموذج احذار خطي بسيط. في حالة كون المتغير التفسيري عشوائي فان القيم المختلفة له تحدث وفق احتمالات معينة، بمعنى اخر انها تتحدد مع قيم متغير الاستجابة المناظرة لها وفق توزيع احتمالي معين بدلا من كونها مسيطر عليها من قبل القائم بعملية التجربة. المتغير التفسيري العشوائي يمكن ان يتبع التوزيع الطبيعي أي نفس التوزيع الاحتمالي لمتغير الاستجابة وفق افتراضات انموذج الانحدار، أو تتحد قيمه وفق أي توزيع احتمالي اخر، [2].

بافتراض ان التوزيع الاحتمالي للمتغير التفسيري  $X$  هو توزيع القيمة المتطرفة وفق الدالة الاحتمالية الآتية، [9]:

$$f(x) = \frac{1}{\theta} \exp \left[ - \left( \frac{x - \alpha}{\theta} \right) - \exp \left( - \frac{x - \alpha}{\theta} \right) \right], \quad -\infty < x < \infty \quad \dots \dots (2)$$

اذ أن،  $\alpha$ : تمثل معلمة الموقع Location parameter،  $\theta$ : تمثل معلمة القياس Scale parameter. وعندما تكون  $\alpha = 0$  و  $\theta = 1$  نحصل على توزيع القيمة المتطرفة القياسي. اما التوزيع الاحتمالي الشرطي لمتغير الاستجابة  $Y$  فيكون التوزيع الطبيعي وفق الدالة الاحتمالية الشرطية الآتية:

$$f(Y|X = x) = [2\pi\sigma^2(1 - \rho^2)]^{-\frac{1}{2}} \exp \left[ - \frac{1}{2\sigma^2(1 - \rho^2)} \left( y - \beta_0 - \rho \frac{\sigma}{\theta} (x - \alpha) \right)^2 \right] \quad -\infty < y < \infty \quad (3)$$

اذ أن،  $\beta_0 \in R$ ،  $\beta_1 = \rho \frac{\sigma}{\theta}$ ،  $\sigma > 0$ ،  $-1 < \rho < 1$ ،  $u = y - \beta_0 - \rho \frac{\sigma}{\theta} (x - \alpha)$  اذ ان  $u$  تمثل حدود الخطأ العشوائي التي تتبع التوزيع الطبيعي بمتوسط شرطي يساوي الصفر ( $E(u|X = x) = 0$ ) وتباين شرطي ثابت  $Var(u|X) = \sigma^2(1 - \rho^2)$  اذ أن، [2]:

$$E(Y|X = x) = \beta_0 - \rho \frac{\sigma}{\theta} (x - \alpha) \quad \dots \dots (4)$$

### 3- طرائق التقدير الحصينة لمعاملات انموذج الانحدار الخطي العام بمتغير تفسيري عشوائي:

تتضمن هذه الفقرة عرض لطريقتي التقدير الحصينتين موضوع البحث والتي اعتمدت في تقدير انموذج الانحدار الخطي البسيط بمتغير تفسيري عشوائي.

#### 1-3 طريقة الامكان الاعظم المعدلة Modified Maximum Likelihood Method (MMLE)

عملية تقدير معاملات انموذج الانحدار بمتغير تفسيري عشوائي وفق المعادلة (1) باستعمال طريقة الإمكان الأعظم وبالاعتماد على الدالتين الاحتمالية (2) و (3)، تكون صعبة وغير مجدية،

اذ ان المعادلات الناتجة من اشتقاق دالة الإمكان ليس لها حلول صريحة وعملية حلها بالطرق التكرارية يكون محفوف بالصعوبات لعدم تحقق التقارب بشكل سريع او التقارب لقيم خاطئة علاوة عن ان لها جذور متعددة. لذا يتم استعمال طريقة الإمكان الأعظم المعدلة والمقترحة من قبل Tiku و Vaughan، [9].

دالة الكثافة الاحتمالية المشتركة لمتغير الاستجابة  $Y$  والمتغير التفسيري  $X$  هي كالآتي:

$$f(y, x) = f(x)f(y|x) = \left(\frac{1}{\theta}\right)\left(\frac{1}{\sigma}\right) (1 - \rho^2)^{-\frac{1}{2}} \exp\left[-\left(\frac{x-\alpha}{\theta}\right) - \exp\left(-\frac{x-\alpha}{\theta}\right)\right] \exp\left[-\frac{1}{2\sigma^2(1-\rho^2)}(u_i)^2\right] \dots (5)$$

وعليه فإن دالة الإمكان تكون وفق الصيغة الآتية:

$$L = (\theta)^{-n}(\sigma)^{-n}(1 - \rho^2)^{-\frac{n}{2}} \exp\left[-\sum_{i=1}^n \left(\left(\frac{x_i-\alpha}{\theta}\right) + \exp\left(-\frac{x_i-\alpha}{\theta}\right)\right) - \frac{1}{2\sigma^2(1-\rho^2)} \sum_{i=1}^n u_i^2\right] \dots (6)$$

وان لوغاريتم دالة الإمكان يكون كما في الصيغة الآتية، بافتراض أن  $(z_i = \left(\frac{x_i-\alpha}{\theta}\right))$

$\ln L$

$$= -n \ln \theta - n \ln \sigma - \frac{n}{2} \ln(1 - \rho^2) - \sum_{i=1}^n (z_i + \exp(-z_i)) - \frac{1}{2\sigma^2(1-\rho^2)} \sum_{i=1}^n u_i^2 \dots (7)$$

بالاشتقاق الجزئي للمعادلة (7) بالنسبة للمعاملات  $(\theta, \beta_1, \beta_0, \sigma, \alpha, \rho)$  ومساواتها بالصفر نحصل على معادلات الإمكان الأعظم التالية التي بحلها نحصل على تقديرات تلك المعاملات:

$$\frac{\partial \ln L}{\partial \alpha} = \frac{n}{\theta} - \frac{1}{\theta} \sum_{i=1}^n \exp(-z_i) - \frac{\rho}{\theta \sigma(1-\rho^2)} \sum_{i=1}^n u_i = 0 \dots (8)$$

$$\frac{\partial \ln L}{\partial \theta} = -\frac{n}{\theta} - \frac{1}{\theta} \sum_{i=1}^n z_i \exp(-z_i) + \frac{1}{\theta} \sum_{i=1}^n z_i - \frac{\rho}{\theta \sigma(1-\rho^2)} \sum_{i=1}^n z_i u_i = 0 \dots (9)$$

$$\frac{\partial \ln L}{\partial \beta_0} = \frac{1}{\sigma^2(1-\rho^2)} \sum_{i=1}^n u_i = 0 \dots (10)$$

$$\frac{\partial \ln L}{\partial \sigma} = -\frac{n}{\sigma} + \frac{1}{\sigma^3(1-\rho^2)} \sum_{i=1}^n u_i^2 + \frac{\rho}{\sigma^2(1-\rho^2)} \sum_{i=1}^n z_i u_i = 0 \dots (11)$$

$$\frac{\partial \ln L}{\partial \rho} = \frac{n\rho}{(1-\rho^2)} - \frac{\rho}{\sigma^2(1-\rho^2)^2} \sum_{i=1}^n u_i^2 + \frac{1}{\sigma(1-\rho^2)} \sum_{i=1}^n z_i u_i = 0 \dots (12)$$

$$\frac{\partial \ln L}{\partial \beta_1} = \frac{\theta}{\sigma^2(1-\rho^2)} \sum_{i=1}^n z_i u_i = 0 \dots (13)$$

المعادلتين (8, 9) تتضمن حدود صعبة الحل مما يؤدي الى ان مجموعة المعادلات (8-13) ليس لها حلول صريحة والتي يجب حلها باستعمال الطرائق التكرارية, لذا فان تقديرات الإمكان الأعظم تكون صعبة الحصول. لمعالجة هذه المشكلة نلجأ الى تحويل المتغير التفسيري الى الإحصاءات المرتبة  $x_{(i)}$ ,  $1 \leq i \leq n$ , من خلال ترتيب قيمه تصاعديا وكالآتي:

$$x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n)} \dots (14)$$

وبالاعتماد على القيم المرتبة للمتغير التفسيري فان  $y_{[i]}$  تمثل قيم متغير الاستجابة المناظرة للقيم المرتبة  $x_{(i)}$ , كذلك الصيغة القياسية  $z_i$  والاختفاء العشوائية  $u_i$  يعاد كتابتها كالآتي:

$$z_{(i)} = \frac{x_{(i)} - \alpha}{\theta} \dots \dots \dots (15)$$

$$u_{[i]} = y_{[i]} - \beta_0 - \rho \frac{\sigma}{\theta} (x_{(i)} - \alpha) \dots \dots \dots (16)$$

بما ان المجموع الكلي للأخطاء لا يتأثر بترتيب القيم مما يؤدي الى ان  $\sum_{i=1}^n u_{[i]} = 0$  وبالاعتماد على ما تقدم يعاد كتابة المعادلات (13-8) كالآتي:

$$\frac{\partial \ln L^*}{\partial \alpha} \equiv \frac{n}{\theta} - \frac{1}{\theta} \sum_{i=1}^n \exp(-z_{(i)}) = 0 \dots \dots \dots (17)$$

$$\frac{\partial \ln L^*}{\partial \theta} \equiv -\frac{n}{\theta} + \frac{1}{\theta} \sum_{i=1}^n z_{(i)} - \frac{1}{\theta} \sum_{i=1}^n z_{(i)} \exp(-z_{(i)}) = 0 \dots \dots \dots (18)$$

$$\frac{\partial \ln L^*}{\partial \beta_0} \equiv \frac{1}{\sigma^2(1-\rho^2)} \sum_{i=1}^n u_{[i]} = 0 \dots \dots \dots (19)$$

$$\frac{\partial \ln L^*}{\partial \sigma} \equiv -\frac{n}{\sigma} + \frac{1}{\sigma^3(1-\rho^2)} \sum_{i=1}^n u_{[i]}^2 = 0 \dots \dots \dots (20)$$

$$\frac{\partial \ln L^*}{\partial \rho} \equiv \frac{n\rho}{(1-\rho^2)} - \frac{\rho}{\sigma^2(1-\rho^2)^2} \sum_{i=1}^n u_{[i]}^2 = 0 \dots \dots \dots (21)$$

$$\frac{\partial \ln L^*}{\partial \beta_1} \equiv \frac{\theta}{\sigma^2(1-\rho^2)} \sum_{i=1}^n z_{(i)} u_{[i]} = 0 \dots \dots \dots (22)$$

المعادلات (17-22) تمثل معادلات الإمكان الأعظم المعدلة. لتحقيق الصيغة الخطية ل  $\exp(-z_{(i)})$  يمكن استعمال اول حدين من متسلسلة تايلر بافتراض ان  $t_{(i)} = E[z_{(i)}]$  كالآتي:

$$\exp(-z_{(i)}) \cong \exp(-t_{(i)}) + [z_{(i)} - t_{(i)}] \left\{ \frac{d}{dz} e^{-z} \right\}_{z=t_{(i)}} \dots \dots \dots (23)$$

والتي يمكن إعادة كتابتها كالآتي:

$$\exp(-z_{(i)}) = a_i + b_i z_{(i)} \quad , \quad 1 \leq i \leq n \quad (0 \dots \dots \dots (24)$$

اذ أن:

$$a_i = e^{-t_{(i)}} (1 + t_{(i)}) \dots \dots \dots (25)$$

$$b_i = - \left\{ \frac{d}{dz} e^{-z} \right\}_{z=t_{(i)}} \dots \dots \dots (26)$$

مع ملاحظة أن  $b_i \geq 0$  لكل  $i = 1, 2, \dots, n$ . المعادلتين (17) و (18) يعاد كتابتهما بعد تعويض المعادلة (24) في كل منهما لتصبحا:

$$\frac{\partial \ln L^*}{\partial \alpha} \equiv \frac{n}{\theta} - \frac{1}{\theta} \sum_{i=1}^n (a_i + b_i z_{(i)}) = 0 \dots \dots \dots (27)$$

$$\frac{\partial \ln L^*}{\partial \theta} \equiv -\frac{n}{\theta} + \frac{1}{\theta} \sum_{i=1}^n z_{(i)} - \frac{1}{\theta} \sum_{i=1}^n z_{(i)} (a_i + b_i z_{(i)}) = 0 \dots \dots \dots (28)$$

بحل المعادلات (27), (28), و (19) - (22) نحصل على ما يسمى بتقديرات الإمكان الأعظم المعدلة والتي تتصف في حالة العينات الكبيرة بأنها غير متحيزة ولها تباين يساوي اقل تباين محدود MVB, وانها كفؤة بالكامل fully efficient, وفي حالة العينات الصغيرة بانها على الاغلب كفؤة بالكامل, كما انها دوال صريحة بمشاهدات العينة وسهلة الحساب. من المعادلة (27) وبعد التعويض عن  $z_{(i)}$  نحصل على تقدير المعلمة  $\alpha$  كالآتي:

$$\frac{n}{\theta} - \frac{1}{\theta} \sum_{i=1}^n \left( \frac{\theta a_i + b_i (x_{(i)} - \alpha)}{\theta} \right) = 0 \dots \dots \dots (29)$$

وبالقيام ببعض العمليات الرياضية نحصل على تقدير  $\alpha$

$$\hat{\alpha} = \frac{\sum_{i=1}^n b_i x_{(i)} + \sum_{i=1}^n (1 - a_i) \hat{\theta}}{\sum_{i=1}^n b_i} \dots \dots \dots (30)$$

التي يمكن إعادة كتابتها كالآتي:

$$\hat{\alpha} = K + D\hat{\theta} \quad \dots\dots\dots (31)$$

اذ ان:

$$K = \frac{\sum_{i=1}^n b_i x_{(i)}}{\sum_{i=1}^n b_i} \quad \text{and} \quad D = \frac{1}{\sum_{i=1}^n b_i} \sum_{i=1}^n (1 - a_i)$$

تقدير المعلمة  $\theta$  نحصل عليه من المعادلة (28) بعد التعويض عن قيمة  $\mathbf{z}_{(i)}$  واجراء بعض العمليات الحسابية وكما مبين ادناه:

$$\sum_{i=1}^n b_i (x_{(i)} - \hat{\alpha})^2 - [\hat{\theta}(n\hat{\theta} + \sum_{i=1}^n (x_{(i)} - \hat{\alpha})) - \hat{\theta} \sum_{i=1}^n a_i (x_{(i)} - \hat{\alpha}) \dots\dots\dots (32)$$

المعادلة (32) يمكن إعادة كتابتها بعد التعويض عن  $\hat{\alpha}$  بما يساويها وفق المعادلة (31) كالآتي:  
 $n\hat{\theta}^2 - \hat{\theta} \sum_{i=1}^n (x_{(i)} - K) (1 - a_i) + \sum_{i=1}^n b_i (x_{(i)} - K)^2 = 0 \quad \dots\dots\dots (33)$

المعادلة (32) يمكن كتابتها بصيغة أبسط كالآتي:

$$A\hat{\theta}^2 - B\hat{\theta} + C = 0 \quad \dots\dots\dots (34)$$

بافتراض ان:  $A = n$ ,  $C = \sum_{i=1}^n b_i (x_{(i)} - k)^2$ ,  $B = \sum_{i=1}^n (x_{(i)} - k) (1 - a_i)$ :  
 فأن الحل للمعادلة (34) يمثل تقدير المعلمة  $\theta$ , أي أن:

$$\hat{\theta} = \frac{B + \sqrt{B^2 + 4aC}}{2a} \quad \dots\dots\dots (35)$$

وللحصول على تقدير المعلمة  $\beta_0$  نعوض عن قيمة  $u_{[i]}$  في المعادلة (19) واجراء بعض العمليات الحسابية كالآتي:

$$\frac{1}{\sigma^2(1-\rho^2)} \sum_{i=1}^n y_{[i]} - \hat{\beta}_0 - \rho \frac{\sigma}{\hat{\theta}} (x_{(i)} - \alpha) = 0$$

$$\hat{\beta}_0 = \bar{y} - \hat{\rho} \frac{\hat{\sigma}}{\hat{\theta}} (\bar{x} - \hat{\alpha}) \quad \dots\dots\dots (36)$$

لإيجاد تقدير  $\sigma$  نعيد كتابة  $u_{[i]}$  بالتعويض عن  $\hat{\beta}_0$  في المعادلة (16) بما يساويها بموجب المعادلة (36) وكالآتي:

$$u_{[i]} = (y_{[i]} - \bar{y}) - \hat{\rho} \left( \frac{\hat{\sigma}}{\hat{\theta}} \right) (x_{(i)} - \bar{x}) \quad \dots\dots\dots (37)$$

بتربيع الطرفين وأخذ المجموع لكل قيم حجم العينة نحصل على:

$$\sum_{i=1}^n u_{[i]}^2 = \sum_{i=1}^n \left[ (y_{[i]} - \bar{y}) - \hat{\rho} \left( \frac{\hat{\sigma}}{\hat{\theta}} \right) (x_{(i)} - \bar{x}) \right]^2 \quad \dots\dots\dots (38)$$

بتعويض المعادلة (38) في المعادلة (20) واجراء بعض العمليات الحسابية نحصل على المعادلة الآتية:

$$\sum_{i=1}^n (y_{[i]} - \bar{y})^2 - 2 \sum_{i=1}^n (y_{[i]} - \bar{y}) (x_{(i)} - \bar{x}) \hat{\rho} \frac{\hat{\sigma}}{\hat{\theta}} + \frac{\hat{\rho}^2 \hat{\sigma}^2}{\hat{\theta}^2} \sum_{i=1}^n (x_{(i)} - \bar{x})^2 - n\sigma^2 + n\sigma^2 \rho^2 = 0 \quad \dots\dots\dots (39)$$

المعادلة الأخيرة يمكن إعادة كتابتها بدلالة التباين والتباين المشترك للعينة ولكل من  $y_{[i]}$  و  $x_{(i)}$  وكالآتي:

$$ns^2_y - 2ns_{xy} \frac{s_{xy} \hat{\theta}}{s^2_x \hat{\sigma}} \cdot \frac{\hat{\sigma}}{\hat{\theta}} + \frac{s^2_{xy} \hat{\theta}^2}{s^4_x \hat{\sigma}^2} \cdot \frac{\hat{\sigma}^2}{\hat{\theta}^2} ns^2_x - n\sigma^2 + n\sigma^2 \frac{s^2_{xy} \hat{\theta}^2}{s^4_x \hat{\sigma}^2} = 0 \quad \dots\dots\dots (40)$$

والحل للمعادلة الأخيرة يمثل تقدير  $\sigma$  وكما مبين ادناه:

$$\hat{\sigma} = \left[ s^2_y + \frac{s^2_{xy}}{s^2_x} \left( \frac{\hat{\theta}^2}{s^2_x} - 1 \right) \right]^{\frac{1}{2}} \dots \dots \dots (41)$$

وللحصول على تقدير  $\beta_1$  نعوض عن قيمة  $u_{[i]}$  و  $z_{(i)}$  في المعادلة رقم (22) واجراء بعض العمليات الحسابية لنحصل على المعادلة الآتية:

$$\sum_{i=1}^n (x_{(i)} - \alpha)(y_{[i]} - \hat{\beta}_0) - (x_{(i)} - \alpha)^2 \hat{\beta}_1 = 0 \dots \dots \dots (42)$$

وبحل تلك المعادلة نحصل على تقدير  $\beta_1$  :

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n (x_{(i)} - \hat{\alpha})(y_{[i]} - \hat{\beta}_0)}{(x_{(i)} - \hat{\alpha})^2} \dots \dots \dots (43)$$

تقدير معلمة الارتباط  $\rho$  نحصل عليه من المعادلة (21) بعد التعويض عن مجموع مربعات الأخطاء العشوائية  $u_{[i]}$  , اذ يمكن إعادة كتابتها بدلالة التباين والتباين المشترك للعينة ولكل من  $x_{(i)}$  و  $y_{[i]}$  بعد اجراء بعض العمليات الحسابية لنحصل على الصيغة الآتية:

$$\left[ \left( \frac{\hat{\sigma}}{\hat{\theta}} \right)^2 S_x^2 + \hat{\sigma}^2 \right] \hat{\rho}^2 - 2 \left[ \left( \frac{\hat{\sigma}}{\hat{\theta}} \right) S_{xy} \right] \hat{\rho} - \left[ \frac{S_{xy}^2}{S_x^2} \left( \frac{\hat{\theta}}{S_x^2} - 1 \right) \right] = 0 \dots \dots \dots (44)$$

والتي يمكن إعادة كتابتها كالآتي:

$$a\hat{\rho}^2 - b\hat{\rho} - c = 0 \dots \dots \dots (45)$$

اذ أن:

$$a = \left[ \left( \frac{\hat{\sigma}}{\hat{\theta}} \right)^2 S_x^2 + \hat{\sigma}^2 \right] \quad b = -2 \left[ \left( \frac{\hat{\sigma}}{\hat{\theta}} \right) S_{xy} \right] \quad , \quad c = - \left[ \frac{S_{xy}^2}{S_x^2} \left( \frac{\hat{\theta}}{S_x^2} - 1 \right) \right]$$

وبحل المعادلة (45) نحصل على تقدير المعلمة  $\rho$  وفق الصيغة الآتية:

$$\hat{\rho} = \frac{\hat{\theta} S_{xy}}{\hat{\sigma} S_x^2} \dots \dots \dots (46)$$

تقديرات طريقة الإمكان الأعظم المعدلة تم اشتقاقها بافتراض ان التوزيع الاحتمالي للمتغير التفسيري X هو توزيع القيمة المتطرفة ولكن يمكن ان يكون أي توزيع احتمالي اخر تفرضه علينا البيانات قيد البحث.

### 3- 2 طريقة الامكان التجريبية الحصينة Robust Empirical Likelihood Method (REL)

طريقة الإمكان التجريبية احدى طرائق التقدير اللامعلمية المقترحة من قبل Owen , 1988 , [5], التي تعد بديل لطريقة الإمكان الأعظم في حالة عدم وجود افتراضات حول التوزيع الاحتمالي لحدود الأخطاء العشوائية لأنموذج الانحدار. هذه الطريقة تفترض أن هناك اوزان احتمالية  $p_i$  لكل مشاهدة من مشاهدات العينة ( $i=1,2,\dots,n$ ) , وتهدف الى تقدير معاملات انموذج الانحدار من خلال تعظيم دالة الإمكان التجريبية التي تعرف على انها مضروب تلك الاوزان الاحتمالية في ظل بعض القيود المتعلقة بمعاملات الانموذج التي يراد تقديرها, تلك القيود تكون مشابهة للمعادلات الطبيعية لطريقة المربعات الصغرى أو لمعادلات الإمكان في ظل فرضية التوزيع الطبيعي, الامر الذي يؤدي الى ان تكون هذه الطريقة حساسة جدا لعدم تحقق فرضية التوزيع الطبيعي أو وجود قيم شاذة أو متطرفة, لذلك اقترح كل من Ozdemir و Arslan , 2018 , [6], استعمال قيود محصنة مستعارة من دالة الهدف الخاصة بطريقة التقدير الحصينة M, بمعنى اجراء توفيق بين طريقة الإمكان التجريبية وطريقة M الحصينة.

لأنموذج الانحدار الخطي المبين بالمعادلة (1) وبافتراض  $p_i$  لكل  $i=1,2,\dots,n$  تمثل الأوزان احتمالية لكل مشاهدة من مشاهدات العينة والتي تكون غير المعلومة وذات قيم مختلفة لكل مشاهدة وتحتاج الى التقدير، إذ أن  $p_i \geq 0$ ، فإن طريقة الإمكان التجريبية المستعملة لتقدير معاملي الانموذج  $\underline{\beta}$  و تباين المجتمع  $\sigma^2$  تقضي بتعظيم دالة الإمكان التجريبية تحت توفر بعض القيود على عملية التعظيم وكما مبين ادناه:

$$L_{EL}(\beta, \sigma^2) = \prod_{i=1}^n p_i m \dots \dots \dots (47)$$

تحت تحقق القيود الآتية:

$$\sum_{i=1}^n p_i = 1 \dots \dots \dots (48)$$

$$\sum_{i=1}^n p_i (Y_i - \beta_0 - \beta_1 X_i) X_i = 0 \dots \dots \dots (49)$$

$$\sum_{i=1}^n p_i [(Y_i - \beta_0 - \beta_1 X_i)^2 - \sigma^2] = 0 \dots \dots \dots (50)$$

بأخذ اللوغاريتم لطرفي المعادلة (47) نحصل على لوغاريتم دالة الإمكان التجريبية كما في المعادلة الآتية:

$$\text{Log} L(\beta, \sigma^2) = \text{Log} \prod_{i=1}^n p_i = \sum_{i=1}^n \text{Log} (p_i) \dots \dots \dots (51)$$

تقدير الأوزان الاحتمالية  $\underline{p}^T = [p_1 p_2 \dots p_n]$  ومتجه المعاملات  $\underline{\beta}^T = [\beta_0 \beta_1]$  والتباين  $\sigma^2$ ، يتم من خلال تعظيم لوغاريتم دالة الإمكان التجريبية تحت تحقق القيود الثلاثة وفق المعادلات (48-50)، مشكلة التعظيم هذه يمكن حلها باستعمال طريقة مضاعف لاكرانج الامر الذي يتطلب صياغة دالة لاكرانج وفق الصيغة الآتية:

$$L(\underline{p}, \underline{\beta}, \lambda_0, \lambda_1, \lambda_2) = \sum_{i=1}^n \text{Log} (p_i) - \lambda_0 (\sum_{i=1}^n p_i - 1) - n\lambda_1 \sum_{i=1}^n p_i (Y_i - X_i^T \underline{\beta}) X_i - n\lambda_2 \sum_{i=1}^n p_i [(Y_i - X_i^T \underline{\beta})^2 - \sigma^2] \dots \dots \dots (52)$$

إذ أن  $\lambda_0, \lambda_1, \lambda_2 \in R$  التي تمثل مضاعفات لاكرانج، عملية التقدير تتطلب أولاً تقدير متجه الأوزان الاحتمالية من خلال اخذ المشتقة لدالة لاكرانج (52) بالنسبة لكل  $p_i$  ومساواتها بالصفر لنحصل على تقدير تلك الأوزان وفق الصيغة الآتية:

$$p_i = \frac{1}{\lambda_0 + n\lambda_1 (Y_i - X_i^T \underline{\beta}) X_i + n\lambda_2 [(Y_i - X_i^T \underline{\beta})^2 - \sigma^2]} \quad \text{for } i = 1, 2, \dots, n \dots \dots (53)$$

بأخذ المجموع لطرفي المعادلة (53) لكل قيم  $i$ ، نحصل على أن  $\lambda_0 = n$ ، لذا يمكن إعادة كتابة تلك المعادلة بعد التعويض عن  $\lambda_0$  وكالآتي:

$$p_i = \frac{1}{n(1 + \lambda_1 (Y_i - X_i^T \underline{\beta}) X_i + \lambda_2 [(Y_i - X_i^T \underline{\beta})^2 - \sigma^2])} \quad \text{for } i = 1, 2, \dots, n \dots \dots (54)$$

بالتعويض عن  $p_i$  بموجب الصيغة (54) في دالة الإمكان التجريبية (51) لنحصل على دالة الهدف التالية بدلالة  $\beta, \lambda_1, \lambda_2, \sigma^2$  فقط:

$$L(\underline{\beta}, \lambda_1, \lambda_2, \sigma^2) = - \sum_{i=1}^n \text{Log} (1 + \lambda_1 (Y_i - X_i^T \underline{\beta}) X_i + \lambda_2 [(Y_i - X_i^T \underline{\beta})^2 - \sigma^2]) - n \text{Log} n \dots \dots \dots (55)$$

قيم مضاعفات لاكرانج  $\lambda_1$  و  $\lambda_2$  بالنسبة لكل من متجه معاملات انموذج الانحدار  $\underline{\beta}$  و  $\sigma^2$  يمكن ان نحصل عليها من خلال مشكلة التصغير الآتية:

$$\hat{\lambda}(\underline{\beta}, \sigma^2) = \arg \min_{\lambda} [- \sum_{i=1}^n \text{Log} (\lambda_0 + \lambda_1 (Y_i - X_i^T \underline{\beta}) X_i + \lambda_2 [(Y_i - X_i^T \underline{\beta})^2 - \sigma^2]) - n \text{Log} n] \dots \dots \dots (56)$$

اذ أن  $\underline{\lambda} = [\lambda_1 \lambda_2]$  مع ملاحظة ان مشكلة التصغير (56) ليس لها حلول صريحة مما نحتاج الى الطرائق العددية لإيجاد الحلول لتلك المشكلة، وبتعويض تلك الحلول في  $L(\underline{\beta}, \lambda_1, \lambda_2, \sigma^2)$  نحصل على لوغاريتم دالة الإمكان التجريبية  $L(\hat{\underline{\lambda}}(\underline{\beta}, \sigma^2), \underline{\beta}, \sigma^2)$  ، هذه الدالة بدلالة متجه معاملات الانموذج  $\underline{\beta}$  والتباين  $\sigma^2$  ، وتقدير الإمكان الأعظم التجريبية لتلك المعلمات نحصل عليها بحل مشكلة التعظيم الآتية:

$$(\hat{\underline{\beta}}, \hat{\sigma}^2) = \arg \max_{(\underline{\beta}, \sigma^2)} L(\hat{\underline{\lambda}}(\underline{\beta}, \sigma^2), \underline{\beta}, \sigma^2) \dots \dots \dots (57)$$

مشكلة التعظيم الأخيرة تحل باستعمال الطرائق العددية لعدم وجود حلول صريحة لها. اقترح الباحثين Ozdemir و Arslan , 2018 , [6]، الدمج بين طريقة التقدير M و طريقة الإمكان التجريبية باستبدال القيود الاعتيادية للطريقة الأخيرة بالقيود اللامعلمية وذلك باستعمال معادلة طريقة M

$$\sum_{i=1}^n X_{ij} \psi \left( \frac{Y_i - \beta_0 - \beta_1 X_i}{\hat{\sigma}} \right) = 0 \quad , \quad j = 0, 1, \dots, k \dots \dots \dots (58)$$

هذا التعديل اعتمد من قبل الباحثين لزيادة فاعلية طريقة الإمكان التجريبية في معالجة تأثير القيم الشاذة على تقدير معاملات الانموذج، وأطلق عليها بطريقة الإمكان التجريبية الحصينة، وهنا سنوظف هذه الطريقة لتقدير معاملات انموذج الانحدار الخطي في حالة كون المتغير التفسيري عشوائي.

دالة الإمكان التجريبية المبينة بالعلاقة (47) سيتم تعظيمها بعد استبدال القيد الثاني والثالث العلاقتين (49) و (50) بالقيد الحصين الآتية:

$$\sum_{i=1}^n p_i \varphi (Y_i - \beta_0 - \beta_1 X_i) X_i = 0 \dots \dots \dots (59)$$

$$\sum_{i=1}^n p_i \varphi [(Y_i - \beta_0 - \beta_1 X_i)^2 - \sigma^2] = 0 \dots \dots \dots (60)$$

اذ ان  $\phi$  هي دالة في الأخطاء العشوائية وهي دالة غير متزايدة في حالة دالة Huber ، ومتناقصة في حالة دالة Tukey. التقدير الحصين لكل من معاملات الانموذج والتباين نحصل عليه من خلال حل مشكلة التعظيم الآتية وباستعمال دالة لاكرانج:

$$L(\underline{p}, \underline{\beta}, \lambda_0, \lambda_1, \lambda_2) = \sum_{i=1}^n \text{Log}(p_i) - \lambda_0 (\sum_{i=1}^n p_i - 1) - n \lambda_1 \left( \sum_{i=1}^n p_i \varphi(Y_i - X_i^T \underline{\beta}) \right) X_i - n \lambda_2 \left( \sum_{i=1}^n p_i [\varphi(Y_i - X_i^T \underline{\beta})^2 - \sigma^2] \right) \dots \dots \dots (61)$$

وبنفس الخطوات السابقة التي تم اتباعها في طريقة الإمكان التجريبية فان عملية التقدير تتطلب أولاً تقدير متجه الاوزان الاحتمالية من خلال اخذ المشتقة لدالة لاكرانج (61) بالنسبة لكل  $p_i$  ومساواتها بالصفر لنحصل على تقدير تلك الاوزان وفق الصيغة الآتية:

$$p_i = \frac{1}{\lambda_0 + n \lambda_1 \varphi(Y_i - X_i^T \underline{\beta}) X_i + n \lambda_2 [\varphi(Y_i - X_i^T \underline{\beta})^2 - \sigma^2]} \quad \text{for } i = 1, 2, \dots, n \dots \dots (62)$$

بأخذ المجموع لطرفي المعادلة (62) لكل قيم  $i$  ، نحصل على أن  $\lambda_0 = n$  ، لذا يمكن إعادة كتابة تلك المعادلة بعد التعويض عن  $\lambda_0$  وكالاتي:

$$p_i = \frac{1}{n(1 + \lambda_1 \varphi(Y_i - X_i^T \underline{\beta}) X_i + \lambda_2 [\varphi(Y_i - X_i^T \underline{\beta})^2 - \sigma^2])} \quad \text{for } i = 1, 2, \dots, n \dots \dots (63)$$

بالتعويض عن  $p_i$  بموجب الصيغة (63) في دالة الإمكان التجريبية (61) لنحصل على دالة الهدف التالية بدلالة  $\underline{\beta}, \lambda_1, \lambda_2, \sigma^2$  فقط:

$$L(\underline{\beta}, \lambda_1, \lambda_2, \sigma^2) = - \sum_{i=1}^n \text{Log} (1 + \lambda_1 \varphi(Y_i - X_i^T \underline{\beta}) X_i + \lambda_2 [\varphi(Y_i - X_i^T \underline{\beta})^2 - \sigma^2]) - n \text{Log} n \dots \dots \dots (64)$$

قيم مضاعفات لاجرانج  $\lambda_1$  و  $\lambda_2$  بالنسبة لكل من متجه معاملات انموذج الانحدار  $\underline{\beta}$  و  $\sigma^2$  يمكن ان نحصل عليها من خلال مشكلة التصغير الآتية:

$$\hat{\lambda}(\underline{\beta}, \sigma^2) = \arg \min_{\lambda} [-\sum_{i=1}^n \text{Log} (\lambda_0 + \lambda_1(\varphi(Y_i - X_i^T \underline{\beta})X_i) + \lambda_2[\varphi(Y_i - X_i^T \underline{\beta})^2 - \sigma^2]) - n \text{Log} n] \dots (65)$$

مع ملاحظة ان مشكلة التصغير (65) ليس لها حلول صريحة مما نحتاج الى الطرائق العددية لإيجاد الحلول لتلك المشكلة، وبتعويض تلك الحلول في  $L(\underline{\beta}, \lambda_1, \lambda_2, \sigma^2)$  نحصل على لوغاريتم دالة الإمكان التجريبية  $L(\hat{\lambda}(\underline{\beta}, \sigma^2), \underline{\beta}, \sigma^2)$ ، هذه الدالة بدلالة متجه معاملات الانموذج  $\underline{\beta}$  والتباين  $\sigma^2$ ، وتقدير الإمكان الأعظم التجريبية الحصينة لتلك المعاملات نحصل عليها بحل مشكلة التعظيم الآتية:

$$(\hat{\underline{\beta}}, \hat{\sigma}^2) = \arg \max_{(\underline{\beta}, \sigma^2)} L(\hat{\lambda}(\underline{\beta}, \sigma^2), \underline{\beta}, \sigma^2) \dots (66)$$

مشكلة التعظيم الأخيرة تحل باستعمال الطرائق العددية لعدم وجود حلول صريحة لها.

#### 4- التطبيق العملي

البيانات التي تم اعتمادها في الجانب التطبيقي تتعلق بالمصابين بأمراض القلب اذ تم اختيار عينة عشوائية بحجم ( 30 ) مريض من الراقدين في مستشفى غازي الحريري للجراحات التخصصية ولكل مريض تم تسجيل قراءات كل من حجم كريات الدم الحمراء ( PCV ) وسكر الدم ( RBS ) وكما مبين بالجدول ( 1 ) والتي استعملت في تقدير العلاقة الخطية البسيطة بين ( PCV ) كمتغير استجابة و ( RBS ) كمتغير تفسيري اذ تم بناء انموذج انحدار خطي بسيط بمتغير تفسيري عشوائي وكما مبين في الصيغة (1) في الجانب النظري اذ وصفت علاقة الانحدار بالأنموذج الاتي :

$$PCV_i = \beta_0 + \beta_1 RBS_i + u_i \dots \dots \dots (67)$$

جدول (1) حجم كريات الدم الحمراء ( PCV ) وسكر الدم ( RBS )

No.	PCV	RBS	No.	PCV	RBS	No.	PCV	RBC
1	36	170	11	26	149	21	31	191
2	23	136	12	27	348	22	30	182
3	39	233	13	46	387	23	30	184
4	22	96	14	32	202	24	27	124
5	21	145	15	31	200	25	29	218
6	25	188	16	29	199	26	27	143
7	21	136	17	26	115	27	34	150
8	25	141	18	33	158	28	29	153
9	22	147	19	31	152	29	34	161
10	41	310	20	29	193	30	32	209

انموذج الانحدار العشوائي وفق العلاقة (67) تم بناءه على افتراض ان متغير الاستجابة PCV يتوزع وفق التوزيع الطبيعي وان المتغير التفسيري RBS يتوزع وفق توزيع القيمة المتطرفة اما حدود الخطأ العشوائي فأنها تتبع التوزيع الطبيعي وبشكل مستقل بمتوسط يساوي الصفر وتباين ثابت يساوي  $\sigma^2(1 - \rho^2)$ . للتحقق من هذين الافتراضين تم اجراء اختبار حسن المطابقة لكل من المتغيرين الاستجابة والتفسيري، اذ تم استعمال اختبار مربع كاي لحسن

المطابقة Chi-Squared Goodness of Fit. فبالنسبة لبيانات متغير الاستجابة (PCV) تم اختبار هل تخضع للتوزيع الطبيعي مقابل عدم خضوعها لهذا التوزيع اختبار الفرضية الآتية:  
 $H_0: PCV \sim Normal \text{ dist.}$

$H_1: PCV \not\sim Normal \text{ dist.}$

نتائج الاختبار مبينة في الجدول (2) أذ تشير تلك النتائج الى ان القيمة الاحتمالية P-Value المرافقة لهذا الاختبار هي أكبر من مستوى المعنوية 0.05، وبذلك لا ترفض فرضية العدم بمعنى أن بيانات المتغير PCV تتوزع توزيعاً طبيعياً، والشكل (1) يوضح المدرج التكراري ومنحنى التوزيع الطبيعي لبيانات هذا المتغير.

أما بالنسبة لبيانات المتغير التفسيري RBS فقد تم اختبار فيما إذا كانت تتبع توزيع القيمة المتطرفة وحسب الفرضية الآتية:

$H_0: RBS \sim Extreme \text{ Value dist.}$

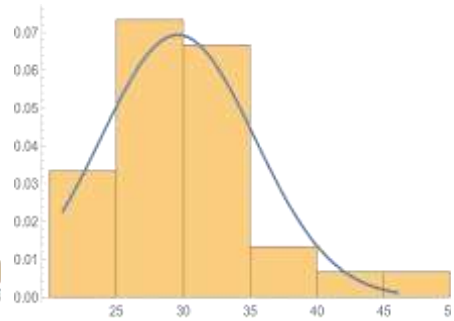
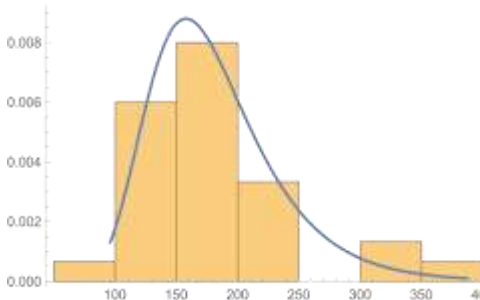
$H_1: RBS \not\sim Extreme \text{ Valuedist.}$

نتائج هذا الاختبار مبينة في الجدول (2) اذ تشير القيمة الاحتمالية لهذا الاختبار والتي هي أكبر من مستوى المعنوية 0.05 الى قبول فرضية العدم بمعنى ان بيانات هذا المتغير تتبع توزيع القيمة المتطرفة والشكل (2) يوضح المدرج التكراري ومنحنى التوزيع لبيانات هذا المتغير.

جدول (2): اختبار مربع كاي لحسن المطابقة لمتغير الاستجابة PCV

وللمتغير التفسيري RBS

	PCV	RBS
Statistic	1.46667	6.26667
P-Value	0.916884	0.281129
Parameters	$\mu = 29.6; \sigma = 5.75384$	$\alpha = 157.785; \theta = 41.8356$



شكل (2): المدرج التكراري ومنحنى توزيع القيمة المتطرفة للمتغير التفسيري RBS

شكل (1): المدرج التكراري ومنحنى التوزيع الطبيعي لمتغير الاستجابة PCV

لتقدير معاملي انموذج الانحدار العشوائي المبين بالعلاقة (67) سيتم الاعتماد على طرائق التقدير موضوع البحث. الجدول (2) يلخص نتائج التقدير، من تلك النتائج نجد ان الطريقتين افرزت نماذج مقدره معنوية وذلك بالاعتماد على المؤشر الاحصائي F ولمستوى معنوية 0.05، كما انها اتفقت على ان هناك تأثير طردي للمتغير التفسيري العشوائي RBS على متغير الاستجابة PCV وان قيمة هذا التأثير متقاربة بين النماذج المقدره، أما معياري حسن المطابقة ودقة التقدير R2 و RMSE فقد اشارت الى افضلية طريقة الإمكان الأعظم المعدلة. من الجدير بالملاحظة أن القيمة المنخفضة لمعامل التحديد R2 كانت نتيجة لاعتماد متغير تفسيري واحد

والمتمثل بالمتغير العشوائي سكر الدم (RBS) اذ هناك عوامل أخرى لم يتم تضمينها في انموذج الانحدار والتي لها تأثير على تحديد حجم كريات الدم الحمراء، لذا يمكن اعتماد قيمة هذا المعيار مقبولة بالنسبة لنموذج الانحدار المقدر موضوع البحث.  
بناء على ما تقدم وبالاغتماد على تقديرات طريقة الإمكان الأعظم فان انموذج الانحدار العشوائي المقدر يكون كالآتي:

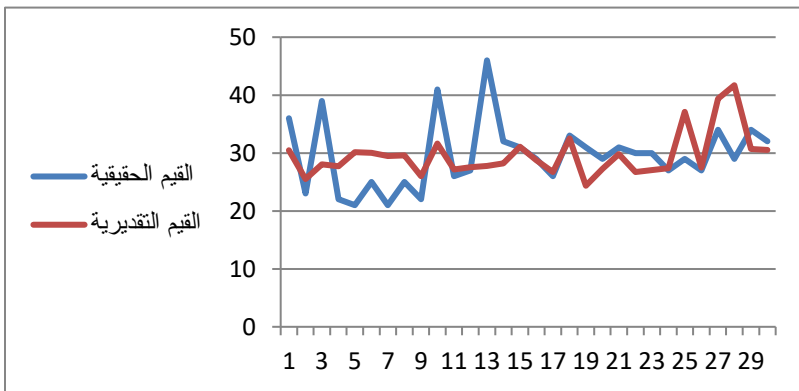
$$PCV_i = 28.0954 + 0.05963RBS_i \dots\dots\dots (68)$$

جدول (3) : نتائج تقدير طرائق التقدير لأنموذج الانحدار العشوائي RELE , MMLE

طريقة التقدير	Estimated parameters		F statistic	R <sup>2</sup>	RMSE
MMLE	$\beta_0$	28.0954	21.71965*	43.68%	4.469
	$\beta_1$	0.05963			
RELE	$\beta_0$	16.09	19.369*	40.89%	4.578
	$\beta_1$	0.07443			

جدول (4): القيم الحقيقية والتقديرية بطريقة الإمكان الأعظم المعدلة لمتغير الاستجابة PCV

No.	Real value	Estimated value	No.	Real value	Estimated value	No.	Real value	Estimated value
1	36	30.494	11	26	27.155	21	31	29.839
2	23	25.485	12	27	27.573	22	30	26.738
3	39	28.050	13	46	27.751	23	30	27.036
4	22	27.692	14	32	28.228	24	27	27.394
5	21	30.137	15	31	31.091	25	29	37.113
6	25	30.017	16	29	28.765	26	27	27.513
7	21	29.481	17	26	26.738	27	34	39.379
8	25	29.600	18	33	32.522	28	29	41.705
9	22	26.022	19	31	24.353	29	34	30.673
10	41	31.627	20	29	27.274	30	32	30.554



شكل (3) : الرسم البياني لمنحنى الانحدار الحقيقي والمقدر

## مناقشة نتائج التطبيق العملي

تم مناقشة تقدير نموذج الانحدار الخطي البسيط بمتغير توضيحي عشوائي والذي يمثل العلاقة الخطية بين حجم كريات الدم الحمراء ( PCV ) كتغير استجابة و وسكر الدم ( RBS ) كمتغير تفسيري عشوائي, ولكون هذا النموذج يخالف الافتراضات الأساسية لأنموذج الانحدار تم استعمال طريقتين حصينتين الأولى طريقة الإمكان الأعظم المعدلة والتي تم تطبيقها من قبل الباحثين في تقدير معاملات نموذج الانحدار العشوائي، والطريقة الثانية هي طريقة الإمكان التجريبية الحصينة والتي تم توظيفها من قبلنا في تقدير معاملات نموذج الانحدار موضوع البحث، إذ حسب علمنا لم يتم استعمالها سابقا في عملية تقدير هذا النموذج وإنما استعملت في تقدير نماذج الانحدار التي تعاني من بعض المشاكل القياسية وفي حالة احتواء مشاهدات العينة على قيم شاذة او متطرفة.

تبين من خلال التطبيق العملي وفيما يتعلق بعينة البحث تفوق طريقة الإمكان الأعظم المعدلة بالاعتماد على معياري حسن المطابقة ودقة التقدير  $R^2$  و  $RMSE$ , وبالرغم من هذا التفوق نجد ان هناك تقارب كبير بين طرائق التقدير بالنسبة لقيمة معياري المفاضلة وكذلك بالنسبة للقيمة التقديرية لتأثير سكر الدم على حجم كريات الدم الحمراء وطبيعة العلاقة الطردية بين هذين المتغيرين، هذا التقارب يقودنا الى الاستنتاج بفاعلية طريقة الإمكان التجريبية الحصينة. استنادا على ما تقدم، يمكن استعمال طريقة الإمكان التجريبية الحصينة في تقدير نموذج الانحدار العشوائي كونها أسهل من حيث التطبيق من طريقة الإمكان الأعظم المعدلة، كما انها لاتستند على افتراضات خاصة بالتوزيع الاحتمالي لكل من حدود الخطأ العشوائي والمتغير التوضيحي العشوائي.

## المصادر

1. Islam, M. Q., & Tiku, M. L. (2005). Multiple linear regression model under nonnormality. Communications in Statistics-Theory and Methods, 33(10), 2443-2467.
2. Islam, M. Q., & Tiku, M. L. (2010). Multiple linear regression model with stochastic design variables. Journal of Applied Statistics, 37(6), 923-943.
3. Kerridge, D. (1967). Errors of prediction in multiple regression with stochastic regressor variables. Technometrics, 9(2), 309-311.
4. Lai, T. L., & Wei, C. Z. (1982). Least squares estimates in stochastic regression models with applications to identification and control of dynamic systems. The Annals of Statistics, 10(1), 154-166.
5. Owen, A. B. (1988). Empirical likelihood ratio confidence intervals for a single functional. Biometrika, 75(2), 237-249.
6. Özdemir, Ş., & Arslan, O. (2020). Combining empirical likelihood and robust estimation methods for linear regression models. Communications in Statistics-Simulation and Computation, 51(3), 941-954
7. Sazak, H. S., Tiku, M. L., & Islam, M. Q. (2006). Regression analysis with a stochastic design variable. International Statistical Review/Revue Internationale de Statistique,
8. Tiku, M. L., & Vaughan, D. C. (1997). Logistic and nonlogistic density functions in binary regression with nonstochastic covariates. Biometrical Journal, 39(8), 883-898.
9. Vaughan, D. C., & Tiku, M. L. (2000). Estimation and hypothesis testing for a nonnormal bivariate distribution with applications. Mathematical and Computer Modelling, 32(1-2), 53-67.

## نموذج لندلي ذي المعلمتين للانحدار الذاتي-الأوساط المتحركة من الرتبة الأولى مع التطبيق

د. بشير جميل خليل / مدرس في وزارة التربية العراقية  
د. علي محمد علي / مدرس في جامعة ميسان، كلية الإدارة والاقتصاد، قسم الاقتصاد  
ازهار سالم احمد / مديرية إحصاء نينوى

تاريخ استلام البحث: 2023/09/19

تاريخ قبول البحوث: 2023/11/23

نشر البحث في العدد الثالث والعشرين: حزيران / يونيو 2024

رمز التصنيف ديوي / النسخة الالكترونية (Online): 2522-64X/515.7

رمز التصنيف ديوي / النسخة الورقية (Print): 2519-948X/515.7

## نموذج لندي ذي المعلمتين للانحدار الذاتي-الأوساط المتحركة من الرتبة الأولى مع التطبيق

د. بشير جميل خليل / مدرس في وزارة التربية العراقية

د. علي محمد علي / مدرس في جامعة ميسان، كلية الإدارة والاقتصاد، قسم الاقتصاد

ازهار سالم احمد / مديرية إحصاء نينوى

### المستخلص

تم في هذا البحث اقتراح نموذج سلاسل زمنية جديد من نوع  $ARMA(1,1)$  عندما تتبع البواقي توزيع لندي ذي المعلمتين، حيث تم تقديم فكرة عامة عن توزيع لندي ذي المعلمتين ونماذج بوكس جنكز، وتقدير معلمات النموذج المقترح بالاعتماد على طريقة الإمكان الأعظم، ومن ثم تمت مقارنة النموذج الجديد مع نماذج أخرى هي كل من النموذج الطبيعي والأسّي والأسّي ذي المعلمتين ولندي بواسطة مجموعتين من البيانات الحقيقية وتمثلت هذه البيانات في سرعة الرياح في العراق وتدفق نهر Somerset Yeo في المملكة المتحدة وقد أثبتت النتائج تفوق النموذج المقترح على كافة النماذج الأخرى وبالنسبة لمجموعي البيانات.

### Two-Parameter Lindley $ARMA(1,1)$ Model with Applications

Dr. Basheer Jameel Khaleel Lecturer in Ministry of Education, Iraq

Dr. Ali Mohammed Ali Lecturer in Misan University, college of

Administration & Economics, Economics department

Azhar Salim Ahmed Nineveh Statistics Directorate

### Abstract

In this research, a new  $ARMA(1,1)$  time series model was proposed when residuals follow the two-parameter Lindley distribution. A general idea of the two-parameter Lindley distribution and Box-Jenkins models was presented. The parameters of the proposed model were estimated using the method of maximum likelihood. Then, the new model was compared with other models, namely the normal, exponential, two-parameter exponential, and Lindley models, using two sets of real-world data. These data sets represented wind speed in Iraq and the flow of the Somerset Yeo river in the United Kingdom. The results demonstrated the superiority of the proposed model over all other models for both data sets.

### 1- المقدمة

تعتبر السلاسل الزمنية من الأساليب الإحصائية المستخدمة في تحليل الكثير من الظواهر التي تتصف بصفة التغير عبر الزمن والتي من خلالها يمكن معرفة طبيعة التغيرات التي تطرأ على قيم تلك الظواهر مع الزمن ومن ثم تحديد النتائج بهدف التنبؤ بالقيم المستقبلية لها والذي يعتبر الهدف الأساس من تحليل تلك الظواهر، وقد تم دراسة السلاسل الزمنية في الكثير من التخصصات والتطبيقات الحديثة مثل التنبؤ بالسلسلة الزمنية المالية، والطبية والهندسية

ومعالجة الإشارات، والأحدث في مجالات التعلم الآلي والتعرف على الأنماط (Privalsky, 2023).

تستخدم التوزيعات الملتوية كبديل عن التوزيع الطبيعي في كثير من التطبيقات في الحياة العملية، وذلك لأن هذه الظواهر في الكثير من الأحيان لا تخضع للتوزيع الطبيعي الذي يمتاز بأنه توزيع متماثل، ومن التوزيعات الشائعة في هذا الجانب التوزيع الأسّي ذي المعلمة والتوزيع الأسّي ذي المعلمتين وتوزيع غاما وغيرها الكثير من التوزيعات الاحتمالية. بالإضافة إلى هذه التوزيعات هناك ما يسمى التوزيعات المختلطة Mixed Distribution والتي تعرف بأنها خليط من توزيعين أو أكثر، وسبب ظهور هذه العائلة من التوزيعات هو وجود ظواهر مقسمة إلى عدة أقسام وكل قسم يخضع لتوزيع معين، حيث يكون المجتمع النهائي هو خليط من هذه التوزيعات بنسب معينة لكل توزيع احتمالي (Rachev et al., 2005).

إن افتراض التوزيع الطبيعي في تحليل البيانات الإحصائية عموماً والسلاسل الزمنية خصوصاً هي مسألة بعيدة عن الواقع في العديد من الحالات. نتيجة لذلك، تم إدخال العديد من نماذج السلاسل الزمنية غير الطبيعية في الأدبيات خلال العقود الأربعة الماضية، لوصف السمات البارزة للبيانات، والتي لم يتم التقاطها من خلال التوزيع الطبيعي (Balakrishna, 2021). نذكر هنا على سبيل المثال بعض الأمثلة على استخدام نماذج ARMA مع توزيعات غير طبيعية كالتوزيع الأسّي (Lawrance & Lewis, 1980) وتوزيع غاما (Sim, 1987) وتوزيع لابلاس (Damsleth & El-Shaarawi, 1989). ومؤخراً التوزيع الأسّي ذي المعلمتين (حسوني و غني، 2021) وتوزيع لندلي ((Bakouch & Popović, 2016) و(حسوني و غني، 2022)) وتوزيع بور (de Araújo et al., 2023).

يهدف البحث إلى بناء نموذج انحدار ذاتي-الأوساط المتحركة (ARMA) Autoregressive-Moving Average، عندما يكون للأخطاء توزيع لندلي ذي المعلمتين، ولتحقيق هدف البحث تم تقسيم ما تبقى منه إلى عدة أقسام، حيث يتناول القسم الثاني نماذج الانحدار ذاتي-الأوساط المتحركة (ARMA)، بينما يضم القسم الثالث توزيع لندلي ذي المعلمتين، والقسم الرابع لتقدير معالم النموذج باستخدام طريقة الإمكان الأعظم، والقسم الخامس الجانب التطبيقي، فيما اختص القسم السادس بأهم الاستنتاجات التي توصل إليها البحث.

## 2- نماذج الانحدار ذاتي-الأوساط المتحركة

تلعب السلاسل الزمنية المستقرة دوراً بارزاً في تحليل بيانات السلاسل الزمنية حيث قدم بوكس جينكز Box-Jenkins نهجاً علمياً ومتكاملاً لبناء النماذج الذي تم استخدامه على مدار نصف القرن الماضي. إذا كانت السلسلة  $Y_t$  فإن نموذج انحدار ذاتي-الأوساط المتحركة  $ARMA(p,q)$ ، حيث  $p$  رتبة الانحدار الذاتي و  $q$  رتبة الأوساط المتحركة يكون كما يأتي (Woodward et al., 2022):

$$Y_t = \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q} \quad (1)$$

إذ إن كل من  $\phi_k, k = 1, \dots, p$  و  $\theta_k, k = 1, \dots, q$  هي ثوابت حقيقية، وإن  $\varepsilon_t$  هو الخطأ العشوائي. وكحالة خاصة من هذا النموذج نموذج  $ARMA(1,1)$  بمعنى أن رتبة كل من الانحدار الذاتي ورتبة الأوساط المتحركة تساوي 1، أي إن  $p=q=1$ ، فيمكن كتابة المعادلة (1) كما يأتي:

$$Y_t = \phi Y_{t-1} + \varepsilon_t - \theta \varepsilon_{t-1} \quad (2)$$

وقد أعاد الباحثان (Kasraie & Sayyareh, 2020) صياغة المعادلة (2) بدلالة الأخطاء العشوائية وكما يأتي:

$$\varepsilon_t = \sum_{j=0}^{t-2} \theta^j (Y_{t-j} - \phi Y_{t-j-1}) + \theta^{t-1} \varepsilon_1, t = 2, 3, \dots, n \quad (3)$$

إذ إن  $\varepsilon_t = E(\varepsilon_t)$ ، عندما  $E(\varepsilon_t)$  يمثل توقع الخطأ (الجدول 1).  
تمتاز المعادلة (3) بأن هناك استقلالية بين قيم السلسلة  $\varepsilon_t$ .

### 3- توزيع لندلي ذي المعلمتين

يمتاز توزيع لندلي Lindley Distribution بأنه أحد التوزيعات المختلطة المستمرة، وقد سُمي بهذا الاسم تيمناً بالباحث لندلي Lindley الذي قدم هذا التوزيع في عام 1958. يتكون هذا التوزيع من دمج متغيرين عشوائيين، حيث يتبع أحدهما التوزيع الأسّي بمعلمة  $(\lambda)$ ، بينما يتبع المتغير الآخر توزيع گاما بالمعلمتين  $(\alpha = 2, \lambda)$  بوزن  $(w_1 = \frac{\lambda}{1+\lambda})$  وتعرف دالة الكثافة الاحتمالية Probability Density Function (PDF) للمتغير العشوائي  $X$  الذي يتبع هذا التوزيع كما في الصيغة الآتية (Ghitany et al., 2008):

$$f(x; \lambda) = \frac{\lambda^2}{\lambda + 1} (1 + x) e^{-\theta x}, ; x > 0, \lambda \geq 0 \quad (4)$$

بينما يتم التعبير عن دالة الكثافة الاحتمالية لتوزيع لندلي ذي المعلمتين Two-Parameter Lindley Distribution الذي اقترح من قبل (Shanker, Sharma & Shanker) في عام 2013 كما يأتي (Shanker et al., 2013):

$$f(x; \alpha, \lambda) = \frac{\lambda^2}{\lambda + \alpha} (1 + \alpha x) e^{-\lambda x}, x > 0, \lambda \geq 0, \alpha > -\lambda \quad (5)$$

وقد استعمل الباحثون الوزن  $w_1 = \frac{\theta}{\theta + \alpha}$  بنفس التوزيعات المستعملة في توزيع لندلي ذي معلمة واحدة.

وقد وضع الباحثون في هذا البحث جدولاً يحتوي على دالة الكثافة الاحتمالية والتوقع لكل توزيع من التوزيعات التي سيتم نمذجة ARMA بها بالنسبة للخطأ العشوائي، وكما يأتي:  
جدول (1): دالة الكثافة الاحتمالية للتوزيعات المدروسة

Distributions	PDF	Expectation
Normal	$f(\varepsilon; \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{\varepsilon^2}{2\sigma^2}}$	$E(\varepsilon) = 0$
Exponential	$f(\varepsilon; \lambda) = \lambda e^{-\lambda\varepsilon}$	$E(\varepsilon) = 1/\lambda$
Two-Par. Exponential	$f(\varepsilon; \lambda, \eta) = \lambda e^{-\lambda(\varepsilon-\eta)}$	$E(\varepsilon) = \frac{\lambda\eta + 1}{\lambda}$
Lindley	$f(\varepsilon; \lambda) = \frac{\lambda^2}{\lambda + 1} (1 + \varepsilon) e^{-\lambda\varepsilon}$	$E(\varepsilon) = \frac{(\lambda + 2)}{\lambda(\lambda + 1)}$
Two-Par. Lindley	$f(\varepsilon; \alpha, \lambda) = \frac{\lambda^2}{\lambda + \alpha} (1 + \alpha\varepsilon) e^{-\lambda\varepsilon}$	$E(\varepsilon) = \frac{(\lambda + 2\alpha)}{\lambda(\lambda + \alpha)}$

#### 4- مقدرات طريقة الإمكان الأعظم

تستخدم طريقة الإمكان الأعظم Maximum Likelihood Method في تقدير معالم النماذج الإحصائية بحيث تجعل هذه المقدرات دالة الإمكان Likelihood Function في نهايتها العظمى (Mavrakakis & Penzer, 2021). وفيما يأتي دالة الإمكان ولوغاريتمها الطبيعي للتوزيعات المدروسة:

##### أولاً: التوزيع الطبيعي

يمكن تعريف دالة الإمكان بالنسبة للتوزيع الطبيعي كما يأتي (العزاوي، 2001):

$$L(\phi, \theta, \sigma^2 | \varepsilon) = [2\pi\sigma^2]^{-\frac{n}{2}} \exp\left(-\frac{1}{2\sigma^2} \sum_{t=1}^n \varepsilon_t^2\right) \quad (6)$$

أما اللوغاريتم الطبيعي لهذه الدالة فيعرف كما يأتي:

$$\ell(\phi, \theta, \sigma^2 | \varepsilon) = -\frac{n}{2} \ln[2\pi\sigma^2] - \frac{1}{2\sigma^2} \sum_{t=1}^n \varepsilon_t^2 \quad (7)$$

##### ثانياً: التوزيع الأسّي

وتعرف دالة الإمكان للتوزيع الأسّي على النحو الآتي (العزاوي، 2001):

$$L(\phi, \theta, \lambda | \varepsilon) = \lambda^n \exp\left(-\lambda \sum_{t=1}^n \varepsilon_t\right) \quad (8)$$

وبالنسبة للوغاريتم الطبيعي لهذه الدالة فيعرف كما يأتي:

$$\ell(\phi, \theta, \lambda | \varepsilon) = n \ln[\lambda] - \lambda \sum_{t=1}^n \varepsilon_t \quad (9)$$

##### ثالثاً: التوزيع الأسّي ذي المعلمتين

كما تعرف دالة الإمكان للتوزيع الأسّي ذي المعلمتين بالشكل الآتي (حسوني و غني، 2021):

$$L(\phi, \theta, \lambda, \eta | \varepsilon) = \lambda^n \exp\left(-\lambda \sum_{t=1}^n \varepsilon_t + \lambda \eta\right) \quad (10)$$

واللوغاريتم الطبيعي لهذه الدالة يعرف كما يأتي:

$$\ell(\phi, \theta, \lambda, \eta | \varepsilon) = n \ln[\lambda] - \lambda \sum_{t=1}^n \varepsilon_t + \lambda \eta \quad (11)$$

##### رابعاً: توزيع لندلي

يمكن تعريف دالة الإمكان لتوزيع لندلي كما يأتي (حسوني و غني، 2022):

$$L(\phi, \theta, \lambda | \varepsilon) = \left[\frac{\lambda^2}{\lambda + 1}\right]^n \prod_{t=1}^n (1 + \varepsilon_t) \exp\left(-\lambda \sum_{t=1}^n \varepsilon_t\right) \quad (12)$$

واللوغاريتم الطبيعي لها يعرف كما يأتي:

$$\ell(\phi, \theta, \lambda | \varepsilon) = 2n \ln(\lambda) - n \ln(\lambda + 1) + \sum_{t=1}^n \ln(1 + \varepsilon_t) - \lambda \sum_{t=1}^n \varepsilon_t \quad (13)$$

خامساً: توزيع لندلي ذي المعلمتين

بالنسبة لتوزيع لندلي ذي المعلمتين فيمكن تعريف دالة الإمكان كما يأتي (هذا البحث):

$$L(\phi, \theta, \lambda, \alpha | \varepsilon) = \left[ \frac{\lambda^2}{\lambda + \alpha} \right]^n \prod_{t=1}^n (1 + \alpha \varepsilon_t) \exp\left(-\lambda \sum_{t=1}^n \varepsilon_t\right) \quad (14)$$

واللوغاريتم الطبيعي لها فيعرف كما يأتي:

$$\ell(\phi, \theta, \lambda, \alpha | \varepsilon) = 2n \ln(\lambda) - n \ln(\lambda + \alpha) + \sum_{t=1}^n \ln(1 + \alpha \varepsilon_t) - \lambda \sum_{t=1}^n \varepsilon_t \quad (15)$$

حيث إن  $\varepsilon_t$  معرفة في المعادلة (3).

ونتيجة للتطور الكبير في مجال الحوسبة الإحصائية وظهور برمجيات تحتوي على أدوات كبيرة تمتاز بالسرعة والدقة فقد تم استخدام دالة optim ضمن حزمة stats في بيئة R (R Core Team, 2023) لتعظيم دوال الإمكان بالنسبة لجميع النماذج المدروسة بدلاً من الحسابات اليدوية.

## 5- الجانب التطبيقي

في هذا الجزء من البحث، تم تقدير نموذج ARMA(1,1) لمجموعتين من البيانات بالاعتماد على النماذج الخمسة التي تم التطرق لها في الجانب النظري، وتتمثل هذه البيانات في بيانات سرعة الرياح من الأول من كانون الثاني عام 2020 إلى 23 شباط عام 2021، والتي عمل عليها (حسوني وغني، 2021)، وقد أثبت الباحثان أن هذه البيانات مستقرة بالاستناد إلى اختباري ديكي فولر الموسع وفيليبس بيرون. ومن خلال رسم دالة الارتباط الذاتي والارتباط الذاتي الجزئي اقترحا استخدام نموذج ARMA(1,1). والمجموعة الثانية تمثل تدفق نهر Somerset Yeo في المملكة المتحدة للفترة من 10 كانون الثاني و28 شباط عام 2021، والتي عمل عليها الباحثان (حسوني وغني، 2022)، واثبتا استقراراً هذه السلسلة واقترحا نموذج ARMA(1,1) أيضاً من خلال رسم دالة الارتباط الذاتي والارتباط الذاتي الجزئي. وقد تم اختيار هذه المجموعتين من البيانات لقرب موضوع البحثين إلى موضوع الدراسة. والجدول الآتي يوضح بعض الإحصاءات الوصفية للسلسلتين:

جدول (2): إحصاءات وصفية لسلسلتى الدراسة

السلسلة	العدد	أقل قيمة	الوسيط	أعلى قيمة	المتوسط الحسابي	الانحراف المعياري	معامل الالتواء	معامل التفلطح
سرعة الرياح	85	7.4	14.8	51.2	17.441	8.75	1.715	6.171
تدفق النهر	50	1.52	5.185	37.9	6.47	6.175	3.153	15.308

حيث يلاحظ أن كلا السلسلتين يعاني من التواء موجب عالي، ويعاني من تدبب أيضاً. وفيما يأتي تقدير النماذج لمجموعتي البيانات، مع ملاحظة أن جميع العمليات جرت في بيئة R:

## 1-5- بيانات سرعة الرياح

تم تقدير معاملات نموذج  $ARMA(1,1)$  لبيانات سرعة الرياح في العراق، وكانت نتائج تقدير المعلمات بالنسبة للنماذج الخمسة كما يأتي:

جدول (3): تقديرات المعلمات لبيانات سرعة الرياح

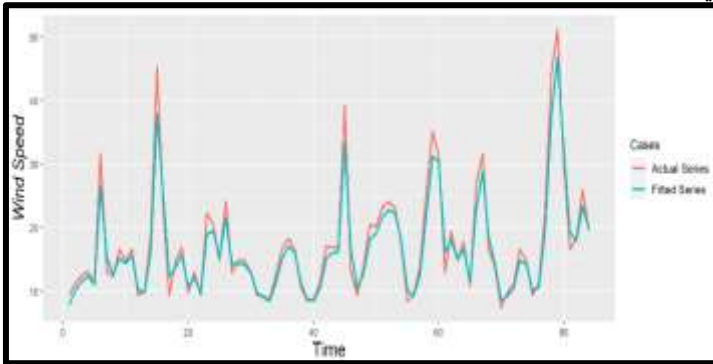
Model	$\hat{\phi}$	$\hat{\theta}$	$\hat{\sigma}^2$	$\hat{\alpha}$	$\hat{\lambda}$	$\hat{\eta}$
Normal	0.93479	0.21052	78.79803	-	-	-
Exponential	0.96180	0.16566	-	-	0.16103	-
Two-Par. Exponential	0.96180	0.16567	-	-	0.16135	0.90754
Lindley	0.91932	0.33776	-	-	0.28033	-
Two-Par. Lindley	0.96688	0.20256	-	0.02108	0.17811	-

وبناءً على هذه المعلمات تم حساب قيم بعض المعايير ليتسنى لنا المقارنة بين النماذج الخمسة المدروسة، وهذه المعايير هي جذر متوسط مربعات الخطأ (RMSE) ومعياري أكايكي للمعلومات (AIC)، ومعياري بيز للمعلومات (BIC)، والجدول (4) يبين قيم هذه المعايير، حيث إن النموذج المقترح هو أفضل نموذج لأنه يمتلك أقل المعايير، وبشكل عام فأن جميع النماذج غير الطبيعية هي أفضل من النموذج الطبيعي، والسبب واضح وهو أن البيانات تعاني من التواء وتفلطح.

جدول (4): معايير المقارنة لبيانات سرعة الرياح

Model	RMSE	AIC	BIC
Normal	9.489228	618.3984	625.7264
Exponential	8.993684	491.1108	500.8814
Two-Par. Exponential	8.951509	490.8426	500.6132
Lindley	8.942945	488.8946	496.2226
Two-Par. Lindley	8.942731	488.7846	496.1126

وقد تم رسم السلسلة مع السلسلة التقديرية باستخدام نموذج لندي ذي المعلمتين، وكما في الشكل الآتي:



شكل (3): رسم سلسلة سرعة الرياح

نلاحظ من الشكل اقتراب السلسلة المقدرة من الحقيقية.

## 2-5- بيانات تدفق النهر

تم تقدير معاملات نموذج ARMA(1,1) لبيانات تدفق نهر Somerset Yeo في المملكة المتحدة، وهكذا كانت نتائج تقييم المعلمات للنماذج الخمسة كما يلي:

جدول (5): تقديرات المعلمات لبيانات تدفق النهر

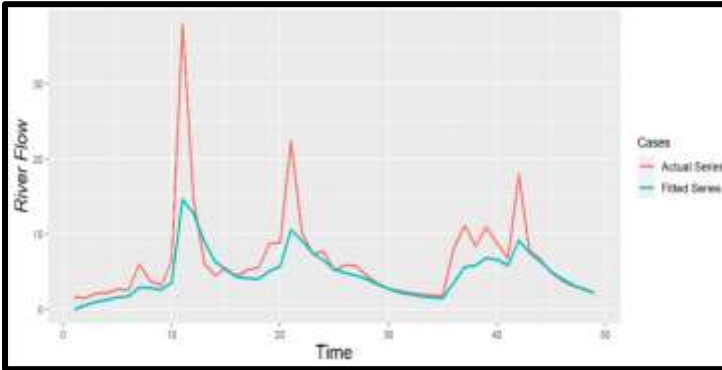
Model	$\hat{\phi}$	$\hat{\theta}$	$\hat{\sigma}^2$	$\hat{\alpha}$	$\hat{\lambda}$	$\hat{\eta}$
Normal	0.91056	0.46918	37.08243	-	-	-
Exponential	0.83307	0.49157	-	-	0.36434	-
Two-Par. Exponential	0.83310	0.49159	-	-	0.36381	0.60408
Lindley	0.89404	0.52482	-	-	0.56616	-
Two-Par. Lindley	0.87536	0.54228	-	0.10883	0.46958	-

بالاستناد إلى هذه المعطيات، تم حساب قيم بعض المعايير ليسهل علينا المقارنة بين النماذج الخمسة المدروسة. الجدول (6) يظهر قيم هذه المعايير. والواضح أن النموذج المقترح هو أفضل من بقية النماذج لأنه يمتلك أقل المعايير، والسبب هو أن البيانات تعاني من التواء وتفلطح.

جدول (6): معايير المقارنة لبيانات تدفق النهر

Model	RMSE	AIC	BIC
Normal	6.318732	328.5562	334.2923
Exponential	6.25384	232.7631	238.4992
Two-Par. Exponential	6.244801	217.8597	225.5078
Lindley	6.140641	213.2416	220.8897
Two-Par. Lindley	6.095835	210.8171	216.5532

وقد تم رسم السلسلة مع السلسلة التقديرية باستخدام نموذج لندلي ذي المعلمتين، وكما في الشكل الآتي:



شكل (4): رسم سلسلة تدفق النهر

نلاحظ من الشكل اقتراب السلسلة المقدر من الحقيقية.

## 6- الاستنتاجات

بناءً على ما سبق نستنتج أن مجموعتي البيانات هي بيانات تعاني من تفلطح والتواء، مما يدل على بعدها عن النمط الطبيعي، وبالتالي لا يكون نموذج التوزيع الطبيعي ملائماً. كما أثبتت نتائج النموذج المقترح أفضلية نموذج لندلي ذي المعلمتين بالاعتماد على مجموعتي بيانات الدراسة، وظهرت جميع نتائج التوزيعات غير الطبيعية أفضلية على النموذج الطبيعي.

## المصادر

1. حسوني، رواء مالك، وغني، علي ياسين، (2021)، تقدير معلمات أنموذج ARMA عندما يتبع الخطأ العشوائي التوزيع الاسي ذو المعلمتين، مجلة الإدارة والاقتصاد، العدد ١٣٠، ص ١٩٠-٢١٢.
2. حسوني، رواء مالك، وغني، علي ياسين، (2022)، تقدير معلمات أنموذج ARMA عندما يتبع الخطأ العشوائي توزيع لندلي، مجلة الإدارة والاقتصاد، العدد 132، ص 260-278.
3. العزاوي، ماجد رشيد، (٢٠٠١)، حول بعض خصائص الانموذج المختلط  $ARMA(1,1)$  غير الطبيعي، أطروحة دكتوراه، جامعة بغداد.
4. Bakouch, H. S., & Popović, B. V. (2016). Lindley first-order autoregressive model with applications. *Communications in Statistics-Theory and Methods*, 45(17), 4988-5006.
5. Balakrishna, N. (2021). *Non-Gaussian autoregressive-type time Series*. Singapore: Springer.
6. Damsleth, E., & El-Shaarawi, A. H. (1989). ARMA models with double-exponentially distributed noise. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 51(1), 61-69.
7. de Araújo, F. J. M., Guerra, R. R., & Peña-Ramírez, F. A. (2023). The Burr XII Autoregressive Moving Average Model. In *Computer Sciences & Mathematics Forum* (Vol. 7, No. 1, p. 46). MDPI.
8. Ghitany, M. E., Atieh, B., & Nadarajah, S. (2008). Lindley distribution and its application. *Mathematics and computers in simulation*, 78(4), 493-506.
9. Kasraie, M., & Sayyareh, A. (2020). Modified Maximum Likelihood Estimation in First-Order Autoregressive Moving Average Models with some Non-Normal Residuals. *Journal of The Iranian Statistical Society*, 19(2), 33-66.
10. Lawrance, A. J., & Lewis, P. A. W. (1980). The exponential autoregressive-moving average EARMA (p, q) process. *Journal of the Royal Statistical Society: Series B (Methodological)*, 42(2), 150-161.
11. Mavrakakis, M. C., & Penzer, J. (2021). *Probability and statistical inference: from basic principles to advanced models*. CRC Press.
12. Prival'sky, V. (2023). *Practical Time Series Analysis in Natural Sciences*. Springer Nature.
13. R Core Team (2023). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org>
14. Rachev, S. T., Menn, C., & Fabozzi, F. J. (2005). *Fat-tailed and skewed asset return distributions: implications for risk management, portfolio selection, and option pricing*. John Wiley & Sons.

15. Rigby, R. A., Stasinopoulos, M. D., Heller, G. Z., & De Bastiani, F. (2019). *Distributions for modeling location, scale, and shape: Using GAMLSS in R*. CRC press.
16. Sim, C. H. (1987). A mixed gamma ARMA (1, 1) model for river flow time series. *Water Resources Research*, 23(1), 32-36.
17. Shanker, R., Sharma, S., & Shanker, R. (2013). A two-parameter Lindley distribution for modeling waiting and survival times data.
18. Woodward, W. A., Sadler, B. P., & Robertson, S. (2022). *Time series for data science: Analysis and forecasting*. CRC Press

## القطاع غير الرسمي : خصائصه وطرق قياسه

جمال قاسم حسن

خير اقتصادي

تاريخ استلام البحث: 2024/04/15

تاريخ قبول البحث: 2024/05/05

نشر البحث في العدد الثالث والعشرين: حزيران / يونيو 2024

رمز التصنيف ديوي / النسخة الالكترونية (Online): 2522-64X/519.5

رمز التصنيف ديوي / النسخة الورقية (Print): 2519-948X/519.5

## القطاع غير الرسمي : خصائصه وطرق قياسه

جمال قاسم حسن

خبير اقتصادي

### مقدمة

يُعتبر تقدير حجم القطاع غير الرسمي "Informal Sector" أحد أبرز التحديات الرئيسية التي تواجه الدول، وتتعلق هذه التحديات بخصائص هذا القطاع وكيفية قياس حجم العاملين في هذا القطاع، ومدى مساهمته في الناتج المحلي الإجمالي.

أشارت منظمة العمل الدولية، إلى أن العديد من الدول خاصة النامية منها لا تقوم بإجراء مسوحات منتظمة لقياس العمالة في القطاع غير الرسمي، ولا تقدير لحجم هذا القطاع وكيفية إدراجه ضمن الناتج المحلي الإجمالي. كما أشارت أيضا بتعدد منهجيات طرق قياس القطاع غير الرسمي، وتعتمد اختيار المنهجيات الدولية لقياس القطاع غير الرسمي على الأهداف المرجوة من المسح، حيث تتراوح هذه الأهداف بين البسيطة مثل الحصول على معلومات حول تطور عدد وخصائص العاملين في القطاع غير الرسمي، أو الحصول على معلومات مُفصلة حول خصائص المؤسسات المعنية، والأنشطة الرئيسية التي يتم تضمينها عند احتساب توليد الدخل في الحسابات القومية.

بلغت نسبة العاملين في القطاع غير الرسمي على مستوى العالم نحو 61.2 في المائة استنادا لأخر إحصائية صدرت عن منظمة العمل الدولية، وتعتبر دول أفريقيا من أكثر الدول التي لديها عمالة غير رسمية، تليها الدول العربية ودول آسيا والمحيط الهادئ. وفي ذات السياق، يستقطب القطاع الزراعي أكبر عمالة غير رسمية.

تهدف هذه الدراسة إلى التعرف على إحصاءات القطاع غير الرسمي وخصائصه ومنهجيات قياسه، إضافة إلى استعراض تجارب بعض الدول والوقوف على أبرز الإجراءات والتدابير المتخذة لتنظيم القطاع غير الرسمي.

### أولاً: تعريف القطاع غير الرسمي

أشارت منظمة الأمم المتحدة إلى أن القطاع غير الرسمي<sup>1</sup> يتضمن الوحدات التي تعمل في إنتاج السلع أو تقديم الخدمات غير الرسمية ويعمل بها عادة عدداً قليلاً من العاملين في حدود 10 عمال، وتعتمد علاقات العمل بها على الأعمال العرضية والموسمية أو القرابة أو العلاقات الشخصية والاجتماعية بدلاً من الترتيبات التعاقدية ذات الضمانات الرسمية.

كما حدد المؤتمر الدولي الخامس عشر لمنظمة العمل الدولية<sup>2</sup> المنعقد سنة 1993 القطاع غير الرسمي بأنه مجموعة من وحدات الإنتاج التي تتكون من شركات مملوكة للأسر عادة تكون شركات صغيرة الحجم غير مُدمجة في الاقتصاد الوطني. فيما تم في المؤتمر الدولي السابع عشر لمنظمة العمل الدولية، تحديد الطابع غير الرسمي من منظور الوظائف، وتعريف مفهوم العمل

<sup>1</sup> United Nations, Economic Commission for Africa (2014), "Operational Guidebook on Accounting for Informal Sector in National Accounts".

<sup>2</sup> International Labour Organisation (1993), "Resolution concerning statistics of employment in the informal sector, adopted by the 15th International Conference of Labour Statisticians".

غير الرسمي على أنه كل عمل مقابل أجر لا يتم تسجيله أو تنظيمه أو حمايته من خلال الأطر القانونية أو التنظيمية القائمة، حيث لا يمتلك العمال غير الرسميين أية عقود عمل، ولا يتمتعون بأي استحقاقات مضمونة قانونياً في إطار نظم الحماية الاجتماعية المتبعة. تتسم وحدات الإنتاج في القطاع غير الرسمي بالسمات المميزة للمنشآت المنزلية (القطاع العائلي)، حيث إن الأصول الثابتة والأخرى المستخدمة لا تنتمي إلى وحدات الإنتاج في حد ذاتها ولكن لأصحابها، ويتعين على مالكي هذه الوحدات جمع التمويل اللازم على نفقاتهم الخاصة ويكون مسؤولين بشكل شخصي عن أية ديون أو التزامات يتم تحملها خلال عملية الإنتاج. إضافة إلى دافع التهرب من دفع ضرائب أو اشتراكات الضمان الاجتماعي، بسبب عدم وجود دفاتر نظامية ولا سجلات تدل على قانونية هذا النشاط، والذي يعتبر انتهاكاً لقوانين العمل والتشريعات لكافة الأنشطة الاقتصادية التي يتم تداولها والتعامل بها.

## 1. هيكل العاملين في القطاع غير الرسمي

يُعتبر قياس حجم العاملين في القطاع غير الرسمي من أبرز التحديات التي تواجه الأجهزة الإحصائية، باعتبار أن هذا القطاع أحد المكونات الأساسية في الهيكل الاقتصادي للعديد من الدول النامية، ويُمثل شكلاً رئيساً من أشكال الإنتاج وموفرأ أساسياً لفرص العمل والدخل في كل من المناطق داخل الدولة سواءً الحضرية منها أو الريفية. ينقسم العاملين في القطاع غير الرسمي إلى ثلاثة شرائح، الشريحة الأولى تعتبر أنشطتهم محصور في القطاع غير الرسمي مثل الباعة الجائلين أو العاملين في التدريس خارج منظومة التعليم أو في خدمات التمريض في المنازل (مشار إليها تحت البند 1). أما الشريحة الثانية فهم العاملين في مؤسسات قانونية ومسجلين على أساس عمالة غير رسمية كمزارعين الموسميّين، والمرشدين في القطاع السياحي. الشريحة الثالثة فهي معنية بالأسر (القطاع العائلي) العاملين في إنتاج سلع أو تقديم خدمات من منازلهم، جدول رقم (1).

### جدول رقم (1)

#### تصنيف وظائف العاملين في القطاع غير الرسمي على أساس الوظائف

على أساس الوظائف الرئيسية والثانوية	على أساس الوظيفة الرئيسية فقط
1- العاملون الذين تكون وظائفهم الرئيسية والثانوية في القطاع غير الرسمي (الباعة الجائلين، التدريس خارج منظومة التعليم، خدمات التمريض،...الخ).	1- العاملون الذين وظيفتهم الأساسية في القطاع غير الرسمي أو القطاع الرسمي.
2- العاملون الذين تكون وظيفتهم الأساسية أو الثانوية في أحد القطاعات الرسمية (عاملين في أحد القطاعات الرسمية كالعامل ضمن شركات قطاع السياحة والزراعة والمسجلين كعمالة غير رسمية).	
3- العاملون ضمن القطاع العائلي (الإنتاج السلي داخل العائلة مثل أنواع التجارة والخدمات داخل المنزل).	2. العاملون الذين وظيفتهم الرئيسية ضمن القطاع العائلي وهم عمالة غير رسمية.
= إجمالي الأشخاص العاملين في القطاع غير الرسمي	

المصدر: منظمة العمل الدولية، دليل تقدير العاملين في القطاع غير الرسمي.

## 2. طبيعة القطاع غير الرسمي

يشمل القطاع غير الرسمي العديد من الأنشطة، وتتراوح هذه الأنشطة ما بين الحرف التقليدية وغير الرسمية مثل الباعة الجائلين، وغيرها من الأنشطة التي ترتبط ببيع السلع وتقديم الخدمات كأنشطة تجهيز الأغذية أو تقديم الخدمات لعامة الشعب، حيث أن هذه الأنشطة لا تتطلب رأسمال كبير وتعتبر أنشطة صغيرة ذات إنتاج ضعيف، إضافة إلى العمالة الموسمية في بعض المؤسسات المرخصة<sup>3</sup>، كما لا تتطلب أنشطة القطاع غير الرسمي قدراً معيناً من المهارات ورأس المال خاصة للوحدات ذات الإنتاجية الأعلى مثل التصنيع والخياطة وإصلاح السيارات. أما أسباب الانخراط في القطاع غير الرسمي، فيعزى إلى حالة التعطل عن العمل، أو من الذين يواجهون عدم توفر وظائف ملائمة لقدراتهم أو بهدف زيادة مستويات الدخل. تجدر الإشارة، إلى أن الغالبية العظمى من أنشطة القطاع غير الرسمي توفر السلع والخدمات، ويتم إخفاؤها عمداً عن السلطات العامة لتجنب دفع ضرائب الدخل أو ضريبة القيمة المضافة أو أية ضرائب أخرى، أو دفع اشتراكات الضمان الاجتماعي، أو التقيد بالإجراءات الإدارية، أو عدم تلبية المتطلبات القانونية مثل الحد الأدنى للأجور، والحد الأقصى لعدد ساعات العمل، ومعايير السلامة أو الصحة، وما إلى ذلك. تشمل أنشطة القطاع غير الرسمي على الأنشطة التالية:

- أ- **الأنشطة غير المرخصة التي يقوم بها الأفراد، وتشمل ما يلي:**
- أنشطة يقوم بها أشخاص غير مصرح لهم (مثل ممارسة الطب غير المرخص، خدمات التمريض والعناية بكبار السن غير المرخصة).
  - نشاط التدريس، تقديم الدروس خارج نطاق المدارس والمراكز المختصة.
  - أنشطة التجارة التي يقوم بها الأشخاص (الباعة الجائلين).
- ب- **أنشطة مرخصة، لكن لا يمكن ملاحظتها من الجانب الإحصائي بسبب عدم توفر عنوان مستقر أو ثابت (مثل على ذلك: سيارات الأجرة والمطاعم المتنقلة).**
- ج- **إنتاج القطاع العائلي للسلع وتقديم الخدمات غير المرخصة لاستخدامها كمصدر دخل، وتشمل:**
- السلع: التي يتم انتاجها وتزويد الجهات أخرى بها، بهدف استخدامها في عمليات الإنتاج أو للاستهلاك الخاص، بما يشمل:
    - إنتاج المحاصيل الزراعية ومعالجتها وتخزينها، وجمع بعض أنواع الفواكه وحفظها بالتجفيف وغيرها من المحاصيل.
    - جمع قطع الأخشاب والحطب، والصيد بما فيها صيد الأسماك.
    - إنتاج السلع الأولية التي تدخل في عمليات لإنتاج منتجات أخرى.
    - إنتاج الحبوب والدقيق (الطحين).
    - إنتاج ودباغة الجلود<sup>4</sup>.
    - إنتاج اللحوم ومنتجاتها، وما إلى ذلك.
    - إنتاج الألبان مثل الزبدة أو الجبن.
    - إنتاج الحلي.

<sup>3</sup> قطاع السياحة والزراعية والبناء والتشييد والتجارة تستقطب عمالة موسمية.

<sup>4</sup>المقصود بها الدباغة، حيث يتم استخدام الجلود المعالجة "المدبوغة" في صناعات العديد من السلع النهائية مثل الأحذية والحقائب.

- إنتاج الملابس والخياطة والنسيج.
- صناعة الفخار والأواني والأثاث والمفروشات،...إلخ.
- الخدمات، وتشمل ما يلي:
  - الخدمات الزراعية التي تقدم للقطاع العائلي والتي تتعلق بالحدائق والبساتين.
  - خدمات تقدم للزبائن بناءً على طلبهم (كالطبخ في المنازل).
  - خدمات العناية بالأطفال والرضع في المنازل.
  - الخدمات التي يتم تقديمها إلى جهات أخرى وتستخدم في إنتاج سلع (كتوفير عمالة وتوظيفها بأجر).
  - خدمات السلع التي تستحوذ على المعرفة كالقيام بالبحوث وتحليل البيانات الاقتصادية أو تكوين رأس المال باستثناء المنتجات التي تنتجها الأسر لاستخدامها الخاص.
  - خدمات الإسكان والمنازل من قبل المالكين (الإيجارات المحتسبة لأصحاب المنازل).
  - خدمات المنازل المأجورة (خارج الأطر القانونية).
- د- الأنشطة الأخرى ضمن القطاع العائلي، بما يشمل البنود التالية:
  - توسعة المباني السكنية التي تتم من خلال صغار المقاولين لغرض استخدامها في الاستثمار
  - إنتاج بعض الأصول الثابتة مثل بناء المساكن وتطوير البرمجيات والتنقيب عن المعادن.
  - المتاجرة من البيوت لبعض السلع التي يتم عرضها للبيع.
  - بعض أنواع أنشطة العناية الشخصية التي تقدم للزبائن (مثل صالونات العناية والتزيين).

### 3. طرق حصر العمالة في القطاع غير الرسمي<sup>5</sup>

تعتمد الجهات المعنية على استبيانات لقياس حجم القطاع غير الرسمي بما في ذلك العاملين في هذا القطاع سواء كانوا في المناطق الحضرية أو الريفية. في حين تقتصر بعض الدول على تنفيذ المسوحات في المناطق الحضرية الكبرى فقط، بسبب صعوبة إجراء المسوحات في بعض المناطق الريفية. إضافة إلى ذلك يعتمد تقدير القطاع غير الرسمي على طبيعة ونوعية البيانات المطلوب توفرها من هذا القطاع وعلى الموارد المالية والبشرية المتوفرة، وعلى الأهداف المرجوة، حيث تنحصر هذه الأهداف بين الحصول على معلومات حول أعداد العاملين في القطاع غير الرسمي، أو أكثر تفصيلاً كالحصول على معلومات حول خصائص المؤسسات المعنية، والأنشطة الرئيسية التي يتم تنفيذها، وعدد الموظفين، وحجم الإنتاج، وتكوين الرأسمالي...إلخ.

<sup>5</sup> United Nations, Economic Commission for Africa (2014), "Operational Guidebook on Accounting for Informal Sector in National Accounts."

### 1.3 مسح العمالة في القطاع غير الرسمي

يتم حصر حجم العمالة في القطاع غير الرسمي، من خلال المسح المباشر، وذلك بالإعتماد على استبيان يتعلق بالعاملين في هذا القطاع، حيث يتم تحديد حجم العينة والتعرف على المؤشرات التفصيلية المتعلقة بالبيانات التشغيلية والنقدية ومستوى الدخل والانفاق على المواد الخام المستخدمة في العملية الانتاجية، وقيمة الاستهلاك الوسيط وتعويضات العاملين. فيما يلي نستعرض بعض البنود التي يجب أن تأخذ في الاعتبار عند اعداد استبيان القطاع غير الرسمي.

#### (1) أسئلة تتعلق بالعاملين

1. توفير المعلومات الأساسية حول العاملة:

- النوع (ذكر أو أنثى) والعمر والحالة الاجتماعية.
- تكوين الأسرة (وجود أطفال).
- التحصيل العلمي.
- نوع التدريب المهني الذي تم تلقيه.
- فترة وأسباب التعطل.

2. أسئلة توضيحية حول طبيعة العمل في القطاع المعني:

- الموقع الجغرافي (حضري أو ريفي)
- نوع العقد المبرم (كتابي أو شفهي).
- طبيعة العمل (ثابت، عارض،..الخ).
- مكان العمل (مكان ثابت معروف، داخل المسكن،..الخ).
- نوع العمل (أساسي أو ثانوي).
- مدة العمل (من تاريخ البدء أو كما هو متفق عليه).

3. الأجر المُتلقى:

- الراتب.
- المنافع الأخرى إن وجدت.
- استحقاق مكافأة نهاية الخدمة إن وجدت.
- ساعات العمل.

#### (2) أسئلة تتعلق بالمنشأة

1. بيانات عن العمالة في المنشأة:

- عدد العاملين.
- طبيعة عملهم.
- اشتراكات التأمين الاجتماعي لأصحاب العمل، إذا كانت متوفرة.
- قيمة الضرائب، إن وجدت.
- المكافآت والبدلات الأخرى، إذا كانت متوفرة.

2. أسئلة تهدف إلى التعرف على حجم الإنتاج والمبيعات

- حجم الإنتاج والخدمات المقدمة.
- قيمة المبيعات والإيرادات الأخرى.

3. أسئلة تهدف إلى التعرف على حجم الانفاق على المواد الخام والمخزون (لقياس الاستهلاك الوسيط)، إضافة إلى المعلومات التالية:

- شراء البضائع لإعادة بيعها حسب المنتج (الكمية والقيمة الإجمالية).
- نفقات أخرى (وقود وماء وكهرباء ومساحات إيجارية).
- تكلفة المواد الخام.

4. أسئلة تخص تكوين رأس المال، وقيمة الأصول الثابتة والنفقات الرأسمالية، بما يشمل:
- قيمة الأصول الثابتة في بداية العام.
  - بنود الإنفاق الرأسمالي على الأصول الثابتة الجديدة والمستعملة خلال العام.
  - قيمة الأصول الثابتة المُباعة خلال العام.
  - من المهم تحديد ما إذا كان الأصل مملوكاً أو مستأجراً (فردياً أو مشتركاً).
  - الأراضي والمباني والمنشآت الأخرى ومعدات النقل والآلات والمعدات والأثاث.
5. إدراج الأصل الثابت في بداية السنة وتاريخ الشراء والقيمة الحالية، حيث يتم تقدير الإهلاك باستخدام المعلومات المتعلقة بالبيانات المدرجة والقيم الحالية إذا توفر.
6. أسئلة حول أية أوجه أخرى للدعم متوفرة.

### 2.3 عملية جمع بيانات القطاع غير الرسمي

أعتمدت منظمة العمل الدولية والمنظمات الدولية ثلاث طرق لجمع بيانات القطاع غير الرسمي بما يشمل العاملين في هذا القطاع:

#### 1. الأسلوب المباشر لجمع بيانات العمالة غير الرسمية

تعتمد طريقة الأسلوب المباشر على المسوحات الاستقصائية المعدة لغرض تقدير حجم القطاع غير الرسمي وقياس العمالة في هذا القطاع، حيث يتكون الأسلوب المباشر من أربعة طرق معنية في قياس القطاع غير الرسمي (المسوحات الخاصة، ومسوحات الأسر، واستطلاعات المؤسسات، والمسح المختلط).

#### 2. الأسلوب غير المباشر لقياس القطاع غير الرسمي

يعتبر هذا الأسلوب ذو طبيعة اقتصادية في الغالب، ويجمع بين المتغيرات الاقتصادية المختلفة ومجموعة من الافتراضات المتعلقة بتقدير إنتاج النشاط الاقتصادي، إلا أنه لا يُمكن من الحصول على المعلومات والمؤشرات الأخرى ذات الصلة بالقطاع غير الرسمي مثل مستوى الحماية الاجتماعية، والوصول إلى التعليم، والصحة، والائتمان، والاختلافات في الأجور وظروف العمل وانتشار الفقر. تقوم طريقة الأسلوب غير المباشر على عدد من الطرق، وهي منهجية التباين، والأسلوب النقدي، وطريقة المدخلات المادية.

#### 3. المنهجية النموذجية لتقدير القطاع غير الرسمي

تتضمن المنهجية النموذجية لتقدير القطاع غير الرسمي، على استخدام نماذج هيكلية لربط المتغيرات المعنية بالقطاع غير الرسمي (التي يمكن أن تشمل على الرواتب والبطالة)، وتشير جميع هذه المؤشرات إلى العملة المتداولة في الاقتصاد والنتائج المحلي الإجمالي الحقيقي بالمؤشرات المرصودة لذلك. أما المكونان الرئيسيان لهذه المنهجية فهما معادلتين، المعادلة القياسية والمعادلة الهيكلية، حيث يوفران حجم الاقتصاد غير الرسمي والمؤشرات المدرجة في التقديرات.

في هذا الإطار، أشار الخبراء الإحصائيون إلى صعوبة استخدام المنهجية النموذجية لتجميع الحسابات القومية<sup>6</sup>، حيث تعاني هذه الطريقة من تحديات تلقي بظلالها على تقدير مؤشرات الحسابات القومية، ويعزى ذلك إلى أن تقديرات القطاع غير الرسمي قد تكون مبنية على معادلات هيكلية غير دقيقة وبالتالي يتم تضمينها في الحسابات القومية استناداً على هذه التقديرات وليس من خلال جمع البيانات التي ذكرت في الطريقتين السابقتين، وعليه يكون تقدير الناتج المحلي الإجمالي غير دقيق.

<sup>6</sup> Eurostat (2010), "Essential SNA: Building the basics".

## ثانياً: إحصاءات العمالة في القطاع غير الرسمي

أشارت منظمة العمل الدولية، إلى أن أكثر من 2 مليار شخص تبلغ أعمارهم 15 سنة وما فوق من السكان العاملين في العالم يعملون في القطاع غير الرسمي وفق أحدث التقديرات<sup>7</sup>، حيث يمثلون حوالي 61 في المائة من العمالة وتختلف نسبة العمالة غير الرسمية باختلاف المناطق. تسجل الدول الأفريقية أعلى حصة من العمالة غير الرسمية على مستوى العالم بنسبة تبلغ نحو 85.8 في المائة من إجمالي العمالة، في حين تبلغ نسبة العمالة في دول آسيا والمحيط الهادئ حوالي 68.2 في المائة. بشكل عام يعتبر القطاع الزراعي من أكثر القطاعات الاقتصادية التي تتسم بهيمنة العمالة غير الرسمية في جميع المناطق.

## (2) جدول رقم

## حصة العمالة غير الرسمية من إجمالي العمالة

(تقديرات بحسب أحدث بيان صادر عن منظمة العمل الدولية) (%)

نسبة العاملين في القطاع غير الرسمي حسب أنواع العمل غير الرسمي <sup>8</sup>				الأقاليم
إجمالي العاملين في القطاع غير الرسمي (1)+(2)+(3)=	العاملين ضمن القطاع العائلي (3)	العاملين في القطاع الرسمي* (2)	العاملين في القطاع غير الرسمي (1)	
85.8	4.3	5.5	76.0	أفريقيا
40.0	2.7	7.9	29.3	الأمريكتين، منها
13.3	0.2	1.7	11.4	الولايات المتحدة الأمريكية
68.6	0.2	7.5	60.9	الدول العربية
68.2	2.5	6.9	58.8	آسيا والمحيط الهادئ
25.1	0.5	5.3	19.4	أوروبا وآسيا الوسطى
61.2	2.5	6.7	51.9	العالم

Source: ILO (2019), "Women and men in the informal economy: A statistical picture", third edition.

(\*) العاملون ضمن المؤسسات الرسمية ويعملون كذلك في القطاع غير الرسمي.

في نفس السياق، أشارت منظمة العمل الدولية<sup>9</sup> إلى أن نسبة العمالة في القطاع غير الرسمي بعد استبعاد القطاع الزراعي تبقى مرتفعة في ثلاث مناطق جغرافية وهي أفريقيا والدول العربية وآسيا والمحيط الهادئ، في حين تنخفض نسبة العمالة في القطاع غير الرسمي في البلدان المتقدمة ذات الدخل المرتفع مثل الولايات المتحدة الأمريكية.

تظهر إحصاءات العمالة غير الرسمية<sup>10</sup> نمواً سريعاً في حجم هذا القطاع خلال السنوات الأخيرة خاصة في العديد من البلدان النامية، حيث يسهم هذا القطاع بنحو 41 في المائة من الناتج المحلي الإجمالي في أفريقيا وحوالي 31 في المائة في آسيا، ونحو 29 في المائة في أمريكا

<sup>7</sup> ILO (2019), "Women and men in the informal economy: A statistical picture", third edition.

<sup>8</sup> ILO (2019), Op cit.

<sup>9</sup> ILO (2019), "Women and men in the informal economy: A statistical picture", third edition.

<sup>10</sup> ILO, Database Labor force, World Bank.

اللاتينية<sup>11</sup>، ويستقطب أكثر من ثلثي العمالة في الدول النامية وتقدر بنحو 69.6 في المائة. أما في الدول العربية، فقد بلغت مساهمة القطاع غير الرسمي حوالي 68.6 في المائة من إجمالي العمالة في الدول التي توفرت لديها بيانات. وبلغت حصة القطاع الزراعي الذي يعتبر أكثر القطاعات استقطاباً للعمالة غير الرسمية نحو 95.6 في المائة من إجمالي العاملين، جدول رقم (2) ورقم (3). في حين تنخفض العمالة غير الرسمية في قطاع الخدمات إلى نحو 57.3 في المائة من إجمالي العمالة في هذا القطاع.

أما بالنسبة للدول العربية، فقد سجلت العمالة غير الرسمية في تونس نحو 44 في المائة من إجمالي العمالة، بينما سجلت نحو 36.3 و25 في المائة في كل من المغرب وليبيا على التوالي<sup>12</sup>.

### (3) جدول رقم

حصة العمالة في القطاع غير الرسمي من إجمالي العمالة حسب القطاعات الاقتصادية (نسبة % تقديرات بحسب أحدث بيان صادر عن منظمة العمل الدولية)

القطاعات الاقتصادية			الأقاليم
الخدمات	الصناعية	الزراعية	
70.2	77.4	97.9	أفريقيا
35.5	38.4	77.5	الأمريكتين، منها:
18.4	17.6	52.3	الولايات المتحدة الأمريكية
57.3	80.9	95.6	الدول العربية
54.1	68.8	94.7	آسيا ومنطقة المحيط الهادي
20.2	21.9	71.6	أوروبا وآسيا الوسطى
47.2	57.2	93.6	العالم

Source: ILO (2019), Women and men in the informal economy: A statistical picture, third edition.

أما حسب الفئات العمرية، فقد أشارت منظمة العمل الدولية إلى أن نسبة الشباب العاملين في القطاع غير الرسمي تبلغ نحو 77.1 في المائة من إجمالي العمالة في هذه الفئة العمرية على مستوى العالم. حيث تنخفض هذه النسبة لدى الشباب من إجمالي الفئات العمرية العاملة في القطاع غير الرسمي في أوروبا وآسيا الوسطى إلى نحو 35.7 في المائة<sup>13</sup>، في حين ترتفع إلى ما يفوق نسبة 94 في المائة في الدول الإفريقية، بينما تبلغ فئة الشباب العاملين بالقطاع غير الرسمي في الدول العربية نحو 85.1 في المائة من إجمالي العاملين، جدول رقم (4).

<sup>11</sup> IMF, René Piché (2017), The Informal Sector Some Measurement Challenges, Fourth Meeting of the Steering Committee of the Arab Statistics Initiative "Arabstat" November 8-9.

<sup>12</sup> صندوق النقد العربي (2017). "دراسة حول إحصاءات القطاع غير الرسمي في الدول العربية"، الاجتماع الرابع لمبادرة الإحصاءات العربية.

<sup>13</sup> منظمة العمل الدولية، قاعدة بيانات إحصاءات العمل.

## جدول رقم (4)

حصة العمالة غير الرسمية من إجمالي العمالة حسب الفئات العمرية  
(نسبة % تقديرات بحسب أحدث بيان صادر عن منظمة العمل الدولية)

القطاع						الأقاليم
65+	55-64	35-54	30-34	25-29	الشباب (24-15)	
96.0	84.2	79.7	83.1	85.5	94.9	أفريقيا
54.4	42.1	39.9	37.6	37.0	46.2	الأمريكيتين
71.7	62.8	56.9	61.0	68.8	85.1	الدول العربية
86.3	72.4	63.8	67.3	70.8	86.3	آسيا والمحيط الهادي
40.8	21.5	20.0	21.8	26.4	35.7	أوروبا وآسيا الوسطى
77.9	59.2	55.7	58.6	62.6	77.1	العالم

المصدر: منظمة العمل الدولية and Men in the Informal Economy: A Statistical Picture Women، الطبعة الثالثة، 2017

## ثالثاً: تجارب الدول في قياس القطاع غير الرسمي

1. تجارب بعض الدول غير العربية<sup>14</sup>

يلعب القطاع غير الرسمي دوراً ورئيساً مهماً في اقتصادات الدول النامية وخاصة دول إفريقيا، وأشارت الإحصائيات المتاحة لدى منظمة العمل الدولية أن هذا القطاع يعتبر مصدر دخل بالنسبة لشريحة كبيرة من السكان. لذلك من الضروري قياس هذا القطاع بالنسبة لمتخذي القرار في هذه الدول، إلا أنها تواجه عدة تحديات، بسبب نقص الموارد المالية والبشرية، حيث إن معظم البلدان في القارة الأفريقية غير قادرة على إجراء المسوحات السنوية اللازمة لتقدير القطاع غير الرسمي.

**الكاميرون**، يعتمد الاقتصاد الكاميروني على القطاع الزراعي والصادرات من هذا القطاع خاصة السلع الغذائية من الكاكاو، والبن، والقطن، والعديد من المنتجات الزراعية، ويتأثر الاقتصاد الكاميروني بطريقة مباشرة بتراجع أسعار هذه السلع في الأسواق العالمية. في عام 1994 انهارت عائدات الصادرات نتيجة انخفاض أسعار السلع من بعض المنتجات الغذائية، وقد أدى ذلك إلى نزوح العديد من السكان من الريف إلى المدن الرئيسية، حيث أدى ذلك إلى زيادة معدلات البطالة، وساهم ذلك بارتفاع كبير في القطاع غير الرسمي، الذي يتميز به اقتصاد الكاميرون. قامت الحكومة من خلال المعهد الوطني للإحصاء الكاميروني (NIS) بإجراء مسح من ثلاث مراحل لقياس مدى انتشار ظاهر العمل غير الرسمي، المرحلة الأولى خصصت لقياس حجم العمالة، في حين ركزت المرحلة الثانية على تقدير القطاع غير الرسمي، أما المرحلة الثالثة فكانت موجهة نحو استهلاك الأسر. وقد تم إجراء المسح على مستوى الوطني بهدف تزويد متخذي القرار بمجموعة من المؤشرات الخاصة بظروف سوق العمل ودخل العمل والقطاع غير الرسمي ومدى مساهمته في الاقتصاد من حيث القيمة المضافة، والعمالة وخلق فرص العمل وبالتالي زيادة الثروة في الكاميرون.

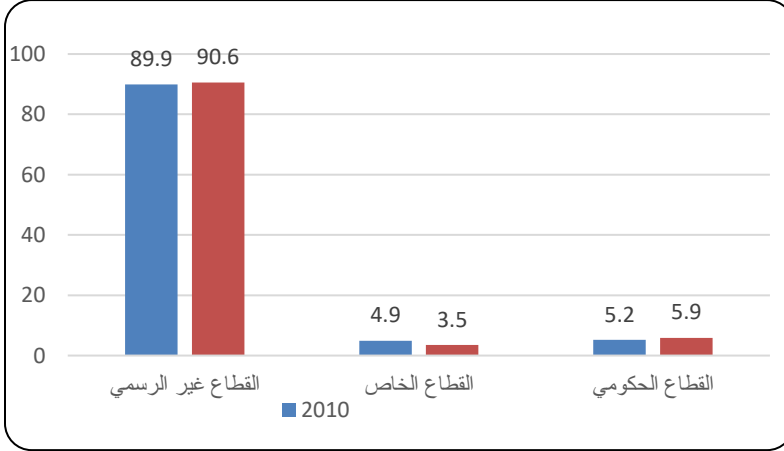
تم تنفيذ المرحلتين الثانية والثالثة بين عامي 2005 و2010 بعنوان "مسح العمالة والقطاع غير الرسمي" لقياس مدى التقدم نحو أهداف الإنمائية الألفية، وهو عبار عن مسح إحصائي

<sup>14</sup> United Nations, Economic Commission for Africa (2014), "Operational Guidebook on Accounting for Informal Sector in National Accounts"

بمرحلتين، المرحلة الأولى (مسح العمالة) تم تنفيذه للتعرف على طبيعة العمالة، حيث ركز المسح على الأفراد الذين تبلغ أعمارهم 10 سنوات فأكثر في الاقتصاد الكاميروني. أما المرحلة الثانية فخصص لتقييم الأنشطة القطاع غير الرسمي غير الزراعي (مسح القطاع غير الرسمي)، فقد تم مسح القطاع غير الرسمي لشركات ذات المستوطنات العشوائية في الكاميرون، وتسمى "وحدات الإنتاج غير الرسمية"، نظرًا لأن العديد منها ليس لديها مباني خاصة بها (كورش عمل، أو متجر،... إلخ). خلال المسح الذي أجري، تم تعريف القطاع غير الرسمي بأنه جميع وحدات الإنتاجية التي ليس لديها رقم تعريفى لدافع الضرائب وليس لديها نظام محاسبة رسمي. أظهرت نتائج المسح المختلطة للحسابات القومية في الكاميرون، أن مساهمة القطاع غير الرسمي في الناتج المحلي الإجمالي بلغ حوالي 37.3 في المائة عام 2010 و38.4 في المائة عام 2015. فيما يتعلق بالعمالة غير الرسمية، فقد مثل حوالي 89.9 في المائة من إجمالي العمالة في هذا القطاع عام 2010 وحوالي 90.6 في المائة من إجمالي العمالة عام 2015.

### شكل رقم (1)

توزيع العمالة في الاقتصاد الكاميروني حسب أحدث مسح ميداني



المصدر: United Nations, Economic Commission for Africa , "Operational Guidebook on Accounting for Informal Sector in National Accounts"

**غانا**، يعتبر قطاع الزراعة أكبر القطاعات مساهمة في الناتج المحلي الإجمالي، وتساهم الأنشطة الاقتصادية المنزلية غير الزراعية بدرجة كبيرة في الناتج المحلي الإجمالي، ولكن المشكلة تكمن في عدم احتسابها بالطريقة السليمة في الناتج المحلي الإجمالي. أشارت إحصاءات منظمة العمل الدولية، أن حصة القيمة المضافة للمؤسسات المنزلية غير الزراعية من إجمالي القيمة المضافة للقطاع غير الرسمي في غانا بلغت نحو 62.7 في المائة حسب آخر إحصائية.

تبنت دائرة الإحصاء في غانا المعيار والارشادات الدولية بالاعتماد على قرارات منظمة العمل الدولية في المؤتمر الدولي الخامس عشر لخبراء إحصاءات العمل المتعلقة، بأن يتم جمع البيانات الخاصة بالمسوحات الخاصة بالنشاط الاقتصادي للقطاع غير الرسمي المستخدمة

في مسح مستويات المعيشة في غانا، وقد تم تقدير الأنشطة غير الزراعية للقطاع غير الرسمي باستخدام بيانات مسح مستويات المعيشة، بما يشمل:

- مؤسسات القطاع غير الرسمي وهي المؤسسات الخاصة والشركات المملوكة للأفراد أو الأسر المعيشية التي لم يتم تشكيلها ككيانات قانونية منفصلة بشكل مستقل عن مالكيها، والتي لا يوجد لها حسابات كاملة
- التركيز على جميع السلع أو الخدمات المنتجة أو على الأقل بعضها معدة للبيع أو المقايضة حسب نظام الحسابات القومية 2008
- تقدير حجم القطاع غير الرسمي من إجمالي العمالة وتحديدهم وفقاً للظروف الوطنية. تم اعداد استبيان خاصة لهذا الغرض، بهدف تقدير مختلف المؤشرات والتعرف مختلف جوانب الظروف المعيشية لهذا القطاع، بما في ذلك استهلاك الأسرة والإنفاق، والصحة، والتعليم، والعمالة، والهجرة، وظروف الإسكان، والمشاريع المنزلية غير الزراعية والأنشطة الزراعية والتحويلات والمدخرات والائتمان والأصول.

## 2. تجارب الدول العربية

تباين كبير في مساهمة القطاع غير الرسمي إلى الناتج المحلي الإجمالي في بعض الدول العربية التي توفرت حولها إحصاءات، حيث بلغ في كل من ليبيا ومصر وتونس نحو 24.5 و 23.1 و 22.0 في المائة على التوالي عام 2019، في حين بلغت مساهمة هذا القطاع في المغرب حوالي 11.3 في المائة. إضافة إلى ذلك أشارت النتائج إلى أن نسبة العاملين في القطاع غير الرسمي في فلسطين بلغ نحو 51 في المائة عام 2019، بينما بلغ حوالي في مصر 63.4 في المائة عام 2018. بينما بلغ نسبة العاملين في موريتانيا حوالي 89.4 حسب أحدث إحصائية لمنظمة العمل الدولية عام 2017، جدول رقم (5).

### جدول رقم (5)

نسبة العاملين في القطاع غير الرسمي في بعض الدول العربية 2010 - 2019

السنة الدول	2019	2018	2017	2016	2015	2014	2013	2012	2011	2010
تونس	..	..	..	..	..	..	44.0	..	..	..
السودان	..	..	..	..	..	..	..	..	77.3	..
فلسطين	50.6	50.8	51.9	51.1	52.0	50.6	51.9	50.7	51.5	51.4
ليبيا	..	..	..	..	..	..	..	25.0	..	..
مصر	..	63.4	52.9	50.5	47.2	..	42.9	41.6	43.8	43.1
المغرب	..	..	..	..	..	..	36.0	..	..	..
موريتانيا	..	..	89.4	..	..	..	..	..	..	..
اليمن	..	..	..	..	..	68.4	..	..	..	..

المصدر: منظمة العمل الدولية.

فيما يلي نستعرض تجارب بعض الدول العربية في كيفية تقدير وقياس القطاع غير الرسمي<sup>15</sup>: الجزائر، يتم تقدير إحصاءات القطاع غير الرسمي عن طريق بيانات مسوحات التشغيل لدى الأسر المعيشية، وتبني الطريقة غير مباشرة المشار إليها في دليل الأمم المتحدة. أما فيما يخص

<sup>15</sup> أمحمد موعش (2017). " دراسة حول إحصاءات القطاع غير الرسمي في الدول العربية"، الاجتماع الرابع لمبادرة الإحصاءات العربية، صندوق النقد العربي

مسح التشغيل فيمكن من خلاله استنباط حجم العمالة في القطاع غير الرسمي، وذلك بناءً على التحليل المتعدد المعايير:

- عدم وجود أي تسجيل إداري للمستخدمين وأصحاب المهن الحرة الذين يعتبرون كوحدات إنتاجية. وإذا لم تتوفر هذه الفئة من المشتغلين على أي نوع من تسجيل إداري (سجل تجاري، رخصة القيام بنشاط، بطاقة الحرفي،...الخ) فتعتبر هذه الوحدات الإنتاجية غير شرعية.
- عدم الانخراط في الضمان الاجتماعي سواء كان ذلك للأجراء أو المستخدمين وأصحاب المهن الحرة.
- غياب محاسبة كاملة على مستوى المستخدمين وأصحاب المهن الحرة.
- بالنسبة لحجم الوحدات الإنتاجية، فيعتبر غير معلوم باعتبار أن الوحدات التابعة للقطاع غير الرسمي عادة ما تكون ذات حجم صغير (أقل من عشرة عمال) مقارنة مع تلك التي نجدتها في القطاع الرسمي.

تونس، استناداً للمسح الوطني حول المؤسسات الصغيرة، يتحصل المعهد الوطني للإحصاء على مسح بالعينة دورية خماسية (كل خمس سنوات)، وذلك بالاتصال المباشر مع أصحاب المؤسسات الصغيرة، ومن ثم تقديرات حول الجزء الأهم في القطاع غير الرسمي وفي المؤسسات التي لا تمتلك محاسبة، حيث يتم تقدير مساهمة القطاع غير الرسمي استناداً للمسح حول المؤسسات الصغيرة بدورية خماسية وذلك عن طريق جملة من المؤشرات: الإنتاج، والقيمة المضافة، والاستهلاك، والتشغيل، والأجور والاستثمار.

السودان، تعتبر وزارة العمل والقوى العاملة الجهة المخولة لإعداد إحصاءات القطاع غير الرسمي، ولم تقم الوزارة بعمل تقديرات لإحصاءات القطاع غير الرسمي، يعزى ذلك صعوبة قياس القطاع غير الرسمي لعدم وجود مسوحات في هذا المجال، وعدم توفير الدعم الفني والمادي لإجراء المسوحات.

فلسطين، يقوم الجهاز المركزي للإحصاء الفلسطيني، بتضمين الإحصائيات حول القطاع غير الرسمي ضمن خطته الرسمية لإعداد الإحصاءات الرسمية لفلسطين منذ عام 1994، ويعطي الجهاز أولوية لقطاع البناء والتشييد والنقل والاتصالات والتخزين وذلك بسبب انتشار وحدات القطاع غير الرسمي ضمن هذه الأنشطة بشكل بارز، إضافة إلى أهميتها للاقتصاد الفلسطيني ومساهمتها الكبيرة في الناتج المحلي الإجمالي. لذا يقوم الجهاز بإجراء العديد من المسوح المتخصصة في هذه الأنشطة المشار إليها مثل مسح الأبنية القائمة ومسح النقل غير الرسمي.

ليبيا، يعتمد على المنهج الوصفي والتحليلي لقياس حجم الاقتصاد غير الرسمي باستخدام الطرق والأساليب المعتمدة وهي:

- طريقة مباشرة، تعتمد على إعداد استبيان كأداة للحصول على البيانات اللازمة لقياس القطاع غير الرسمي في ليبيا.
- طريقة غير مباشرة، تعتمد على البيانات المالية والنقدية المنشورة رسمياً لتقدير وقياس حجم الاقتصاد غير الرسمي.

فيما يتعلق بالمعايير المتبعة في تصنيف القطاع غير الرسمي، فيتم من خلال المعايير التي يتم تصنيف النشاط محل البحث بأنه نشاط غير رسمي:

- حسب المنشأة أو الوحدة الاقتصادية، حيث إن تصنف الوحدة الاقتصادية تعتبر غير رسمية في حالة عدم التزامها بالتشريعات والقوانين المنظمة لمجال عملها.
- حسب نوعية تشغيل العمالة، حيث يصبح النشاط غير المنظم إذا كانت العمالة غير مسجلة ولا تحظى بالحماية التي يوفرها قانون العمل واللوائح المنظمة له وقانون التأمين الاجتماعي.
- حسب موقع النشاط والمتعاملين به، إذا كان موقع النشاط ثابت ومرخص، متجول ومتنقل أو يتم داخل المنزل أو في الأقبية والأسطح.
- حسب مستوى دخل العاملين بالنشاط ومدى استمراره.
- القدرة على التنقل بين الأنشطة الاقتصادية المختلفة.

مصر، يتم تقدير إحصاءات القطاع غير الرسمي في الحسابات القومية باستخدام مسوحات خارج المنشأة والتعداد الاقتصادي ومسح القوى العاملة لكافة الأنشطة الاقتصادية غير الرسمية التي تقع في حدود الإنتاج، ويمكن اعتبارها جزء من الإنتاج القومي أو الاستهلاك القومي وكذلك مسح نفقات ودخل الأسر. إضافة إلى ذلك يتم إجراء مسح دوري وفقاً لمعيار حجم العمالة في مجال الإحصاءات الصناعية والتجارية.

المغرب، يعتمد المعايير الدولية والتوصيات الصادرة عن منظمة العمل الدولية في مجال تعريف وقياس القطاع غير الرسمي في قياس هذا القطاع، حيث يتم الحصول على البيانات على أساس مسح بالمعاينة لوحدات الإنتاج غير الرسمية، ويتم إجراء هذا المسح بصفة دورية عند تحيين (تعديل) سنة الأساس للمحاسبة الوطنية. كما يتم تقدير إحصاءات القطاع غير الرسمي بالاعتماد على تعميم النتائج المحصل عليها من المسح. في حين يتم ادراج بيانات القطاع غير الرسمي في الحسابات القومية من خلال تصميم استمارة البحث (استبيان المسح)، حيث يتم توفير المعطيات الضرورية للمحاسبة الوطنية لإعداد حسابات الإنتاج والتوزيع الأولي للقطاع غير الرسمي.

#### رابعاً: منهجية تقدير القطاع غير الرسمي في منظومة الحسابات القومية

يوصي نظام الحسابات القومية 2008، أن يتم تجميع جدولين تكميلين للقطاع غير الرسمي، بحيث يغطي أحدهما الإنتاج وتوليد الدخل والآخر يغطي العمالة<sup>16</sup>. كما هو معلوم لا يعتبر القطاع غير الرسمي قطاعاً مؤسسياً حسب تعريف نظام الحسابات القومية للقطاعات المؤسسية، بل يشمل مجموعة من الوحدات الإنتاجية التي تندرج تحت القيمة المضافة ومكوناتها في القطاع المنزلي.

ينقسم القطاع المؤسسي للاقتصاد الوطني في الحسابات القومية وفقاً لتوصيات نظام الحسابات القومية إلى وحدتين:

- **الوحدات المقيمة ويرمز لها بالرمز (S1)، وتشمل على خمس قطاعات تنظيمية:**
  - المشروعات غير المالية.
  - المشروعات المالية.
  - الحكومة العامة.

<sup>16</sup> Operational Guidebook on Accounting for Informal Sector in National Accounts

- الهيئات التي لا تهدف إلى الربح وتخدم العائلات
- قطاع العائلات.
- الوحدة غير المقيمة ويرمز لها بالرمز (S2)، وتشمل على:
  - العالم الخارجي (بقية العالم الخارجي).

ويندرج القطاع غير الرسمي<sup>17</sup> تحت القطاع العائلي ضمن القطاع المؤسسي في نظام الحسابات القومية، وتعتبر الوحدات المنزلية في القطاع العائلي مستهلكين ومنتجين في نفس الوقت، وذلك من خلال كسب دخل وتقديم خدمات عمل خاصة بهم إلى منتجين آخرين، أو من خلال تلقي تحويلات من قطاعات مؤسسية أخرى مثل الحكومة والشركات غير الحكومية والشركات المالية والأسر الأخرى. أما الوحدات المنزلية المنتجة وغير المسجلة في الأطر القانونية، فهي الأسر التي تنتج سلع وخدمات للاستهلاك الخاص وللسوق. وقد تكون هذه الوحدات في القطاع العائلي ضمن المزارعين، أو صغار تجار التجزئة والمصنعين، أو مقرضي الأموال، أو أصحاب المساكن التي يشغلها مالكوها، أو الأسر التي بها موظفين محليين يعملون بأجر.

### 1. كيفية تضمين القطاع غير الرسمي في نظام الحسابات القومية

يتطلب لتضمين قيمة القطاع غير الرسمي في الناتج المحلي الإجمالي، التأكد من الوحدات الإنتاجية في الإطار المؤسسي لنظام الحسابات القومية، وتصنيف المعايير المستخدمة لتحديد القطاع الغيرمنظم في منظومة الحسابات القومية على النحو التالي:

- تحديد المعايير الأساسية العامة:
  - التنظيم القانوني للمشروع/المؤسسة غير المسجلة (ضمن القطاع العائلي).
  - طبيعة ملكية المشروع/المؤسسة (في حالة ملكيتها إلى الأسرة).
  - طبيعة ونوع الحسابات (عدم وجود حسابات كاملة منفصلة).
  - وجهة الإنتاج: على الأقل بعض الإنتاج موجه للبيع أو المقايضة، والمؤسسات المنزلية التي ليس لها إنتاج في السوق (الزراعة أو البناء لحسابها الخاص)، وخدمات العمال المنزليين بأجر، واستبعاد الخدمات من المساكن التي يشغلها مالكوها.
- تحديد المعايير التشغيلية الإضافية:
  - تحديد حجم المؤسسة: عدد الموظفين المشاركين في الإنتاج (بالمؤسسات صغيرة الحجم يعمل بها أقل من خمسة موظفين).
  - ضمان عدم تسجيل الموظفين في مؤسسة ضمن بعض فروع الحكومة.
  - النشاط الاقتصادي: يشمل الوحدات التي تعمل بشكل رئيسي في القطاع الزراعي وتؤدي أنشطة ثانوية غير زراعية.
  - الموقع الجغرافي للنشاط، (المناطق الحضرية أو الريفية).

ويتم تضمين كل الأنشطة الاقتصادية التي يتم الاضطلاع بها في الاقتصاد ضمن قطاع الإنتاج في الحسابات القومية، وذلك بعد أن يتم استبعاد الإنتاج غير القانوني في تقديرات الناتج المحلي الإجمالي. إضافة إلى الأخذ بالاعتبار ما يلي<sup>18</sup> :

<sup>17</sup> الأمم المتحدة، دليل الحسابات القومية 2008.

<sup>18</sup> United Nations, Economic Commission for Africa (2014), Operational Guidebook on Accounting for Informal Sector in National Accounts.

1. الأنشطة غير المسجلة في الضرائب والضمان الاجتماعي والإحصاء والسلطات العامة الأخرى.
2. يستثنى إنتاج الخدمات المنزلية والشخصية التي يتم إنتاجها واستهلاكها داخل نفس المنزل (باستثناء توظيف موظفين محللين بأجر وخدمات المساكن التي يشغلها مالكوها).
3. الأنشطة التطوعية التي ينتج عنها سلع، على سبيل المثال إنشاء مسكن أو أي مبنى آخر، وكذلك الأنشطة التطوعية على سبيل العناية والتنظيف بدون مدفوعات.

## 2. منهجية احتساب القطاع غير الرسمي في منظومة الحسابات القومية

تواجه تقدير القيمة المضافة للقطاع غير الرسمي عدد من التحديات بما يشمل، تسجيل مدفوعات العمالة غير المنضمة في دفاتر حسابات المؤسسات الرسمية، حيث تظهر في هذه الدفاتر حسابات التوظيف الرسمي فقط، أما غير الرسمية فلا تظهر في دفاتر هذه المؤسسات الرسمية، بسبب أن العاملين في هذه المؤسسات قد تحصلوا على هذه الوظائف مدفوعة الأجر، ولا يتم تغطيتها ضمن تشريعات العمل الرسمية، التي تندرج تحت فئة العمالة غير الرسمية في (المؤسسات الرسمية) وتكون ذات طبيعة "مقدمي الخدمات" لهذه المؤسسات، حيث يعتبروا في الأساس من العمالة المتعاقدين أو عمال يتم الاستعانة بهم من مصادر خارجية<sup>19</sup>.

بينما يمثل التحدي الأخر في تقدير القيمة المضافة للقطاع المنظم، خاصة أن حساب العمالة غير الرسمية لا يتم تضمينها في حساب القطاعات الرسمية ولا تنعكس في منظومة الحسابات القومية. وبالتالي سيؤثر ذلك على تقديرات القيمة المضافة للقطاع الرسمي، حيث تُظهر في دفاتر حسابات المؤسسات الرسمية تعويضات الموظفين الذين يتم دفع رواتبهم للعمل الرسمي وكذلك شراء مدفوعات الخدمات التي يتم دفعها إلى العمالة غير الرسمية مقابل خدماتهم، وهو بند مصروف للمؤسسات وجزء من الاستهلاك الوسيط، على الرغم من أن الموظفين الرسميين وغير الرسميين يعملون في هذه المؤسسات.

نتيجة لهذه الممارسة المحاسبية المتمثلة في معالجة المدفوعات المقدمة إلى العمالة غير الرسمية من قبل المؤسسات الرسمية، يتم تجاوز قيم الاستهلاك الوسيط لهذه المؤسسات، مما يؤدي إلى انخفاض القيمة المضافة للمؤسسات الرسمية، وبالتالي يؤدي إلى عدم الاتساق بين مدخلات العمل والمخرجات. إضافة إلى ذلك، قد يكون من المفيد أيضاً توفير معلومات إضافية عن توزيع الإنتاج حسب نوع النشاط، أو نسبة الإنتاج الإجمالي في الصناعة التي تنتجها المؤسسات غير المنظمة<sup>20</sup>.

وبالتالي يتم تجميع البيانات حول هذا القطاع بحيث تغطي الإنتاج وتوليد الدخل وفي حين تغطي البيانات الأخرى العمالة، وذلك كما هو موضح:

■ الإنتاج وتوليد الدخل، ويشمل على ما يلي:

- الإنتاج، ومنها الاستخدام الخاص
- الاستهلاك الوسيط

<sup>19</sup> قد تكون البيانات حول هذه العمالة غير متاحة في حسابات المؤسسات الرسمية، ولكنها قد تكون متاحة من خلال بعض مسوحات القوى العاملة التي تتوفر بيانات حول إحصاءات العمالة غير المنظمة في القطاع الرسمي. <sup>20</sup> في بعض البلدان يتم تغطية بعض الوحدات الصغيرة التي يمكن اعتبارها جزءاً من القطاع غير الرسمي من خلال استقصاءات المؤسسة وتضمينها في قطاع الشركات.

- القيمة المضافة
- تعويضات الموظفين
- إجمالي الدخل المختلط
- استهلاك رأس المال الثابت
- صافي الدخل المختلط.
- التوظيف
- 1. العمالة في القطاع غير الرسمي
  - الوظائف الرسمية
  - الوظائف غير الرسمية
- 2. العمالة خارج القطاع غير الرسمي
  - في القطاع الرسمي
  - في المؤسسات المنزلية الأخرى الفردية.

يمكن احتساب الأنشطة المعنية بالقطاع غير الرسمي في الحسابات القومية<sup>21</sup> في حالة توفر بيانات حول المخرجات (Output) والاستهلاك الوسيط، والتغيرات في المخزونات، وتعويضات الموظفين، والدخل المختلط، واستهلاك رأس المال الثابت والضرائب أخرى على الإنتاج من مؤسسات القطاع غير المنظم. في حين تعتبر أهم القضايا الرئيسية لتقدير القيمة المضافة للقطاع غير الرسمي هو تسجيل العمالة غير الرسمية والمدفوعات للعمالة غير الرسمية في دفاتر حسابات المؤسسات الرسمية، حيث تظهر لدى الشركات فقط حسابات التوظيف الرسمية، أما غير الرسمية فلا تظهر في دفاتر هذه المؤسسات.

وبالتالي، قد لا تتوفر بيانات حول العمالة غير الرسمية في القطاع الرسمي في حسابات المؤسسات الرسمية، ولكنها قد تكون متاحة في مسح القوى العاملة، والتي توفرها هذه المسوحات الخاصة بإحصاءات العمالة غير الرسمية في القطاع الرسمي. لذلك، أثناء استخدام مصادر مختلفة في بناء مصفوفة مدخلات العمل، من المهم أن تضمن فئات العمالة الرسمية وغير الرسمية في كل من القطاعات المؤسسية في مصفوفة خاصة لذلك، ومصنفة بشكل أكبر حسب (القطاعات)، بحيث يتم عرض تقديرات "مقدمي الخدمات" غير الرسميين في القطاع الرسمي بشكل منفصل ويمكن حسابهم بشكل مناسب في الحسابات القومية وفقاً لمفاهيم نظام الحسابات القومية.

المسألة الثانية التي تنشأ على حساب العمالة غير الرسمية في القطاعات الرسمية هي الطريقة التي تسدد بها المؤسسات الرسمية المدفوعات مقابل الخدمات التي يقدمها الموظفون غير الرسميين، مما يؤثر على تقديرات القيمة المضافة للقطاع الرسمي وكذلك الاقتصاد الكلي. إضافة إلى ذلك، تُظهر في دفاتر حسابات المؤسسات الرسمية تعويضات الموظفين الذين يتم دفع رواتبهم للعمل الرسمي وكذلك شراء مدفوعات الخدمات التي يتم دفعها إلى العمالة غير الرسمية مقابل خدماتهم، وهو بند مصروف للمؤسسات وجزء من الاستهلاك الوسيط، على الرغم من الموظفين الرسميين وغير الرسميين يعملون في هذه المؤسسات الرسمية. نتيجة لهذه الممارسة المحاسبية المتمثلة في معالجة المدفوعات المقدمة إلى العمالة غير الرسمية من قبل المؤسسات الرسمية، يتم تجاوز قيم الاستهلاك الوسيط للمؤسسات الرسمية، مما

<sup>21</sup> من خلال الطريقة المباشرة الخاصة بالمسوحات المؤسسية أو مسح الشركات المنزلية المختلطة التي تمت مناقشتها في بند الطرق التقديرية في اختيار وقياس حجم القطاع غير المنظم.

يؤدي إلى انخفاض القيمة المضافة للمؤسسات الرسمية، وبالتالي يؤدي إلى عدم الاتساق بين مدخلات العمل والمخرجات.

قد يكون من المفيد توفير معلومات إضافية عن توزيع الإنتاج حسب نوع النشاط، أو نسبة الإنتاج الإجمالي في الصناعة التي تنتجها المؤسسات غير الرسمية، حيث يتم في بعض البلدان تغطية بعض الوحدات الصغيرة التي يمكن اعتبارها جزءاً من القطاع غير الرسمي من خلال استقصاءات المؤسسة وتضمينها في قطاع الشركات.

- منهجية احتساب القيمة المضافة لمؤسسة القطاع الرسمي التي لديها عمالة غير رسمية<sup>22</sup>

ينتهي حساب الإنتاج الذي يرمز في الحسابات القومية بالكود (P.1) برصيد القيمة المضافة، ويمثل انتاج كافة الوحدات الإنتاجية. في الجانب الأيسر من مصفوفة حساب الدخل وتتضمن جميع ما تم استهلاكه واستخدامه للحصول على الإنتاج، كاستهلاك الوسيط والإعانات والضرائب واهتلاك رأس المال. وعليه يتم الحصول على القيمة المضافة الإجمالية (B.1g)، جدول (6):

1. القيمة المضافة (الناتج المحلي الإجمالي بأسعار السوق (GDP) = الإنتاج (يشمل جميع الأنشطة الاقتصادية بما فيها القطاع العائلي) + الضرائب ناقصاً الإعانات على المنتجات - (الاستهلاك الوسيط (جميع المواد الخام وجميع المواد التي تدخل في عملية الإنتاج)
2. صافي القيمة المضافة = القيمة المضافة يطرح منها اهتلاك رأس المال.

الاستهلاك الوسيط، يشمل كل القيم المجمعة من خلال الاستمارة المعدة لذلك، حيث يتم تقدير مشتريات المواد الخام، والمواد المستهلكة في عملية الإنتاج، إضافة إلى جميع المواد الأخرى التي تدخل في عملية الإنتاج كالكهرباء والماء... الخ. في حين ينتهي حساب توليد الدخل برصيد فائض التشغيل، تكون القيمة المضافة الإجمالية (أو الصافية) في جانب الموارد، وفي جانب الاستخدامات تكون تعويضات العاملين، والضرائب على الإنتاج والواردات يطرح منها الإعانات.

#### جدول رقم (6)

#### مثال : حساب الإنتاج (Production Account)

الإنتاج		الاستخدام	
P.1	الإنتاج	P.2	الاستهلاك الوسيط
		D.21	الضرائب
		D.31	الإعانات على المنتجات
		B.1g	القيمة المضافة الإجمالية
		K.1	إهلاك رأس المال
		B.1N	القيمة المضافة الصافية
			B.1g-K.1

المصدر: الأمم المتحدة (1993) دليل الحسابات القومية، جزء (Overview) صفحة (40).

على ضوء البيانات المتاحة التي تم توفيرها من المسوحات الاقتصادية السنوية مثل (مسح التجارة الخارجية، ومسح المنشأة الاقتصادية، ومسح الأسعار... الخ)، وموازنات الوحدات الاقتصادية وحساب المنشأة المالية وشركات التأمين، وكذلك المسوحات الخاصة بالقطاع

<sup>22</sup> Operational Guidebook on Accounting for Informal Sector in National Accounts

العائلي بما فيها القطاع غير الرسمي. وبالتالي فإن طريقة تقدير الناتج المحلي باستخدام طريقة الإنتاج أو القيمة المضافة تحسب على أساس جمع القيم المضافة لبند الموازنة في حساب الإنتاج لكل الوحدات المؤسسية على النحو التالي:

1. القيمة المضافة (الناتج المحلي الإجمالي بأسعار السوق (GDP) = الإنتاج (يشمل جميع الأنشطة الاقتصادية بما فيها القطاع العائلي) + الضرائب ناقصاً الإعانات على المنتجات - الاستهلاك الوسيط (جميع المواد الخام وجميع المواد التي تدخل في عملية الإنتاج)
  2. صافي القيمة المضافة = القيمة المضافة يطرح منها إهلاك رأس المال.
- حساب توليد الدخل: ينتهي حساب توليد الدخل برصيد فائض التشغيل، ففي جانب الموارد تكون القيمة المضافة الإجمالية (أو الصافية)، وفي جانب الاستخدام تكون تعويضات العاملين، والضرائب على الإنتاج والواردات ويطرح منها الإعانات، بعد ذلك يتم طرح القيمة المضافة الصافية من بنود الاستخدام للحصول على فائض التشغيل/الدخل المختلط، جدول رقم (7).

### جدول رقم (7)

مثال: حساب توليد الدخل (Income Generation)

الاستخدام		الموارد	
	تعويضات العاملين	D.1	B.1N القيمة المضافة الصافية
	الضرائب على الإنتاج والواردات	D.2	
	الإعانات على المنتجات (-)	D.3	
B.1N-(D.1+D.2-D.3)	فائض التشغيل/الدخل المختلط	B.2/B.3g	

المصدر: الأمم المتحدة (1993) دليل الحسابات القومية، صفحة (42).

يتم حساب توليد الدخل للقطاع غير الرسمي، من خلال تقدير القيمة المضافة لكل فرع معني في هذا القطاع، وكما تم الإشارة فإن حساب توليد الدخل مقسم إلى ثلاث مكونات: تعويضات العاملين، وتشمل على الأجور مضاف إليها اشتراكات الأخرى بما فيها الضمان الاجتماعي التي تدفع من قبل صاحب العمل، وكذلك الضرائب على الإنتاج، وعليه يتم احتساب الدخل المختلط/فائض التشغيل لكل وحدة إنتاجية، وذلك من خلال طرح تعويضات العاملين والضرائب الأخرى من صافي القيمة المضافة، وعليه يتم احتساب حساب توليد الدخل كما هو موضح أدناه:

1. فائض التشغيل/الدخل المختلط = صافي القيمة المضافة - تعويضات العاملين (الرواتب والأجور) + الضرائب ناقصاً الإعانات على المنتجات
2. الدخل المحلي الإجمالي (مكونات الدخل للناتج المحلي الإجمالي) = تعويضات العاملين (الرواتب والأجور) + الضرائب ناقصاً الإعانات على المنتجات - فائض التشغيل/الدخل المختلط.
3. الدخل القومي الإجمالي = الناتج المحلي الإجمالي بأسعار السوق + الضرائب ناقصاً الإعانات على المنتجات + صافي دخل الملكية.

### ● معالجة القيمة المضافة للعمالة غير الرسمية

تعتبر الأسرة الوحدة التنظيمية في القطاع العائلي، بما يشمل الأفراد الذين يعيشون في نفس المسكن ولهم دخل يتم من خلاله تغطية تكاليف المعيشية ويستهلكون بعض السلع والخدمات بصورة جماعية كالمسكن والطعام، واستناداً لدليل الحسابات القومية فإن هذا

- القطاع يتضمن أيضا المشاريع غير المنظمة وهي الوحدات الصغيرة التي لا تملك حسابات مستقلة عن حسابات الأسرة المالكة لها. مع مراعاة البنود التالية في القطاع غير المنظم:
- قياس الإنتاج في القطاع غير الرسمي، يتم من خلال جمع كل قيم المعاملات المنجزة للوحدات التي تنتمي إلى هذا الفرع، كذلك احتساب الإنتاج الذي يعطى مجانا للجهات الأخرى، في حين يتم إهمال الجزء الخاص بتغير المخزون، وذلك لأن قياس هذا الجانب غير دقيق، حيث إن المنتجين في القطاع غير المنظم لا يقومون في العادة بتخزين الإنتاج الخاص بهم.
  - الاستهلاك الوسيط، كما تمت الإشارة سابقا يتم جمع كل القيم من خلال استمارة يتم من خلالها تقدير المشتريات من المواد الخام، والمواد المستهلكة في عملية الإنتاج وجميع المشتريات الأخرى التي تدخل في عملية الإنتاج كالكهرباء والماء،... الخ.
- فيما يلي أمثلة تطبيقية تتضمن كيفية التعامل مع القطاع غير المنظم. في هذا الإطار، تم ادراج العمالة غير المنظمة ضمن الاستخدام في جانب حساب الإنتاج (الاستهلاك الوسيط تجاه الخدمات التي تقدمها العمالة غير المنظمة).

## جدول رقم (8)

مثال : حساب الإنتاج (Production Account) (وحدة نقدية)

رمز	الموارد	رمز	الاستخدام
P.1	100 الإنتاج	20	الاستهلاك الوسيط تجاه الخدمات التي تقدمها العمالة غير المنظم.
		10	الاستهلاك الوسيط فيما يتعلق بالتعاقد من الباطن للعمل
		15	نفقات الاستهلاك الوسيط الأخرى
		45	مجموع الاستهلاك الوسيط
		55	القيمة المضافة الإجمالية
			B.1g

## جدول رقم (9)

مثال : حساب توليد الدخل (Income Generation) (وحدة نقدية)

رمز	الموارد	رمز	الاستخدام
B.1g	55 القيمة المضافة الإجمالية	D.1	تعويضات العاملين المدفوعة مقابل العمل
		25	فائض التشغيل/الدخل المختلط
			B.2/B.3g

يمكن أن تظهر بعض مشاكل التقدير في حساب الإنتاج وتوليد الدخل الخاص بالعمالة غير المنظمة، حيث تعامل المبالغ المدفوعة للعمل غير المنظم على أنها استهلاك وسيط، وبالتالي فإن تقدير القيمة المضافة تعتبر أقل، إضافة إلى عدم وجود مخرجات، تُقاس مقابل الاستهلاك الوسيط، حيث توضح في مدخلات عمل العمالة غير المنظمة كموظفين في مؤسسة القطاع المنظم، وعليه فإن الناتج المحلي الإجمالي للاقتصاد الإجمالي يُعتبر أقل تقديرًا أيضًا. يمكن

التغلب على مشكلة نقص التقدير في القيمة المضافة إذا تم تطبيق أحد من الخيارين التاليين، حيث تبقى القيمة المضافة كما هي بالرغم من وجود زيادة في قيمة الإنتاج في الحالة الثانية.

- (المعالجة رقم 1): مُعاملة المدفوعات للعمل غير المنظم كتعويضات للعاملين بدلاً من الاستهلاك الوسيط.
  - (المعالجة رقم 2): مُعاملة العمالة غير المنظمة كمقدمي للخدمات وينتمون إلى القطاع غير المنظم ويتم تقدير مقابل إنتاجهم بالمقابل، بما يعادل المدفوعات التي يتلقونها من القطاع المنظم.
  - المعالجة رقم (1): معاملة المدفوعات كتعويضات للعاملين
- يتم استبدال الاستهلاك الوسيط للخدمات التي تقدمها العمالة غير الرسمية إلى تعويضات العاملين المدفوعة مقابل العمل غير المنظم.

#### جدول رقم (10)

القيمة المضافة لمؤسسة قطاع رسمي ولديها عمالة غير رسمية  
(إظهار المدفوعات إلى العمالة غير الرسمية كتعويض للموظفين)

القيمة	البند
100	0. الإنتاج
10	1. الاستهلاك الوسيط فيما يتعلق بالتعاقد من الباطن للعمل
15	2. نفقات الاستهلاك الوسيط الأخرى
75	القيمة المضافة (2-1-0)
20	4. تعويضات العاملين المدفوعة مقابل العمل غير المنظم
30	5. تعويضات العاملين المدفوعة مقابل العمل المنظم
50	6. إجمالي تعويضات العاملين (5+4)
25	7. فائض التشغيل/الدخل المختلط (القيمة المضافة - إجمالي تعويضات العاملين)

Source: United Nations, Economic Commission for Africa (2014), Operational Guidebook on Accounting for Informal Sector in National Accounts, page (60).

- المعالجة رقم (2): معاملة العمالة غير الرسمية كتعويضات للعاملين

## جدول رقم (11)

## القيمة المضافة للقطاع المنظم وغير المنظم

(معاملة الموظفين غير المنظمة في القطاع المنظم كمقدمي خدمة ينتمون إلى القطاع غير المنظم)

إجمالي الاقتصاد	مؤسسات القطاع غير المنظم*	مؤسسات القطاع المنظم	البنود
القيمة	القيمة	القيمة	
120	20	100	3. الإنتاج
20		20	4. الاستهلاك الوسيط تجاه الخدمات التي تقدمها العمالة غير المنظمة
10		10	5. الاستهلاك الوسيط فيما يتعلق بالتعاقد من الباطن للعمل
15		15	6. نفقات الاستهلاك الوسيط الأخرى
45		45	7. مجموع الاستهلاك الوسيط (4+3+2)
75	20	55	8. القيمة المضافة (5-1)
50	20	30	9. تعويضات العاملين المدفوعة للوظائف الرسمية / الإيراد المختلط
25		25	10. فائض التشغيل/الدخل المختلط

Source: United Nations, Economic Commission for Africa (2014), Operational Guidebook on Accounting for Informal Sector in National Accounts, page (60).

(\* تقديم الخدمات للقطاع المنظم (يعامل كعمالة غير منظمة في القطاع المنظم)

## خامساً: الإجراءات والتدابير لتنظيم القطاع غير الرسمي

بناءً على ما تقدم، فيما يلي نستعرض أهم الإجراءات التي يمكن أن تساهم في تنظيم القطاع غير الرسمي في الدول العربية:

1. تكليف جهة وطنية من الجهات المعنية داخل الدولة لدراسة واقع القطاع غير الرسمي والوقوف على التحديات التي تواجهه وتقدير حجم القطاع.
2. إعداد إطار مفاهيمي على مستوى الدولة لتحديد القطاع غير الرسمي وتعريف أهم المصطلحات المتعلقة بهذه القطاع وتبني المعايير الدولية لدمج القطاع في منظومة الاقتصاد الرسمي والإحصاءات الرسمية.
3. تبسيط الإجراءات والتسهيلات الإدارية للتسجيل في السجلات التجارية للشركات.
4. دعم القدرات الفنية للهياكل الإحصائية في الدول العربية خاصة فيما يتعلق بدمج أنشطة القطاع غير الرسمي في منظومة الإحصاءات الوطنية.
5. التنسيق مع المؤسسات الرسمية داخل الدولة لتطوير قاعدة بيانات المنشآت الاقتصادية لتسجيل الأعمال لتكون مصدراً للمعلومات وإطار معاينه للدراسات المتخصصة في هذا المجال.
6. تطوير إطار لسياسات التشغيل (برنامج تأهيل وبرنامج إدماج في السوق العمل) تهدف لتشغيل الشباب في القطاع الرسمي.
7. نشر روح المبادرة لدى المواطنين خاصة الشباب وتشجيع القطاع غير الرسمي للاندماج في الاقتصاد الوطني.
8. تنفيذ مسوحات ميدانية حول اتجاهات العمالة والبطالة حسب المعايير المعتمدة من قبل منظمة العمل الدولية والمؤسسات الدولية، بهدف تحديد ساعات العمل وتقدير

- أعداد العمالة المنتجة فعلياً لتبني سياسات من شأنها تحفيز انخراط العمالة في القطاع الرسمي وتقليل البطالة.
9. تنفيذ التعدادات الاقتصادية بصفة دورية كل خمس سنوات والتركيز على المسوحات الخاصة بالحسابات القومية والحسابات التابعة لها.
10. استخدام بيانات المسوحات المعدة من الأجهزة الإحصائية، كتعداد المنشآت الاقتصادية ومسوحات الأسر التي تساهم في التعرف بشكل أكثر تفصيلاً حول واقع هذا القطاع في الدول العربية، إضافة إلى تنفيذ مسوحات متخصصة لتقدير بيانات الإنتاج للقطاعات المعنية بطريقة مباشرة.
11. تشجيع انضمام مؤسسات القطاع غير الرسمي في أنظمة الضمان الاجتماعي وتوضيح أهميته بالنسبة لأفراد المجتمع، وتعزيز الثقة بين مؤسسات القطاع غير الرسمي والجهات المعنية في الدولة.
12. تسهيل فرص النفاذ إلى التمويل لاسيما فيما يتعلق بالإناث والشباب لتشجيع إقامة المنشآت في إطار القطاع الرسمي.
13. تبني سياسات تحفيزية لدمج مشروعات القطاع الرسمي في إطار منظومة الاقتصاد الرسمي.

## المراجع

### المراجع العربية

1. أمحمد موعش (2017)، دراسة حول إحصاءات القطاع غير الرسمي في الدول العربية، الاجتماع الرابع لمبادرة الإحصاءات العربية، صندوق النقد العربي.

### المراجع الانجليزية

2. Eurostat (2010), "Essential SNA: Building the basics.
3. ILO, Labor force Database.
4. ILO (2019), Women and men in the informal economy: A statistical picture, third edition.
5. ILO (1993), "Resolution concerning statistics of employment in the informal sector, adopted by the 15th International Conference of Labour Statisticians", January.
6. ILO, (2013). "Measuring informality: A statistical manual on the informal sector and informal employment.
7. IMF, René Piché (2017), The Informal Sector Some Measurement Challenges, Fourth Meeting of the Steering Committee of the Arab Statistics Initiative "Arabstat".
8. United Nations, Economic Commission for Africa (2014), "Operational Guidebook on Accounting for Informal Sector in National Accounts".
9. United Nations, System National Account 1993 and 2008.
10. World Bank, World Development Indicators, Database.

## Maternal mortality and associated Statistical analysis factors in Iraq

Dr.Arjwan Marwan Shaban

Family medicine physician at MOH.

Dr. Dawood Ahmed Dawood Sulaiman

Specialist in Anesthesia & ICU Nineveh health directorate

Nineveh health Family medicine specialist Dr. Amina Mamon Akram Alsawaf

تاريخ استلام البحث: 2023/11/14

تاريخ قبول البحوث: 2024/01/29

نشر البحث في العدد الثالث والعشرين: يونيو / حزيران 2024

2522-64X/312.563

رمز التصنيف ديوي / النسخة الالكترونية (Online):

2519-948X/312.563

رمز التصنيف ديوي / النسخة الورقية (Print):

## Maternal mortality and associated Statistical analysis factors in Iraq

Dr.Arjwan Marwan Shaban

Family medicine physician at MOH.

Dr. Dawood Ahmed Dawood Sulaiman

Specialist in Anesthesia & ICU Nineveh health directorate

Dr. Amina Mamon Akram Alsawaf Family medicine specialist Nineveh health

### Abstract:

**Introduction:** Maternal mortality is described as the death of a woman during pregnancy or within six weeks of delivery. The Maternal Mortality Ratio is an important health marker that reflects the strength of health care system. Maternal mortality caused by a variety of causes. COVID19 pandemic affect the Maternal Mortality Ratio. **Patients and method:** The study conducted in Iraq, the data were collect from health directorate's reports that were send periodically to the Ministry of health for the year of 2021. **Results:** Maternal mortality ratio in the country was (46.1), and COVID19 is the main causative agent for maternal mortality (40%). Most of the dead women were housewives, live in urban places, at bearing age 20-39 years, not complete the primary schools and the place for their delivery was at governmental hospitals. **Discussion:** The maternal mortality ratio is higher than the previous year and this may be due to COVID pandemics.

**Keywords:** COVID19, maternal mortality, live birth.

### وفيات الأمهات وعوامل التحليل الإحصائي المرتبطة بها في العراق

#### الخلاصة

مقدمة: توصف وفيات الأمهات بأنها وفاة المرأة أثناء الحمل أو خلال ستة أسابيع من الولادة. نسبة وفيات الامهات هي عدد الامهات المتوفيات لكل 100000 ولادة طفل حي. تعد نسبة وفيات الأمهات علامة صحية مهمة تعكس قوة نظام الرعاية الصحية. وفيات الأمهات ناجمة عن مجموعة متنوعة من الأسباب. تؤثر جائحة كوفيد 19 على نسبة وفيات الأمهات. المرضى وطريقة العلاج: الدراسة الوصفية التي أجريت في العراق لايجاد نسبة وفيات الامهات والعوامل المرتبطة بها، تم جمع البيانات من تقارير مديريات الصحة التي ترسل بشكل دوري إلى وزارة الصحة لعام 2021. النتائج: عدد الولادات الحية 1059781 , عدد الامهات المتوفيات 489 بلغت نسبة وفيات الأمهات في الدولة (46.1) وكانت اعلى نسبة في محافظة ذي قار و اقل نسبة في صلاح الدين و الانبار، ويعتبر فيروس كورونا المسبب الرئيسي لوفيات الأمهات (40%). معظم الوفيات هن ربات البيوت، يعشن في المناطق الحضرية، في سن الإنجاب 20-39 سنة، لم يكملن الدراسة الابتدائية ومكان ولادتهن كان في المستشفيات الحكومية. المناقشة: نسبة وفيات الأمهات أعلى من العام السابق وقد يكون ذلك بسبب جائحة فيروس كورونا.

**الكلمات المفتاحية:** كوفيد19، وفيات الأمهات، المواليد الأحياء.

**Introduction:**

Maternal mortality is described as the death of a woman during pregnancy or within six weeks of delivery, whatever the gestational age was. <sup>(1)</sup> The Maternal Mortality Ratio (MMR) is defined as the number of maternal deaths per 100,000 live births, MMR is an important health marker that reflects the strength of health care system. <sup>(2-5)</sup> Maternal mortality has a big burden on the community and families specially the infants. <sup>(6,7)</sup> The World Health Organization records a worldwide ratio of 400 maternal deaths per 100 000 live births. <sup>(8)</sup> Globally more than half million woman die each year due to complications of pregnancy or labor; half of them from Sub-Saharan Africa. <sup>(9)</sup> Maternal mortality caused by a variety of direct and indirect causes, the main direct causes in undeveloped countries are hemorrhage (34%), infection (10%) and hypertension (9%), while in developed countries the major cause of death is hypertension followed by hemorrhage. <sup>(10)</sup> Indirect causes form 20% of total causes; they are already present before pregnancy but exaggerated during it. <sup>(10)</sup> Many of the causes of maternal mortality is preventable as treatment is known, <sup>(11-15)</sup> the millennium declaration of the United Nations was consent in 2000: MMR reduction by three quarters between 1990 and 2015. <sup>(16,17)</sup> MMR globally show a reduction of about one third from 2007 to 2017, by the UN inter-agency estimates, most of the deaths where in the undeveloped countries. <sup>(18)</sup> Despite all efforts to improve the maternal care in United States( U.S. ), the MMR had been doubled in the last twenty years. <sup>(19-21)</sup> Meanwhile, most of the other developed countries show decrease in MMR. <sup>(22)</sup> This increase of MMR in U.S. may be due to the use of death certificate coding in last two decades, <sup>(21)</sup> sociodemographic factors also play role in this increase of MMR in US like advanced maternal age, rural residence and illiteracy. <sup>(23,24)</sup> MMR remained elevated in many undeveloped countries. <sup>(17)</sup> COVID19 pandemic affect the MMR assessment. <sup>(25,26)</sup> As the pandemic of COVID19 begin in 2019, many strategies were adopted by the health organizations like tele-consultation to protect the pregnant ladies. <sup>(27, 28)</sup> The compromise of immune and respiratory systems make them more risky to have complication when get infected with COVID19, <sup>(29)</sup> also they are more at risk of getting the infection because they have to go to health centers for their antenatal care visits despite that many of them suffer from severe infection with the virus. <sup>(30)</sup>

### Patients and method:

The study conducted in Iraq, the data were collected from health directorate's reports that were sent periodically to the Ministry of health. Descriptive study was done to set the MMR and the causes behind it together with the associated factors in 2021 were COVID 19 still in pandemic. Lenovo computer did data collection and Minitab system (19th edition) were used to find out the statistical relations.

### Results:

I. Maternal mortality ratio: Results shows that number of live birth was (1059781) ,the number of dead women was ( 489) and the average MMR in the country was (46.1), with highest rate seen in Thi-Qar governorate (75.8) ,and the lowest in both Al-Anbar and Salah Al- Deen governorates (0.0) as shown in table -1-.

Table -1- Maternal mortality rate in Iraqi governorates

Governorate	Number of live births	Number of maternal mortality	MMR
Baghdad	203810	138	67.7
Basrah	91568	44	48.1
Nineveh	90167	34	37.7
Maysan	35295	25	70.8
Al- Dewaniya	33340	12	36.0
Diala	45690	19	41.6
Al-Anbar	48345	0	0.0
Babylon	61932	21	33.9
Kerbela	42476	28	65.9
Kirkuk	39911	14	35.1
Wasit	40250	17	42.2
Thi-Qar	56749	43	75.8
Al-Muthanna	27333	18	65.9
Salah Al- Deen	41281	0	0.0
Al-Najaf	46426	33	71.1
Erbil	63446	15	23.6
Duhouk	43360	8	18.5
Al- Sulaimani	48402	20	41.3
Total	1059781	489	46.1

II. Causes of maternal mortality: Results shows that indirect causes for maternal death form the highest group (54%) and COVID19 is the main causative agent for maternal mortality (40%) in this group of causes, on the other hand ; Postpartum hemorrhage is the main direct cause for maternal mortality (8.9%) as shown in table -2- .

Table – 2- Causes of maternal mortality

	Causes	Percentage %
Direct causes	Postpartum hemorrhage	8.9
	Gestational hypertension	8.7
	Obstetric thromboembolism	6.2
	Antepartum hemorrhage	2.7
	Complications of anesthesia during labor and delivery	2.0
	Infection	1.8
	Uterine rupture during childbirth	0.7
	Amniotic fluid embolism	0.7
	Ectopic pregnancy	0.7
	Abortion	0.5
	Total	32.9
Indirect causes	Covid-19	40.0
	Suspect corona	5.3
	Cancers	2.3
	Preexisting hypertension disorder superimposed proteinuria	1.8
	Heart diseases	1.6
	DM	0.9
	Renal failure	0.9
	Liver diseases	0.4
	Lung infections	0.4
	Nervous system complications	0.4
	Total	54.0
Other causes	Like car accident ,trauma ... etc.	13.1
	Total	100.0

### III. Factors associated with maternal mortality:

#### 1. Residence :

Study results illustrate that (68.7%) dead mothers were live in urban places, figure -1-.

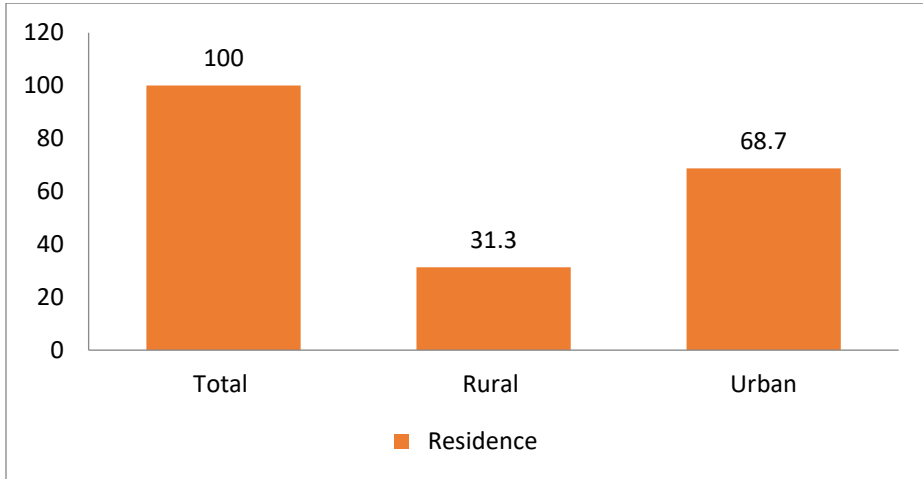


Figure -1- Residence

2. Level of education:

Results shows that (65.2%) of the dead mothers not complete, the primary schools figure -2-.

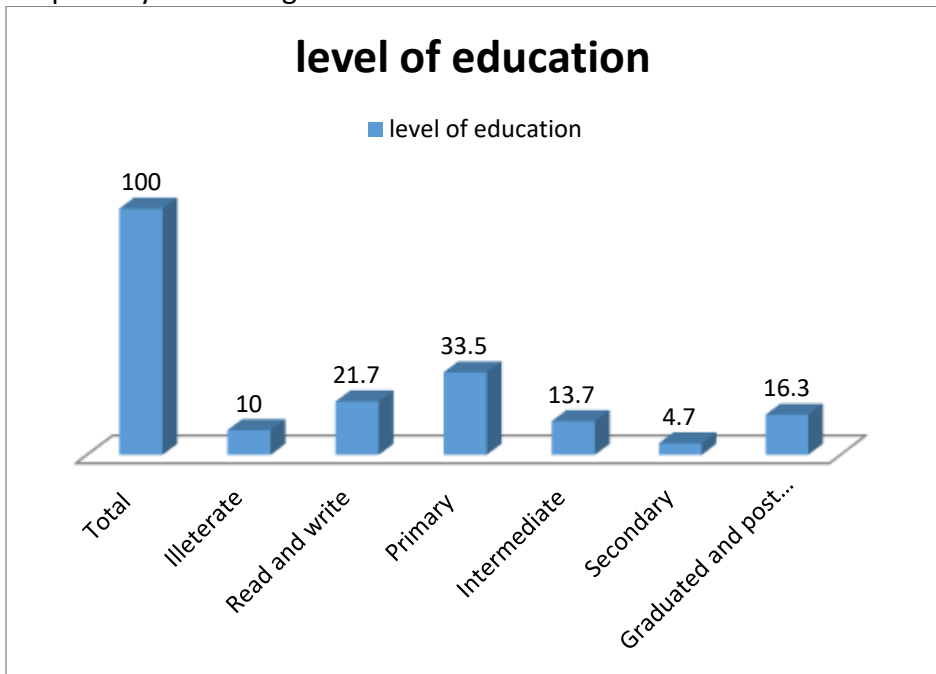


Figure -2- Educational level.

3. Mothers age:

Results shows that (78.1%) of the dead mothers were aged 20-39 years old figure -3-.

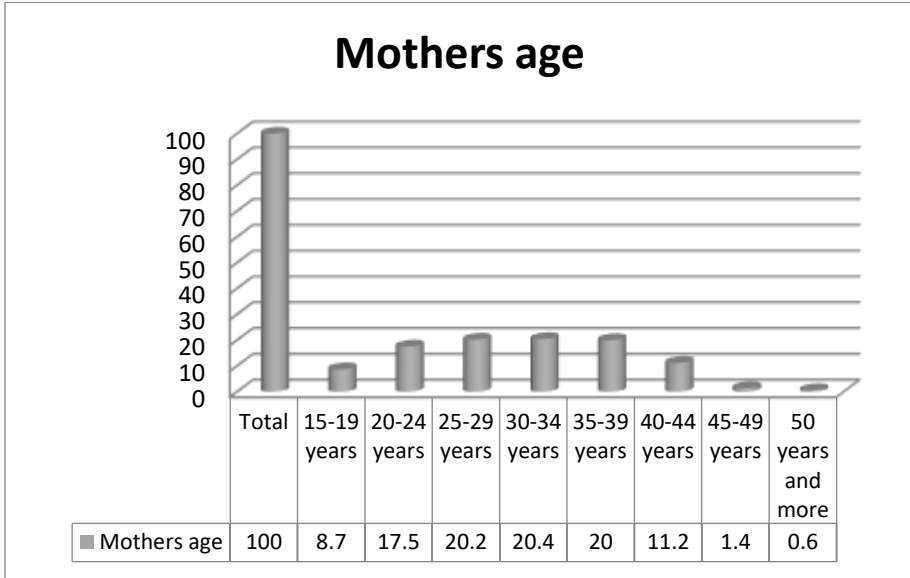


Figure -3- Mothers age

4. Occupation:

Results shows that (87.7%) of the dead mothers were housewife, and only (9.2%) were employee, figure -4-.

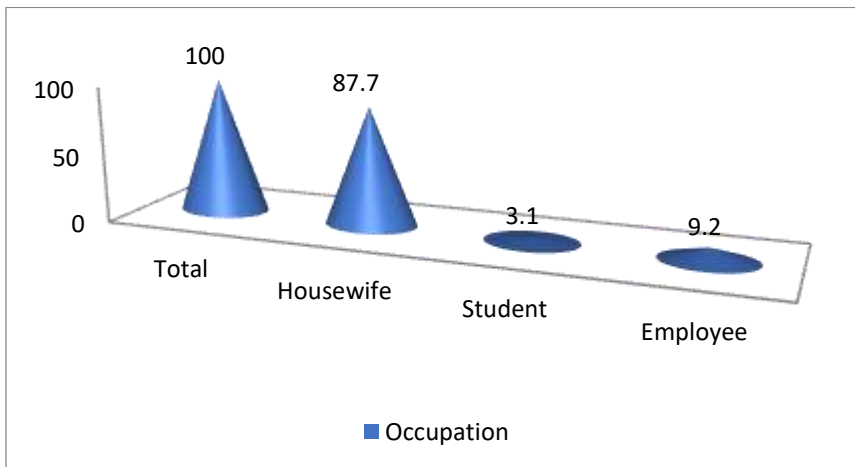


Figure -4- occupation.

### 5. Site of birth:

Results shows that most of the dead mothers were delivered at governmental hospitals (65.8 %), figure -5-.

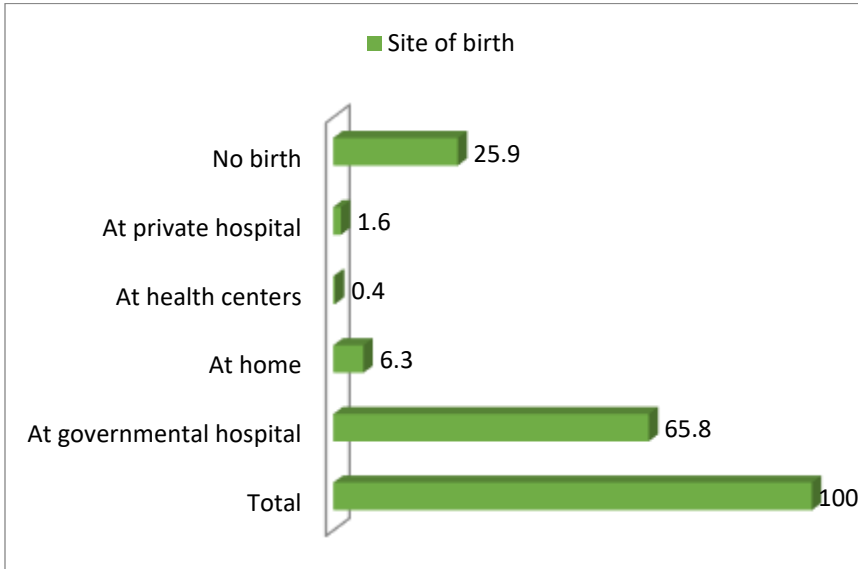


Figure -5- Site of birth.

### Discussion:

The average MMR in this study was (46.1) for the year 2021 , while in 2020 was (34.2), this may be due to COVID pandemics, in Mexico MMR was 46.6 in 2020, as shown in Mendez-Dominguez et al study<sup>(31)</sup> (which is highest ratio in last 10 year in that country). The MMR was less in northern governorates; this may be related to geographic factors or the level of health services. Results shows that COVID19 is the main causative agent for maternal mortality, like results of Mendez-Dominguez et al <sup>(31)</sup>. More than two –third of the dead mothers were lived in urban areas, the same results of Shahidi et al study <sup>(32)</sup>, and opposite to results of both Karimzaei et al and Tirkesh et al studies <sup>(33,34)</sup>, the authors suggests the cause may be due the fact that number of urban population is more and the recording programs more accurate there. Again, about two third of dead mothers not complete the primary schools, near results of both Karimzaei et al and Jamshidi et al studies <sup>(33,35)</sup>, and opposite to Farzianpour et al results <sup>(36)</sup>, this may be due to more early marriage in low educated females which associated with high mortality, and of course the educated females have more health care informations than

the others. Results shows that most of the dead mothers were aged between 20 and 39 years (78.1%), exactly the same percentage in Mendez-Dominguez et al study<sup>(31)</sup>, this may be due to the fact that this is the bearing age in women. Results shows that only (9.2%) of dead mothers were employee, same results of Vahiddastjerdy M et al study<sup>(37)</sup> in which (9.0%) were employee, opposite to Farzianpour et al results<sup>(36)</sup>, in which 71% of dead mothers were employee, this may be due to that house wife prefer more home delivery and have less antenatal care visits. Results shows that most of the dead mothers were delivered at governmental hospitals, this may be due to many reasons, first the number of hospital deliveries is more, second the recording system is better in hospital than other delivery places, the results is same like in Nepal in which death in hospital was 59.6% as shown in Sitaula, S. et al study<sup>(38)</sup>.

### References:

1. WHO. Maternal Mortality: 1990 to 2010. Organization. 2010; 32(5):1–55.
2. Moazzeni MS. Maternal mortality in the Islamic Republic of Iran: on track and in transition. *Matern Child Health J.* 2013; 17(4):577–80.
3. World Health Organization. Trends in maternal mortality: 1990 to 2013. Estimates by WHO, UNICEF, UNFPA, The World Bank and the United Nations Population Division. 2014.
4. World Health Organization. Maternal Mortality ratio (per 100 000 live births). *Health Statistics and Health Information Systems.* 2013; 1.
5. Kassebaum NJ, Bertozzi-Villa A, Coggeshall MS, Shackelford KA, Steiner C, Heuton KR, et al. Global, regional, and national levels and causes of maternal mortality during 1990-2013: a systematic analysis for the global burden of disease study 2013. *Lancet.* 2014; 384(9947):980–1004.
6. Sliwa K, Anthony J. Late maternal deaths: a neglected responsibility. *Lancet.* 2016; 387(10033):2072.
7. National Research Council. *The Consequences of Maternal Morbidity and Maternal Mortality: Report of a Workshop.* Washington, DC: The National Academies Press; 2000.
8. WHO, authors: *The World Health Report 2005: Make Every Mother and Child Count.*
9. WHO: *Maternal mortality in 2005: estimates developed by WHO, UNICEF, UNFPA, and the World Bank.* Geneva: WHO Press; 2007:16.
10. Khan K, Wojdyla D, Say L, Gülmezoglu A, Van Look P: WHO analysis of causes of maternal death: a systematic review. *Lancet* 2006, 367:1066-1074.
11. National Center for Health Statistics. *Maternal mortality and related concepts.* Hyattsville: U.S. Department of Health and Human Services,

- Centers for Disease Control and Prevention, National Center for Health Statistics; 2007.
12. Davis GG, Onaka AT. Report on the 2003 revision of the US standard certificate of death. *Am J Foren Med Path.*2001; 22(1):38–42.
  13. MacKay AP, Berg CJ, Liu X, Duran C, Hoyert DL. Changes in pregnancy mortality ascertainment: United States, 1999-2005. *Obstet Gynecol.* 2011; 118(1):104–10.
  14. Campbell KH, Savitz D, Werner EF, et al. Maternal morbidity and risk of death at delivery hospitalization. *Obstet Gynecol.* 2013; 122(3):627–33.
  15. Frolich MA, Banks C, Brooks A, Sellers A, Swain R, Cooper L. Why do pregnant women die? A review of maternal deaths from 1990 to 2010 at the University of Alabama at Birmingham. *Anesth Analg.* 2014; 119(5):1135–9.
  16. United Nations Millennium Declaration: eighth plenary meeting 2000.
  17. Kassebaum NJ, Bertozzi-Villa A, Coggeshall MS, Shackelford KA, Steiner C, Heuton KR, et al. Global, regional, and national levels and causes of maternal mortality during 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013. *Lancet.* 2014; 384(9947):980-1004.
  18. World Health Organization. Trends in Maternal Mortality: 2000 to 2017. Estimates by WHO, UNICEF, UNFPA, World Bank Group and the United Nations Population Division; 2019.
  19. Berg CJ, Atrash HK, Koonin LM, Tucker M. Pregnancy-related mortality in the United States, 1987-1990. *Obstet Gynecol.* 1996; 88(2):161–7.
  20. Creanga AA, Berg CJ, Syverson C, Seed K, Bruce FC, Callaghan WM. Pregnancy-related mortality in the United States, 2006-2010. *Obstet Gynecol.* 2015; 125(1):5–12.
  21. MacDorman MF, Declercq E, Cabral H, Morton C. Recent increases in the US maternal mortality rate disentangling trends from measurement issues. *Obstet Gynecol.* 2016; 128(3):447–55.
  22. Trends in Maternal Mortality: 1990 to 2013—estimates by WHO, UNICEF, UNFPA, The World Bank and the United Nations Population Division. 2014.
  23. Wang Y, Tanbo T, Abyholm T, Henriksen T. The impact of advanced maternal age and parity on obstetric and perinatal outcomes in singleton gestations. *Arch Gynecol Obstet.* 2011; 284(1):31–7.
  24. Aguilar-Vargas E, Azcorra H, Novelo-Pérez JP, et al. Descripción de la variación estacional en la mortalidad materna de México y el estado de Yucatán 2002–2016. *Ginecol Obstet Mex.* 2018; 86(10):640–9.
  25. Lumbreras-Marquez MI, Campos-Zamora M, Lizaola-Diaz de Leon H, Farber MK. Maternal mortality from COVID-19 in Mexico. *Int J Gynaecol Obstet.* 2020; 150(2):266–7.
  26. Robertson T, Carter ED, Chou VB, Stegmuller AR, Jackson BD, Tam Y, et al. Early estimates of the indirect effects of the COVID-19 pandemic on

- maternal and child mortality in low-income and middle-income countries: a modelling study. *Lancet Glob Health*. 2020; 8(7):e901–8.
27. Mendez-Dominguez N, Alvarez-Baeza A, Carrillo G. Demographic and health indicators in correlation to interstate variability of incidence, confirmation, hospitalization, and lethality in Mexico: preliminary analysis from imported and community acquired cases during COVID-19 outbreak. *Int J Environ Res Public Health*. 2020; 17(12):4281.
  28. Zaigham M, Andersson O. Maternal and perinatal outcomes with COVID-19: a systematic review of 108 pregnancies. *Acta Obstet Gynecol Scand*. 2020; 99(7):823–9.
  29. Wastnedge EAN, Reynolds RM, van Boeckel SR, Stock SJ, Denison FC, Maybin JA, et al. Pregnancy and COVID-19. *Physiol Rev*. 2021; 101(1):303–18.
  30. Panamerican Health Organization. Alerta Epidemiológica COVID-19 Durante el embarazo 13 de agosto de 2020.
  31. Mendez-Dominguez, N., Santos-Zaldívar, K., Gomez-Carro, S. et al. Maternal mortality during the COVID-19 pandemic in Mexico: a preliminary analysis during the first year. *BMC Public Health* 21, 1297 (2021).
  32. Shahidi M et al. prevalence and risk factors of maternal mortality in a 10-year period (2003–2013) in Sabzevar. 2017.
  33. Karimzaei T et al. frequency of maternal mortality in urban and rural areas of Iranshahr county (Southeast of Iran) in 2009–2013: A retrospective study. *Journal of clinical and diagnostic research: JCDR*, 2016. 10(8): p. QC14.
  34. Tirkesh F, Bahrami N, Bahrami S. Assessment of Achievement to Improving Maternal Health from Third Millennium Development Goal in Dezful University of Medical Sciences. *Community Health Journal*. 2015; 2(2):98–105.
  35. Jamshidi F, Ghorbani A, Jafarri RM. The causes of maternal mortality and associated risk factors in Khuzestan province (2009–2013). *Int J Pharmaceutical Research & Allied Sciences*. 2016; 5(3):150–6.
  36. Farzianpour F, Emami AH, Ramezani KH. Causes of maternal death during childbirth in Tehran-Iran in 2011-2015. *Pak J Nutr*. 2017; 16:253–60.
  37. Vahiddastjerdy M, et al. Maternal mortality ratio and causes of death in IRI between 2009 and 2012. *J Family Reprod Health*. 2016; 10(3):154.
  38. Sitaula, S., Basnet, T., Agrawal, A. et al. Prevalence and risk factors for maternal mortality at a tertiary care centre in Eastern Nepal- retrospective cross sectional study. *BMC Pregnancy Childbirth* 21, 471 (2021).

## Determine Sample Size to Estimate the Average Parameter of a Heavy Tails Distribution Using Bayesian Methodology

Sarmad Abdulkhaleq Salih  
University of Al-Hamdaniya , Mosul  
Omar Ramzi Jasim  
University of Al-Hamdaniya , Mosul

تاريخ استلام البحث: 2024/04/24

تاريخ قبول البحث: 2024/05/22

نشر البحث في العدد الثالث والعشرين: يونيو / حزيران 2024

2522-64X/512

2519-948X/512

رمز التصنيف ديوي / النسخة الالكترونية (Online):

رمز التصنيف ديوي / النسخة الورقية (Print):

## Determine Sample Size to Estimate the Average Parameter of a Heavy Tails Distribution Using Bayesian Methodology

Sarmad Abdulkhaleq Salih  
University of Al-Hamdaniya , Mosul  
Omar Ramzi Jasim  
University of Al-Hamdaniya , Mosul

### ABSTRACT:

The generalized modified Bessel distribution is one of the most suitable mixed distributions. It is the result of mixing the normal distribution with the generalized inverse Gaussian distribution.

In this paper, The optimal sample size Analysis has been taken from the generalized modified Bessel population to estimate the mean parameter when the variance and shape parameters are known, using the informative prior information to estimate the mean parameter under the quadratic loss function. Then sampling and non-sampling approaches are used for the estimate of the parameter. Also, it has been noted that the posterior probability distribution for a mean parameter is following a generalized modified Bessel distribution. Through the simulation, we note Bayesian sample size is inversely proportional to the sampling cost (c) per unit.

**Keywords:** Generalized Modified Bessel Distribution, Quadratic Loss Function, Cost Function, Bayesian Sample Size.

### الملخص:

توزيع Bessel المعدل المعمم هو أحد أنسب التوزيعات المختلطة. إنه نتيجة خلط التوزيع الطبيعي مع التوزيع الغاوسي العكسي المعمم. في هذا البحث ، تم أخذ تحليل حجم العينة الأمثل من مجتمع بيسل المعدل المعمم لتقدير متوسط المعلمة عند معرفة معلمات التباين والشكل، باستخدام المعلومات السابقة المفيدة لتقدير المعلمة المتوسطة تحت دالة الخسارة التربيعية. ثم يتم استخدام طرق أخذ العينات وغير أخذ العينات لتقدير المعلمة. أيضا، لوحظ أن التوزيع الاحتمالي الخلفي لمعلمة متوسطة يتبع توزيع بيسل المعدل المعمم. من خلال المحاكاة ، نلاحظ أن حجم عينة بايزي يتناسب عكسيا مع تكلفة أخذ العينات (c) لكل وحدة.

الكلمات المفتاحية: توزيع بيسل المعدل المعمم ، دالة الخسارة التربيعية ، دالة التكلفة ، حجم عينة بايزي.

## 1. INTRODUCTION

The generalized modified Bessel distribution is considered one of the most suitable mixed distributions, as this distribution is more general than the two distributions (normal and T) as special cases, and it is considered that the generalized modified Bessel distribution as a special case of the symmetrical hyperbolic distributions, and this distribution has practical applications in a variety of areas which includes stock market data presentation, quality control data and filtering random-sign analysis.

To conduct any study on data that follows this distribution, its parameters must be estimated to be able to predict and obtain accurate results to solve a particular problem. The characteristics of the community are determined by studying a sample drawn from it, provided that the sample bears the maximum degree of accuracy to represent the community, and this leads to reducing the costs of the field study due to the small size of the sample about the size of the community. There are several sampling methods to determine the sample size to be used in statistical inference, and one of these methods is using the Bayesian sampling method, which in its study depends on taking into account the loss functions with the cost function.

(Thabane and Haq, 2003) was the first to use the Bayesian method to estimate the linear regression model, assuming that the random error distribution is a generalized multivariate modified Bessel distribution when the prior distribution of parameters is a proper distribution, which is the result of mixing two continuous probability distributions. They are the normal distribution and the generalized inverse Gaussian distribution. [14]

(Saiful Islam, 2011) Determine the Bayesian sample size to estimate the mean and the difference between the mean of the normal distribution when the variance is known and estimate the variance when the mean is known, as well as determine the Bayesian sample size to estimate the parameter of the Poisson and Exponential distribution using different loss functions. [11]

The research provided a general introduction to the subject and a description of the generalized modified Bessel distribution, as well as the Bayesian sample size determination to estimate the mean parameter when the prior distribution is proper and using a quadratic loss function.

## 2. Search Objectives

This paper aims to analyze the optimal Bayesian sample size taken from a population followed by is generalized modified Bessel distribution to estimate the mean parameter using the quadratic loss function.

## 3. Generalized Modified Bessel distribution

The generalized modified Bessel distribution is expressed in form of a mixed distribution from the normal distribution and the generalized inverse Gaussian distribution as follows: [6][13] [14]

$$f(Y) = \frac{\left(\frac{\lambda}{\psi}\right)^{\frac{1}{4}}}{(2\pi \sigma^2)^{\frac{1}{2}} K_\nu(\sqrt{\lambda\psi})} \left[1 + \frac{(Y - \theta)^2}{\psi\sigma^2}\right]^{\frac{2\nu-1}{4}} \times K_{\frac{2\nu-1}{2}}\left(\sqrt{\lambda\psi \left(1 + \frac{(Y - \theta)^2}{\psi\sigma^2}\right)}\right) \quad (1)$$

Whereas,  $K_\nu(\cdot)$  represents a modified Bessel function of the third type with order ( $\nu$ ) [1][9]. This distribution is descriptively expressed as:

$$Y \sim GMBD_1(\theta, \sigma^2, \lambda, \psi, \nu) \quad (2)$$

## 4. Quadratic Loss Function

The quadratic loss function is the most commonly used in symmetric loss functions, when the parameter ( $\theta$ ) is univariate, the general formula of the function is as follows:[4][11]

$$L(\hat{\theta}, \theta) = k(\hat{\theta} - \theta)^2 \quad (3)$$

Since ( $k$ ) positive real constant is often equal to an integer one.

Therefore, the Bayesian estimator in the case that information is available for this function represents the average of the posterior probability distribution and the Bayesian estimator in the absence of information represents the average of the prior distribution. [10]

$$\hat{\theta}_{sq} = \hat{\theta}_{Bayes} = E(\theta|Y) \quad (4)$$

## 5. Posterior Risk Function of The Quadratic Loss Function

If  $L(\hat{\theta}, \theta)$  is a loss function and  $P(\theta|Y)$  is posterior distribution of ( $\theta$ ), then the expectation of the loss function is called the posterior risk function as follows: [4]

$$PR = \int_{\forall} L(\hat{\theta}, \theta) * P(\theta|Y) d\theta \quad (5)$$

$(PR)$  is the posterior risk. The posterior risk function will be calculated for the quadratic loss function as follows:

$$PR_S = \int_{\forall \theta} k(\hat{\theta}_{sq} - \theta)^2 * P(\theta|Y) d\theta = k \text{Var}(\theta|Y) \quad (6)$$

This constant ( $k$ ) does not affect calculating the Bayesian estimator, but it affects determining the Bayesian sample size.

## 6. The Cost Function

The cost function is one of the important issues in any study that aim to make it as minimal as possible, and there are different formulas for the cost function, and the linear cost function has been used in this research and it's as follows:

$$C_{(n)} = c_0 + cn \quad (7)$$

Where  $C_{(0)} = 0$  and  $n > 0$

And ( $c_0$ ) is the cost of preparing the sampling or any cost related to sampling and ( $c$ ) is the sampling cost per unit and it is one of the important components for obtaining the Bayesian sample size. By using the Bayes method to estimate the sample size, the total cost ( $TC_{(n)}$ ) must be calculated, which can be done by the following formula when the posterior risk function does not depend on the observations  $Y$ . [8]

$$TC_{(n)} = C_{(n)} + PR \quad (8)$$

But if the posterior risk function depends on ( $Y$ ) observations, in this case, we take the expectation of the posterior risk function and add to it the linear cost function to get the average total cost as follows:

$$E(TC_{(n)}/n) = C_{(n)} + E(PR) \quad (9)$$

When two samples are drawn from both groups, the linear cost function will be in the following form:

$$C_{(n_1, n_2)} = c_0 + c_1 n_1 + c_2 n_2 \quad (10)$$

As ( $c_1$ ) is the sampling cost per unit of ( $n_1$ ), which represents the sample size for the first population and ( $c_2$ ) is the sampling cost per unit of ( $n_2$ ), which represents the sample size for the second population. The costs ( $c_1$  &  $c_2$ ) can be equal to the two populations, or the sample sizes ( $n_1$  &  $n_2$ ) can be equal to the two populations. To get the sample size that gives the lowest cost, one must derive the total cost function relative to ( $n$ ) and by equating the derivative to zero, to get the Bayesian sample size, which is a positive integer. [4]

### 7. Determining the Bayesian sample size to estimate the mean parameter of the generalized modified Bessel distribution for one population using initial information belonging to the conjugate family

Assuming to have a random sample  $(Y_1, Y_2, \dots, Y_n)$  conditioned by the random variable  $(\tau)$  taken from normal population  $N(\theta, \sigma^2 \tau)$ . And that the probability distribution of  $(\tau)$  follows the generalized inverse Gaussian distribution, the probability function is defined as follows: [3][7] [9]

$$P(\tau) = \frac{\left(\frac{\lambda}{\psi}\right)^{\frac{v}{2}}}{2K_v(\sqrt{\lambda\psi})} \tau^{v-1} \exp\left[-\frac{1}{2}\left\{\left(\frac{\psi}{\tau}\right) + \lambda\tau\right\}\right], \tau > 0 \quad (11)$$

Whereas:

$(\lambda, \psi)$ : The scale parameters.

$v$ : The shape parameter.

The Bayesian sample size is found to estimate the mean parameter conditioned by the random variable  $(\tau)$  by calculating the total cost function as follows:

$$TC_{(n)} = C_{(n)} + PR$$

By substituting, the linear cost function  $C_{(n)}$  defined in equation (7) and the posterior risk function (PR) (expected the loss function), the total cost function becomes as follows:

$$TC_{(n)} = c_0 + cn + E_{\theta}(L(\hat{\theta}, \theta))$$

Substituting for the quadratic loss function defined in equation (3), to get:

$$TC_{(n)} = c_0 + cn + E_{\theta}(k(\hat{\theta}_{sq} - \theta)^2)$$

And by substituting for the subsequent risk function (PR) shown in equation (6), the total cost function becomes as follows:

$$TC_{(n)} = c_0 + cn + k \text{Var}(\theta|Y) \quad (12)$$

To find  $\text{Var}(\theta|Y)$  the posterior probability distribution of the mean parameter conditional of the random variable  $(\tau)$  must be calculate as follows:

The prior distribution of the mean parameter  $(\theta)$  conditioned by  $(\tau)$  and  $(\sigma^2)$  is describe as follows: [2]

$$(\theta|\sigma^2, \tau) \sim N(\theta_0, \sigma_0^2 \sigma^2 \tau)$$

Since the prior probability density function for  $(\theta|\sigma^2, \tau)$  is as follows:

$$P(\theta|\sigma^2, \tau) \propto e^{-\frac{(\theta-\theta_0)^2}{2\sigma_0^2 \sigma^2 \tau}} \tag{13}$$

After combining the sample information with the initial information about the parameter ( $\theta$ ), the posterior distribution of ( $\theta$ ) conditional on the variable ( $\tau$ ) and ( $\sigma^2$ ) is found using the following equation:

$$P(\theta|\underline{Y}, \sigma^2, \tau) \propto P(\theta|\sigma^2, \tau) f(\underline{Y}|\theta, \sigma^2, \tau) \tag{14}$$

After substituting the probability function for the random vector ( $\underline{Y}$ ) conditional by ( $\tau$ ), defined according to the following equation:

$$f(\underline{Y}|\theta, \sigma^2, \tau) = (2\pi\sigma^2\tau)^{-\frac{n}{2}} e^{-\frac{1}{2\sigma^2\tau}\sum(Y_i-\theta)^2} \tag{15}$$

With the prior distribution of the parameter ( $\theta$ ) conditional by the variables ( $\tau$ ) and ( $\sigma^2$ ); in the equation (14) above to get the kernel of the posterior distribution of ( $\theta$ ) conditional by ( $\tau$ ) and ( $\sigma^2$ ) ( $P(\theta|\underline{Y}, \sigma^2, \tau)$ ) which is defined below:

$$P(\theta|\underline{Y}, \sigma^2, \tau) \propto e^{-\frac{1}{2\sigma^2\tau}\left[n(\theta-\bar{Y})^2 + \frac{(\theta-\theta_0)^2}{\sigma_0^2}\right]} \tag{16}$$

The quadratic form in equation (16) above is similar to the following quadratic form: [2]

$$A(z - a)^2 + B(z - b)^2 = (A + B)(Z - m)^2 + \frac{AB}{A + B} (a - b)^2 \tag{17}$$

Whereas

$$m = \frac{aA + bB}{A + B} \tag{18}$$

Therefore, the posterior probability distribution for the mean parameter conditioned by ( $\tau$ ) and ( $\sigma^2$ ) is normal and is as follows:

$$P(\theta|\underline{Y}, \sigma^2, \tau) = \frac{\left(\frac{1}{\sigma_0^2} + n\right)^{0.5}}{\sqrt{2\pi\sigma^2\tau}} e^{-\frac{\frac{1}{\sigma_0^2} + n}{2\sigma^2\tau}[\theta - m]^2} \tag{19}$$

Since:

$$m = \frac{\theta_0 + n\bar{Y} \sigma_0^2}{1 + n \sigma_0^2} \tag{20}$$

The distribution is descriptively described by:

$$(\theta|\underline{Y}, \sigma^2, \tau) \sim N\left(m, \frac{\sigma^2\tau}{\frac{1}{\sigma_0^2} + n}\right)$$

Using the concept of mixed distributions, the posterior probability distribution for the mean parameter unconditional of ( $\tau$ ) is as follows:

$$P(\theta|\underline{Y}, \sigma^2) = \int_0^{\infty} P(\theta|\underline{Y}, \sigma^2, \tau) P(\tau) d\tau \quad (21)$$

After substituting the components of equation (21) and performing the integration concerning ( $\tau$ ), yielding the following equation:

$$P(\theta|\underline{Y}, \sigma^2) = \frac{\left(\frac{1}{\sigma_0^2} + n\right)^{0.5} \left(\frac{\lambda}{\psi}\right)^{0.25}}{\sqrt{2\pi\sigma^2} K_v(\sqrt{\lambda\psi})} \left[ 1 + \frac{\left(\frac{1}{\sigma_0^2} + n\right) (\theta - m)^2}{\psi\sigma^2} \right]^{\frac{2v-1}{4}} \\ \times K_{\frac{2v-1}{2}} \left( \sqrt{\lambda\psi \left( 1 + \frac{\left(\frac{1}{\sigma_0^2} + n\right) (\theta - m)^2}{\psi\sigma^2} \right)} \right) \quad (22)$$

The probability density function of  $(\theta|\underline{Y}, \sigma^2)$  defined in equation (22) represents the generalized modified Bessel distribution and is descriptively expressed as:

$$(\theta|\underline{Y}, \sigma^2) \sim GMBD \left( m, \frac{\sigma^2}{\frac{1}{\sigma_0^2} + n}, \lambda, \psi, v \right)$$

Since

$$m = \frac{\theta_0 + n\bar{Y} \sigma_0^2}{1 + n \sigma_0^2} \quad (23)$$

The unconditional Bayes estimator for the parameter ( $\theta$ ) is:

$$E(\theta|\underline{Y}, \sigma^2) = E_{\tau} E(\theta|\underline{Y}, \sigma^2, \tau) = m \quad (24)$$

Since ( $m$ ) defined in equation (23) represents a Bayes estimator under the quadratic loss function  $\hat{\theta}_{sq}$ .

$$\hat{\theta}_{sq} = \frac{\theta_0 + n\bar{Y} \sigma_0^2}{1 + n \sigma_0^2} \quad (25)$$

Using the properties of mathematical expectation, it is possible to find the variance for the posterior distribution of  $(\theta|\underline{Y}, \sigma^2)$  as follows: [5]

$$Var(\theta|\underline{Y}, \sigma^2) = E[Var(\theta|\underline{Y}, \sigma^2, \tau)] + Var[E(\theta|\underline{Y}, \sigma^2, \tau)] \\ = \frac{\sigma^2}{\frac{1}{\sigma_0^2} + n} E(\tau) \quad (26)$$

$$= \frac{\sigma^2}{\frac{1}{\sigma_0^2} + n} \frac{K_{v+1}(\sqrt{\lambda\psi})}{K_v(\sqrt{\lambda\psi})} \left(\frac{\lambda}{\psi}\right)^{-0.5} \quad (27)$$

By substituting equation (27) above into equation (12), the total cost function is as follows:

$$TC_{(n)} = c_0 + cn + k \frac{\sigma_0^2 \sigma^2}{1 + n\sigma_0^2} \frac{K_{v+1}(\sqrt{\lambda\psi})}{K_v(\sqrt{\lambda\psi})} \left(\frac{\lambda}{\psi}\right)^{-0.5} \quad (28)$$

Now by finding the derivative of equation (28) relative to  $(n)$  and equal the derivative to zero, we get the Bayesian sample size under the quadratic loss function:

$$\frac{\partial TC_{(n)}}{\partial n} = c - \frac{k \sigma_0^4 \sigma^2}{(1 + n\sigma_0^2)^2} \frac{K_{v+1}(\sqrt{\lambda\psi})}{K_v(\sqrt{\lambda\psi})} \left(\frac{\lambda}{\psi}\right)^{-0.5}$$

$$\frac{\partial TC_{(n)}}{\partial n} = 0$$

$$c = \frac{k \sigma_0^4 \sigma^2}{(1 + n^*_{sq} \sigma_0^2)^2} \frac{K_{v+1}(\sqrt{\lambda\psi})}{K_v(\sqrt{\lambda\psi})} \left(\frac{\lambda}{\psi}\right)^{-0.5}$$

$$1 + n^*_{sq} \sigma_0^2 = \sigma_0^2 \sqrt{\frac{k \sigma^2}{c} \frac{K_{v+1}(\sqrt{\lambda\psi})}{K_v(\sqrt{\lambda\psi})} \left(\frac{\lambda}{\psi}\right)^{-0.5}}$$

Where

$$n^*_{sq} = \sqrt{\frac{k \sigma^2}{c} \frac{K_{v+1}(\sqrt{\lambda\psi})}{K_v(\sqrt{\lambda\psi})} \left(\frac{\lambda}{\psi}\right)^{-0.5}} - \frac{1}{\sigma_0^2} \quad (29)$$

Since  $(n^*_{sq})$  represents the Bayesian sample size under the quadratic loss function.

## 8. The sampling and non-sampling status of estimating the mean parameter

In the case that is not sample taken ( $n^* = 0$ ), the total cost under the quadratic loss function shown in equation (28) is: [2]

$$TC_{(0)} = k \frac{\sigma_0^2 \sigma^2}{K_v(\sqrt{\lambda\psi})} \left(\frac{\lambda}{\psi}\right)^{-0.5} \quad (30)$$

When taking a sample ( $n^* = n^*_{sq}$ ), the total cost will be:

$$TC_{(n)} = c_0 + c n^*_{sq} + k \frac{\sigma_0^2 \sigma^2}{1 + n^*_{sq} \sigma_0^2} \frac{K_{v+1}(\sqrt{\lambda\psi})}{K_v(\sqrt{\lambda\psi})} \left(\frac{\lambda}{\psi}\right)^{-0.5}$$

Substituting for  $(n^*_{sq})$  previously defined in equation (29), we get:

$$TC_{(n)} = c_0 + c \left[ \sqrt{\frac{k \sigma^2 K_{v+1}(\sqrt{\lambda\psi})}{c K_v(\sqrt{\lambda\psi})} \left(\frac{\lambda}{\psi}\right)^{-0.5} - \frac{1}{\sigma_0^2}} \right] + k \frac{\sigma_0^2 \sigma^2}{1 + \left[ \sqrt{\frac{k \sigma^2 K_{v+1}(\sqrt{\lambda\psi})}{c K_v(\sqrt{\lambda\psi})} \left(\frac{\lambda}{\psi}\right)^{-0.5} - \frac{1}{\sigma_0^2}} \right] \sigma_0^2} \frac{K_{v+1}(\sqrt{\lambda\psi})}{K_v(\sqrt{\lambda\psi})} \left(\frac{\lambda}{\psi}\right)^{-0.5}$$

Performing some algebraic operations for the above equation, we get:

$$TC_{(n)} = c_0 - \frac{c}{\sigma_0^2} + 2 \sqrt{\frac{c k \sigma^2 K_{v+1}(\sqrt{\lambda\psi})}{K_v(\sqrt{\lambda\psi})} \left(\frac{\lambda}{\psi}\right)^{-0.5}} \tag{31}$$

Therefore, the optimal Bayes sample size is generally:

$$n^* = \max \left[ 0, \sqrt{\frac{k \sigma^2 K_{v+1}(\sqrt{\lambda\psi})}{c K_v(\sqrt{\lambda\psi})} \left(\frac{\lambda}{\psi}\right)^{-0.5} - \frac{1}{\sigma_0^2}} \right] \tag{32}$$

Therefore, it is decided to take a sample in the case of sampling when the

$$\left[ \sqrt{\frac{k \sigma^2 K_{v+1}(\sqrt{\lambda\psi})}{c K_v(\sqrt{\lambda\psi})} \left(\frac{\lambda}{\psi}\right)^{-0.5} - \frac{1}{\sigma_0^2}} \right]$$

is positive, and the total cost in the case of sampling is less than the total cost in the case of non-sampling.

That is:

$$k \frac{\sigma_0^2 \sigma^2}{K_v(\sqrt{\lambda\psi})} \left(\frac{\lambda}{\psi}\right)^{-0.5} > c_0 - \frac{c}{\sigma_0^2} + 2 \sqrt{\frac{c k \sigma^2 K_{v+1}(\sqrt{\lambda\psi})}{K_v(\sqrt{\lambda\psi})} \left(\frac{\lambda}{\psi}\right)^{-0.5}} \ \& \ \sqrt{\frac{k \sigma^2 K_{v+1}(\sqrt{\lambda\psi})}{c K_v(\sqrt{\lambda\psi})} \left(\frac{\lambda}{\psi}\right)^{-0.5}} > \frac{1}{\sigma_0^2}$$

And it is decided not to take a sample in the case of sampling when the

$$\left[ \sqrt{\frac{k \sigma^2 K_{v+1}(\sqrt{\lambda\psi})}{c K_v(\sqrt{\lambda\psi})} \left(\frac{\lambda}{\psi}\right)^{-0.5} - \frac{1}{\sigma_0^2}} \right]$$

is negative, and the total cost in the case of sampling is greater than the total cost in the case of non-sampling.

That is:

$$k \frac{\sigma_0^2 \sigma^2 K_{v+1}(\sqrt{\lambda\psi})}{K_v(\sqrt{\lambda\psi})} \left(\frac{\lambda}{\psi}\right)^{-0.5} < c_0 - \frac{c}{\sigma_0^2} + 2 \sqrt{\frac{c k \sigma^2 K_{v+1}(\sqrt{\lambda\psi})}{K_v(\sqrt{\lambda\psi})} \left(\frac{\lambda}{\psi}\right)^{-0.5}} \& \sqrt{\frac{k \sigma^2 K_{v+1}(\sqrt{\lambda\psi})}{c K_v(\sqrt{\lambda\psi})} \left(\frac{\lambda}{\psi}\right)^{-0.5}} > \frac{1}{\sigma_0^2}$$

By solving the inequality and substituting for each  $\sqrt{c}$  of  $(\vartheta)$  in the sampling case, we get the following formula:

$$k \frac{\sigma_0^2 \sigma^2 K_{v+1}(\sqrt{\lambda\psi})}{K_v(\sqrt{\lambda\psi})} \left(\frac{\lambda}{\psi}\right)^{-0.5} > c_0 - \frac{c}{\sigma_0^2} + 2 \sqrt{\frac{c k \sigma^2 K_{v+1}(\sqrt{\lambda\psi})}{K_v(\sqrt{\lambda\psi})} \left(\frac{\lambda}{\psi}\right)^{-0.5}} \& \sqrt{\frac{k \sigma^2 K_{v+1}(\sqrt{\lambda\psi})}{c K_v(\sqrt{\lambda\psi})} \left(\frac{\lambda}{\psi}\right)^{-0.5}} > \frac{1}{\sigma_0^2}$$

$$\frac{\vartheta^2}{\sigma_0^2} - 2\vartheta \sqrt{\frac{k \sigma^2 K_{v+1}(\sqrt{\lambda\psi})}{K_v(\sqrt{\lambda\psi})} \left(\frac{\lambda}{\psi}\right)^{-0.5}} + k \frac{\sigma_0^2 \sigma^2 K_{v+1}(\sqrt{\lambda\psi})}{K_v(\sqrt{\lambda\psi})} \left(\frac{\lambda}{\psi}\right)^{-0.5} - c_0 > 0$$

$$\& \vartheta < \sqrt{\frac{k \sigma^2 \sigma_0^4 K_{v+1}(\sqrt{\lambda\psi})}{K_v(\sqrt{\lambda\psi})} \left(\frac{\lambda}{\psi}\right)^{-0.5}}$$

$$\frac{\vartheta^2}{\sigma_0^2} - 2\vartheta \sqrt{\frac{k \sigma^2 K_{v+1}(\sqrt{\lambda\psi})}{K_v(\sqrt{\lambda\psi})} \left(\frac{\lambda}{\psi}\right)^{-0.5}} + k \frac{\sigma_0^2 \sigma^2 K_{v+1}(\sqrt{\lambda\psi})}{K_v(\sqrt{\lambda\psi})} \left(\frac{\lambda}{\psi}\right)^{-0.5} > c_0$$

$$\& \vartheta < \sqrt{\frac{k \sigma^2 \sigma_0^4 K_{v+1}(\sqrt{\lambda\psi})}{K_v(\sqrt{\lambda\psi})} \left(\frac{\lambda}{\psi}\right)^{-0.5}}$$

Completing the square for the left side of the inequality, to get:

$$\left( \frac{\vartheta}{\sqrt{\sigma_0^2}} - \sqrt{k \frac{\sigma_0^2 \sigma^2 K_{v+1}(\sqrt{\lambda\psi})}{K_v(\sqrt{\lambda\psi})} \left(\frac{\lambda}{\psi}\right)^{-0.5}} \right)^2 > c_0$$

Taking the square root of both sides of the inequality, to get:

$$\mp \left( \frac{\vartheta}{\sqrt{\sigma_0^2}} - \sqrt{k \frac{\sigma_0^2 \sigma^2 K_{v+1}(\sqrt{\lambda\psi})}{K_v(\sqrt{\lambda\psi})} \left(\frac{\lambda}{\psi}\right)^{-0.5}} \right) > \sqrt{c_0}$$

The roots of the above inequality are as follows:

$$\vartheta > \sqrt{k \frac{\sigma_0^4 \sigma^2 K_{v+1}(\sqrt{\lambda\psi})}{K_v(\sqrt{\lambda\psi})} \left(\frac{\lambda}{\psi}\right)^{-0.5}} + \sqrt{c_0 \sigma_0^2}$$

Or:

$$\vartheta < \sqrt{k \frac{\sigma_0^4 \sigma^2 K_{v+1}(\sqrt{\lambda\psi})}{K_v(\sqrt{\lambda\psi})} \left(\frac{\lambda}{\psi}\right)^{-0.5}} - \sqrt{c_0 \sigma_0^2}$$

But work interest is in the lowest cost, so taking the second root to show when the sampling takes place or not, as follows:

$$\sqrt{c} < \sqrt{k \frac{\sigma_0^4 \sigma^2 K_{v+1}(\sqrt{\lambda\psi})}{K_v(\sqrt{\lambda\psi})} \left(\frac{\lambda}{\psi}\right)^{-0.5}} - \sqrt{c_0 \sigma_0^2}$$

So it is sampling when:

$$c < \left[ \sqrt{k \frac{\sigma_0^4 \sigma^2 K_{v+1}(\sqrt{\lambda\psi})}{K_v(\sqrt{\lambda\psi})} \left(\frac{\lambda}{\psi}\right)^{-0.5}} - \sqrt{c_0 \sigma_0^2} \right]^2$$

And non-sampling when:

$$c > \left[ \sqrt{k \frac{\sigma_0^4 \sigma^2 K_{v+1}(\sqrt{\lambda\psi})}{K_v(\sqrt{\lambda\psi})} \left(\frac{\lambda}{\psi}\right)^{-0.5}} - \sqrt{c_0 \sigma_0^2} \right]^2$$

If the sampling cost per unit exceeds the

$$\left[ \sqrt{k \frac{\sigma_0^4 \sigma^2 K_{v+1}(\sqrt{\lambda\psi})}{K_v(\sqrt{\lambda\psi})} \left(\frac{\lambda}{\psi}\right)^{-0.5}} - \sqrt{c_0 \sigma_0^2} \right],$$

the sampling is not of value due to the high cost of sampling per unit, and in this case, only the prior information is relied on to estimate the mean parameter for the generalized modified Bessel distribution.

## 9. Simulation

In this section, random data was generated that follow the generalized modified Bessel distribution through the concept of mixed distribution in the programming language (Matlab R2015a) [12], After generating the observations, it was tested whether the data followed the generalized modified Bessel distribution by Kolmogorov-Smirnov test if it was found that the (p-value) is greater than the level of significance ( $\alpha = 0.05$ ) at different values of the shape parameters, and therefore the hypothesis that the data follows the generalized modified Bessel distribution is accepted. And by executing the program, the estimator of the mean

parameter is calculated under the quadratic loss function and defined in equation (25) at different initial values and the comparison through (MSE):

$$MSE(\hat{\theta}_{sq}) = \frac{\sum_{i=1}^R (\hat{\theta}_{sq} - \theta)^2}{R} \tag{33}$$

Where is the number of iterations ( $R = 100$ ).

This comparison will be done in two cases. In the first case, the optimal Bayesian sample size was used to calculate the mean parameter defined in the equation (29) and in the second case, different sample sizes were used ( $n = 25,75,125$ ). Table (1) shows the assumed values of generation and calculates estimators:

**Table 1. Assumed values for generation and for calculating estimators**

Shape parameters			$\theta$	$\sigma^2$	$\theta_0$	$\sigma_0^2$	$c$	$k$
$\lambda = 2$	$\psi = 1.5$	$v = 3$	33.5	0.5	3	1.5	0.01	2
$\lambda = 8$	$\psi = 3$	$v = 10$		4	7		0.005	5

Table (2) and (3) shows the (MSE) values for an estimate of the mean parameter at different shape parameters and different sample sizes when ( $\sigma^2 = 0.5,4$ ).

**Table 2. MSE values for different shape parameters ( $\lambda, \psi, v$ ) and different sample sizes at ( $\sigma^2 = 0.5$ )**

$(\lambda, \psi, v)$	n	$\theta_0$	$\hat{\theta}_{sq}$	MSE(n)
<b>(2,1.5,3)</b>	25	3	3.3001	0.0502
		7	3.0054	0.0681
	75	3	3.3615	0.0453
		7	3.1242	0.0601
	125	3	3.5123	0.0398
		7	3.4458	0.0495
<b>(8,3,10)</b>	25	3	3.3102	0.0428
		7	3.1254	0.0599
	75	3	3.3984	0.0414
		7	3.2451	0.0576
	125	3	3.5012	0.0202
		7	3.4746	0.0399

**Table 3. MSE values for different shape parameters  $(\lambda, \psi, v)$  and different sample sizes at  $(\sigma^2 = 4)$**

$(\lambda, \psi, v)$	N	$\theta_0$	$\hat{\theta}_{sq}$	MSE(n)
(2,1.5,3)	25	3	3.9101	0.0880
		7	4.0604	0.0989
	75	3	3.1923	0.0778
		7	3.9781	0.0928
	125	3	3.7825	0.0682
		7	3.1632	0.0789
(8,3,10)	25	3	3.2524	0.0732
		7	3.8942	0.0900
	75	3	3.2998	0.0712
		7	3.9321	0.0892
	125	3	3.6123	0.0532
		7	3.3549	0.0692

We notice from tables (2) & (3), that the (MSE) value decreases with the increase in the shape parameters and sample size at  $(\sigma^2 = 0.5)$  and when the value of  $(\theta_0)$  is close to  $(\theta)$ .

Table (4) represents the Bayesian sample size values for different shape parameters  $(\lambda, \psi, v)$  when  $(\sigma^2 = 0.5, \theta_0 = 3)$ .

**Table 4. Bayesian sample size values for different shape parameters  $(\lambda, \psi, v)$**

$(\lambda, \psi, v)$	k	c	$n^*_{sq}$	$\hat{\theta}_{sq}$	MSE( $n^*_{sq}$ )
(2,1.5,3)	2	0.01	18	2.9171	0.2345
		0.005	23	3.0014	0.1982
	5	0.01	36	3.2045	0.1156
		0.005	41	3.2204	0.0968
(8,3,10)	2	0.01	27	3.1121	0.1756
		0.005	31	3.1984	0.1536
	5	0.01	55	3.7956	0.0864
		0.005	72	3.6548	0.0645

We notice from table (4) that the Bayesian sample size increases with the increase in the shape parameters and the (MSE) value decreases with the increase in the shape parameters, as well as the Bayesian sample size is inversely proportional to the sampling cost (c) per unit.

## 10. Conclusions

- The posterior probability distribution for the mean parameter follows the generalized modified Bessel distribution.
- The Bayesian sample size was determined to estimate the mean parameter of the generalized modified Bessel distribution when the parameters  $(\lambda, \psi, \nu, \sigma^2)$  are known, and the sampling decision is made under quadratic loss function when the sampling cost per unit (c) is less than the

$$\left[ \sqrt{k \frac{\sigma_0^4 \sigma^2 K_{\nu+1}(\sqrt{\lambda\psi})}{K_{\nu}(\sqrt{\lambda\psi})} \left(\frac{\lambda}{\psi}\right)^{-0.5} - \sqrt{c_0 \sigma_0^2}} \right]^2.$$

- Through simulation, we notice that when the values of the shape parameters are increased, the values of the estimators of the mean parameter using the Bayesian sample size are close to the values of the estimators of the mean parameter using a relatively large sample size.

## 11. Recommendations

- Using a non-linear cost function to determine the Bayesian sample size to estimate the mean parameter of the generalized modified Bessel distribution under the quadratic loss function.
- Traditional sample size analysis for estimating parameters of the generalized modified Bessel distribution.
- Using other distributions with heavy tails to determine the Bayesian sample size, including the double exponential distribution (Laplace distribution).
- Using other symmetric and asymmetric loss functions to determine the Bayesian sample size to estimate the mean parameter of the generalized modified Bessel distribution.

## 12. REFERENCES

1. Barndorff-Nielsen, O., "Hyperbolic distributions and distributions on hyperbolae", Scand.J.Statist.5, P.P.151-157. 1978.
2. Box, G.E.P. and Taio, G.C., "Bayesian inference in statistical Analysis", Addition-Wesley Publishing Company, California, London, 1973.
3. Butler, R. W., "Generalized inverse Gaussian distribution and their wishart connection", Colorado state university, Vol.25, P.P.69-75, USA, 1998.
4. DeGroot, M.H., "Optimal statistical decision", McGraw-Hill.1970.

5. Hogg R. et al., "Introduction to Mathematical Statistics (6th edition)", Macmillan publishers company, New york, 2005.
6. Hu, W., "Calibration of multivariate generalized hyperbolic distribution using the EM algorithm, with applications in risk management, Portfolio optimization and portfolio credit risk", Florida State University, Electronic theses, Treatises and Dissertation, The Graduate School, 2005.
7. Koudou, A. E. and Ley, C., " Characterizations of GIG laws: a survey complemented with two new results", *Proba. Surv.* , vol. 11, p.p. 161-176, 2014.
8. Lindley, D.V., "Bayesian statistics, a Review", Philadelphia Society for Industrial and Applied Mathematics, 1972.
9. Mora, J. A. N. & Mata, L. M., "Numerical aspects to estimate the generalized hyperbolic probability distribution", *Journal of Finance & Economics*, vol. 1(4), p.p. 1-9, 2013.
10. Press, S. J., "Subjective and objective Bayesian statistics", A John Wiley and Sons, 2003.
11. Saiful Islam, A. F. M., "Loss function, utility function and Bayesian sample size determination", PhD Thesis, University of London, 2011.
12. Salih, S. A. and Aboudi, E. H., "Bayesian Inference of a Non-normal Multivariate Partial Linear Regression Model" *Iraqi Journal of Statistical Science* (34), pp. 91-115, 2021.
13. Thabane, L. and Drekić, S., "Hypothesis testing for the generalized multivariate modified Bessel model", *Journal of Multivariate Analysis*, 86, P.P.360-374, 2003.
14. Thabane, L. & Haq, M. S., "The generalized multivariate modified Bessel distribution and its Bayesian applications", *Journal of Applied Statistical Science*, vol. 11(3), p.p.225-267, 2003.

**Publication Prerequisites and terms**

- 1- The journal publishes scientific research and studies in statistics and informatics written in Arabic, English and French, to make it clear that research submitted for publication has been published or submitted for publication in magazines or other periodicals or presented and published in periodicals for conferences or seminars.
- 2- Send electronic copies (word & PDF) of the research and studies to the editor should include the name of the researcher or researchers and their scientific titles and places of work with the address of the correspondence, the numbers of telephones and e-mail. The research to be published should be sent electronically in accordance with the specifications below:
  - a. To be printed on A4 paper and be in the form of a single column and use the Type simplified Arabic and Times New Roman for English and French and with a font size (12). Using Microsoft Word and on one face of the paper.
  - b. The margin is 2.5 cm for all sides of the paper.
  - c. The researcher will attach a summary of his research in Arabic, English, or French in no more than one page.
  - d. Place references at the end of the paper and separated page. It is recommended to use the Harvard system of referencing, which (author's name, year of publication, source address, publishing house, country).
  - e. Numbered tables, illustrations, and others as they are received in the research, documents as aliases of the original sources.
  - f. The number of search or study pages should not exceed (25) pages.
- 3- Authors will notified of receiving their research within two working days from the date of receipt of the research.
- 4- Referees will evaluate all submitted research, the Authors will informed of the proposed evaluation and modifications if any within two weeks of receipt of the research.
- 5- The editorial board of the Journal has the right to accept or reject the research and has the right to make any modification or partial redrafting of the material submitted for publication in accordance with the format adopted in its publication after the approval of the researcher.
- 6- Published research becomes the property of the Journal and may not republished elsewhere.
- 7- The articles published in the magazine reflect the opinions of the authors, and do not necessarily reflect the view of the Journal or the Arab Institute for Training and Research in Statistics.
- 8- The research is sent to the magazine's e-mail address:  
[journal@aitrs.org](mailto:journal@aitrs.org) or [Info@aitrs.org](mailto:Info@aitrs.org)

# Journal of Statistical Sciences

Scientific Referred Journal

## Editorial Board

### Editor-in Chief

Mr. Hedi Saidi

### Editorial Secretary

Dr. Bachioua Lahcene

## Editorial Board Members

Prof. Dr. Faisal Al-Sharabi  
Dr. Salwa Mahmoud Assar  
Dr. Hassan Abuhassan

Prof. Dr. Abed Khaliq Tohami  
Prof Dr. Ahmed Shaker Almutwali  
Dr. Hamid Bouzida

Prof. Dr. Mukhtar Al-Kouki  
Prof. Dr. Issa Masarweh

## Scientific Consulting Committee

Dr. Qassim Al-Zoubi  
Dr. Daaa Awad  
Dr. luay shabaneh

Dr. Nabeel M. Shams  
Dr. Khalifa Al-Barwani  
Prof. Dr. Ghazi Raho  
Dr. Ola Awad

Prof. Dr. Awad Haje Ali  
Prof. Dr. Maytham Elaibi Ismael  
Dr. Mohammed Husain Ali Al-Janabi

Listed in Ulrich's website

[www.ulrichsweb.com](http://www.ulrichsweb.com)

Classified in The Arab Citation & Impact Factor (Arcif)

[www.emarefa.net/arcif/](http://www.emarefa.net/arcif/)

ISSN 2522-64X (Online), ISSN 2519-948X (Print)