مقارنة طرائق توليد بيانات كلا من توزيع كاما وتوزيع بيتا

أ. م. د خالد ضاري الطائي
كلية الإدارة والاقتصاد / جامعة بغداد
قسم الاحصاء

المستخلص

يتمتع توزعي كاما وبيتا باهمية بالغة من الناحية التطبيقية في مختلف المجالات الاحصائية وفي تطبيقات المعولية ومراقبة جودة الانتاج .وهناك عدد من الطرائق لتوليد بيانات تسلك على وفق هذين التوزيعين وهذه الطرائق تعتمد بالدرجة الاولى على معلمات الشكل لكل توزيع والعلاقة بين التوزيعين بالدرجة الاساس وعلاقتهما ببعض التوزيعات الاحتمالية الاخرى .

يهدف هذا البحث الى تحديد الطريقة الاكفأ لتوليد بيانات تسلك على وفق توزيع كاما Gamma بناءً على قيمة معلمة الشكل α وعلى وفق توزيع بيتا Beta على اساس معلمتي الشكل α, β .

الشائع ان المقارنات بين طرائق تولبد المتغيرات العشوائية حاسوبيا تعتمد على عدد الارقام العشوائية $U\sim U(0,1)$, فضلا عن الوقت المستغرق لتوليد متغير عشوائي واحد بغض النظر عن كفاءته الاحصائية.

في هذا البحث تم لاول مرة استخدام متوسط مربعات الخطأ MSE لتقديرات المتوسطات والتباينات للعينات المولدة باحجام مختلفة كموشر أحصائي لكفاءة الطرائق .

فضلا عن أقتراح طريقتين لتوليد بيانات تتوزع حسب توزيع بيتا باستعمال أقل عدد من الارقام العشوائية وباعلى كفاءة .

تم أثبات ذلك بتنفيذ عدد من تجارب المحاكاة، ببرامج كتبت من قبل الباحث بلغة البرمجة MATLAB.

مجلة العلوم
الاقتصادية والإدارية
المجلد ١٩
العدد ٧٢
الصفحات ٢١٤-٢٤٢

المقدمة

يتمتع توزعي كاما وبيتا باهمية بالغة من الناحية التطبيقية في مختلف المجالات الاحصائية وفي تطبيقات المعولية ومراقبة جودة الانتاج .هناك عدد من الطرائق لتوليد بيانات تسلك على وفق هذين التوزيعين وهذه الطرائق تعتمد بالدرجة الاولى على معلمات الشكل لكل توزيع والعلاقة بين التوزيعين بالدرجة الاساس وعلاقتهما ببعض التوزيعات الاحتمالية الاخرى.

من خلال بحثنا في محاكاة هذين التوزعين وتدريسنا لمادة المحاكاة في الدراستين الاولية والعليا وجدنا من المناسب ان نبحث في طرائق توليد هذين التوزعين ومقارنتها لتحديد الاكفأ منها ولكل انواع المعلمات للتوزيعين .

الجانب النظري للبحث تضمن عرض مختصر للتوزيعين والعلاقة بينهما وبين بعض التوزيعات الاحتمالية الاخرى فضلا عن الاشتقاقات والبراهين الضرورية للتمهيد لمناقشة الخوارزميات المستخدمة لتوليد بيانات كل توزيع وبحسب الحالة .

تم تقديم مقترحين لخوارزميتين لتوليد بيانات توزيع بيتا وبحسب معلمات الشكل للتوزيع بهدف زيادة الكفاءة وتقليل عدد الارقام العشوائية المستخدة في الخوارزمية الاولى وتقليل الارقام العشوائية المستعملة لتوليد قيمة واحدة من بيانات التوزيع في الخوارزمية الثانية .

الجانب التجريبي تضمن عدد من تجارب المحاكاة لمقارنة طرائق توليد بيانات كل توزيع في ضوء المعطيات المتوفرة لكلا التوزيعين .

الشائع في المقارنات لتحديد الطريقة الفضلى حاسوبيا بالأعتماد على عدد الارقام العشوائية **Random Number** التي تستعمل لتوليد قيمة واحدة من المتغير العشوائي الذي يسلك بحسب التوزيع المطلوب وعلى الرغم من أهمية ذلك ألا أن الهدف من توليد عينات التوزع هو كفاءتها من الناحية الاحصائية لذلك أستعملنا الموشر الاحصائي متوسط مربعات الخطأ **mse** هنا لاول مرة لتحديد الطرائق الاكفأ واثبات كفائة الطرائق المقترحة ، وتم التوصل لعدد من الاستنتاجات توجت هذا البحث .

هدف البحث

يهدف البحث الى مناقشة محاكاة توزيعي كاما وبيتا ومقارنة طرائق توليد كل توزيع فضلا عن الاجابة على عدد من التسائلات حول طرائق توليد بيانات هذين التوزيعين واختيار اكثر الطرق كفاءة حسب المعطيات المتوفرة .فضلا عن اقتراح حلول لبعض المشكلات الخاصة بتوليد بيانات التوزيعين.

اولا- الجانب النظري

يتضمن هذا الجانب عرض ومناقشة عدد من طرائق توليد بيانات كلا من توزيعي كاما وبيتا فضلا عن أثارة عدد من التساؤلات بغية الاجابة عنها واقتراح خوارزميتين لتوليد بيانات بيتا .

١- توزيع كاما Gamma Distribution [11,14]

ان توزيع كاما له دالة كثافة احتمالية مشتقة أصلا من دالة كاما **Gamma Function** وتسمى في بعض الاحيان بتكامل كاما .
اهمية هذا التوزيع تكمن في دراسة التوقفات في المصانع وفي مجال المعولية **Reliability**
وان دالة الكثافة الاحتمالية **p.d.f** للتوزيع هي :-

$$f(x) = \begin{cases} \frac{1}{\Gamma\alpha\ \beta^{\alpha}}\ e^{\frac{-x}{\beta}} x^{\alpha-1} & 0 \leq\ x \leq\ \infty \\ 0 & elswhere \end{cases} \qquad (1)$$

حيث

$\alpha,\beta > 0$

α تمثل معلمة الشكل **(Shape Parameter)**
β تمثل معلمة القياس **(Scale Parameter)**
$\Gamma\alpha$ تمثت دالة كاما الكاملة **(Complete Gamma Function)**
نرمز للتوزيع بـ $ga(\alpha,\beta)$ ، وسط التوزيع هو $\mu = \alpha\beta$ وتباينه $\sigma^2 = \alpha\beta^2$
في هذا المبحث سنحاول ان نناقش طرائق توليد متغيرات عشوائية تسلك على وفق توزيع كاما وهذه الطرائق تعتمد في كفاءتها على قيمة وطبيعة معلمة الشكل للتوزيع. في حالتيها الصحيحة **Integer** وغير الصحيحة **Noninteger**

١-١ طرائق محاكاة توزيع كاما عندما تكون معلمة الشكل صحيحة Integer

هناك عدة طرائق لتوليد بيانات توزيع كاما عندما تكون معلمة الشكل α عدد صحيح موجب تم أختيار ثلاث طرائق هي :-

١-١-١ طريقة (1-١) The Convolution Method [8 ,9,11,12,13]

اذا كان $x \sim ga(\alpha,\beta)$ وكانت α *__عدد صحيح موجب و صغير__*. أي $c \leq \alpha$ *باعتبار* c ثابت صحيح .
بالامكان الاعتماد على حقيقة احصائية تقول اذا كان

$y \sim \exp(\beta)$

يتوزع توزيعا اسيأ فان مجموع $i=1.....\alpha$ y_i من المتغيرات المستقلة يتوزع حسب توزيع كاما بالمعلمات (α,β). وفي بعض الاحيان يسمى التوزيع في هذه الحالة **(Erlang Distribution)**

أي أن بالاعتماد على طريقة التحويل العكسي (Inverse Transformation Method) فأن $y_i = -\frac{1}{\beta}\log(U_i)$

$$x = \sum_{i=1}^{\alpha} y_i \sim \text{ga}(\alpha,\beta) \quad \ldots\ldots\ldots \ (2)$$

طريقة التوليد تعتمد على هذه الحقيقة وتعد من ابسط طرائق توليد المتغيرات العشوائية التي تسلك على وفق توزيع كاما. مما تقدم نخلص الى الخوارزمية الاتية :

خوارزمية (1-1)

الخطوة الاولى : ولد α من المتغيرات العشوائية التي تتوزع حسب التوزيع المنتظم

$U_1\ U_2\ U_3 \ldots\ldots\ldots\ldots U_\alpha$

الخطوة الثانية : ولد α من المتغيرات التي تتبع التوزيع الآسي باستعمال الصيغة الاتية

$$y_i = -\frac{1}{\beta}\log(U_i)$$

الخطوة الثالثة : أجعل

$$x = \sum_{i=1}^{\alpha} y_i \qquad i = 1 \ldots\ldots\alpha$$

أي أن

$$x = -\sum_{i=1}^{\alpha}\frac{1}{\beta}\log(U_i) = -\frac{1}{\beta}\log\left(\prod_{i=1}^{\alpha}U_i\right)$$

هذه الطريقة تتطلب α من المتغيرات العشوائية $U \sim Uniform(0,1)$. لتوليد قيمة واحدة من x .
فاذا كانت قيمة α كبيرة فان ذلك يتطلب عدد كبير من الارقام العشوائية لتوليد قيمة واحدة من x .
لذلك تكون هذه الطريقة غير كفوءة وغير عملية فنلجأ الى طرائق أخرى تتطلب اقل عدد من الارقام العشوائية.

بالامكان استعمال طريقة ثانية لتوليد بيانات تتوزع حسب توزيع كاما بعدد اقل من الارقام العشوائية وان واحد من الاساليب الكفوءة لمحاكاة توزيع كاما هو اللجوء الى طريقة الرفض والقبول .

٢-١-١ طريقة (٢-١) طريقة الرفض والقبول Acceptance-Rejection Method [12]

طريقة الرفض والقبول لتوليد بيانات كاما عندما تكون α عدد صحيح وكبير أي $\alpha > c$.
تتلخص الطريقة بالخطوات الاتية:-

١- توليد متغير عشوائي x يتوزع حسب توزيع كوشي المبتور بالمعلمات $\beta = \sqrt{2\alpha - 1}$ و $\theta = \alpha - 1$ بشرط ان يكون المتغير موجبا ويقبل باحتمال محدد.

الان نبدأ بمحاكاة الدالة الاتية بعد التعويض عن قيمة β وθ

$$h(x) = \frac{\beta}{A\pi[\beta^2 + (x-\theta)^2]}$$

$$= \frac{\sqrt{\alpha - 1}}{A\pi[2\alpha - 1 + (x - \alpha + 1)^2]}$$

حيث

$$A = P\{Gauhy(\sqrt{2\alpha - 1}\ \ , \alpha - 1) > 0\}$$

$$= \frac{\sqrt{2\alpha - 1}}{\pi} \int_0^\infty \frac{dx}{2\alpha - 1 + (x - \alpha + 1)^2}$$

والان بالامكان أن نلاحظ ان كل قيم x>0

$$\leq (2\alpha - 1)\frac{(\alpha-1)^{\alpha-1}}{e^{\alpha-1}} e^{x} x^{\alpha-1}[2\alpha - 1 + (x - \alpha + 1)^2]$$

وكذالك

$$\frac{f((x)}{h(x)} = \frac{e^{-x}x^{\alpha-1}\ [2\alpha - 1 + (x - \alpha + 1)^2]A\pi}{(\alpha - 1)!\ \sqrt{2\alpha - 1}}$$

$$\leq \frac{A\pi\sqrt{2\alpha - 1}\ (\alpha - 1)^{\alpha-1}}{(\alpha - 1)!\ e^{\alpha-1}} \equiv c$$

وحسب طريقة الرفض (Von – Neumann) يمكن توليد قيمة لـ (X) من الدالة h ويتم قبول X اذا كان الرقم العشوائي U يحقق العلاقة الاتية:-

$$U \leq \frac{f(X)}{Ch(X)} = e^{-(X-\alpha+1)}\left(\frac{X}{\alpha-1}\right)^{\alpha-1}\left[1 + \frac{(X-\alpha+2)^2}{2\alpha-1}\right]$$

أو

$$-\log U \geq X-\alpha+1-(\alpha-1)\log\left(\frac{X}{\alpha-1}\right)-\log\left[1+\frac{(X-\alpha+1)^2}{2\alpha-1}\right]$$

وبما ان الطرف الايمن من العلاقة السابقة هو متغير عشوائي يتوزع توزيعا أسيا

- log (U) ~ exp (1)

بذلك يمكن بناء ا لخوارزمية الاتية لمحاكاة المتغير العشوائي x الذي يتوزع على وفق توزيع كاما $x \sim ga(\alpha,1)$

الخوارزمية (٢-١)

الخطوه الاولى : ولد رقمان عشوائيان U_2 U_1 , ثم اجعل

$= U_1-\frac{1}{2}$, $Y_2 = U_2-\frac{1}{2}Y_1$

الخطوة الثانية : أذا كان $Y_1^2+Y_2^2 \leq \frac{1}{4}$ أجعل $W = {}^{Y_1}/_{Y_2}$

بذلك يكون W متغير عشوائي يتوزع حسب توزيع كوشي القياسي **Standard Cauchy Distribution.**

الخطوة الثالثة : اجعل

$$X = \left(\sqrt{2\alpha-1}\right)W+\alpha-1$$

بذلك يكون X متغير عشوائي يتوزع حسب توزيع كوشي بالمعلمات

$\beta = \sqrt{2\alpha-1}$,$\theta = \alpha-1$

اذا كان **X < 0** ارجع الى الخطوة الاولى .

الخطوة الرابعة : ولد المتغير العشوائي v الذي يتوزع توزيعا أسيا

V= - log (u)

أذا كان

$$V \geq X-\alpha+1-(\alpha-1)\log\left(\frac{X}{\alpha-1}\right)-\log\left[1+\frac{(X-\alpha-1)^2}{2\alpha-1}\right]$$

توقف واعتمد **X ~ ga(α ,1)** عدا ذالك أرجع للخطوة الاولى .

عندما تكون α كبيرة فان هذه الطريقة تكون في قمة كفائتها لمحاكاة **ga (α ,1)** أ ذا كان عدد التكرارات في الخطوة الرابعة هي $\sqrt{\pi}$=1.77 [12]

مما تقدم نجد ان :

الخوارزمية (1-1) تستعمل عندما تكون معلمة الشكل صغيرة أي $\alpha \leq c$.

الخوارزمية (2-1) تستعمل عندما تكون معلمة الشكل كبيرة أي $\alpha > c$.

هنا يبرز سؤال : ماهي قيمة c ؟

لم نعثر على أجابة عن هذا السؤال في المصادر العلمية لذلك كان واحدا من اهداف هذا البحث الاجابة عنه .

2-1- **توزيع كاماعندما تكون معلمة الشكل غير صحيحة** Gamma Distribution with No integral Shape Parameter

1-2-1 **الطريقة**(3-1) [8]

أفرض ان x يتوزع حسب توزيع كاما بالمعالم (α, β) وان معلمة الشكل α عددا غير صحيحا . في هذه الحالة نفرض ان x هو عبارة عن مجموع k+1 من المشاهدات المستقلة تتوزع حسب توزيع كاما ، جميعها لها معلمة قياس واحدة هي β ولكن اول k منها لها معلمة شكل قيمتها (١) واحد ، وان المشاهدة k+1 لها معلمة شكل هي

$$\gamma = \alpha - [\alpha]$$

حيث $k=[\alpha]$ يمثل الجزء الصحيح من (α) و γ يمثل الكسر منها.

الان نفرض ان Y,Z هي متغيرات مستقلة تتوزع حسب التوزيعات الاتية:-

$$Z \sim ga(1,1) \quad \text{و} \quad Y \backsim \beta e(\gamma\ ,1-\gamma\ )$$

فان $W=\beta YZ \sim ga(\gamma, \beta)$ ولاثبات ذالك نلاحظ ان

$$f_{y,z}(y,x) = \begin{cases} \dfrac{1}{\Gamma\gamma\Gamma(\gamma-1)} \quad y^{\gamma-1}\,(1-y)^{\gamma} \quad e^{-z} & 0 \leq y \leq 1, 0 \leq z \leq \infty \\ 0 & elsewhere \end{cases}$$

فان

$$f_{w,z}(w,z) = \begin{cases} \dfrac{\beta^{-\gamma}}{\Gamma(\gamma)\Gamma(1-\gamma)} w^{\gamma-1} \left(z - \dfrac{w}{\beta}\right)^{-\gamma} \quad e^{-z} & 0 \leq w, z \leq \infty \\ 0 \end{cases}$$

$$f_w(w) = \int_0^{\infty} f_{w,z}(w,z)\, dz = \begin{cases} \dfrac{\beta^{-\gamma}}{\Gamma(\gamma)} w^{\gamma-1} e^{-w/\beta} & 0 \leq w, z \leq \infty \\ 0 & elsewhere \end{cases}$$

لتوليد متغير يتوزع $ga\ (\alpha,\beta)$ في حالة α عدد غير صحيح نستعمل الخوارزمية الاتية :

الخوارزمية (1-3) ***[8]***

الخطوة الاولى : اجعل k يمثل اكبر عدد صحيح في α أي $k=[\alpha]$

الخطوة الثانية : اجعل $\gamma=\alpha-k$

الخطوة الثالثة : ولد v الذي يتوزع على وفق توزيع كاما $v\sim ga\ (k,1)$

أي $v=-\log\left(\prod_{i=1}^{k}u_i\right)$

الخطوة الرابعة : ولد المتغيرات y , z بحيث

$z\sim exp(1)$ $\quad y\sim\beta a(\gamma\,,1-\gamma)$

الخطوة الخامسة : أجعل $x=\beta(v+yz)$ بذلك $x\sim ga\ (\alpha,\beta)$

في هذه الخوارزمية نجد ان العدد المتوقع من الارقام العشوائية الذي يتطلبه توليد قيمة واحدة من x هو $k+3$ [8]

الخوارزمية (1-٤) خوارزمية (Ahrens-Dieter) [15]

بالاعتماد على طريقة الرفض والقبول سنعرض خوارزمية أخرى لتوليد متغير يتوزع على وفق توزيع كاما$\xi\sim ga\,(\gamma,1)$ $\quad 0<\gamma<1$

الخطوة الاولى : أجعل m=1 .

الخطوة الثانية : ولد V_{3m-1} V_{3m-2}, و V_{3m} كمتغيرات مستقلة تتوزع U(0,1) .

الخطوة الثالثة : **اجعل $v_0=\frac{e}{e+\gamma}$ اذا كان $V_{3m-2}\ \le v_0$ اذهب للخطوة الرابعة عدا ذالك**

ا ذهب للخطوة الخامسة .

الخطوة الرابعة: **اجعل $\xi_m=(V_{3m-1}\ )^{\frac{1}{\gamma}}$ و $\eta_m=V_{3m}\xi_m^{\gamma-1}$ أذهب للخطوة السادسة .**

الخطوة الخامسة : أجعل $\xi_m=1-LnV_{3m-1}$ و $\eta_m=V_{3m}\ e^{-\xi_m}$.

الخطوة السادسة : أذا $\eta_m>\xi_m^{\gamma-1}e^{-\xi_m}$ اجعل m=m+1 ثم أذهب الى الخطوة الثانية .

الخطوة السابعة: أجعل $\xi = \xi_m$ حيث $\xi \sim ga\,(\gamma, 1)$.

الان

يمكن ان نولد $x \sim ga\,(\alpha, \beta)$ عندما تكون α عددا غير صحيح وذلك من خلال العلاقة الاتية:-

$$x = \beta\left(\xi - \sum_{i=1}^{k} \log\,(U_i)\right)$$

حيث

$k=[\alpha]$ ***يمثل الجزء الصحيح من α .***

$\gamma = \alpha - k$ ***يمثل الجزء الكسري من α***

لم نعثر على عدد للارقام العشوائية U المستعملة لتوليد قيمة واحدة من x . هذا ما سنحاول تحديده من خلال تجارب المحاكاةلهذه الخوارزمية .

٢- توزيع بيتا Beta Distribution

ان توزيع بيتا مشتق من دالة بيتا beta function أو ما يسمى في بعض الاحيان بتكامل بيتا . للتوزيع أهمية في تطبيقات الرقابة على جودة الانتاج من خلال تكوين جداول عينات القبول Acceptance Sampling Tables والتي تستخدم في أتخاذ القرار بشأن قبول وجبات الانتاج أستنادا الى نسب الوحدات المعيبة في العينة فضلا عن التطبيقات الاخرى.

دالة الكثافة الاحتمالية pdf للتوزيع هي :-

$$f(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma\alpha\;\Gamma\beta}\; x^{\alpha-1}(1-x)^{\beta-1} \qquad 0 \le x \le 1 \qquad (3)$$

في هذه الحالة يقال أن المتغير العشوائي x يتوزع وفق دالة توزيع بيتا بالمعلمتين α ، β ويرمز له بالرمز $x \sim \mathfrak{Be}\,(\alpha , \beta)$.

وللتوزيع وسط هو $\mu = \frac{\alpha}{\alpha+\beta}$ وتباين هو $\sigma^2 = \frac{\alpha\beta}{(\alpha+\beta+1)(\alpha+\beta)^2}$

٢-١ محاكاة توزيع بيتا عندما تكون α و β عددا صحيح

هناك عدة طرائق لتوليد متغيرات عشوائية تتوزع على وفق توزيع بيتا عندما تكون معلمتي الشكل β ، α للتوزيع اعدادا صحيحة .

٢-١-١ الطريقة الاولى (2-1), [8,9,12]

هذه الطريقة تعتمد على العلاقة بين توزعي كاما وبيتا

لو فرضنا أن x_1 , x_2 متغيرين عشوائين مستقلين يتوزعان حسب توزيع كاما بدوال كثافة أحتمالية هي:-

$$f_{x_1}(x)=\begin{cases} \frac{1}{\Gamma(\alpha)}\, e^{-x} x^{\alpha-1} & 0 \le x \le \infty \\ 0 & x < 0 \end{cases}$$

$$f_{x_2}(y)=\begin{cases} \frac{1}{\Gamma(\alpha)}\, e^{-y} y^{\alpha-1} & 0 \le y \le \infty \\ 0 & y < 0 \end{cases}$$

على التوالي فان

$$\mathrm{x} = x_1 / (x_1 + x_2) \qquad \text{.....} \quad \textbf{(4)}$$

x: متغير عشوائي يتوزع على وفق توزيع بيتا بدالة كثافة احتمالية كما في العلاقة (٣)

لإثبات ذلك نأخذ الدالة المشتركة لـ x_1 , x_2

$$f_{x_1x_1}(x,y) = \begin{cases} \frac{1}{\Gamma(\alpha)\Gamma(\beta)}\, e^{-(x+y)}\, x^{\alpha-1} y^{\beta-1} & 0 \le x \le \infty \\ 0 & x, y < 0 \end{cases}$$

نفرض أن $S = x_1$ و $T = x_1 + x_2$ فان الدالة الاحتمالية المشتركة لـ S و T هي

$$f_{S,T}(s,t) = J\, f_{x_1x_2}(x,y)$$

$$f_{S,T}(s,t) = \begin{cases} \frac{1}{\Gamma(\alpha)\Gamma(\beta)}\, e^{-t}\, s^{\alpha-1} (t-s)^{\beta-1} & 0 \le s,t \le \infty \\ 0 & elsewhere \end{cases}$$

وبما ان جوكوبيان J هي :-

$$J=\left\|\begin{matrix} \frac{\partial x}{\partial s} & \frac{\partial x}{\partial t} \\ \frac{\partial y}{\partial s} & \frac{\partial y}{\partial t} \end{matrix}\right\|$$

نفرض أن

W=S و X= S/T $=X_1/(X_1+X_2)$

فان دالة الكثافة الاحتمالية المشتركة لـ W و X هي :-

$$f_{W,X}(w,x) = \begin{cases} \frac{1}{\Gamma(\alpha)\Gamma(\beta)} \; e^{-\frac{w}{x}} \, w^{\alpha+\beta-1} \, x^{-(\beta-1)} (1-x)^{\beta-1} & 0 \le w \le \infty \;\; 0 \le x \le 1 \; \ldots \; (5) \\ 0 & elsewhere \end{cases}$$

وبتكامل العلاقة (٥) بالنسبة الى W نستخرج العلاقة (٣) .

استنادا لما سبق أذا كان

x_1 ~ ga (α,1)

x_2 ~ ga (β ,1)

و

$$x = \frac{x_1}{x_1+x_2}$$

فان

x ~ 𝔅e (α , β)

بالأعتماد على هذه الحقيقة يمكن بناء الخوارزمية التالية لتوليد متغيرات عشوائية تتوزع على وفق توزيع بيتا بالمعلمات α وβ

الخوارزمية(1- 2)

الخطوة الاولى : اجعل i=1 و $x_1 = 0$ و $x_2 = 0$

الخطوة الثانية : ولد متغير عشوائي U_i ~ U (0, 1)

الخطوة الثالثة : اذا كان $i \le \alpha$ أجعل $x_1 = x_1 - \log(U_i)$

عدا ذالك اجعل $x_2 = x_2 - \log(U_i)$

الخطوة الرابعة: اجعل i = i+1

الخطوة الخامسة: أذا كان $i \le \alpha + \beta$ أرجع للخطوة الثانية .

الخطوة السادسة : أجعل $x = \frac{x_1}{(x_1+x_2)}$

ان ملاحظتنا حول هذه الطريقة هي أن حساب قيمة x_2 و x_1 أعتمدت على الخوارزمية (1-1) كما هو واضح في الخطوة الثالثة أعلاه بغض النظر عن قيمة معلمة الشكل الخاصة بكل متغير أذا كانت صغيرة أم كبيرة بما ان هذه الخوارزمية تستعمل في حالة كون معلمة الشكل صغيرة نعتقد ان ذلك يوثر على كفاءة الطريقة اذا كان
$\alpha \& \beta > c$ أو $\alpha\ or \beta > c$

٢-١-2 **الطريقة** (2-2) **لتوليد بيانات بيتا عندما تكون معلمات الشكل اعداداً صحيحة** , [5,12]

هناك طريقة أخرى لتوليد المتغيرات العشوائية التي تتوزع على وفق توزيع بيتا عندما تكون معلمات التوزيع أعدادا صحيحة، و فكرة هذه الطريقة مستمدة من نظرية الاحصائات المرتبة Statistic Theory Order. تتلخص الطريقة بما ياتي :

اذا كان لدينا $U_1\ U_2\ U_3 \dots\dots\dots\dots. U_{\alpha+\beta-1}$ **من المتغيرات العشوائية المستقلة** U_i **U(0,1) ~**
ورتبت بشكل تصاعدي بحيث يكون $U_1 < U_2 < \cdots\ U_{\alpha+\beta-1}$

نقسم المتغيرات المرتبة الى ثلاثة مجموعات هي :-

١- المجموعة الاولى : ١ –α من القيم أقل من x
٢- المجموعة الثانية : قيمة واحدة تساوي x
٣- المجموعة الثالثة : β من القيم أكبر من x

والان اذا 1 - α +β من المتغيرات العشوائية المنتظمة قسمت الى ثلاثة مجموعات جزئية بالاحجام ١ –α و١ و β-1
فان احتمال كل متغير في المجموعة الاولى هو اقل من x والمتغير في المجموعة الثانية يساوي x ، وجميع المتغيرت في المجموعة الثالثة اكبر من x
هو

$$(P\{U < x\})^{\alpha-1}\ f_u(x)(P\{U > x\})^{\beta-1} = x^{\alpha-1}\ (1-x)^{\beta-1} \qquad (6)$$

وعلى ذلك فان U_α **يتوزع على وفق توزيع بيتا بالمعالم** (α,β)
لذلك فان

$$f(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma\alpha\Gamma\beta} x^{\alpha-1}\ (1-x)^{\beta-1} \qquad 0 \le x \le 1$$

وهذه مطابقة الى الصيغة رقم (3)
والان يمكن ان نخلص الى الخوارزمية الاتية:

الخوارزمية (2-2)

الخطوة الاولى: ولد $U_1, U_2, \dots\dots\dots, U_{\alpha+\beta-1}$ من المتغيرات العشوائية المستقلة التي تتوزع U(0,1) .

الخطوة الثانية: رتب U_i .من القيم ترتيبا تصاعديا $i = 1\dots\dots \alpha + \beta - 1$.

الخطوة الثالثة : أحعل $x = U_\alpha$.

ان الطريقة أعلاه والمتمثلة بالخوارزمية (2-2) تحتاج $\alpha + \beta - 1$ من الارقام العشوائية وعدد من المقارنات هو
$\sum (\alpha + \beta - 1 - J) = (\alpha/2)(\alpha + 2\beta - 3)$ توليد قيمة واحدة من x .

2-1-3 الطريقة (2-3) طريقة مقترحة

سنقدم هنا المقترح الاول في هذا البحث والمتمثل بتعديل بالخوارزمية رقم(2-1) .

لتوليد $x \sim \mathfrak{Be}(\alpha_1, \alpha_2)$ بما أن x في العلاقة رقم (٤) هو نتجة لعلاقة بين x_1 و x_2 والتي تتوزع $x_1 \sim ga(\alpha_1, 1)$ و $x_2 \sim ga(\alpha_1, 1)$ على التوالي .

للحصول على x بأقل عدد من الارقام العشوائية وباعلى كفائة ياتي من كون كلا من x_1 و x_2 تحسب باقل عدد من الارقام العشوائية وباعلى كفاءة .
الخوارزمية (١-٢) تستعمل الخوارزمية (١-١) لحساب قيمة x_i , i=1,2 بغض النظر عن قيمة معلمة الشكل α_i

ونعتقد أن ذلك يقلل من كفائة الخوارزمية لان الخوارزمية (١-١) كفوٌة عندما تكون معلمة الشكل صغيرة ، أي $\alpha_i \le c$ أما في حالة كون معلمة الشكل كبيرة أي $\alpha_i > c$ يمكن استخدام خوارزميات أخرى لحساب x_i مثل
الخوارزمية (٢- ١) أو (٣-١) .
بناء على ما تقدم نقترح الاتي :-
أستعمال الخوارزمية رقم(1-١) لتوليد قيمة x_i , i=1,2 اذا كانت $\alpha_i \le c$
و أستعمال الخوارمية رقم (1-2) لتوليد قيمة x_i , i=1,2 أذا كانت $\alpha_i > c$
بذلك يمكن صياغة خوارزمية خاصة بالطريقة كما يأتي :-

الخوارزمية المقترحة (2-3)

الخطوة الاولى : أجعل i=1 .

الخطوة الثانية: اذا كان $\alpha_i \leq c$ استخرج قيمة x_i بتنفيذ الخوارزمية رقم(١-1) .
عدا ذلك استخرج قيمة x_i بتنفيذ الخوارزمية رقم(٢-1) .

الخطوة الثالثة: أجعل i= i+1 .

الخطوة الرابعة: أذا كان i≤ 2 أرجع للخطوة الثانية .

الخطوة الخامسة: أجعل $x = x_1 / (x_1 + x_2)$

ملاحظة :الخوارزمية المقترحة اعلاه لايمكن تطبيقها عمليا الا في حالة تحديد قيمة c والتي تم تحديدها لاول مرة في هذا البحث حيث كانت c=5 لذلك تم أختيار الخوارزمية (١-1) عندما تكون . تكون$\alpha \leq c$.كما تم استعمال الخوارزمية (٢-١) عندما تكون $\alpha > c$

٢-٢ محاكاة توزيع بيتا عندما تكون α أو β عددا غير صحيحا Beta Distribution with No integral *Parameters*

في بعض الاحيان تكون α أو β أوكلاهماعددا غير صحيح وبذلك تكون الطرائق السابقة لتوليد متغيرات على وفق توزيع بيتا غير قابلة للتطبيق . فنستعمل طرائق اخرى تتناسب وقيمة معلمات الشكل وأهمها الطريقتين التاليتين:

١-٢-٢ الطريقة (2-4)

لتوليد البيانات لهذا التوزيع عندما تكون α ،β أعدادا غير صحيحة أو احداهما وذلك من خلال الخوارزمية الآتية:-

الخوارزمية (2-4) [7, 8,9]

الخطوة الاولى : **أجعل** $k_1 = [\alpha]$ **و** k_2 $[\beta]$= **الذان يمثلان الجزء الصحيح من** β، α .

الخطوة الثانية : **أجعل** $\gamma_1 = \alpha - k_1$ **و** $\gamma_2 = \beta - k_2$ **الذان يمثلان الجزء الكسري من** β، α

الخطوة الثالثة : **ولد** $h_2 \sim ga\,(k_2, 1)$ و $h_1 \sim ga\,(k_1, 1)$

الخطوة الرابعة: **ولد** $y_1 \sim \mathfrak{Be}\,(\gamma_1, 1-\gamma_1)$ **و** $y_2 \sim \mathfrak{Be}\,(\gamma_2, 1-\gamma_2)$
$z_1 \sim \exp(\beta)$ $z_2 \sim \exp(\beta)$

الخطوة الخامسة : أحسب المتغير x الذي يتوزع حسب توزيع بيتا من العلاقة الاتية

$$x = (h_1 + y_1 z_1)/(h_1 + h_2 + y_1 z_1 + y_2 z_2)$$

$$x \sim \mathfrak{Be}\,(\alpha, \beta)$$

ان عدد الارقام العشوائية المستعملة لتوليد قيمة واحدة من x هي **k1+k2+6**

٢-٢-٢ الطريقة(2-5) طريقة Jöhnk [8 ,10 , 11 ,12]

في بعض الاحيان تكون α أو β أو كلاهما عددا غير صحيح وبذلك تكون الطرائق السابقة لتوليد متغيرات على وفق توزيع بيتا غير قابلة للتطبيق .

لذلك سنحاول ان نناقش طريقة تستخدم لاي قيمة لـ α, β .

نفرض ان

$y = U_1^{1/\alpha}$ $\qquad$ $z = U_2^{1/\beta}$

فاذا كان $y+z \le 1$ فان $x = y/(y+z)$ يتوزع حسب توزيع بيتا

$x \sim \mathfrak{Be}\,(\alpha , \beta)$

<u>البرهان</u>

نلاحظ ان

$$f_y(y) = \begin{cases} \alpha\, y^{\alpha-1} & 0 \le y \le 1 \\ 0 & elsewhere \end{cases}$$

$$f_z(z) = \begin{cases} \beta z^{\beta-1} & 0 \le z \le 1 \\ 0 & elsewhere \end{cases}$$

$$f_{y,z}(y,z) = \begin{cases} \alpha\, \beta y^{\alpha-1} z^{\beta-1} & 0 \le y, z \le 1 \\ 0 & elsewhere \end{cases}$$

نفرض ان

$$w = y + z \text{ و } x = y/(y+z)$$

بذلك فان

$$f_{x,w}(x,w) = \begin{cases} \alpha\beta x^{\alpha-1}(1-x)^{b-1}w^{\alpha-\beta-1} & 0 \le x \le 1, 0 \le w \le 2 \\ 0 & elsewhere \end{cases}$$

$$f_{x,w}(x,\ \ 0 \le w \le 1) = \int_0^1 f_{x,w}\ \ (x,w)dw$$

$$= \begin{cases} \frac{\propto\beta}{\alpha+\beta}\ \ x^{\alpha-1}\ (1\ -x\ )^{\beta\neg1} & 0 \le x \le 1 \\ 0 & esewhere \end{cases}$$

$$Prob(0 \le w \le 1)\ = \int_0^1 f_{x,w}\ \ (x, 0 \le w \le 1)\ \ dx$$

$$= \frac{\alpha\beta}{\alpha+\beta}\ \frac{\Gamma\alpha\Gamma\beta}{\Gamma(\alpha+\beta)}$$

وكذالك

$$fx\ (x|0 \le w\ \le 1) = \frac{f_{x,w}\ (x, 0 \le w \le 1)}{Prob(0 \le w \le 1)}$$

$$= \begin{cases} \frac{\Gamma(\alpha+\beta)}{\Gamma\alpha\Gamma\beta} x^{\alpha-1}\ (1-x)^{\beta-1} & 0 \le x \le 1 \\ 0 & elsewhere \end{cases}$$

بناء على ما تقدم يمكن اعتماد الخوارزمية الاتية :

الخوارزمية (2-5) Jöhnk's Algorithm

الخطوة الاولى: ولد متغيرين عشوائيين مستقلين U_1 و U_2 يتوزعان حسب التوزيع المنتظم U(0,1) .

الخطوة الثانية : اجعل $v_1 = U_1^{(1/\alpha)}$ و $v_2 = U_2^{(1/\beta)}$.

الخطوة الثالثة : أجعل $w = v_1 + v_2$

ا لخطوة الرابعة : اذا كان $w \le 1$ اجعل $x = \frac{v_1}{w}$ عدا ذلك ارجع الى الخطوة الاولى .

ولكون عدد محاولات النجاح تتبع التوزيع الهندسي .فان العدد المتوقع لـ U_i لتوليد قيمة واحدة من x الذي بتوزع على وفق توزيع بيتا هو $\frac{2(\alpha+\beta)\Gamma(\alpha+\beta)}{\alpha\beta\Gamma\alpha\Gamma\beta}$

على سبيل المثال اذا كان $\alpha = 2, \beta = 7.6$ فان الارقام العشوائية المتوقعة هي 25 وان العدد المتوقع عندما $\alpha = 4.5 , \beta = 7.5$ هو 110 .

يتبين ان عدد الارقام العشوائية كبير جدا عندما تكون المعلمات كبيرة.ولذلك أقترحنا طريقة أخرى نهدف منها أستعمال أقل عدد من الارقام العشوائية وسنحاول عرضها في ادناه .

٣-٢-٢ الطريقة المقترحة (2-6)

المقترح الثاني في هذا البحث يتمثل بتعديل الخوارزمية (٥-٢).
تهدف الطريقة الى تقليل عدد الارقام العشوائية الستخدمة لتوليد قيمة واحدة من x الى رقم عشوائي واحد بدلامن $\frac{2(\alpha+\beta)\Gamma(\alpha+\beta)}{\alpha\beta\Gamma\alpha\Gamma\beta}$ من الارقام العشوائ ة

اذا كان لدينا $x_1, x_2, x_3 \dots \dots, x_n$ **من قيم** $x \sim \mathfrak{Be}(\alpha, \beta)$
وان الوسط الحسابي لها هو $\mu = \frac{\alpha}{\alpha+\beta}$ و $0 < \mu < 1$

فاذا كان $U \le \mu$ يتوزع على وفق توزيع $U(0,\mu)$ واذا كان $U \ge \mu$ فان $U(\mu,1)$
مما تقدم يمكن صياغة الخوارزمية الاتية:-

الخوارزمية المقترحة (2-6).

الخطوة الاولى : أجعل $\mu = \frac{\alpha}{\alpha+\beta}$ و $v = \frac{1}{\mu}$ و $w = \frac{1}{(1-\mu)}$.

الخطوة الثانية : اجعل i=1 ولد متغير عشوائي U_i=U (0,1) .

الخطوة الثالثة : اذا كان $U_i \le \mu$ أجعل $U_{i+1} = vU_i$
عدا ذلك أجعل $U_{i+1} = w(U_i - \mu)$.

الخطوة الرابعة : أجعل أجعل $Z = U_i^{\frac{1}{\alpha}}$ و $Y = U_{i+1}^{\frac{1}{\beta}}$.

الخطوة الخامسة : أذا كان $Z + Y \leq 1$ أجعل $X = \frac{Y}{Y+Z}$
عدا ذلك أجعل i=i+1 ثم أرجع الى الخطوة الثالثة .
تم اختبار الارقام العشوائية الناتجة من تطبيق الخوارزمية اعلاه وذلك باختبار حسن المطابقة وكانت مطابقة للتوزيع المنتظم $U\sim$ U(0,1) .
من الخوارزمية اعلاه نجد اننا استعملنا متغير عشوائي واحد $U\sim$ U(0,1) في الخطوة الثانية

بدلا من $\frac{2(\alpha+\beta)\,\Gamma(\alpha+\beta)}{\alpha\beta\Gamma\,\alpha\Gamma\,\beta}$ من الارقام العشوائية لتوليد متغير عشوائي $x \sim \mathfrak{B}\,(\alpha, \beta)$

ثانيا- الجانب التجريبي

في هذا الجانب سنحاول مقارنة طرائق توليد كلا من توزيعي كاما وبيتا وحسب المعطيات المختلفة للوقوف على اكثر الطرائق كفاءة .
الشائع ان المقارنة بين طرائق توليد المتغيرات العشوائية تعتمد حاسوبيا على أي الطرائق تستحدم أقل عدد من الارقام العشوائية فضلا عن الوقت المستغرق لتوليد قيمة واحدة من المتغير العشوائي x بغض النظر عن كفائتها الاحصائية .
تم في هذا البحث أعتماد متوسط مربعات الخطأ mse لمتوسط وتباين العينات التي تم تولدها وللاحجام المختلفة للعينات لاول مرة كموشر أحصائي للمقارنة فضلا عن عدد الارقام العشوائية المستعملة وذلك من خلال عدد من تجارب المحاكاة . التي كتبت برامجها بلغة البرمجة MATLAB من قبل الباحث كما في الملحق رقم (١).

١- التجربة الاولى

كما بينا في الجانب النظري في حالة كون معلمة الشكل α عدد موجب وصغير أي $\alpha \leq c$ نستعمل الخوارزمية (١-١) لتوليد بيانات كاما عدا ذلك نستعمل خوارزميات أخرى تتطلب عدد أقل من الارقام العشوائية، مثل والخوارزمية رقم (2-1) ولكن لم نعثر على مصدر يحدد قيمة c .
تهدف هذه التجربة الى تحديد الخوارزمية الافضل لتوليد بيانات تسلك على وفق توزيع كاما على اساس قيمة معلمة الشكل α .عندما تكون عدد صحيح موجب وتحديد قيمة c .
في هذه التجربة تم حساب عدد الارقام العشوائية التي يتطلبها توليد قيمة واحدة من المتغير العشوائي x حيث $x \sim \text{ga}\,(\alpha, 1)$
لكلا من الخوارزمية رقم (١-١) و الخوارزمية رقم (2-1) على ضوء قيمة معلمة الشكل α وحساب mse لمقدرات
كلا من الوسط الحسابي والتباين لعينات التوزيع .
تم اختيار احجام العينات (١٠و ٢٥و٥٠و١٠٠) و$\alpha = 1,2,3 \ldots \ldots 11$.
تم تكرارالتجربة لـ ١٠٠٠0 مرة و كانت النتائج كما في الجدول رقم (١) .

الجدول (١) عدد الارقام العشوائية ومتوسط مربعات الخطأ للوسط والتباين للخوارزميات (1-1) و(2-1)

n	α	عدد الارقام العشوائية لكل قيمة		MSE للوسط		MSE للتباين	
		الخوارزمية (1-1)	الخوارزمية رقم (1-2)	الخوارزمية (1'1)	الخوارزمية (1- ٢)	الخوارزمية (1-1)	الخوارزمية (1- 2
10	٢	2	6.2559	0.1906	0.0342	1.7996	0.2424
	٣	3	6.1950	0.2777	0.0385	3.2670	0.4239
	٤	4	6.2660	0.3802	0.0469	5.5338	0.6007
	٥	5	6.2982	0.5414	0.0513	7.6387	0.8269
	٦	6	6.1350	0.5889	0.0605	11.0293	1.1254
	٧	7	6.1841	0.7362	0.0725	14.4808	1.5732
	٨	8	6.2592	0.7247	0.0798	20.1798	1.8431
	٩	9	6.1688	0.08836	0.0798	20.1798	1.8431
	١٠	10	6.2101	0.9686	0.1024	29.5230	2.6410
	١١	11	6.1409	1.1224	0.1202	35.1871	3.3478
25	2	2	6.2559	0.0655	0.0313	0.6634	0.2235
	3	3	6.1950	0.1027	0.0373	1.7900	0.4085
	4	4	6.2660	0.1385	0.0426	1.9423	0.6139
	5	5	6.2982	0.1623	0.0498	2.7481	0.8121
	6	6	6.1350	0.1976	0.0576	3.9388	1.1192
	7	7	6.1841	0.2395	0.0684	5.2696	1.4365
	8	8	6.2592	0.2721	0.0861	6.3023	1.6835
	9	9	6.1688	0.2953	0.0870	8.1361	2.1890
	10	10	6.2101	0.31[illegible]3	0.0958	9.5450	2.9598
	11	11	6.1409	0.3669	0.1128	10.3060	3.2779
50	2	2	6.2559	0.0367	0.0332	0.3727	0.2511
	3	3	6.1950	0.0599	0.0370	0.7408	0.4250
	4	4	6.2660	0.0794	0.0433	1.1610	0.6073
	5	5	6.2982	0.0978	0.0568	1.1662	0.8915
	6	6	6.1350	0.1312	0.0653	2.0662	1.0562
	7	7	6.1841	0.1358	0.0721	2.7779	1.4625
	8	8	6.2592	0.1557	0.0891	3.4768	1.9397
	9	9	6.1688	0.1764	0.0873	4.2698	2.0750
	10	10	6.2101	0.2048	0.1074	5.1284	2.9120
	11	11	6.1409	0.2240	0.1124	6.5918	3.01880.
100	2	2	6.2559	0.0189	0.0325	0.1945	0.2593
	3	3	6.1950	0.0326	0.0380	0.3513	0.4268
	4	4	6.2660	0.0397	0.0444	0.5568	0.5902
	5	5	6.2982	0.0457	0.0506	0.8290	0.7886
	6	6	6.1350	0.0631	0.0569	1.1865	1.1326
	7	7	6.1841	0.0631	0.0696	1.5435	1.4568
	8	8	6.2592	0.0768	0.0844	1.7978	1.7809
	9	9	6.1688	0.0841	0.0925	2.0992	2.2177
	10	10	6.2101	0.0903	0.1032	2.6032	2.5990
	11	11	6.1409	0.1080	0.1177	3.1287	2.7348

من النتائج في الجدول رقم(١) تبين ما ياتي :-

١- ان عدد الارقام العشوائية لتوليد قيمة واحدة من X مساوية الى قيمة α في الخوارزمية رقم (١-1) وتقريبا (٦) ارقام عشوائية لكل قية في الخوارزمية رقم(2-1)

بذلك ***نجد ان الخوارزمية الاولى هي الافضل حاسوبيا عندما تكون $\alpha \le 5$ والخوارزمية الثانية هي الافضل عندما تكون $\alpha > 5$ وهذا تم تحديده من قبلنا من خلال تجربة المحاكاة الاولى وبذلك اجبنا على التساؤل باعتبار ان $\alpha \le 5$ تكون صغيرة وعدا ذلك تكون كبيرة أي ان C=5 ونكون قد حققنا أحد أهداف هذا البحث .***

٢- من مقارنة قيم mse لمقدرات المتوسط نجد ان الخوارزمية رقم(٢-1) هي الافضل بالنسبة لاحجام العينات (١٠و٢٥و٥٠) والخوارزمية (١-١) هي الآفضل عندما n=100 .

٣- من مقارنة قيم mse لمقدرات التباين نجد ان الخوارزمية رقم(٢-1) هي الافضل للعينات بحجم(١٠ و٢٥ و٥٠) اما بحجم(١٠٠) فأن الخوارزمية (١-١) هي الافضل عندما تكون $\alpha \le 5$
والخوارزمية رقم (2-1) هي الافضل عندما $\alpha \ge 5$

بذلك يمكن ان نخلص الى القول ان الخوارزمية رقم(2-1) هي الافضل من الخوارزمية رقم (1-1) لجميع قيم α للعينات الصغيرة والمتوسطة من الناحية الاحصائية أما الخوارزمية رقم(1-1) تكون الافضل في حالة $\alpha \le 5$ *للعينات الكبيرة ولتقديرات التباين والمتوسط .*

٢—التجربة الثانية(٢)

وضفت هذه التجربة لتحديد أي الطرائق أكفأ لتوليد متغيرات عشوائية تسلك على وفق توزيع كاما عندما تكون معلمة الشكل α عدد غير صحيح أي بمقارنة الخوارزمية (3-1) مع الخوارزمية (4-1).
تم تنفيذ التجربة لتوليد بيانات توزيع كاما لاحجام العينات (١٠ و٢٥ و٥٠ و١٠٠) .
وباختيار $\alpha = 0.1, 0.2, 0.3, \dots \dots \dots 0.9$.
وتم تكرار التجربة (١٠٠٠٠) مرة.

وتم حساب mse لتقديرات المتوسط والتباين لعينات التجربة فكانت النتائج كما في الجدول رقم(٢)

الجدول (٢) نتائج متوسط مربعات الخطأ للوسط والتباين للخوارزميات (1-3) و (1-4)

n	α	MSE للوسط		MSE للتباين	
		الخوارزمية رقم (1-3)	الخوارزمية رقم (1-4)	الخوارزمية رقم (1-3)	الخوارزمية رقم (1-4)
10	0.1	0.0123	0.0010	0.1118	0.0061
	0.3	0.0266	0.0031	0.1902	0.0209
	0.4	0.0432	0.0041	0.3135	0.0258
	0.5	0.0498	0.0050	0.4633	0.0355
	0.6	0.0570	0.0060	0.3516	0.0478
	0.7	0.0683	0.0071	0.4464	0.0472
	0.8	0.0718	0.0084	0.5095	0.0610
	0.9	0.0927	0.0098	0.7487	0.0688
25	0.1	0.0040	0.0010	0.0291	0.0053
	0.2	0.0076	0.0020	0.0457	0.0128
	0.3	0.0113	0.0032	0.0715	0.0200
	0.4	0.0155	0.0040	0.1066	0.0289
	0.5	0.0205	0.0051	0.1470	0.0334
	0.6	0.0224	0.0059	0.1380	0.0489
	0.7	0.0283	0.0070	0.1979	0.0496
	0.8	0.0305	0.0077	0.2019	0.0541
	0.9	0.0346	0.0084	0.2951	0.0668
50	0.1	0.0023	0.0010	0.01235	0.0075
	0.2	0.0038	0.0019	0.0269	0.0114
	0.3	0.0057	0.0031	0.0428	0.0204
	0.4	0.0083	0.0040	0.0506	0.0280
	0.5	0.0099	0.0051	0.0727	0.0338
	0.6	0.0123	0.0061	0.0900	0.0453
	0.7	0.0145	0.0067	0.1121	0.0492
	0.8	0.0161	0.0077	0.1039	0.0556
	0.9	0.0183	0.0087	0.1430	0.0708
١٠٠	0.1	0.0011	0.0011	0.0068	0.0073
	0.2	0.0019	0.0020	0.0118	0.0127
	0.3	0.0030	0.0030	0.0191	0.0193
	0.4	0.0041	0.0041	0.0302	0.0270
	0.5	0.0047	0.0052	0.0346	0.0354
	0.6	0.0060	0.0060	0.0436	0.0464
	0.7	0.0070	0.0073	0.0523	0.0522
	0.8	0.0080	0.0079	0.0592	0.0594
	0.9	0.0094	0.0096	0.0737	0.0781

١-من التجربة أعلاه تبين ان عدد الارقام العشوائية المطلوبة لتوليد قيمة واحدة من $x \sim ga(\alpha, 1)$ عندما تكون $0 < \alpha < 1,$ هو (٣)ارقام عشوائية للخوارزمية (1-3) كما هو معروف[8] ورقمين عشوائين للخوارزمية (1-4)، وهذا تم تحديده لاول مرة من خلال التجربة أعلاه . وبشكل عام اذا كان $\alpha > 1$ وعدد غير صحيح فان عدد الارقام العشوائية لتوليد قيمة واحدة من x هي $k+3$ لخوارزمية (1-3) و $k+2$ للخوارزمية (1-4) علما ان k تمثل أكبر عدد صحيح من α . وبذلك تكون الخوارزمية (1-4) هي الافضل حاسوبيا لانها تتطلب اقل عدد من الارقام العشوائية .

٢- الجدول (٢) وبمقارنة قيم mse لمقدرات للمتوسط والتباين نجد أن الخوارزمية (1-4) أفضل من الخوارزمية(٣-١) لجميع قيم $0 < \alpha < 1$ ولاحجام العينات (٥٠,٢٥,١٠) وتتمتع بافضلية نسبية عندما n=10

٣- التجربة الثالثة (3)

صممت هذه التجربة لمقارنة طرائق توليد بيانات تسلك على وفق توزيع بيتا عندما تكون معلمات الشكل للتوزيع أعداد صحيحة أي بمقارنة الخوارزمية (1-2) و(٢-٢) والخوارزمية المقترحة (2-3) .
تم أختيار قيم مختلفة للمعالم ولآحجام مختلفة من العينات.
كررت التجربة ١٠٠٠٠ مرة
فكانت النتائج كما في الجدول

الجدول (٣) الارقام العشوائية المستعملة لتوليد كل قيمة ونتائج متوسط مربعات الخطأ للمتوسط والتباين للخوارزميات ((2-1) و(2-3)

المعلمات		حجم العينة	الارقا(م العشوائية			MSE للوسط			MSE للتباين		
α	β	n	(2-1)	2-(2)	(2-3)	الخوازمية (2-1)	الخوارزمية (٢-٢)	الخوارزمية ألمقترحة (2-3)	الخوازمية (2-1)	الخوازمية (2-3)	الخوارزمية ألمقترحة (2-3)
٤	3	10	٧	٦	٧	0.0030	0.0032	0.0030	0.0002	0.0002	0.0002
		25				0.0012	0.0013	0.0012	0.0001	0.0001	0.0001
		50				0.0006	0.0006	0.0006	0.0000	0.0000	0.0000
		75				0.0004	0.0004	0.0004	0.0000	0.0000	0.0000
		100				0.0003	0.0003	0.0003	0.0000	0.0000	0.0000
3	8	10	11	10	9	0.0851	0.0987	0.0633	0.0006	0.0002	0.0006
		25				0.0578	0.0936	0.0425	0.0010	0.0002	0.0009
		50				0.0261	0.0183	0.0129	0.0012	0.0002	0.0008
		75				0.0071	0.0893	0.0039	0.0006	0.0002	0.0002
		100				0.0002	0.0890	0.0002	0.0000	0.0002	0,0000
8	4	10	12	11	10	0.1150	0.0137	0.1191`	0.0004	0.0002	0.0002
		25				0.0870	0.0124	0.0828	0.0011	0.0002	0.0007
		50				0.0409	0.0101	0.0404	0.0019	0.0003	0.0015
		75				0.0093	0.0084	0.0115	0.0015	0.0002	0.0011

		100				0.0002	0.0093	0.0002		0.0000	0.00002	0.0000
12	10	10	22	21	12	0.0152	0.0003	0.0099		0.0000	0.0005	0.0000
		25				0.0110	0.0009	0.0075		0.0003	0.0004	0.0002
		50				0.0055	0.0011	0.0038		0.0000	0.0004	0.0000
		75				0.0015	0.0006	0.0009		0.0003	0.0005	0.0002
		100				0.0001	0.0009	0.001		0.0000	0.0004	0.00000

من نتائج التجربة في الجدول(٣) اعلاه نجد ما ياتي:-

١- **ان الخوارزمية (٢-٢) تتطلب عدد أقل من الارقام العشوائية عندما** $\alpha \,\&\beta \le 5$ **.**

٢- **أن الخوارزمية المقترحة (٣-٢) تتطلب عدد أقل من الارقام العشوائية من الخوارزميات (١-٢) و (٢-٢) عندما تكون** $\alpha \; or \; \beta \le 5$ **. وفي حالة** $\alpha \,\&\beta > 5$ **وبذلك تكون هي الافضل حاسوبيا .**

٣- **بمقارنة قيم mse لتقديرات المتوسط والتباين نجد ان انها متساوية للخوارزميات (١-٢)و (٢-٢) و(٣-٢) عندما** $\alpha \,\&\beta \le 5$ **بالنسبة للتباين وهناك اختلاف في قيم mse للمتوسط للخوارزمية (٢-٢) وهذا طبيعي لان الخوارزمية المقترحة هي نفسها (١-٢) .**

٤- **من مقارنة mse لتقديرات المتوسط عندما تكون** $\alpha \; or \; \beta \le 5$ **.وفي حالة** $\alpha \,\&\beta > 5$ **نجد ان هناك أفضلية للخوارزمية (٢-٢) على الخوارزميات البقية ولكن الخوارزمة (٣-٢) المقترحة كتعديل للخوارزمية كانت أفضل من الحوارزمية المعدلة (١-٢).**

٥- **من مقارنة mse لتقديرات التباين نجد ان الطريقة المفترحة (٣-٢) هي الافضل نسبيا من الخوارزمية (٢-٢) وافضل بالمطلق من الخوارزمية (١-٢) .**

مما تقدم نخلص الى القول أن الخوارزمية رقم (2-1) يمكن أن تستعمل لتوليد قيم x . عندما $\alpha \,\&\beta \le 5$ **عدا ذلك تستعمل الخوارزمية المقترحة رقم (2-3)**

٤-التجربة الرابعة (٤)

في هذه التجربة تم مقارنة طرائق توليد بيانات توزيع بيتا عندما تكون واحدة من معلمات الشكل أو كلاهما عدد غير صحيح وتم أختيار أعداد مختلفة من المعلمات ولاحجام مختلفة من العينات كررت التجرجة ١٠٠٠٠ مرة.

الجدول (٤) عدد الارقام العشوائية المستعملة للخوارزمية (٥-٢) ونتائج mse للمتوسط والتباين للعينات المستخرجة للخوارزميات في التجربة.

n	المعلمات		R.No	MSE للوسط			MSE للتباين		
	α	β		الخوارزمية (2-4)	الخوارزمية (2-5)	الخوارزميةالمقترحة (2-6)	الخوارزمية (2-4)	الخوارزمية (2-5)	الحوارزمية المقترحة (2-6)
10	4.3	3,2	27	0.0028	0.0000	0.0015	0.0001	0.0008	0.0007
	4.5	5.6	64	0.0023	0.0000	0.0011	0.0001	0.0005	0.0005
	7.5	8.3	473	0.0015	0.0000	0.0004	0.0001	0.0002	0.0002
	9.5	9.6	1503	0.0013	0.0000	0.0003	0.0000	0.0002	0.0002
	10.2	15.6	11045	0.0009	0.0005	0.0000	0.0000	0.0001	0.0001
25	4.3	3.2	26	0.0011	0.0007	0.0013	0.0001	0.0008	0.0007
	4.5	5.6	68	0.0009	0.0003	0.0011	0.0000	0.0005	0.0005
	7.5	8.3	468	0.0006	0.0001	0.0004	0.0000	0.0002	0.0002
	9.5	9.6	1542	0.0005	0.0001	0.0003	0.0000	0.0002	0.0002
	10.2	15.6	12005	0.0003	0.0002	0.0000	0.0000	0.0001	0.0001
50	4.3	3.2	27	0.0006	0.0011	0.0012	0.0000	0.0008	0.0007
	4.5	5.6	63	0.0004	0.0007	0.0011	0.0000	0.0005	0.0005
	7.5	8.3	450	0.0003	0.0002	0.0004	0.0000	0.0002	0.0002
	9.5	9.6	1504	0.0002	0.0002	0.0003	0.0000	0.0002	0.0002
	10.2	15.6	11240	0.0002	0.0001	0.0000	0.0000	0.0001	0.0001
100	4.3	3.2	26	0.0003	0.0012	0.0012	0.0000	0.0007	0.0007
	4.5	5.6	64	0.0002	0.0009	0.0011	0.0000	0.0005	0.0005
	7.5	8.3	417	0.0002	0.0003	0.0004	0.0000	0.0002	0.0002
	9.5	9.6	1461	0.0001	0.0003	0.0003	0.0000	0.0002	0.0002
	10.2	15.6	12103	0.0001	0.0001	0.0000	0.0000	0.0001	0.0001

من الجدول (٤) اعلاه نلاحظ ما ياتي :-

١- لتوليد قيمة واحدة من x نستعمل k1+k2+6 من الارقام العشوائية في الخوارزمية(2-4)

و $\frac{2(\alpha+\beta)\Gamma(\alpha+\beta)}{\alpha\beta\Gamma\alpha\Gamma\beta}$ في الخوارزمية رقم (2-5) كما مبين في العمود (R.No) أما الخوارزمية المقترحة رقم (2-6) تستعمل رقم عشوائي واحد .
وبذالك ان الطريقة المقترحة هي الاقل عددا من الارقام العشوائية وبذلك تكون هي الافضل حاسوبيا.

٢- من مقارنة mse لتقديرأ لمتوسطات للخوارزميات الثلاثة نجد ان الخوارزمية رقم (2-5) هي الافضل .
٣- من مقارنة mse للتباين نجد ان الخوارزمية رقم (2-4) هي الافضل نسبيا.
٤- ان الخوارزمية المقترحة(٦-٢) هي تعديل للخوارزمية (٥-٢) ومن النتائج نجد ان الخوارزمية المقترحة افضل حاسوبيا لانها تتطلب رقم عشوائي واحد بدلا من $\frac{2(\alpha+\beta)\Gamma(\alpha+\beta)}{\alpha\beta\Gamma\alpha\Gamma\beta}$ من الارقام العشوائية ، فضلا انها افضل عند مقارنة mse للتباين من الخوارزمية المعدلة (٥-٢).

وبذلك نخلص الى القول ان الخوارزمية المقترحة للتعديل (٦-٢)هي الافضل من الخوارزمية المعدلة (٥-٢).

ثالثا - الاستنتاجات

خلال الجانب النظري والجانب التجريبي للبحث خلصنا الى الاستنتاجات الخاصة بطرائق توليد بيانات كل توزيع وكما يأتي:-

أولا توزيع كاما

ا- عندما تكون معلمة الشكل عدد صحيح موجب .
١- ان عدد الارقام العشوائية لتوليد قيمة واحدة من X مساوية الى قيمة α في الخوارزمية رقم (١-1) وتقريبا (٦) ارقام عشوائية لكل قيمة في الخوارزمية رقم(2-1)

بذلك نجد ان الخوارزمية الاولى هي الافضل حاسوبيا عندما تكون $\alpha \leq 5$ والخوارزمية الثانية هي الافضل عندما تكون $\alpha > 5$ وهذا تم تحديده من قبلنا من خلال تجربة المحاكاة الاولى وبذلك اجبنا على التساءل باعتبار ان $\alpha \leq 5$ تكون صغيرة وعدا ذلك تكون كبيرة أي ان $C=5$ ونكون قد حققنا أحد أهداف هذا البحث .

٢- من مقارنة قيم mse لمقدرات المتوسط نجد ان الخوارزمية رقم(٢-1) هي الافضل بالنسبة لاحجام العينات (١٠و٢٥و٥٠و١٠٠).

٣- من مقارنة قيم mse لمقدرات التباين نجد ان الخوارزمية رقم(٢-1) هي الافضل للعينات بحجم(١٠ و٢٥ و٥٠) اما بحجم(١٠٠) فأن الخوارزمية (١-١) هي الافضل عندما تكون $\alpha \leq 5$ والخوارزمية رقم (2-1) هي الافضل عندما $\alpha \geq 5$

بذلك يمكن ان نخلص الى القول ان الخوارزمية رقم(2-1) هي الافضل من الخوارزمية رقم (1-1) لجميع قيم α للعينات الصغيرة والمتوسطة من الناحية الاحصائية أما الخوارزمية رقم(1-1) تكون الافضل في حالة $\alpha \leq 5$ *للعينات الكبيرة ولتقديرات التباين فقط .*

ب – عندما تكون معلمة الشكل عدد غير صحيح .

١- عندما تكون معلمة الشكل α عددا غير صحيحا فان الخوارزمية رقم (4-1) افضل من الخوارزمية رقم (3-1) لانها تتطلب $k+2$ من الارقام العشوائية و(هذا تم تحديده لاول مرة من خلال التجربة رقم (٢)) بينما الخوارزية رقم (3-1) تتطلب $k+3$ من الارقام العشوائية .وبذلك تكون الخوارزمية (٤-١) هي الافضل حاسوبيا.

٢- بمقارنة قيم mse لمقدرات للمتوسط والتباين نجد أن الخوارزمية (4-1) أفضل من الخوارزمية (٣-١) لجميع قيم $0 < \alpha < 1$ ولاحجام العينات (١٠,٢٥,٥٠) وتتمتع بافضلية نسبية عندما n=100.

بذلك نستطيع القول أن الخوارزمية (٤-١) أفضل من الخوارزمية (٣-١)

ثانياً: توزيع بيتا

أ - في حالة كون معلمات الشكل لتوزيع بيتا α, β اعداداً صحيحة فأن:- .

١- الارقام العشوائية للخوارزمية رقم (3-2) أقل من الارقام العشوائية المطلوبة في الخوارزمية رقم (2-1) و(٢-٢) عندما تكون $\alpha \; or \; \beta \leq 5$. وفي حالة $\alpha \,\&\beta > 5$ وهذا يعطي افضلية حاسوبيا للخوارزمية المقترحة (3-2).

٢- ومن مقارنة قيم mse لتقديرات المتوسط نجد ان هناك أفضلية نسبية للخوارزمية (٢-٢) تليها الخوارزمية المقترحة (٣-٢) عندما تكون $\alpha \; or \; \beta \leq 5$.وفي حالة $\alpha \,\&\beta > 5$.

٣- من مقارنة mse لتقديرات للتبا ين نجد ان االخوارزمية المقترحة (٣-٢) لتعديل الخوارزمية (١-٢) هي الافضل نسبيا من الخوارزمية (٢-٢) والافضل بالمطلق من الخوارزمية المعدلة (١-٢) .

مما تقدم نخلص الى القول أن الخوارزمية المقترحة (٣-٢) للتعديل افضل من الخوارزمية المعدلة (١-٢).

ب- في حالة كون معلمات الشكل لتوزيع بيتا α, β اعداداً غير صحيحة فأن:-

١- لتوليد قيمة واحدة من x نستعمل k1+k2+6 من الارقام العشوائية في الخوارزمية (٤-٢) و $\frac{2(\alpha+\beta)\Gamma(\alpha+\beta)}{\alpha\beta\Gamma\alpha\Gamma\beta}$ في الخوارزمية رقم (٥-٢) ،أما الخوارزمية المقترحة رقم (٦-٢) تستعمل رقم عشوائي واحد . وبذلك تكون هي الاقل عدد من الارقام العشوائية وبالتالي هي الافضل حاسوبيا.

٢- من مقارنة mse لتقدير لمتوسط الخوارزميات الثلاثة نجد ان الخوارزمية رقم (٥-٢) هي الافضل .

٣- من مقارنة mse للتباين نجد ان الخوارزمية رقم (٤-٢) هي الافضل تليها الخوارزمية(٦-٢) المقترحة للتعديل ومن ثم الخوارزمية المعدلة (٥-٢)

مما تقدم تبين ان الخوارزمية المقترحة(٦-٢) للتعديل هي الآكفأ من الخوارزمية المعدلة (٥-٢).

References

1- Ahrens, J. H. and Dieter, U. (1982). Generating gamma varieties by a modified rejection technique .*Communications of the ACM*, 25, 47–54. Algorithm GD, p. 53

2- A.C.Atkinson (1977), “An Easily Programmed Algorithm for Generating Gamma Random Variables”. JR.Statist.Soc.A.140,Part 2 , pp 232-234 .www.maphysto.dk 3- Athanasios Papoulis & S. Unnikrishna Pillai (2002),”Probability, Random Variables and astic Processes” Fourth Edition,Mc Graw Hill .

4- Cheng, R. C. H., and G. M. Feast (1979), Some simple gamma variate generators, *lied Statistics* 28, 290–295

5 -David, H. A., Nagaraja, H. N. (2003) Order Statistics (3rd Edition). Wiley, New Jerse pp 458. ISBN 0-471-38926-9

6- . Debasis Kundu & Rameshwar D.Gupta(2006) ‘’A Convenient Way of Generating Gamma Random Variables Using Generalized Exponential Distribution . http://home.iitk.ac.in/~kundu

7- George.S. Fishman (1976) , “Sampling from the Gamma Distribution on Computer”,.Mangement Science Operation Research , Vo. 19 Nu.7

8- George.S. Fishman(1973),”Concepts and Methods in Disscrete Event Digital Simulation” JOHN WILE&SO

9- James E.Gentle (2003). " Random Number Generation and Monte Carlo Methods" Secand Edition , Springe

10- J. Keith Gilless and Jeremy S. Fried, (2000)," Generating beta random rate variables from probabilistic estimates of fireline production times" , Annals of Operations 9520 5–215 205,
www.jeremyfried.net/jfried/.../gilless_fried_2000_annals_of_or.pdf

11- Martin Haugh (2004) , " Generating Random Variables and Stochastic Proceses" ,
Monte Carlo Simulation E4703 www.columbia .
edu/~mh2078/MCSO4/MCS_genrte_rv pdf .

12-Sheldon M.Ross (2007),"Introduction to Probability Models"Ninth Edition,Academic Pricss.

13- S.V.N.Vishwanathan (2011)., "Generating Random Variables " ,
learning.stat.purdue.edu/wiki/_media/.../rvgen.pdf - United States

14- van der Warden,B.L.(2002),"Mathematical Statistics "Springer ,ISBN 978-3-540-04507-6

15 – WIKIPEDA .Generating gamma – distribution random variables . http://en.wikipedia.org/wiki/Wikipedia

Compared Methods of Generating Both Gamma Distribution and Beta Distribution

Abstract

Gamma and Beta Distributions has very important in practice in various areas of statistical and applications reliability and quality control of production. and There are a number of methods to generate data behave on according to these distribution. and These methods bassic primarily on the shape parameters of each distribution and the relationship between these distributions and their relationship with some other probability distributions.

This research aims to determine how most efficient method for generating gamma varieties upon the shape parameter α and generating beta varieties upon the two shape parameters α, β.

The common comparison between the methods of generating random variables computerized rely on the less number of random numbers $U \sim U(0,1)$, as well as the time it takes to generate a single value of a random variable x, regardless of statistical efficiency.

In this paper and for the first time,we used the mean square error MSE of the mean and the varianes of the samples that have been generated for the different sizes of the samples as the statistical comparison.

As well as a proposal of two methods to generate data distributed according to beta distribution using the least number of random numbers and the highest efficiency.

We prove by running a number of simulation experiments, programs written by researcher used MATLAB language.

Key word : **Gamma distribution - Beta distribution - random variables - Generating - Simulation**.