

UNIVERSITÉ DE MONTRÉAL

IFT6561 – SIMULATION: ASPECTS STOCHASTIQUES

---

**Devoir 5**

---

par:

**Eugénie Yockell**

(20071932)

**Cong Liu**

(20161998)

Date de remise: 2 décembre 2020

## Question 1

On sait que  $Cov(X, Y) = E[XY] - E[X]E[Y]$ . On définit une fonction  $g(\boldsymbol{\mu})$  où on veut un intervalle de confiance pour  $g(\boldsymbol{\mu})$  où  $\boldsymbol{\mu} = (\mu_1, \dots, \mu_d)$  et  $g$  est continue mais potentiellement non linéaire.

$$Cov(X, Y) = g(\mu_1, \mu_2, \mu_3) = \mu_3 - \mu_1\mu_2$$

où  $\mu_1 = \frac{1}{n} \sum_{i=1}^n X_i$ ,  $\mu_2 = \frac{1}{n} \sum_{i=1}^n Y_i$  et  $\mu_3 = \frac{1}{n} \sum_{i=1}^n X_i Y_i$ .

En utilisant le théorème de delta, pour  $g : \mathbb{R}^d \rightarrow \mathbb{R}$  continûment différentiable dans un voisinage de  $\boldsymbol{\mu}$ , on sait que  $r(n)(\mathbf{Y}_n - \boldsymbol{\mu}) \Rightarrow \mathbf{Y}$  quand  $n \rightarrow \infty$ , alors

$$r(n)(g(\mathbf{Y}_n) - g(\boldsymbol{\mu})) \Rightarrow (\nabla g(\boldsymbol{\mu}))^\top \mathbf{Y}$$

Étant donné que ce sont des moyennes empiriques, on peut appliquer le corollaire sous le théorème central limite où nous avons

$$\sqrt{n}(\mathbf{Y}_n - \boldsymbol{\mu}) \Rightarrow \mathbf{Y} \sim N(0, \Sigma)$$

Le théorème nous dit que  $\sqrt{n}(g(\mathbf{Y}_n) - g(\boldsymbol{\mu})) \sim N(0, \sigma_g)$ . Où  $\sigma_g = (\nabla g(\boldsymbol{\mu}))^\top \Sigma \nabla g(\boldsymbol{\mu})$ . Et l'intervalle de confiance est  $[g(\mathbf{Y}_n) \pm \Phi^{-1}(1 - \alpha/2)\sigma_g/\sqrt{n}]$ . Il nous faut alors  $\sigma_g$  pour définir l'intervalle de confiance.

Définissons la matrice de covariance  $\Sigma$ ,

$$\Sigma = \begin{pmatrix} Var(X) & Cov(X, Y) & Cov(X, XY) \\ Cov(Y, X) & Var(Y) & Cov(Y, XY) \\ Cov(XY, X) & Cov(XY, Y) & Var(XY) \end{pmatrix} = \begin{pmatrix} \sigma_1^2 & \sigma_{12}^2 & \sigma_{13}^2 \\ \sigma_{12}^2 & \sigma_2^2 & \sigma_{23}^2 \\ \sigma_{13}^2 & \sigma_{23}^2 & \sigma_3^2 \end{pmatrix}$$

Calculons  $\nabla g(\boldsymbol{\mu})$ ,

$$(\nabla g(\boldsymbol{\mu}))^\top = \begin{pmatrix} -\mu_2 & -\mu_1 & 1 \end{pmatrix}$$

Ainsi,

$$\sigma_g = \begin{pmatrix} -\mu_2 & -\mu_1 & 1 \end{pmatrix} \begin{pmatrix} \sigma_1^2 & \sigma_{12}^2 & \sigma_{13}^2 \\ \sigma_{12}^2 & \sigma_2^2 & \sigma_{23}^2 \\ \sigma_{13}^2 & \sigma_{23}^2 & \sigma_3^2 \end{pmatrix} \begin{pmatrix} -\mu_2 \\ -\mu_1 \\ 1 \end{pmatrix}$$

$$\begin{aligned} \sigma_g &= \begin{pmatrix} (-\mu_2\sigma_1^2 - \mu_1\sigma_{12}^2 + \sigma_{13}^2) & (-\mu_2\sigma_{12}^2 - \mu_1\sigma_2^2 + \sigma_{23}^2) & (-\mu_2\sigma_{13}^2 - \mu_1\sigma_{23}^2 + \sigma_3^2) \end{pmatrix} \begin{pmatrix} -\mu_2 \\ -\mu_1 \\ 1 \end{pmatrix} \\ &= \mu_2\sigma_1^2 + \mu_1\mu_2\sigma_{12}^2 - \mu_2\sigma_{13}^2 + \mu_1\mu_2\sigma_{12}^2 + \mu_1\sigma_2^2 - \mu_1\sigma_{23}^2 - \mu_2\sigma_{13}^2 - \mu_1\sigma_{23}^2 + \sigma_3^2 \\ &= \mu_2^2\sigma_1^2 + \mu_1^2\sigma_2^2 + \sigma_3^2 - 2\mu_2\sigma_{13}^2 - 2\mu_1\sigma_{23}^2 + 2\mu_1\mu_2\sigma_{12}^2 \end{aligned}$$

Le candidat évident est de remplacer par les estimateurs de variance et covariance,

$$\hat{\sigma}_g = \mu_2^2 \hat{\sigma}_1^2 + \mu_1^2 \hat{\sigma}_2^2 + \hat{\sigma}_3^2 - 2\mu_2 \hat{\sigma}_{13}^2 - 2\mu_1 \hat{\sigma}_{23}^2 + 2\mu_1 \mu_2 \hat{\sigma}_{12}^2$$

où,

$$\begin{aligned}\hat{\sigma}_{12} &= \frac{1}{n-1} \sum_i^n (X_i - \bar{X}_n)(Y_i - \bar{Y}_n) \\ \hat{\sigma}_{13} &= \frac{1}{n-1} \sum_i^n (X_i - \bar{X}_n)(X_i Y_i - \bar{X} \bar{Y}_n) \\ \hat{\sigma}_{23} &= \frac{1}{n-1} \sum_i^n (Y_i - \bar{Y}_n)(X_i Y_i - \bar{X} \bar{Y}_n) \\ \hat{\sigma}_1 &= \frac{1}{n-1} \sum_i^n (X_i - \bar{X}_n)^2 \\ \hat{\sigma}_2 &= \frac{1}{n-1} \sum_i^n (Y_i - \bar{Y}_n)^2 \\ \hat{\sigma}_3 &= \frac{1}{n-1} \sum_i^n (X_i Y_i - \bar{X} \bar{Y}_n)^2\end{aligned}$$

Nous savons  $\sigma_g$  est fortement consistant, puisque les estimateurs de moyennes et variances sont consistants. Puisque  $\sigma_g$  est fortement consistant nous avons le TLC,

$$\frac{\sqrt{n}(g(\mathbf{Y}_n) - g(\boldsymbol{\mu}))}{\sigma_g} \Rightarrow N(0, 1) \quad \text{pour } n \rightarrow \infty$$

Et l'intervalle de confiance à 95% est,

$$\begin{aligned}& \left[ g(\mathbf{Y}_n) \pm \Phi^{-1}(1 - \alpha/2) \sigma_g / \sqrt{n} \right] \\ & \left[ g(\mathbf{Y}_n) \pm 1.96 \frac{(\mu_2^2 \hat{\sigma}_1^2 + \mu_1^2 \hat{\sigma}_2^2 + \hat{\sigma}_3^2 - 2\mu_2 \hat{\sigma}_{13}^2 - 2\mu_1 \hat{\sigma}_{23}^2 + 2\mu_1 \mu_2 \hat{\sigma}_{12}^2)}{\sqrt{n}} \right]\end{aligned}$$

## Question 2

a)

Dans cette question, on fait  $m = 1000$  simulations. À chaque simulation il y a  $n = 100$  observations de variables aléatoires exponentielles. Au seuil du risque  $\alpha = 0.05$ , le résultat obtenu est à la table 1 et la figure 1.

	Valeur
Esperance coverage probability $\hat{\rho}$	0.665
Confidence interval for $\mathbf{p}$	(0.582, 0.766)
Average width of confidence probability $\mathbf{W}$	0.581

Table 1: Statistiques des estimation de IC en utilisant  $\chi^2$  distribution pour  $n = 100$

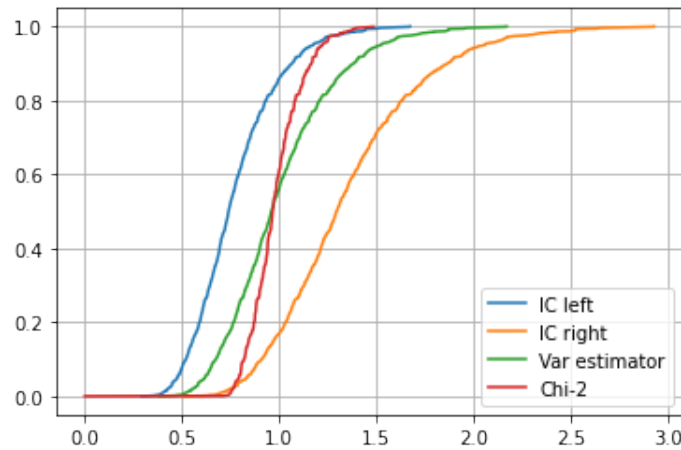


Figure 1: distribution empirique avec  $\chi^2$  et  $m = 1000, n = 100$

b)

Dans la question b , on a fait  $m = 1000$  simulations, à chaque simulation il y a  $n = 100$  observations de variables aléatoire exponentielle. Au seuil du risque  $\alpha = 0.05$ , le résultat obtenu est à la table 2 et à la figure 2

	Valeur
Esperance coverage probability $\hat{\rho}$	0.661
Confidence interval for $\mathbf{p}$	(0.633, 0.689)
Average width of confidence probability $\mathbf{W}$	0.176

Table 2: Statistiques des estimation de IC en utilisant  $\chi^2$  distribution pour  $n=1000$

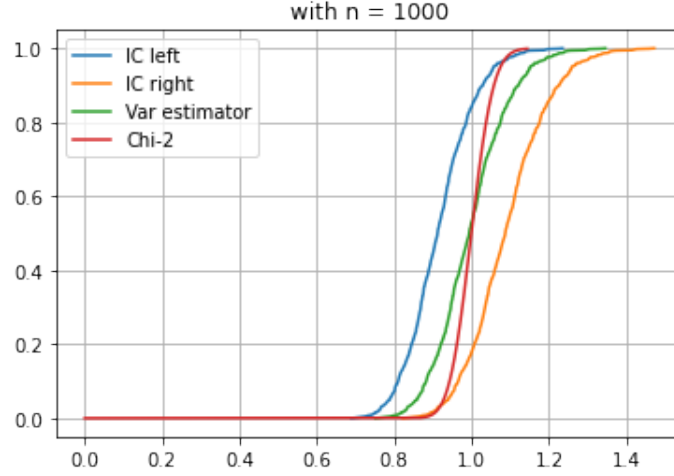


Figure 2: Distribution empirique avec  $\chi^2$   $m = 1000$  et  $n = 1000$

### Discussion

Selon des statistiques, on trouve que lorsqu'on augmente la taille d'échantillon de  $n = 100$  à  $n = 1000$ , la largeur de l'intervalle de confiance diminue, la probabilité de couvrage de  $p$  devient plus petite, la performance devient plus faible. En effet, en calculant l'intervalle de confiance de la variance, on a supposé au préalable que  $\frac{(n-1)S_n^2}{\sigma^2}$  suit une distribution  $\chi^2$ . On remarque sur la figure 1 et 2 que ce n'est pas la vraie distribution. Et donc en grandissant la taille d'échantillon, il y a plus de chance que cette intervalle de confiance ne contient pas la vraie valeur 1.

c)

Dans cette partie, on utilise la méthode bootstrap avec  $m = 1000$  exemples. Au seuil du risque  $\alpha = 0.05$ , le résultat obtenu est à la table 3 et à la figure 3.

	Valeur
Esperance coverage probability $\hat{\rho}$	0.838
Confidance interval for p	(0.797, 0.845)
Average width of confidance probability W	0.943

Table 3: Statistiques des estimation de IC en utilisant basic nonparametric bootstrap

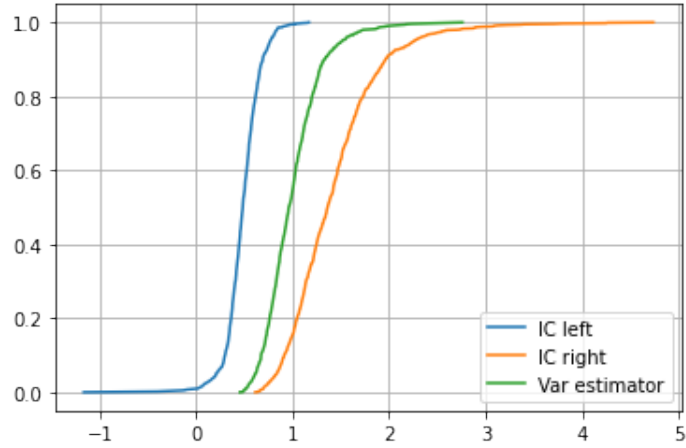


Figure 3: distribution empirique avec basic non-para bootstrap  $m = 1000, n = 100$

### Discussion

En comparant avec les résultats obtenues en **a)** et **b)**, on constate qu'avec la méthode bootstrap non-paramétrique, la probabilité de couverture a beaucoup augmenté, la largeur de l'IC devient plus grande, la performance est significativement améliorée. En effet, en utilisant le méthode Bootstrap non-paramétrique, au lieu de supposer une distribution, on rééchantillonne plusieurs fois pour avoir plus d'information de l'échantillon initiale.

### Question 3

Nous avons,

$$MSE = \text{Biais}^2 + \text{Variance}$$

$$MSE \leq \left(\frac{\kappa_0 \beta^{n_0}}{n - n_0}\right)^2 + \left(\frac{\kappa_1}{n - n_0}\right)$$

Une heuristique simple pour mesurer la valeur optimal de *warm-up* qui minimise le MSE est la solution graphique proposée par Welch (1983). Le but est de trouver une valeur de  $n_0$  tel que le processus du coût moyen semble s'aplatir en fonction du temps. Comme les réalisations individuelles du processus peuvent avoir une grande variance, on génère  $k$  répétitions de réalisation du coût. Ces réalisations sont générées à partir de  $m$  étapes, soit le  $n$ . On doit avoir un  $m$  très grand, beaucoup plus grand que la valeur anticipée de  $n_0$ . Les étapes sont comme suit,

1. On fait  $k$  simulations de longueur  $m$   
Il est préférable de prendre  $k = 5$  ou  $k = 10$   
On définit  $C_{ij}$  comme l'observation  $j$  à la répétition  $i$
2. Pose  $\bar{C}_j = \sum_{i=1}^k C_{ij}/k$
3. On applique un lissage des hautes fréquences par une moyenne mobile de longueur  $w$   
 $\bar{C}_j(w) = \frac{1}{2w+1} \sum_{s=-w}^w \bar{C}_{j+s}$
4. On étudie ensuite le graphique de  $\bar{C}_j(w)$  en fonction de  $j$  où  $n_0$  est le  $j$   
Il est conseillé de tester différentes valeurs de longueur  $w$  et prendre la plus petite valeurs de  $w$  qui correspond à un graphique raisonnablement lisse  
Si aucun résultat n'est satisfaisant, augmenter  $k$  et  $m$
5. On choisit la valeur de  $n_0$  comme le  $j$  au moment où le processus semble avoir convergé graphiquement

Avec cette heuristique, il est possible de déterminer une valeur optimal de  $n_0$  qui minimise le MSE. Une des difficultés est qu'il faut que  $k$  soit assez grand si les réalisations de coût ont une grande variance entre eux. [1]

Dans le cas où  $n \rightarrow \infty$ , comme  $\kappa_0$  et  $\kappa_1$  sont des constantes, il existe une borne supérieure pour chaque terme tel  $\max(\kappa_0^2, \kappa_1)$  et on peut poser,

$$MSE \in \mathcal{O}\left(\frac{\beta^{2n_0}}{(n - n_0)^2} + \frac{1}{n - n_0}\right)$$

Lorsque  $n \rightarrow \infty$  le terme de  $\beta^{2n_0}$  devient négligeable par rapport à son dénominateur en  $n^2$  qui fait tendre rapidement le biais vers 0. De plus, lorsque  $n \gg n_0$ , le terme pour le biais diminue plus rapidement que le terme pour la variance. Nous avons alors,

$$MSE \in \mathcal{O}\left(\frac{1}{n - n_0}\right)$$

Ainsi, quand  $n \rightarrow \infty$ , le  $n_0$  optimal asymptotique est évidemment  $n_0 = 0$ . Ce résultat n'est pas étonnant, par exemple, Blomqvist (1970) a montré que pour le processus  $M/M/1$  et

pour d'autre processus de file d'attentes, avec un  $n$  assez grand, le MSE est minimal lorsque  $n_0 = 0$ . De plus, le *warm-up* permet de particulièrement diminuer le MSE lorsque le biais de l'initialisation est grand et que l'auto-corrélation est forte. Dans ce cas, on retrouve encore que la valeur de  $n_0$  qui minimise le MSE diminue à la même vitesse que  $n$  augmente. On conclue alors que pour de très grande valeur de  $n$ , la *warm-up* n'est pas conseillé lorsque la mesure de performance est le MSE. [1]



## Question 4

Montrons que la somme de  $k$  exponentielle de moyenne  $1/\gamma$  est équivalente à la distribution de Erlang. Étudions les MGF,

$$M_{exp}(t) = \frac{\gamma}{\gamma - t}$$

$$M_{erlang}(t) = \left(1 - \frac{t}{\gamma}\right)^{-k}$$

On sait que pour une variable  $S = \sum_{i=1}^k X_i$  où  $X_i$  sont variables exponentielles iid, la MGF est  $M_S(t) = M_{X_1}(t)M_{X_2}(t)\dots M_{X_k}(t)$ . Nous avons alors,

$$\begin{aligned} M_{\sum exp} &= M_{exp}(t)^k \\ &= \left(\frac{\gamma}{\gamma - t}\right)^k \\ &= \left(\frac{\gamma}{\gamma} \frac{1}{1 - \frac{t}{\gamma}}\right)^k \\ &= \left(\frac{1}{1 - \frac{t}{\gamma}}\right)^k = \left(1 - \frac{t}{\gamma}\right)^{-k} = M_{erlang}(t) \end{aligned}$$

Ainsi, on voit que les deux méthodes sont équivalentes.

Si la méthode de sommation de variables exponentielles est équivalente à  $Erlang(k, \gamma)$ , alors on maximise la corrélation dans les deux systèmes. Posons  $X_1 \sim Erlang(k, \gamma_1)$  et  $X_2 \sim Erlang(k, \gamma_2)$ . On génère  $X_1$  et  $X_2$  avec des sommes d'exponentielles où on utilise l'exponentielle inverse  $F^{-1}(U, \gamma) = \frac{-\ln(1-U)}{\gamma}$ .

$$\begin{array}{ll} Y_{11} = F^{-1}(u_1, \gamma_1) & Y_{21} = F^{-1}(u_1, \gamma_2) \\ Y_{12} = F^{-1}(u_2, \gamma_1) & Y_{22} = F^{-1}(u_2, \gamma_2) \\ \dots & \dots \\ Y_{1k} = F^{-1}(u_K, \gamma_1) & Y_{2k} = F^{-1}(u_K, \gamma_2) \\ X_1 = \sum_{i=1}^k Y_{1i}(U, \gamma_1) & X_2 = \sum_{i=1}^k Y_{2i}(U, \gamma_2) \end{array}$$

En effet, selon le théorème 2.5 du manuel, la pair  $(X, Y) = (F^{-1}(U), G^{-1}(U))$  où  $U \sim U(0, 1)$  maximise la corrélation entre  $X$  et  $Y$ .

Les distributions de Erlang peuvent alors être généré comme,

$$G_{\gamma_1}^{-1}(\bar{U}_1) = \sum_{i=1}^k \frac{-\ln(1 - U_i)}{\gamma_1} = \frac{-1}{\gamma_1} \ln\left(\prod_{i=1}^k (1 - U_i)\right)$$

$$G_{\gamma_2}^{-1}(\bar{U}_2) = \sum_{i=1}^k \frac{-\ln(1 - U_i)}{\gamma_2} = \frac{-1}{\gamma_2} \ln\left(\prod_{i=1}^k (1 - U_i)\right)$$

On peut écrire les équations comme,

$$\begin{aligned}\gamma_1 G_{\gamma_1}^{-1}(\bar{U}_1) &= -\ln\left(\prod_{i=1}^k (1 - U_i)\right) \\ \gamma_2 G_{\gamma_2}^{-1}(\bar{U}_2) &= -\ln\left(\prod_{i=1}^k (1 - U_i)\right)\end{aligned}$$

ainsi,

$$\begin{aligned}\gamma_1 G_{\gamma_1}^{-1}(\bar{U}_1) &= \gamma_2 G_{\gamma_2}^{-1}(\bar{U}_2) \\ \Rightarrow G_{\gamma_1}^{-1}(\bar{U}_1) &= \frac{\gamma_2}{\gamma_1} G_{\gamma_2}^{-1}(\bar{U}_2) \\ \Rightarrow \bar{U}_1 &= G_{\gamma_1}\left(\frac{\gamma_2}{\gamma_1} G_{\gamma_2}^{-1}(\bar{U}_2)\right)\end{aligned}$$

Une propriété de la distribution Erlang, pose que si  $X \sim \text{Erlang}(k, \gamma)$  alors  $aX \sim \text{Erlang}(k, \frac{\gamma}{a})$ . Ainsi,

$$\begin{aligned}G_{\gamma_1}^{-1}(\bar{U}_1) &= \frac{\gamma_2}{\gamma_1} G_{\gamma_2}^{-1}(\bar{U}_2) \\ \Rightarrow G_{\gamma_1}^{-1}(\bar{U}_1) &= G_{\frac{\gamma_1}{\gamma_2} \gamma_2}^{-1}(\bar{U}_2) \\ \Rightarrow G_{\gamma_1}^{-1}(\bar{U}_1) &= G_{\gamma_1}^{-1}(\bar{U}_2) \\ \Rightarrow \bar{U}_1 &= \bar{U}_2\end{aligned}$$

Ainsi, on est bien dans le cas du théorème 2.5 du manuel où on maximise la corrélation.

Définissons maintenant la corrélation maximale,

$$\rho(X_1, X_2) = \frac{\mathbb{E}[(X_1 - \mu_1)(X_2 - \mu_2)]}{\sigma_1 \sigma_2}$$

où

$$X_1 = G_{\gamma_1}^{-1}(\bar{U}) = \frac{-1}{\gamma_1} \ln\left(\prod_{i=1}^k (1 - U_i)\right) \quad X_2 = G_{\gamma_2}^{-1}(\bar{U}) = \frac{-1}{\gamma_2} \ln\left(\prod_{i=1}^k (1 - U_i)\right)$$

$$\mu_1 = \frac{k}{\gamma_1} \quad \mu_2 = \frac{k}{\gamma_2} \quad \sigma_1 = \frac{\sqrt{k}}{\gamma_1} \quad \sigma_2 = \frac{\sqrt{k}}{\gamma_2}$$

Nous avons alors,

$$\begin{aligned}
\rho(X_1, X_2) &= \frac{\mathbb{E}\left[\left(\frac{-1}{\gamma_1}\ln(\prod_{i=1}^k(1-U_i)) - \frac{k}{\gamma_1}\right)\left(\frac{-1}{\gamma_2}\ln(\prod_{i=1}^k(1-U_i)) - \frac{k}{\gamma_2}\right)\right]}{\frac{\sqrt{k}}{\gamma_1}\frac{\sqrt{k}}{\gamma_2}} \\
&= \frac{\gamma_1\gamma_2}{k}\mathbb{E}\left[\left(\frac{-1}{\gamma_1}\ln(\prod_{i=1}^k(1-U_i)) - \frac{k}{\gamma_1}\right)\left(\frac{-1}{\gamma_2}\ln(\prod_{i=1}^k(1-U_i)) - \frac{k}{\gamma_2}\right)\right] \\
&= \frac{\gamma_1\gamma_2}{k}\mathbb{E}\left[\left(\frac{-1}{\gamma_1}\right)\left(\ln(\prod_{i=1}^k(1-U_i)) + k\right)\left(\frac{-1}{\gamma_2}\right)\left(\ln(\prod_{i=1}^k(1-U_i)) + k\right)\right] \\
&= \frac{\gamma_1\gamma_2}{k}\frac{1}{\gamma_1\gamma_2}\mathbb{E}\left[\left(\ln(\prod_{i=1}^k(1-U_i)) + k\right)\left(\ln(\prod_{i=1}^k(1-U_i)) + k\right)\right] \\
&= \frac{1}{k}\mathbb{E}\left[\left(\ln(\prod_{i=1}^k(1-U_i)) + k\right)^2\right]
\end{aligned}$$

On remarque que ça correspond à la définition de la variance d'une Erlang  $\gamma = -1$  où la variance est  $k/\gamma^2 = k$ . Nous avons alors,

$$\rho(X_1, X_2) = \frac{1}{k}k = 1$$

Les deux distributions sont alors corrélées au maximum et sont indépendantes du facteur  $\gamma$ .

## Question 5

On cherche ici à étudier différentes variables de contrôle pour l'option asiatique. On veut évaluer les méthodes les plus efficaces selon les paramètres de l'option asiatique.

i)  $t_j = (110 + j)/365$

	K=80	K=90	K=100	K=110
<b>Without any VRT</b>	21.377	11.656	1.909	-7.821
<b>Geometric variance control variable</b>	21.377	11.637	1.896	-7.427
<b>Sum variance control variable</b>	21.912	12.183	2.429	-7.302
<b>Both variance control variables</b>	21.379	11.623	1.90	-7.837
<b>Antithetic variates</b>	16.468	6.795	1.010	nan

Table 4: Espérance de X pour  $t_j = (110 + j)/365$

	K=80	K=90	K=100	K=110
<b>Without any VRT</b>	3.095	3.107	3.133	3.058
<b>Geometric variance control variable</b>	44129	43897	47998	1.098
<b>Sum variance control variable</b>	87	110	84	97
<b>Both variance control variables</b>	43 654	44 258	46 938	$6.651 \times 10^6$
<b>Antithetic variates</b>	3542	80	80 639	nan

Table 5: Variance reduction factor (VCF) pour  $t_j = (110 + j)/365$

ii)  $t_j = (12j)/365$

	K=80	K=90	K=100	K=110
<b>Without any VRT</b>	259 354	244 996	244 894	265 725
<b>Geometric variance control variable</b>	99 724	63 151	72 333	61 839
<b>Sum variance control variable</b>	232 953	232 943	232 943	232 928
<b>Both variance control variables</b>	232 985	233 900	233 992	234 252
<b>Antithetic variates</b>	232 940	232 938	232 920	232 923

Table 6: Espérance de X pour  $t_j = (12j)/365$

	K=80	K=90	K=100	K=110
<b>Without any VRT</b>	$3.247 \times 10^{12}$	$3.247 \times 10^{12}$	$8.338 \times 10^{12}$	$1.833 \times 10^{12}$
<b>Geometric variance control variable</b>	7.017	2.860	1.114	1.292
<b>Sum variance control variable</b>	$1.908 \times 10^{10}$	$4.101 \times 10^{10}$	$3.689 \times 10^7$	$2.676 \times 10^8$
<b>Both variance control variables</b>	$3.527 \times 10^9$	$3.251 \times 10^6$	$5.949 \times 10^5$	$9.245 \times 10^5$
<b>Antithetic variates</b>	$1.325 \times 10^8$	$4.382 \times 10^8$	$3.125 \times 10^7$	$1.732 \times 10^9$

Table 7: Variance reduction factor (VCF) pour  $t_j = (12j)/365$

iii)  $t_j = j/365$

	K=80	K=90	K=100	K=110
<b>Without any VRT</b>	20.797	11.116	1.335	-8.385
<b>Geometric variance control variable</b>	20.784	11.043	1.485	-5.477
<b>Sum variance control variable</b>	20.783	11.0434	1.303	-8.437
<b>Both variance control variables</b>	20.783	11.043	1.334	-8.498
<b>Antithetic variates</b>	20.390	10.986	3.146	0.387

Table 8: Espérance de X pour  $t_j = j/365$

	K=80	K=90	K=100	K=110
<b>Without any VRT</b>	42.353	42.028	43.116	43.825
<b>Geometric variance control variable</b>	4 080	4 223	4415	2.51
<b>Sum variance control variable</b>	$2.499 \times 10^8$	$4.762 \times 10^{10}$	$1.259 \times 10^8$	$5.624 \times 10^6$
<b>Both variance control variables</b>	$8.133 \times 10^9$	$3.418 \times 10^8$	155 854	$8.535 \times 10^7$
<b>Antithetic variates</b>	25 898	182	147	542

Table 9: Variance reduction factor (VCF) pour  $t_j = j/365$

## Discussion

D'après des statistiques on constate qu'il n'y a pas une méthode qui est toujours robuste peu importe des paramètres. En fait, des performances des différents méthodes de variance de contrôle dépendent vraiment du choix de paramètres. Selon les tableaux des espérances de pay-off, en changeant des valeurs du paramètre  $K$ , l'estimateur VC et l'estimateur arithmétique n'ont plus la même espérance.

## References

- [1] A. M. Law, *Simulation Modeling and Analysis*. McGraw-Hill Education, 2013.