Warith Scientific Journal

التقدير البيزي لمعلمات التوزيع الكروي باستعمال دالة الخسارة التربيعية

Bayesian estimation of the parameters of the spherical distribution using the squared loss function

مروه حيدر غازي[1]
Marwa haider ghazi
Mh761075@gmail.com
كلية الإدارة والاقتصاد /جامعة كربلاء
College of administration and economics ,university of Karbala

ا.د عواد كاظم الخالدي[2]
Awad Kazem al-Khalidi
Awad2005_k@gmail.com
كلية الإدارة والاقتصاد / جامعة الوارث
College of administration and economics university of wraith Anbiyaa And university of Karbala

المستخلص

تعدد الأساليب والطرائق الاحصائية في تقدير معلمات النماذج الاحصائية وتميز هذه التقديرات معايير مهمة لبيان الأفضلية في التقدير أهمها معامل الخطأ . والهدف الأساس من أي عملية تقدير هو الوصول إلى أفضل تقدير أو أقرب تقدير للمعلمة المجهولة من بين كل التقديرات الممكنة ، لذلك على متخذ القرار أن يختار الأسلوب أو الصيغة الأفضل لتقدير المعلمة المجهولة .

في هذا البحث تم استعمال التقدير البيزي لمعلمات التوزيع الكروي اذ تم الحصول على القيم الإفتراضية لتوزيع درشلية الكروي ثلاثي الابعاد تجريبياً من إجراء تجارب عدة واختيار القيم التي إستقرت عندها تقديرات بيز وأعطت افضل النتائج عند المعلمات الافتراضية ($(\alpha_1 = 1, \alpha_2 = 1, \alpha_3 = 1 (\alpha_1 = 1, \alpha_2 = 1, \alpha_3 = 1$ اذ أظهرت النتائج ان دالة الخسارة التربيعية حققت افضلية عند حجم عينة (500) حققت افضلية على باقي حجوم العينات في تقدير معلمات التوزيع الكروي يليها حجم العينة (400) لانها سجلت اقل متوسط مربعات خطأ

عند المعلمات الافتراضية $(\alpha_1 = 1, \alpha_2 = 2, \alpha_3 = 1)(\alpha_1 = 1, \alpha_2 = 2, \alpha_3 = 1)$ ان طريقة بيز في ظل دالة خسارة (التربيعية) عند حجم عينة (500) حققت افضلية على باقي حجوم العينات في تقدير معلمات التوزيع الكروي يليها حجم العينة (400) لانها سجلت اقل متوسط مربعات خطأ.

الكلمات المفتاحية: نظرية بيز , التوزيع الكروي ديريشليه دالة الخسارة

Abstract

There are many statistical methods and methods in estimating the parameters of statistical models, and these estimates distinguish important criteria for indicating preference in the estimation, the most important of which is the error coefficient . The main goal of any estimation process is to reach the best estimate or the closest estimate of the unknown parameter out of all possible estimates, so the decision maker should choose the best method or formula for estimating the unknown parameter .

In this research, a Bayesian estimation of the parameters of the spherical distribution was used, where the default values of the three-dimensional spherical Dirichlet distribution were obtained experimentally from conducting several experiments and selecting the values at which the Bayes estimates stabilized and gave the best results in the default parameters $(\alpha_1 = 1, \alpha_2 = 1, \alpha_3 = 1)(\alpha_1 = 1, \alpha_2 = 1, \alpha_3 = 1)$ where the results showed that the quadratic loss function achieved an advantage at sample size (500) achieved an advantage over the rest of the sample sizes in estimating the parameters of the spherical distribution followed by the sample size (400) because

At the default parameters $(\alpha_1 = 1, \alpha_2 = 2, \alpha_3 = 1)(\alpha_1 = 1, \alpha_2 = 2, \alpha_3 = 1)$ the Bayes method under the loss function (quadrature) at a sample size (500) achieved an advantage over the rest of the sample sizes in estimating the parameters of the spherical distribution, followed by the sample size (400) because it recorded the lowest average error squares.

Keywords: Bayes ' theorem, spherical Dirichlet distribution , loss function

1- المقدمة Introduction

يُعد التقدير البيزي من الطرق الإحصائية المهمة في تحليل البيانات، ويستخدم لتقدير قيمة مجهولة باستخدام المعلومات السابقة حولها والمعلومات الجديدة التي تم جمعها. يستعمل التقدير البيزي توزيع مُسبق لتقدير قيمة المعلمة المجهولة، ويستند هذا التوزيع إلى المعرفة السابقة المتوفرة أو المفترضة حول المعلمة.

يُعد التقدير البيزي للتوزيع الكروي من التقنيات الإحصائية المهمة التي يتم استخدامها في تحليل البيانات، حيث يتم استخدامها لتحديد قيمة المعلمة المجهولة من خلال توزيع كروي يعتمد على البيانات المتوفرة والمعرفة السابقة حول المعلمة المراد تقديرها .

يمكن استعمال التقدير البيزي للتوزيع الكروي في العديد من التطبيقات الإحصائية، بما في ذلك تحليل البيانات الطبية والتحليل الاقتصادي وتحليل النظم البيئية. ويعد التقدير البيزي للتوزيع الكروي أحد اهم الطرق الاساسية التي يمكن استعمالها لتحليل البيانات التي تتوفر فيها بيانات محدودة أو غير كافية.

2- منهجية البحث Research methodology

1.2 مشكلة البحث the research problem

من خلال دراسة البحوث السابقة برزت فشل الكثير من طرائق تقدير المعلمات للتوزيعات الكروية لذا تركز البحث في إيجاد طريقة غير متحيزة وفعالة للتقدير للحصول على نموذج يمثل البيانات بصورة دقيقة تمثلت هذه الطريقة بتقدير بايزي لمعلمات التوزيع الكروي .

2.2هدف البحث Research objective

يهدف البحث الى تقدير معلمات توزيع ديريشليه الكـــروي(The Spherical-Dirichlet Distribution) (باستعمال طريقة بيز في التقدير عند دالة الخسارة التربيعية المتماثلة (Squared Error Loss Function)

3-الجانب النظري Theoretical part

1.3 نظرية بيز Bayes ' theorem [7] [4]

نظرية بيز (Bayesian theory) نظرية إحصائية تستخدم لتحليل البيانات وتقدير المعلمات المجهولة. تعتمد هذه النظرية على مبدأ الاحتمالات الشرطية ، والتي تعني أن الاحتمالية الحقيقية لحدوث حدث ما تعتمد على المعرفة السابقة التي تملكها حول الحدث او حول المعلمة المراد تقديرها.

يعد الفيلسوف والرياضي العالم الأمريكي توماس فرانسيس بيوني (Thomas Francis Jr). أحد أوائل العلماء الذين طرحوا نظرية في التقدير اذ قام بتطوير نظرية التقدير بالاعتماد على الاحتمالات والعينات المعتمدة على الأعداد العشوائية ونشرها بعد وفاته ريتشاردبرايس(Price, R.) عام (1763) اللذان أول من كتبا العلاقة بين الاحتمالات الحدية والشرطية.

توجد مدرستان في التقدير الاولى تسمى المدرسة التقليدية والتي تفترض ان المعلمات المراد تقديرها كميات ثابتة يتم تقديرها بالطرائق التقليدية مثل طريقة الامكان الاعظم وطريقة المربعات الصغرى وطريقة العزوم وغيرها ، والأخرى تسمى المدرسة البيزية والتي تفترض أن المعلمة (المعلمات) المجهولة الخاصة بالتوزيع الاحتمالي تعد كمتغيرات عشوائية وان هنالك معلومات سابقة عنها (اولية) تصاغ تلك المعلومات على شـــــــكل توزيع احتمالي $(Prior\ Probability\ function)$ QUOTE يعرف بدالة الكثافة الاحتمالية السابقة $(Prior\ Probability\ function)$.

تعتمد نظرية بيز على المعلومات الحالية للعينة التي يمكن ان تُمثل بدالة الامكان (Likelihood Function) الخاصة بالمشاهدات الحالية. وبدمج دالة الكثافة الاحتمالية السابقة للمعلمات مع دالة الامكان الاعظم للمشاهدات الحالية نحصل على التوزيع الاحتمالي اللاحق(Posterior) وعند دالة خسارة (Loss Function) التي يمكن عن طريقها قياس الخسارة الناتجة عن اتخاذ القرار بالاعتماد على قيمة ($\hat{\underline{\alpha}}\ \hat{\underline{\alpha}}$ QUOTE) بينما القرار الواجب اتخاذه يعتمد على ($\underline{\alpha}\ \underline{\alpha}$ QUOTE) اي وجود فرق بين المعلمة وتقديرها.

2.3 خطوات ايجاد مقدر بيز (Steps of Bayesian estimator)

لإيجاد مقدر بيز يجب ان نتبع الخطوات الآتية

1- معرفة الدالة الاساسية (دالة الكثافة الاحتمالية للتوزيع المدروس) وهي الدالة الاحتمالية لتوزيع البيانات QUOTE $f(x_1, x_2, \ldots\ldots, x_n|\underline{\alpha})$ $f(x_1, x_2, \ldots\ldots, x_n|\underline{\alpha})$ ومنها نحصل على دالة الامكان (Likelihood

$l(x_1,x_2,\dots\dots,x_n|\underline{\alpha}) = \prod_{i=1}^{n} l(xi|\underline{\alpha})$ QUOTE للمشاهدات الحالية (function

$$l(x_1,x_2,\dots\dots,x_n|\underline{\alpha}) = \prod_{i=1}^{n} l(xi|\underline{\alpha})$$

.

2- حيث الاحتمال السابق (Prior Distribution) QUOTE) وهو التوزيع الاحتمالي الاولي لمعلمات التوزيع الاحتمالي المدروس ويرمز له $\pi(\underline{\alpha})$ $\pi(\underline{\alpha})$ QUOTE

3- نجد الاحتمال اللاحق (Posterior Distribution) ويرمز له $h(\alpha_i|x_1,x_2,\dots\dots,x_n)$ QUOTE يتم الحصول على التوزيع اللاحق باستعمال صيغة بيز العكسية $h(\alpha_i|x_1,x_2,\dots\dots,x_n)$ QUOTE وكالآتي: $(Bayesian\ Inversion\ Formula)$ $(Bayesian\ Inversion\ Formula)$

$$h(\underline{\alpha}|x_1,x_2,\dots\dots,x_n) = \frac{l(x_1,x_2,\dots\dots,x_n|\underline{\alpha})\pi(\underline{\alpha})}{\int_{\forall\underline{\alpha}} l(x_1,x_2,\dots\dots,x_n|\underline{\alpha})\pi(\underline{\alpha})d\underline{\alpha}}$$

QUOTE

$$h(\underline{\alpha}|x_1,x_2,\dots\dots,x_n) = \frac{l(x_1,x_2,\dots\dots,x_n|\underline{\alpha})\pi(\underline{\alpha})}{\int_{\forall\underline{\alpha}} l(x_1,x_2,\dots\dots,x_n|\underline{\alpha})\pi(\underline{\alpha})d\underline{\alpha}} \quad \dots (1\text{-}1)$$

وتُعرف بأنها دالة شَرطية لمتجه المعلمات $(\underline{\alpha})$ $(\underline{\alpha})$ QUOTE بوجود معلومات العينة الحالية وهي دالة تمثل كل المعلومات السابقة والحالية عن $(\underline{\alpha})$ $(\underline{\alpha})$ QUOTE المراد تقديرها .

ومن معادلة (1-2) يمكن كتابة الآتي :

$$f(x) = \int l(x_1,x_2,\dots\dots,x_n|\underline{\alpha})\pi(\underline{\alpha})d\underline{\alpha}$$

QUOTE

$$f(x) = \int l(x_1,x_2,\dots\dots,x_n|\underline{\alpha})\pi(\underline{\alpha})d\underline{\alpha} \quad \dots (1\text{-}2)$$

والتي تمثل دالة الكثافة الاحتمالية الحدية $(Marginal\ p.d.f)$ $(Marginal\ p.d.f)$ QUOTE للمتغير العشوائي المشاهد . (x) (x) QUOTE

وأن :

$$g(x_1,x_2,\dots\dots,x_n|\underline{\alpha}) = l(x_1,x_2,\dots\dots,x_n|\underline{\alpha})\pi(\underline{\alpha})$$

QUOTE

$$g(x_1,x_2,\dots\dots,x_n|\underline{\alpha}) = l(x_1,x_2,\dots\dots,x_n|\underline{\alpha})\pi(\underline{\alpha})$$

QUOTE … (1-3)

تمثل دالة الكثافة الاحتمالية المشتركة ($Joint\ p.d.f$ $Joint\ p.d.f$ QUOTE). إذ يمكن التعبير عن دالة الكثافة الاحتمالية اللاحقة التي تتناسب مع دالة الكثافة الاحتمالية المشتركة وكالآتي :

$$h(\underline{\alpha}|x_1,x_2,\dots\dots,x_n) \propto g(x_1,x_2,\dots\dots,x_n|\underline{\alpha})$$

QUOTE

$$h(\underline{\alpha}|x_1,x_2,\dots\dots,x_n) \propto g(x_1,x_2,\dots\dots,x_n|\underline{\alpha}) \quad \dots (1\text{-}4)$$

ومن ثُم فإن مقدر بيز ($\hat{\underline{\alpha}}_{Bayes}$ $\hat{\underline{\alpha}}_{Bayes}$ QUOTE) لمتجه المعلمات ($\underline{\alpha}$ $\underline{\alpha}$ QUOTE) يمكن الحصول عليه باستعمال اي دالة خسارة ($Loss\ Function$ $Loss\ Function$ QUOTE) إذ يرمز لها بالرمز $l(\hat{\alpha},\alpha)$ $l(\hat{\alpha},\alpha)$ QUOTE التي يعد توفرها أمراً أساسياً في طريقة بيز في التقدير لايجاد دالة المخاطرة البيزية $(Bayesian\ Risk\ Function)$ $(Bayesian\ Risk\ Function)$ QUOTE والتي هي عبارة عن توقع دالة الخسارة $(Expected\ of\ Loss\ Function)$ $(Expected\ of\ Loss\ Function)$ QUOTE وكالآتي:-

QUOTE $V(\hat{\underline{\alpha}},\underline{\alpha}) = E[\, l\,(\hat{\underline{\alpha}},\underline{\alpha})\,]$ $V(\hat{\underline{\alpha}},\underline{\alpha}) = E[\, l\,(\hat{\underline{\alpha}},\underline{\alpha})\,]$

$$= \int_{\forall \underline{\alpha}}^{\square} l\,(\hat{\underline{\alpha}},\underline{\alpha})\, h(\underline{\alpha}\,|x_1,x_2,\ldots\ldots,x_n)\, d\underline{\alpha}$$

QUOTE

$$= \int_{\forall \underline{\alpha}}^{\square} l\,(\hat{\underline{\alpha}},\underline{\alpha})\, h(\underline{\alpha}\,|x_1,x_2,\ldots\ldots,x_n)\, d\underline{\alpha}$$

… (1-5)

إذ إن مقدر بيز (QUOTE $\hat{\underline{\alpha}}_{Bayes}$ $\hat{\underline{\alpha}}_{Bayes}$) لمتجه المعلمات QUOTE $(\underline{\alpha})$ $(\underline{\alpha})$ هو المقدر الذي يجعل دالة المخاطرة في نهايتها الصغرى

[4]

3.3 دالة الخسارة التربيعية (Squared Error Loss Function)

تعد دالة الخسارة التربيعية ((Squared Error Loss Function) واحدة من أكثر الدوال الإحصائية استخدامًا في تحليل البيانات وتقدير المعلمات المجهولة. تعتمد هذه الدالة على حساب الفرق بين التقدير الحالي والقيمة الحقيقية للمتغيرات، وتركز على الحد الأدنى للخسارة التربيعية.

دالة الخسارة التربيعية هي من دوال الخسارة المتماثلة التي تعرف بالصيغة الآتية:

$$\mathbf{L}(\hat{\underline{\alpha}},\underline{\alpha}) = \mathbf{a_0}(\hat{\underline{\alpha}} - \underline{\alpha})^2 \; \mathbf{L}(\hat{\underline{\alpha}},\underline{\alpha}) = \mathbf{a_0}(\hat{\underline{\alpha}} - \underline{\alpha})^2$$

QUOTE … (1-6)

إذ أن QUOTE $(\mathbf{a_0})$ $(\mathbf{a_0})$ هو ثابت موجب نفرض بانه يساوي واحد لأنه لا يؤثر في ايجاد مقدر بيز حيث يُحذف عند الاشتقاق لمقدر بيز بالنسبة لهذه الدالة ، فتصبح صيغة دالة الخسارة التربيعية بالشكل الآتي:[8]

$$\mathbf{L}(\hat{\underline{\alpha}},\underline{\alpha}) = (\hat{\underline{\alpha}} - \underline{\alpha})^2 \; \mathbf{L}(\hat{\underline{\alpha}},\underline{\alpha}) = (\hat{\underline{\alpha}} - \underline{\alpha})^2$$

QUOTE … (1-7)

لذلك فأن مقدر بيز في ظل دالة الخسارة التربيعية الذي يجعل دالة المخاطرة اقل ما يمكن والتي تمثل توقع دالة الخسارة بعد ايجاد المشتقة الاولى بالنسبة للمعلمة المراد تقديرها ومساواتها بالصفر نحصل على :

$$\mathbf{Risk} = \mathbf{E}(\hat{\underline{\alpha}} - \underline{\alpha})^2)$$

$$= \int_{\forall \underline{\alpha}} (\hat{\underline{\alpha}} - \underline{\alpha})^2 \mathbf{h}\,(\underline{\alpha}|\underline{\mathbf{x}})\mathbf{d}\underline{\alpha}$$

$$= \int_{\forall \underline{\alpha}} (\hat{\underline{\alpha}}^2 - 2\underline{\alpha}\,\hat{\underline{\alpha}} + \underline{\alpha}^2)\mathbf{h}\,(\underline{\alpha}|\underline{\mathbf{x}})\mathbf{d}\underline{\alpha}$$

$$= \hat{\underline{\alpha}}^2 - 2\hat{\underline{\alpha}}\mathbf{E}(\underline{\alpha}|\underline{\mathbf{x}}) + \mathbf{E}(\underline{\alpha}^2|\underline{\mathbf{x}})$$

QUOTE

$$= \hat{\underline{\alpha}}^2 - 2\hat{\underline{\alpha}}\mathbf{E}(\underline{\alpha}|\underline{\mathbf{x}}) + \mathbf{E}(\underline{\alpha}^2|\underline{\mathbf{x}})$$

… (1-8)

وباشتقاق المعادلة رقم(2-8) بالنسبة للـ QUOTE $\hat{\underline{\alpha}}$ $\hat{\underline{\alpha}}$ ومساواة المشتقة بالصفر نحصل على :[8]

$$\frac{\partial \mathbf{E}(\hat{\underline{\alpha}} - \underline{\alpha})^2)}{\partial \hat{\underline{\alpha}}} = \mathbf{0}$$

$$= \qquad = \qquad 2\hat{\underline{\alpha}} - 2\mathbf{E}(\underline{\alpha}|\underline{\mathbf{x}}) = \mathbf{0}$$

QUOTE QUOTE

$$2\hat{\underline{\alpha}} - 2\mathbf{E}(\underline{\alpha}|\underline{\mathbf{x}}) = \mathbf{0}$$

$$\therefore \hat{\underline{\alpha}}_{\mathbf{SELBayes}} = \mathbf{E}(\underline{\alpha}|\underline{\mathbf{x}}) \qquad \therefore \hat{\underline{\alpha}}_{\mathbf{SELBayes}} = \mathbf{E}(\underline{\alpha}|\underline{\mathbf{x}})$$

QUOTE … (1-9)

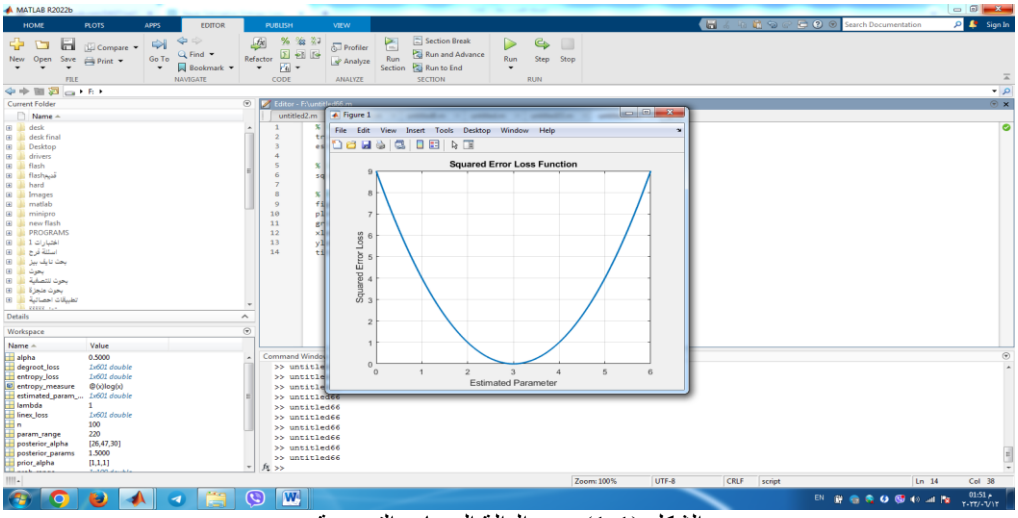
وهذا يعني ان تقدير المعلمات في ظل دالة خسارة تربيعية هو النوقع الشرطي للمعلمات عندما تكونX معلومة إذ أن:

$\underline{\alpha}$ $\underline{\alpha}$ QUOTE : متجه المعلمات المراد تقديرها.

$\hat{\underline{\alpha}}$ $\hat{\underline{\alpha}}$ QUOTE : متجه المعلمات المقدرة.

$\hat{\underline{\alpha}}_{SELBayes}$ $\hat{\underline{\alpha}}_{SELBayes}$ QUOTE : متجه المعلمات المقدرة بموجب طريقة بيز القياسية عند دالة الخسارة التربيعية.

ونلحظ ان مقدر بيز القياسي $\hat{\underline{\alpha}}_{SELBayes}$ $\hat{\underline{\alpha}}_{SELBayes}$ QUOTE لمتجه المعلمات $\underline{\alpha}$ $\underline{\alpha}$ QUOTE ما هو الا متوسط التوزيع اللاحق (Posterior mean) للمعلمة العشوائية ( ) $\underline{\alpha}$ $\underline{\alpha}$ QUOTE
من بين المزايا الرئيسية لاستخدام دالة الخسارة التربيعية هي سهولة استخدامها وفهمها، وقدرتها على توفير تقديرات عالية الدقة للمتغيرات المجهولة .



الشكل (1-1) رسم الدالة الخسارة التربيعية

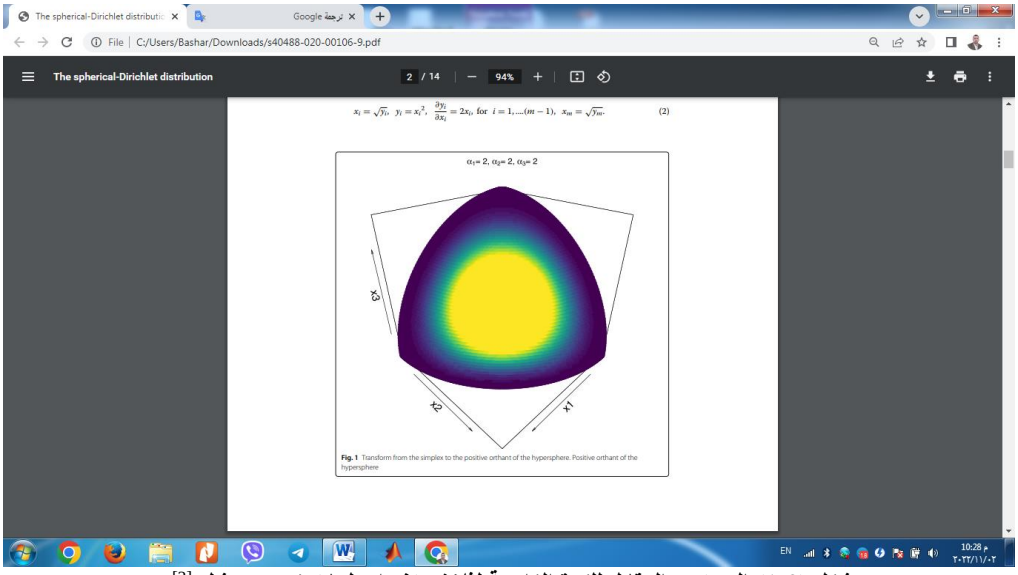
4.3 توزيع ديريشليه الكروي ((The Spherical-Dirichlet Distribut Distribution [2] [3]

هو توزيع احتمالي متعدد الأبعاد يستخدم بشكل شائع في التحليل الإحصائي والتعلم الآلي. ويعد هذا التوزيع إحدى الأدوات المفيدة في تعريف أوزان المخاطر في أنظمة التصنيف الآلي.
يستخدم توزيع ديريشليه الكروي في العديد من التطبيقات، مثل تصنيف الصور والنصوص وتحليل البيانات الجينية والتنبؤ بالطقس. يتم استخدام هذا التوزيع لتحديد توزيع احتمالي لعدة فئات مختلفة بناءً على بعض الخصائص المتعلقة بالبيانات.
يتميز توزيع درشلت الكروي بأنه يعتبر توزيعاً متعدد الأبعاد، حيث يمكن استخدامه لتحديد توزيع احتمالي لعدة فئات. كما أنه يعتبر توزيعاً مستمراً، حيث يمكن أن تأخذ قيم احتمالية في أي نقطة داخل الفضاء متعدد الأبعاد.
نستطيع الحصول على توزيع ديريشليه الكروي باستعمال تحويل توزيع Dirichlet على المستوي المقابل للكرة الزائدية فاذا كان لدينا دالة لكثافة الاحتمالية للتوزيع f_{Dir} f_{Dir} فانه يتم تحويله الى كروي باستعمال المستوي المقابل للكرة الزائدية عن طريق التحويلة الآتية: [3]

$$x_i = \sqrt{y_i} \quad , y_i = x_i^{\,2} \quad , \frac{\partial y_i}{\partial x_i} = 2x_i \, , i = 1, \ldots . (k-1), x_k = \sqrt{y_k}$$

$$x_i = \sqrt{y_i} \quad , y_i = x_i^{\,2} \quad , \frac{\partial y_i}{\partial x_i} = 2x_i \, , i = 1, \ldots . (k-1), x_k = \sqrt{y_k}$$

… (10-1)



شكل (2-1) المستوي المقابل للكرة الزاندية لثلاث متغيرات لها توزيع ديريشلي[2]

ونقوم بحساب محدد الجاكوبيان $|J|\,|J|$ لجميع المتغيرات المستقلة:

$$J = \begin{vmatrix} \frac{\partial y_1}{\partial x_1} = 2x1 & 0 & \cdots & 0 & \ldots. \\ 0 & \frac{\partial y_2}{\partial x_2} = 2x_2 & \cdots & 0 & \ldots. \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & \cdots & \vdots & \vdots \\ 0 & 0 & \cdots & \cdots & \frac{\partial y_{k-1}}{\partial x_{k-1}} = 2x_{k-1} \end{vmatrix}$$

$$= \prod_{i=1}^{k-1} 2x_i = 2^{k-1} \prod_{i=1}^{k-1} x_i \ = \prod_{i=1}^{k-1} 2x_i = 2^{k-1} \prod_{i=1}^{k-1} x_i$$

… (11-1)

وبذلك تكون دالة الكثافة الاحتمالية لتوزيع ديريشليه الكروي كالآتي:

$$f_{SDir} = \frac{2^{k-1}\Gamma(\alpha_0)}{\prod_{i=1}^{k}\Gamma\alpha_i} \prod_{i=1}^{k-1} x_i^{2\alpha i-1} x_k^{2\alpha_k-1} f_{SDir} = \frac{2^{k-1}\Gamma(\alpha_0)}{\prod_{i=1}^{k}\Gamma\alpha_i} \prod_{i=1}^{k-1} x_i^{2\alpha i-1} x_k^{2\alpha_k-1}$$

$$= \frac{2^{k-1}\Gamma(\alpha_0)}{\prod_{i=1}^{k}\Gamma\alpha_i} \prod_{i=1}^{k} x_i^{2\alpha i-1} x_k^{-1} \qquad = \frac{2^{k-1}\Gamma(\alpha_0)}{\prod_{i=1}^{k}\Gamma\alpha_i} \prod_{i=1}^{k} x_i^{2\alpha i-1} x_k^{-1}$$

$$= \frac{2^{k-1}\Gamma(\alpha_0)}{\prod_{i=1}^{k}\Gamma\alpha i} \prod_{i=1}^{k-1} x_i^{2\alpha i-1} (1 - \sum_{i=1}^{k-1} x_i^{\,2})^{\alpha_k-1}$$

$$= \frac{2^{k-1}\Gamma(\alpha_0)}{\prod_{i=1}^{k}\Gamma\alpha i} \prod_{i=1}^{k-1} x_i^{2\alpha i-1} (1 - \sum_{i=1}^{k-1} x_i^{\,2})^{\alpha_k-1}$$

… (12-1)

$$\alpha_i \in R^+, \alpha_0 = \sum_{i=1}^{k} \alpha_i \ , \ 0 \le x_i \le 1, \ \sum_{i=1}^{k} x_i^{\,2} = 1$$

$$\alpha_i \in R^+, \alpha_0 = \sum_{i=1}^{k} \alpha_i \ , \ 0 \le x_i \le 1, \ \sum_{i=1}^{k} x_i^{\,2} = 1$$

$x_i \sim SDD(\alpha i) x_i \sim SDD(\alpha i)$ اذن الصيغة (12-1) هي صيغة التوزيع الكروي ديريتشليه اذ ان

5.3 مقدر بيز القياسي المعلوماتي لتوزيع ديريشليه الكروي عند دالة خسارة تربيعية

(Informative Standard Bayesian Estimator for The Spherical-Dirichlet Distribution under Squares Error Loss function) (INESBSDD)

ليكن $X_i X_i$ عينة عشوائية من المتجهات $X_1, X_2, \dots\dots, X_n X_1, X_2, \dots\dots, X_n$ المستقلة ، اذ ان $X_i \in R^k$

$X_i \in R^k$ التي لها توزيع ديريشليه الكرويSDD)) بدالة الكثافة الاحتمالية المعرفة في المعادلة (12-1) , فان دالة الامكان لتوزيع ديريشليه الكروي (Likelihood function of The Spherical-Dirichlet Distribution) يمكن ان تكتب كالآتي:

$$L(\alpha) = \prod_{i=1}^{n} \left(\frac{2^{k-1}\Gamma\left(\sum_{j=1}^{k} \alpha_j\right)}{\prod_{j=1}^{k} \Gamma\alpha_j} \prod_{j=1}^{k} {x_{ij}}^{2\alpha i - 1} {x_{ik}}^{-1} \right)$$

$$L(\alpha) = \prod_{i=1}^{n} \left(\frac{2^{k-1}\Gamma\left(\sum_{j=1}^{k} \alpha_j\right)}{\prod_{j=1}^{k} \Gamma\alpha_j} \prod_{j=1}^{k} {x_{ij}}^{2\alpha i - 1} {x_{ik}}^{-1} \right)$$

… (13-1)

وباخذ اللوغاريتم الطبيعي للمعادلة (13-1) نحصل على لوغاريتم دالة الامكان لتوزيع دريشليه الكروي وكالآتي:

$$l = lnL(\alpha) = \ln\left(\prod_{i=1}^{n} \frac{2^{k-1}\Gamma\left(\sum_{j=1}^{k} \alpha_j\right)}{\prod_{j=1}^{k} \Gamma\alpha_j} \prod_{j=1}^{k} {x_{ij}}^{2\alpha j - 1} {x_{ik}}^{-1} \right)$$

$$l = lnL(\alpha) = \ln\left(\prod_{i=1}^{n} \frac{2^{k-1}\Gamma\left(\sum_{j=1}^{k} \alpha_j\right)}{\prod_{j=1}^{k} \Gamma\alpha_j} \prod_{j=1}^{k} {x_{ij}}^{2\alpha j - 1} {x_{ik}}^{-1} \right)$$

$$= \sum_{i=1}^{n} \ln\left(\frac{2^{k-1}\Gamma\left(\sum_{j=1}^{k} \alpha_j\right)}{\prod_{j=1}^{k} \Gamma\alpha_j} \right) + \sum_{i=1}^{n} \sum_{j=1}^{k} \ln\left({x_{ij}}^{2\alpha j - 1} {x_{ik}}^{-1} \right)$$

$$= \sum_{i=1}^{n} \ln\left(\frac{2^{k-1}\Gamma\left(\sum_{j=1}^{k} \alpha_j\right)}{\prod_{j=1}^{k} \Gamma\alpha_j} \right) + \sum_{i=1}^{n} \sum_{j=1}^{k} \ln\left({x_{ij}}^{2\alpha j - 1} {x_{ik}}^{-1} \right)$$

$$= \sum_{i=1}^{n} \left(\ln\left(2^{k-1}\right) + \ln\left(\Gamma\left(\sum_{j=1}^{k} \alpha_j\right)\right) \right) - \sum_{i=1}^{n} \sum_{j=1}^{k} \ln\left(\Gamma\alpha_j\right) + \sum_{i=1}^{n} \sum_{j=1}^{k} \left(\ln\left({x_{ij}}^{2\alpha j - 1} \right) - \ln\left({x_{ik}}^{-1} \right) \right)$$

$$= \sum_{i=1}^{n} \left(\ln\left(2^{k-1}\right) + \ln\left(\Gamma\left(\sum_{j=1}^{k} \alpha_j\right)\right) \right) - \sum_{i=1}^{n} \sum_{j=1}^{k} \ln\left(\Gamma\alpha_j\right) + \sum_{i=1}^{n} \sum_{j=1}^{k} \left(\ln\left({x_{ij}}^{2\alpha j - 1} \right) - \ln\left({x_{ik}}^{-1} \right) \right)$$

$$= n(k-1)\ln(2) + nln\left(\Gamma\left(\sum_{j=1}^{k} \alpha_j\right)\right) - n\sum_{j=1}^{k} \ln\left(\Gamma\alpha_j\right) + \sum_{j=1}^{k} (2\alpha j - 1) \frac{1}{n} \sum_{i=1}^{n} \ln \frac{1}{n} \sum_{i=1}^{n} \ln\left(x_{ik}\right)$$

$$= n(k-1)\ln(2) + nln\left(\Gamma\left(\sum_{j=1}^{k} \alpha_j\right)\right) - n\sum_{j=1}^{k} \ln\left(\Gamma\alpha_j\right) + \sum_{j=1}^{k} (2\alpha j - 1) \frac{1}{n} \sum_{i=1}^{n} \ln \frac{1}{n} \sum_{i=1}^{n} \ln\left(x_{ik}\right)$$

… (14-1)

ولنفرض ان هنالك معلومات أولية عن المعلمات المراد تقديرها ($\alpha_j \alpha_j$) والمتمثلة بدالة الكثافة الاحتمالية المشتركة لعدة دول كثافة احتمالية كروية محولة مقترحة بالمعلمات الزائدية ($a_j a_j$ Hyper parameters) و $b_j b_j$ وكما يأتي:

ليكن لدينا دالة الكثافة الاحتمالية لتوزيع كاما التقلدي كما في المعادلة الآتية:

$$\pi(\alpha_j) = \prod_{j=1}^{k} \frac{b_j^{a_j}}{\Gamma a_j} \alpha_j^{a_j-1} e^{-b_j\alpha_j} \pi(\alpha_j) = \prod_{j=1}^{k} \frac{b_j^{a_j}}{\Gamma a_j} \alpha_j^{a_j-1} e^{-b_j\alpha_j}$$

...

(15-1)

سيتم تحويل توزيع كاما التقليدي الى كروي باستعمال المستوي المقابل للكرة الزائدية عن طريق التحويلة الآتية:

$$\alpha_j = \sqrt{x_j} \quad , x_j = \alpha_j^2 \ , \frac{\partial \alpha_j}{\partial x_j} = 2\alpha_j \ , j = 1, \ldots . k, \alpha_k = \sqrt{x_k}$$

$$\alpha_j = \sqrt{x_j} \quad , x_j = \alpha_j^2 \ , \frac{\partial \alpha_j}{\partial x_j} = 2\alpha_j \ , j = 1, \ldots . k, \alpha_k = \sqrt{x_k}$$

... (16-1)

وان محدد الجاكوبيان $|j|$ $|j|$ لجميع المتغيرات المستقلة كالآتي:

$$J = \begin{vmatrix} \frac{\partial x_1}{\partial \alpha_1} = 2\alpha 1 & 0 & \cdots & 0 & \ldots . \\ 0 & \frac{\partial x_2}{\partial \alpha_2} = 2\alpha_2 & \cdots & 0 & \ldots . \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & \cdots & \vdots & \vdots \\ 0 & 0 & \cdots & \cdots & \frac{\partial x_k}{\partial \alpha_k} = 2\alpha_k \end{vmatrix}$$

$$= \prod_{j=1}^{k} 2\alpha_j = 2^k \prod_{j=1}^{k} \alpha_j \ = \prod_{j=1}^{k} 2\alpha_j = 2^k \prod_{j=1}^{k} \alpha_j$$

...

(1-17)

$$\pi^*(\alpha_j) = \prod_{j=1}^{k} \frac{b_j^{a_j}}{\Gamma a_j} (\alpha_j^2)^{a_j-1} e^{-b_j\alpha_j^2} 2^k \prod_{j=1}^{k} \alpha_j$$

$$\pi^*(\alpha_j) = \prod_{j=1}^{k} \frac{b_j^{a_j}}{\Gamma a_j} (\alpha_j^2)^{a_j-1} e^{-b_j\alpha_j^2} 2^k \prod_{j=1}^{k} \alpha_j$$

$$= \left(\frac{2b_j^{a_j}}{\Gamma a_j}\right)^k \prod_{j=1}^{k} \alpha_j^{2(a_j-1)} e^{-b_j\alpha_j^2} \alpha_j$$

$$= \left(\frac{2b_j^{a_j}}{\Gamma a_j}\right)^k \prod_{j=1}^{k} \alpha_j^{2(a_j-1)} e^{-b_j\alpha_j^2} \alpha_j$$

$$= \left(\frac{2b_j^{a_j}}{\Gamma a_j}\right)^k \prod_{j=1}^{k} \alpha_j^{2(a_j-1)+1} e^{-b_j\alpha_j^2}$$

$$= \left(\frac{2b_j^{a_j}}{\Gamma a_j}\right)^k \prod_{j=1}^{k} \alpha_j^{2(a_j-1)+1} e^{-b_j\alpha_j^2} \quad ; \quad \alpha_j > 0 \alpha_j > 0$$

... (1-18)

الآن نجد التوزيع اللاحق (Posterior Distribution) وحسب نظرية بيز نحصل على التوزيع اللاحق باستعمال صيغة بيز العكسية $(Bays\ Inversion\ Formula)(Bays\ Inversion\ Formula)$ وكالآتي:

$$h(\underline{\alpha}\ /X_1, X_2, \ldots . , X_n) = \frac{l(X_1, X_2, \ldots . , X_n/\underline{\alpha})\pi(\underline{\alpha})}{\int_{\forall \underline{\alpha}} l(X_1, X_2, \ldots . , X_n/\underline{\alpha})\pi(\underline{\alpha})\, d\underline{\alpha}}$$

$$h(\underline{\alpha}\ /X_1, X_2, \ldots . , X_n) = \frac{l(X_1, X_2, \ldots . , X_n/\underline{\alpha})\pi(\underline{\alpha})}{\int_{\forall \underline{\alpha}} l(X_1, X_2, \ldots . , X_n/\underline{\alpha})\pi(\underline{\alpha})\, d\underline{\alpha}}$$

$$= \frac{G(X_i,\underline{\alpha})}{M(X_i)} \qquad = \frac{G(X_i,\underline{\alpha})}{M(X_i)}$$

… (1-19)

اذ ان :

$G(X_i,\ \underline{\alpha})G(X_i,\ \underline{\alpha})$ هي الدالة المشتركة التي تعبر عن مجال المعلمات $(\underline{\alpha})\,(\underline{\alpha})$ بوجود معلومات العينة الحالية والمعلومات السابقة والحالية عن $(\underline{\alpha})(\underline{\alpha})$ المراد تقديرها .

$M(X_i)M(X_i)$ الدالة الحدية للمتغير X_iX_i

وعليه فان الدالة المشتركة $G(X_i,\ \underline{\alpha})G(X_i,\ \underline{\alpha})$ تكون كالآتي:

$$G(X_i,\underline{\alpha}) = l(X_1,X_2,\dots\dots,X_n/\underline{\alpha})\pi(\underline{\alpha})\ G(X_i,\underline{\alpha}) = l(X_1,X_2,\dots\dots,X_n/\underline{\alpha})\pi(\underline{\alpha})$$

$$= n(k-1)\ln(2) + nln\left(\Gamma\left(\sum_{j=1}^{k}\alpha_j\right)\right) - n\sum_{j=1}^{k}\ln(\Gamma\alpha_j) + \sum_{j=1}^{k}(2\alpha j -$$

$$1)\frac{1}{n}\sum_{i=1}^{n}\ln(x_{ij}) - \frac{1}{n}\sum_{i=1}^{n}\ln(x_{ik})\left(\frac{2b_j{}^{a_j}}{\Gamma a_j}\right)^k \prod_{j=1}^{k}\alpha_j{}^{2(a_j-1)+1}\, e^{-b_j\alpha_j{}^2}$$

$$= n(k-1)\ln(2) + nln\left(\Gamma\left(\sum_{j=1}^{k}\alpha_j\right)\right) - n\sum_{j=1}^{k}\ln(\Gamma\alpha_j) + \sum_{j=1}^{k}(2\alpha j -$$

$$1)\frac{1}{n}\sum_{i=1}^{n}\ln(x_{ij}) - \frac{1}{n}\sum_{i=1}^{n}\ln(x_{ik})\left(\frac{2b_j{}^{a_j}}{\Gamma a_j}\right)^k \prod_{j=1}^{k}\alpha_j{}^{2(a_j-1)+1}\, e^{-b_j\alpha_j{}^2}$$

$$= \left(\frac{2b_j{}^{a_j}}{\Gamma a_j}\right)^k \prod_{j=1}^{k}\alpha_j{}^{2(a_j-1)+1}\, e^{-b_j\alpha_j{}^2}\left(n(k-1)\ln(2) + nln\left(\Gamma\left(\sum_{j=1}^{k}\alpha_j\right)\right) -\right.$$

$$\left. n\sum_{j=1}^{k}\ln(\Gamma\alpha_j) + \sum_{j=1}^{k}(2\alpha j-1)\frac{1}{n}\sum_{i=1}^{n}\ln(x_{ij}) - \frac{1}{n}\sum_{i=1}^{n}\ln(x_{ik})\right)$$

$$= \left(\frac{2b_j{}^{a_j}}{\Gamma a_j}\right)^k \prod_{j=1}^{k}\alpha_j{}^{2(a_j-1)+1}\, e^{-b_j\alpha_j{}^2}\left(n(k-1)\ln(2) + nln\left(\Gamma\left(\sum_{j=1}^{k}\alpha_j\right)\right) -\right.$$

$$\left. n\sum_{j=1}^{k}\ln(\Gamma\alpha_j) + \sum_{j=1}^{k}(2\alpha j-1)\frac{1}{n}\sum_{i=1}^{n}\ln(x_{ij}) - \frac{1}{n}\sum_{i=1}^{n}\ln(x_{ik})\right)$$

... (1-20)

الدالة الحدية للمتغير $X_i\,X_i \quad M(X_i)M(X_i)$ تكون كالآتي:

$$M(X_i) = \int_0^\infty\int_0^\infty\dots\int_0^\infty\left(\frac{2b_j{}^{a_j}}{\Gamma a_j}\right)^k \prod_{j=1}^{k}\alpha_j{}^{2(a_j-1)+1}\, e^{-b_j\alpha_j{}^2}\left(n(k-1)\ln(2) +\right.$$

$$nln\left(\Gamma\left(\sum_{j=1}^{k}\alpha_j\right)\right) - n\sum_{j=1}^{k}\ln(\Gamma\alpha_j) + \sum_{j=1}^{k}(2\alpha j-1)\frac{1}{n}\sum_{i=1}^{n}\ln(x_{ij}) -$$

$$\left.\frac{1}{n}\sum_{i=1}^{n}\ln(x_{ik})\right)d\alpha_k\ \dots\ d\alpha_1$$

$$M(X_i) = \int_0^\infty\int_0^\infty\dots\int_0^\infty\left(\frac{2b_j{}^{a_j}}{\Gamma a_j}\right)^k \prod_{j=1}^{k}\alpha_j{}^{2(a_j-1)+1}\, e^{-b_j\alpha_j{}^2}\left(n(k-1)\ln(2) +\right.$$

$$nln\left(\Gamma\left(\sum_{j=1}^{k}\alpha_j\right)\right) - n\sum_{j=1}^{k}\ln(\Gamma\alpha_j) + \sum_{j=1}^{k}(2\alpha j-1)\frac{1}{n}\sum_{i=1}^{n}\ln(x_{ij}) -$$

$$\left.\frac{1}{n}\sum_{i=1}^{n}\ln(x_{ik})\right)d\alpha_k\ \dots\ d\alpha_1$$

$$M(X_i) = \left(\frac{2b_j{}^{a_j}}{\Gamma a_j}\right)^k \int_0^\infty \int_0^\infty \dots \int_0^\infty \prod_{j=1}^k \alpha_j{}^{2(a_j-1)+1} e^{-b_j\alpha_j{}^2} \left(n(k-1)\ln(2) + nln\left(\Gamma\left(\sum_{j=1}^k \alpha_j\right)\right) - n\sum_{j=1}^k \ln(\Gamma\alpha_j) + \sum_{j=1}^k (2\alpha j - 1)\frac{1}{n}\sum_{i=1}^n \ln(x_{ij}) - \frac{1}{n}\sum_{i=1}^n \ln(x_{ik})\right) d\alpha_k \dots d\alpha_1$$

$$M(X_i) = \left(\frac{2b_j{}^{a_j}}{\Gamma a_j}\right)^k \int_0^\infty \int_0^\infty \dots \int_0^\infty \prod_{j=1}^k \alpha_j{}^{2(a_j-1)+1} e^{-b_j\alpha_j{}^2} \left(n(k-1)\ln(2) + nln\left(\Gamma\left(\sum_{j=1}^k \alpha_j\right)\right) - n\sum_{j=1}^k \ln(\Gamma\alpha_j) + \sum_{j=1}^k (2\alpha j - 1)\frac{1}{n}\sum_{i=1}^n \ln(x_{ij}) - \frac{1}{n}\sum_{i=1}^n \ln(x_{ik})\right) d\alpha_k \dots d\alpha_1$$

... (1-21)

وعليه فان التوزيع اللاحق الشرطي يكون كما يأتي:

$$h(\underline{\alpha}/X_1, X_2, \dots, X_n) = \frac{\left(\frac{2b_j{}^{a_j}}{\Gamma a_j}\right)^k \prod_{j=1}^k \alpha_j{}^{2(a_j-1)+1} e^{-b_j\alpha_j{}^2}\left(n(k-1)\ln(2) + nln\left(\Gamma\left(\sum_{j=1}^k \alpha_j\right)\right) - n\sum_{j=1}^k \ln(\Gamma\alpha_j) + \sum_{j=1}^k (2\alpha j-1)\frac{1}{n}\sum_{i=1}^n l\right.}{\left(\frac{2b_j{}^{a_j}}{\Gamma a_j}\right)^k \int_0^\infty \int_0^\infty \dots \int_0^\infty \prod_{j=1}^k \alpha_j{}^{2(a_j-1)+1} e^{-b_j\alpha_j{}^2}\left(n(k-1)\ln(2) + nln\left(\Gamma\left(\sum_{j=1}^k \alpha_j\right)\right) - n\sum_{j=1}^k \ln(\Gamma\alpha_j) + \sum_{j=1}^k (2\alpha j-1)\frac{1}{n}\sum_{i=1}^n l\right.}$$

$$h(\underline{\alpha}/X_1, X_2, \dots, X_n) = \frac{\left(\frac{2b_j{}^{a_j}}{\Gamma a_j}\right)^k \prod_{j=1}^k \alpha_j{}^{2(a_j-1)+1} e^{-b_j\alpha_j{}^2}\left(n(k-1)\ln(2) + nln\left(\Gamma\left(\sum_{j=1}^k \alpha_j\right)\right) - n\sum_{j=1}^k \ln(\Gamma\alpha_j) + \sum_{j=1}^k (2\alpha j-1)\frac{1}{n}\sum_{i=1}^n l\right.}{\left(\frac{2b_j{}^{a_j}}{\Gamma a_j}\right)^k \int_0^\infty \int_0^\infty \dots \int_0^\infty \prod_{j=1}^k \alpha_j{}^{2(a_j-1)+1} e^{-b_j\alpha_j{}^2}\left(n(k-1)\ln(2) + nln\left(\Gamma\left(\sum_{j=1}^k \alpha_j\right)\right) - n\sum_{j=1}^k \ln(\Gamma\alpha_j) + \sum_{j=1}^k (2\alpha j-1)\frac{1}{n}\sum_{i=1}^n l\right.}$$

$$= \frac{\prod_{j=1}^k \alpha_j{}^{2(a_j-1)+1} e^{-b_j\alpha_j{}^2}\left(n(k-1)\ln(2) + nln\left(\Gamma\left(\sum_{j=1}^k \alpha_j\right)\right) - n\sum_{j=1}^k \ln(\Gamma\alpha_j) + \sum_{j=1}^k (2\alpha j-1)\frac{1}{n}\sum_{i=1}^n \ln\right.}{\int_0^\infty \int_0^\infty \dots \int_0^\infty \prod_{j=1}^k \alpha_j{}^{2(a_j-1)+1} e^{-b_j\alpha_j{}^2}\left(n(k-1)\ln(2) + nln\left(\Gamma\left(\sum_{j=1}^k \alpha_j\right)\right) - n\sum_{j=1}^k \ln(\Gamma\alpha_j) + \sum_{j=1}^k (2\alpha j-1)\frac{1}{n}\sum_{i=1}^n \ln\right.}$$

$$= \frac{\prod_{j=1}^k \alpha_j{}^{2(a_j-1)+1} e^{-b_j\alpha_j{}^2}\left(n(k-1)\ln(2) + nln\left(\Gamma\left(\sum_{j=1}^k \alpha_j\right)\right) - n\sum_{j=1}^k \ln(\Gamma\alpha_j) + \sum_{j=1}^k (2\alpha j-1)\frac{1}{n}\sum_{i=1}^n \ln\right.}{\int_0^\infty \int_0^\infty \dots \int_0^\infty \prod_{j=1}^k \alpha_j{}^{2(a_j-1)+1} e^{-b_j\alpha_j{}^2}\left(n(k-1)\ln(2) + nln\left(\Gamma\left(\sum_{j=1}^k \alpha_j\right)\right) - n\sum_{j=1}^k \ln(\Gamma\alpha_j) + \sum_{j=1}^k (2\alpha j-1)\frac{1}{n}\sum_{i=1}^n \ln\right.}$$

... (1-22)

وبما ان مقدر بيز القياسي المعلوماتي عند دالة خسارة تربيعية هو متوسط التوزيع اللاحق وكالآتي:

$$\underline{\hat{\alpha}}_{\text{INSSBSDD}} = E\left(h(\underline{\alpha}/X_1, X_2, \dots, X_n)/\underline{\hat{\alpha}}\right) = \int_0^\infty \int_0^\infty \dots \int_0^\infty \underline{\alpha} h(\underline{\alpha}/X_1, X_2, \dots, X_n) d\alpha_k \dots d\alpha_1$$

$$\underline{\hat{\alpha}}_{\text{INSSBSDD}} = E\left(h(\underline{\alpha}/X_1, X_2, \dots, X_n)/\underline{\hat{\alpha}}\right) = \int_0^\infty \int_0^\infty \dots \int_0^\infty \underline{\alpha} h(\underline{\alpha}/X_1, X_2, \dots, X_n) d\alpha_k \dots d\alpha_1$$

$$= \int_0^\infty \int_0^\infty \dots \int_0^\infty \frac{\underline{\alpha}\prod_{j=1}^k \alpha_j{}^{2(a_j-1)+1} e^{-b_j\alpha_j{}^2}\left(n(k-1)\ln(2) + nln\left(\Gamma\left(\sum_{j=1}^k \alpha_j\right)\right) - n\sum_{j=1}^k \ln(\Gamma\alpha_j) + \sum_{j=1}^k (2\alpha j-1)\frac{1}{n}\sum_{i=1}^n \ln(x_{ij}\right.}{\int_0^\infty \int_0^\infty \dots \int_0^\infty \prod_{j=1}^k \alpha_j{}^{2(a_j-1)+1} e^{-b_j\alpha_j{}^2}\left(n(k-1)\ln(2) + nln\left(\Gamma\left(\sum_{j=1}^k \alpha_j\right)\right) - n\sum_{j=1}^k \ln(\Gamma\alpha_j) + \sum_{j=1}^k (2\alpha j-1)\frac{1}{n}\sum_{i=1}^n \ln(x_{ij})\right.}$$

$$= \int_0^\infty \int_0^\infty \dots \int_0^\infty \frac{\underline{\alpha}\prod_{j=1}^k \alpha_j{}^{2(a_j-1)+1} e^{-b_j\alpha_j{}^2}\left(n(k-1)\ln(2) + nln\left(\Gamma\left(\sum_{j=1}^k \alpha_j\right)\right) - n\sum_{j=1}^k \ln(\Gamma\alpha_j) + \sum_{j=1}^k (2\alpha j-1)\frac{1}{n}\sum_{i=1}^n \ln(x_{ij}\right.}{\int_0^\infty \int_0^\infty \dots \int_0^\infty \prod_{j=1}^k \alpha_j{}^{2(a_j-1)+1} e^{-b_j\alpha_j{}^2}\left(n(k-1)\ln(2) + nln\left(\Gamma\left(\sum_{j=1}^k \alpha_j\right)\right) - n\sum_{j=1}^k \ln(\Gamma\alpha_j) + \sum_{j=1}^k (2\alpha j-1)\frac{1}{n}\sum_{i=1}^n \ln(x_{ij})\right.}$$

... (1-22)

وبحل المعادلة (22-1) نحصل المعلمات المراد تقديرها $\alpha_j \alpha_j$ وهي دالة لاخطية لايمكن ان تحل بالطرائق التحليلية الاعتيادية ولذلك سيتم استعمال طريقة ((Metropolis-Hasting عن طريق محاكاة سلسلة ماكروف مونت-كارلوا ((MCMC

4- العملي (الجانب التجريبي) Practical (experimental part)

1.4 وصف تجارب المحاكاة: Description of simulation experiments [6] [5]

اعتمد أسلوب المحاكاة مونت-كارلو (Monte-Carlo Simulation) لغرض اختبار كفاءة طرائق التقدير المستعملة لتقدير معلمات توزيع معلمات التوزيع ومقارنة المقدرات المستخرجة باستعمال معيار متوسط مربعات الخطأ التكاملي (IMSE) وفقاً المراحل الآتية :[6]

2.4 إختيار القيم الافتراضية لمعلمات توزيع ديرشليه الكروي

Choosing default values for the parameters of the spherical Dirichlet distribution

تم الحصول على القيم الإفتراضية لتوزيع درشلية الكروي ثلاثي الابعاد تجريبياً من إجراء تجارب عدة واختيار القيم التي إستقرت عندها تقديرات بيز وأعطت افضل النتائج وكما مبينة في الجدول (1-2) .

جدول (1-2) القيم الإفتراضية لمعلمات توزيع ديريشليه الكروي

Model Parameter	1	2
α_1	1	1
α_2	1	2
α_3	1	1

3.4 توليد البيانات (Data Generating)

تم توليد عينة التوزيع الكروي ديرشلية ثلاثي الابعاد كما في الخوارزمية الآتية:

1. توليد عينات من توزيع درشليه التقليدي ثلاثي الابعاد بالمعلمات ($\alpha_1, \alpha_2, \alpha_3$ $\alpha_1, \alpha_2, \alpha_3$) باحجام عينات (n=100, 200, 300,400,500) باستعمال برنامج ماتلاب وعن طريق الدالة الآتية :

dirichlet_sample_3d(alpha, n)

اذ ان alpha متجه المعلمات الافتراضية كما في النماذج المدرجة في جدول (1-3) اي ان :

Alpha=[$\alpha_0, \alpha_0, \alpha_0 \alpha_0, \alpha_0, \alpha_0$]

ولكل انموذج .

2. تحويل عينات توزيع ديريشليه ثلاثي الابعاد من الاحداثي الكارتيزي الى الاحداثي القطبي (الكروي) باستعمال الدالة cart2sph في برنامج ماتلاب.
3. عمل normalize للعينات للحصول على عينات لها توزيع ديريشليه الكروي ذو الثلاث ابعاد وبثلاث معلمات بالمتجه $\underline{x} = [x,y,z] \underline{x} = [x,y,z]$

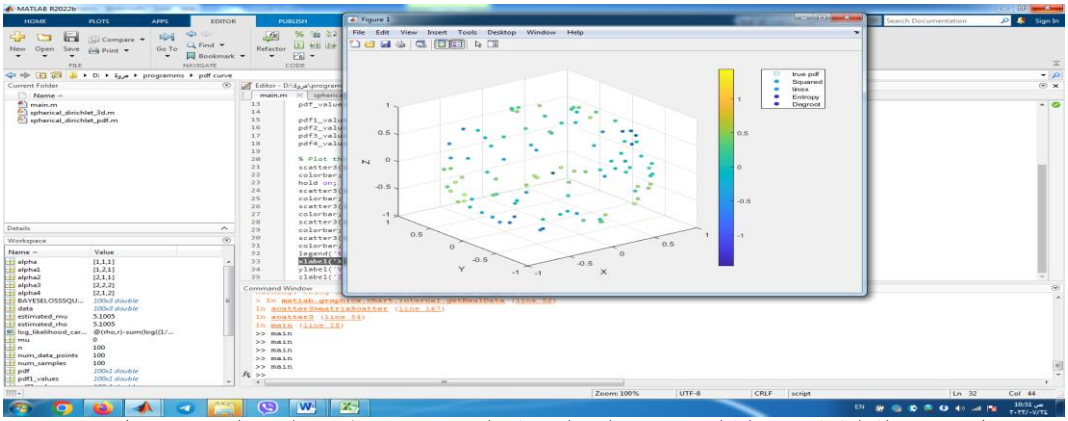
4.4 مناقشة نتائج تجارب المحاكاة : Discussion of the results of simulation experiments

سيتم عرض وتحليل نتائج تجارب المحاكاة لتقدير معلمات توزيع ديريشليه الكروي حسب الطرائق المبينة في الجانب النظري باستعمال محاكاة ماركوف مونت كارلوا (MCMC).

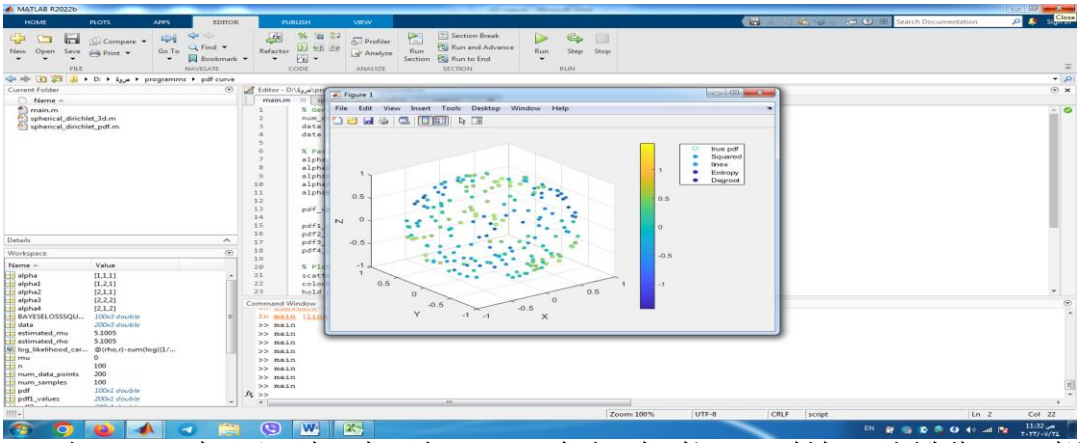
جدول (1-2) تقديرات المعلمات ومتوسط مربعات الخطا التكاملي (IMSE)(IMSE) لكل معلمة لكل طرائق التقدير للانموذج الأول

Method		Squared	
n	Parameter	Estimation	IMSE
100	α_1	1.67671	0.0006767
	α_2	1.56433	0.0005643
	α_3	1.52344	0.0005234

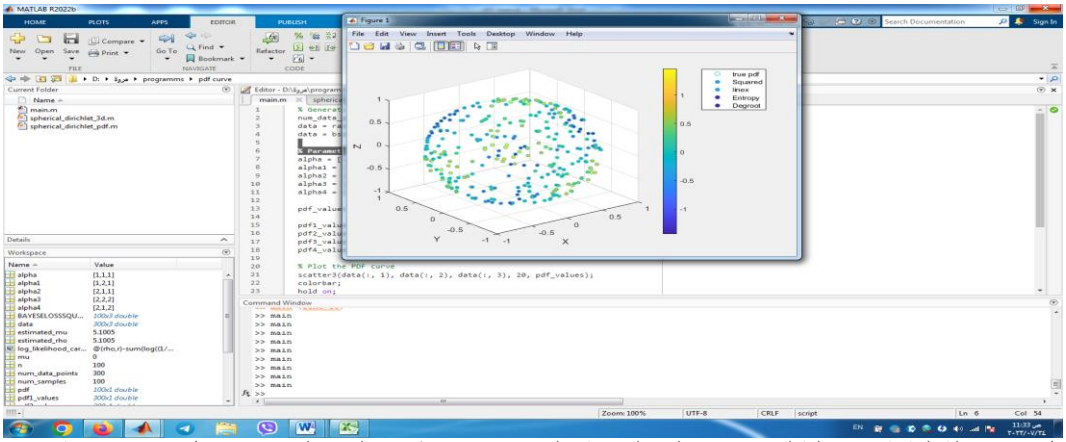
200		α_1	1.56558	0.0005656
		α_2	1.53577	0.0005358
		α_3	1.50575	0.0005058
300		α_1	1.53467	0.0005347
		α_2	1.51689	0.0005169
		α_3	1.46896	0.000469
400		α_1	1.45679	0.0004568
		α_2	1.46454	0.0004645
		α_3	1.44757	0.0004476
500		α_1	1.43578	0.0004358
		α_2	1.44267	0.0004427
		α_3	1.42356	0.0004236



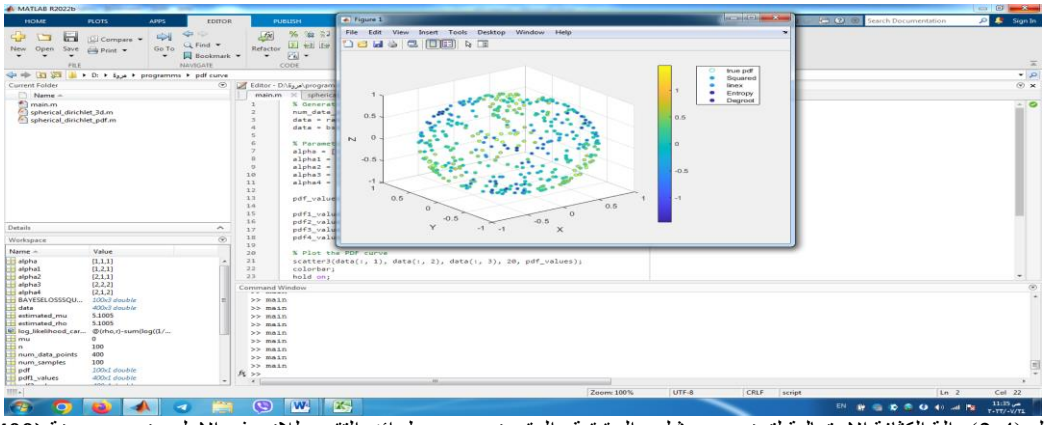
شكل (2-1) دالة الكثافة الاحتمالية لتوزيع ديريشليه الحقيقية والمقدرة بموجب طرائق التقدير للانموذج الاول عند حجم عينة **(100)**



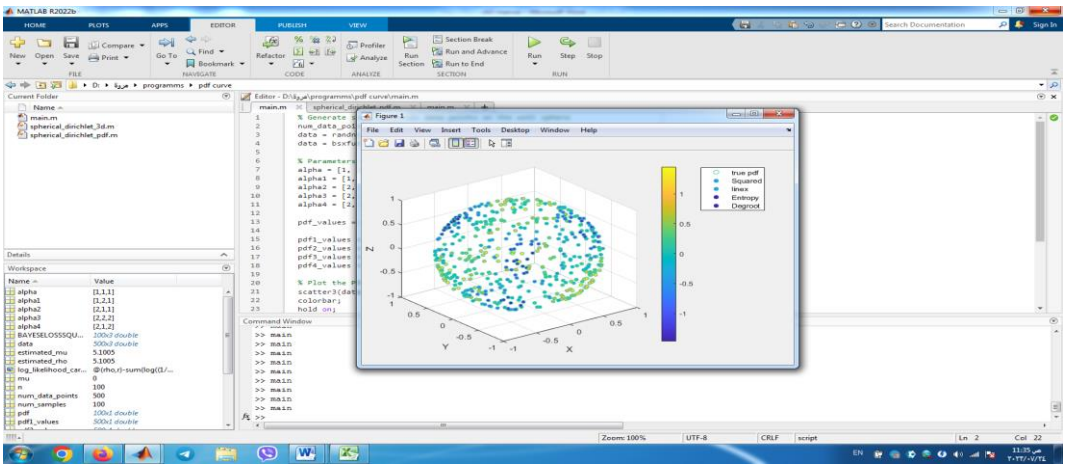
شكل (2-2) دالة الكثافة الاحتمالية لتوزيع ديريشليه الحقيقية والمقدرة بموجب طرائق التقدير للانموذج الاول عند حجم عينة **(200)**



شكل (2-3) دالة الكثافة الاحتمالية لتوزيع ديريشليه الحقيقية والمقدرة بموجب طرائق التقدير للانموذج الاول عند حجم عينة **(300)**



شكل (2-4) دالة الكثافة الاحتمالية لتوزيع ديريشليه الحقيقية والمقدرة بموجب طرائق التقدير للانموذج الاول عند حجم عينة **(400)**



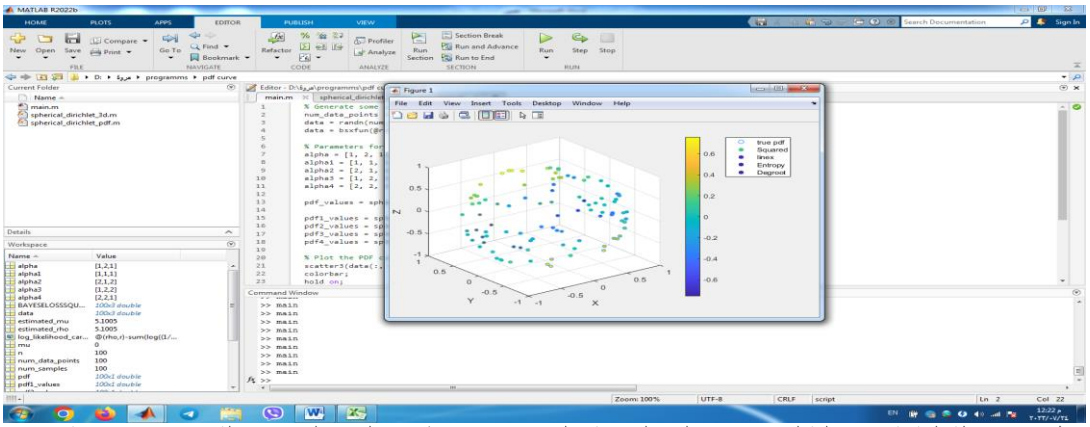
شكل (2-5) دالة الكثافة الاحتمالية لتوزيع ديريشليه الحقيقية والمقدرة بموجب طرائق التقدير للانموذج الاول عند حجم عينة **(500)**

نلاحظ من جدول (2-1) والاشكال (2-1)الى (2-5) مايأتي :

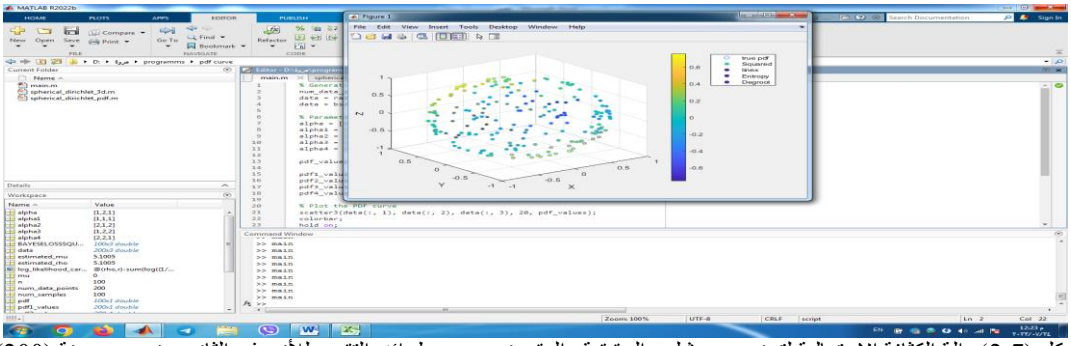
عند المعلمات الافتراضية $(\alpha_1 = 1, \alpha_2 = 1, \alpha_3 = 1)(\alpha_1 = 1, \alpha_2 = 1, \alpha_3 = 1)$ ان طريقة بيز في ظل دالة خسارة (التربيعية) عند حجم عينة (500) حققت افضلية على باقي حجوم العينات في تقدير معلمات التوزيع الكروي يليها حجم العينة (400) لانها سجلت اقل متوسط مربعات خطأ.

جدول (2-2) تقديرات المعلمات ومتوسط مربعات الخطأ التكاملي (IMSE)(IMSE) لكل معلمة لكل طرائق التقدير للأنموذج الثاني

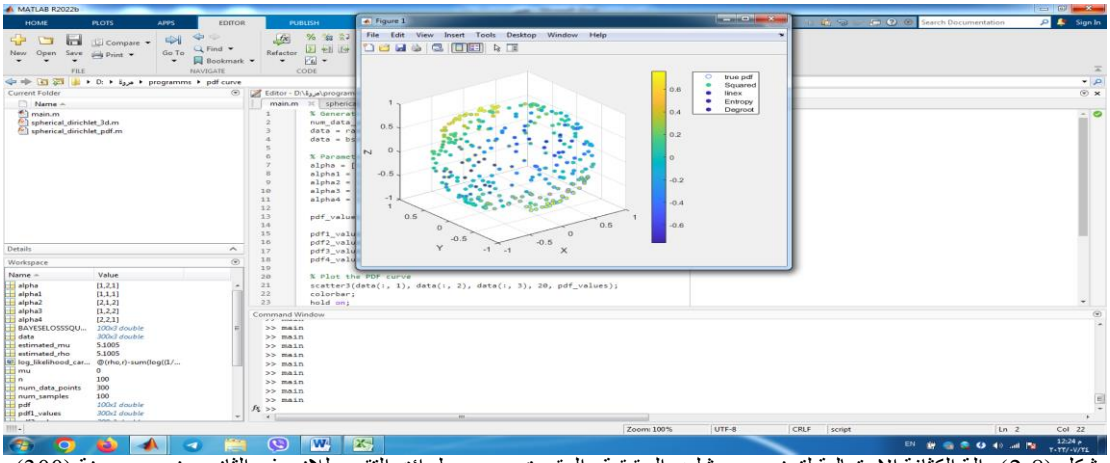
Method		Squared	
n	Parameter	Estimation	IMSE
100	α_1	1.55637	0.0005564
	α_2	2.58474	0.0005847
	α_3	1.54516	0.0005452
200	α_1	1.55344	0.0005534
	α_2	2.58124	0.0005812
	α_3	1.54345	0.0005435
300	α_1	1.54367	0.0005437
	α_2	2.55677	0.0005568
	α_3	1.52354	0.0005235
400	α_1	1.53789	0.0005379
	α_2	2.54566	0.0005457
	α_3	1.51908	0.0005191
500	α_1	1.47833	0.0004783
	α_2	2.45223	0.0004522
	α_3	1.44934	0.0004493



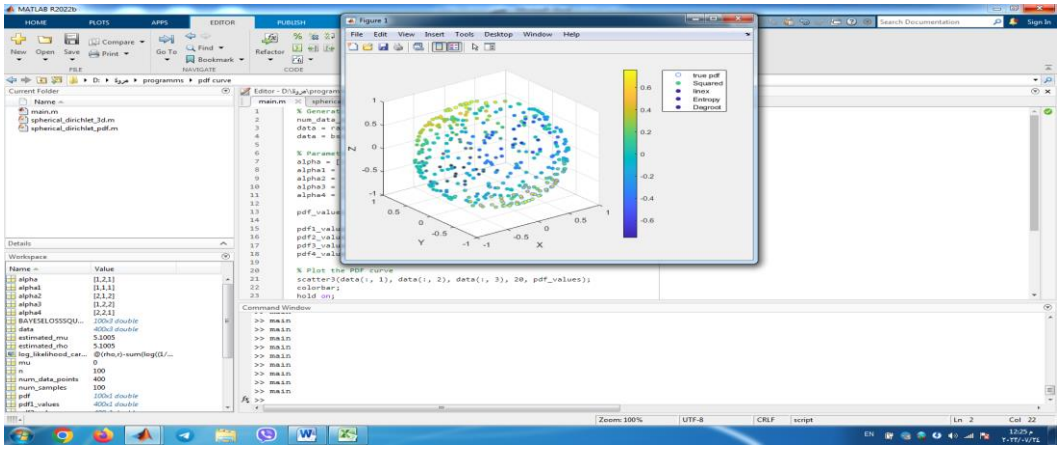
شكل (2-6) دالة الكثافة الاحتمالية لتوزيع ديريشليه الحقيقية والمقدرة بموجب طرائق التقدير للانموذج الثاني عند حجم عينة (100)



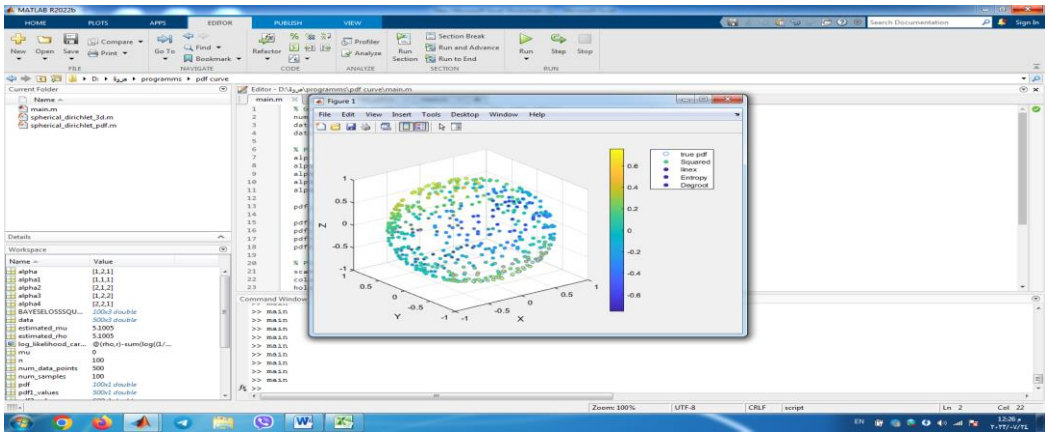
شكل (2-7) دالة الكثافة الاحتمالية لتوزيع ديريشليه الحقيقية والمقدرة بموجب طرائق التقدير للأنموذج الثاني عند حجم عينة (200)



شكل (8-2) دالة الكثافة الاحتمالية لتوزيع ديريشليه الحقيقية والمقدرة بموجب طرائق التقدير للانموذج الثاني عند حجم عينة (300)



(400) شكل (9-2) دالة الكثافة الاحتمالية لتوزيع ديريشليه الحقيقية والمقدرة بموجب طرائق التقدير للانموذج الثاني عند حجم عينة



(500) شكل (10-2) دالة الكثافة الاحتمالية لتوزيع ديريشليه الحقيقية والمقدرة بموجب طرائق التقدير للانموذج الثاني عند حجم عينة

نلاحظ من جدول (2-2) والاشكال (6-2)الى (10-2) مايأتي :

عند المعلمات الافتراضية $(\alpha_1 = 1, \alpha_2 = 2, \alpha_3 = 1)(\alpha_1 = 1, \alpha_2 = 2, \alpha_3 = 1)$ ان طريقة بيز في ظل دالة خسارة (التربيعية) عند حجم عينة (500) حققت افضلية على باقي حجوم العينات في تقدير معلمات التوزيع الكروي يليها حجم العينة (400) لانها سجلت اقل متوسط مربعات خطأ.

5.4 الاستنتاجات

1. أظهرت النتائج من خلال الجانب التجريبي ان دالة الخسارة التربيعية حققت افضلية عند حجم عينة (500) حققت افضلية على باقي حجوم العينات في تقدير معلمات التوزيع الكروي يليها حجم العينة (400) لانها سجلت اقل متوسط مربعات خطأ.

2. عند المعلمات الافتراضية $(\alpha_1 = 1, \alpha_2 = 2, \alpha_3 = 1)(\alpha_1 = 1, \alpha_2 = 2, \alpha_3 = 1)$ ان طريقة بيز في ظل دالة خسارة (التربيعية) عند حجم عينة (500) حققت افضلية على باقي حجوم العينات في تقدير معلمات التوزيع الكروي يليها حجم العينة (400) لانها سجلت اقل متوسط مربعات خطأ تكاملي.
3. التوزيع الكروي تكون تقديراته افضل عندما يكون حجوم العينات كبيرة.

6.4 التوصيات

1. نوصي باستعمال حجوم عينات اعلى من حجم العينة 500 عند التقدير لكونه يعطى نتائج افضل.
2. استعمال دوال خسارة متعددة في تقدير بيز للتوزيع الكروي .
3. يمكن استعمال قيم افتراضية اعلى لمعلمات توزيع ديريشليه الكروي $\alpha_I \alpha_I$ ثلاثي الابعاد

المصادر

1. *der Merwe , Sean van; de Waala, Daan , (2018), " Bayesian Fitting of Dirichlet Type I and II Distributions ", arXiv:1801.02962v2 [math.ST] 5.*
2. Guardiola, Jose H., (2020), " The spherical-Dirichlet distribution ", Guardiola Journal of Statistical Distributions and Applications (2020) 7:6 https://doi.org/10.1186/s40488-020-00106-9
3. Hauberg , Soren, (2018) , "Directional Statistics with the Spherical Normal Distribution" 2018 21st International Conference on Information Fusion (FUSION)
4. Kazem, Amory Hadi , and Muslim, Bassem shalabeh (2002), "advanced economic measurement theory and practice", dunia Al-Amal press, Baghdad-Iraq
5. Maghrabi, Tarek Mansour (2007), "programming in the MATLAB language", modern world of books, Irbid-Jordan..
6. Robert, C. P., & Casella, G. (2010). Monte Carlo statistical methods. Springer Science & Business Media.
7. *Ying-Ying Zhang, (2020), " Bayesian Inference on Complicated Data", Yunnan University, China.*