Induction de corrélation

Conditions suffisantes pour que la corrélation entre deux v.a. soit positive, ou soit négative?

Induction de corrélation

Conditions suffisantes pour que la corrélation entre deux v.a. soit positive, ou soit négative?

Comment maximiser ou minimiser la corrélation pour des lois marginales données?

Induction de corrélation

Conditions suffisantes pour que la corrélation entre deux v.a. soit positive, ou soit négative?

Comment maximiser ou minimiser la corrélation pour des lois marginales données?

Théorème (bornes de Fréchet; rappel).
Parmi les paires de v.a. (X, Y) dont les f.r. marginales sont F et G, la paire $(X, Y) = (F^{-1}(U), G^{-1}(U))$ où $U \sim U(0, 1)$, maximise $\rho[X, Y]$,

Induction de corrélation

Conditions suffisantes pour que la corrélation entre deux v.a. soit positive, ou soit négative?

Comment maximiser ou minimiser la corrélation pour des lois marginales données?

Théorème (bornes de Fréchet; rappel).
Parmi les paires de v.a. (X, Y) dont les f.r. marginales sont F et G, la paire $(X, Y) = (F^{-1}(U), G^{-1}(U))$ où $U \sim U(0, 1)$, maximise $\rho[X, Y]$, et la paire $(X, Y) = (F^{-1}(U), G^{-1}(1 - U))$ minimise $\rho[X, Y]$.

Induction de corrélation

Conditions suffisantes pour que la corrélation entre deux v.a. soit positive, ou soit négative?

Comment maximiser ou minimiser la corrélation pour des lois marginales données?

Théorème (bornes de Fréchet; rappel).
Parmi les paires de v.a. (X, Y) dont les f.r. marginales sont F et G, la paire $(X, Y) = (F^{-1}(U), G^{-1}(U))$ où $U \sim U(0, 1)$, maximise $\rho[X, Y]$, et la paire $(X, Y) = (F^{-1}(U), G^{-1}(1 - U))$ minimise $\rho[X, Y]$.

Par exemple, si la fonction de répartition d'une durée de service est F dans le premier système et G dans le second, et si on génère les durées de service par $X = F^{-1}(U)$ et $Y = G^{-1}(U)$, alors $\mathrm{Cov}[X, Y] \geq 0$, et cette covariance est maximisée.

L'inversion peut être seulement implicite.
Par exemple, on peut générer X par une méthode quelconque, puis poser

$$Y = G^{-1}(F(X)).$$

Équivaut à l'inversion à partir de $U = F(X) = G(Y)$.

L'inversion peut être seulement implicite.
Par exemple, on peut générer X par une méthode quelconque, puis poser

$$Y = G^{-1}(F(X)).$$

Équivaut à l'inversion à partir de $U = F(X) = G(Y)$.
Facile, e.g., si F et G ne diffèrent que par des paramètres d'échelle et de localisation.

L'inversion peut être seulement implicite.
Par exemple, on peut générer X par une méthode quelconque, puis poser

$$Y = G^{-1}(F(X)).$$

Équivaut à l'inversion à partir de $U = F(X) = G(Y)$.
Facile, e.g., si F et G ne diffèrent que par des paramètres d'échelle et de localisation.

Exemple: Si $X \sim N(\mu_x, \sigma_x^2)$ et $Y \sim N(\mu_y, \sigma_y^2)$, il suffit de générer $Z \sim N(0, 1)$ n'importe comment, puis poser $X = \mu_x + \sigma_x Z$ et $Y = \mu_y + \sigma_y Z$.

L'inversion peut être seulement implicite.
Par exemple, on peut générer X par une méthode quelconque, puis poser

$$Y = G^{-1}(F(X)).$$

Équivaut à l'inversion à partir de $U = F(X) = G(Y)$.
Facile, e.g., si F et G ne diffèrent que par des paramètres d'échelle et de localisation.

Exemple: Si $X \sim N(\mu_x, \sigma_x^2)$ et $Y \sim N(\mu_y, \sigma_y^2)$, il suffit de générer $Z \sim N(0,1)$ n'importe comment, puis poser $X = \mu_x + \sigma_x Z$ et $Y = \mu_y + \sigma_y Z$.
Équivalent à l'inversion à partir de $U = \Phi(Z)$.

L'inversion peut être seulement implicite.
Par exemple, on peut générer X par une méthode quelconque, puis poser

$$Y = G^{-1}(F(X)).$$

Équivaut à l'inversion à partir de $U = F(X) = G(Y)$.
Facile, e.g., si F et G ne diffèrent que par des paramètres d'échelle et de localisation.

Exemple: Si $X \sim N(\mu_x, \sigma_x^2)$ et $Y \sim N(\mu_y, \sigma_y^2)$, il suffit de générer $Z \sim N(0, 1)$ n'importe comment, puis poser $X = \mu_x + \sigma_x Z$ et $Y = \mu_y + \sigma_y Z$.
Équivalent à l'inversion à partir de $U = \Phi(Z)$.

Exemple: Le centre d'appels. Inversion pour les durées inter-arrivées, durées de service, etc.

Mais pour maximiser la corrélation entre les $G_i(s)$, pour deux configurations donnés, il faudrait les générer directement par inversion! Trop difficile.

Discussion. On espère qu'une partie de la corrélation sera transmise.
On le vérifie empiriquement.

Soient

$$\begin{aligned} X &= f(U) \text{ où } U \sim U(0,1) \quad \text{et} \\ Y &= g(V) \text{ où } V \sim U(0,1). \end{aligned}$$

Si f et g sont monotones dans le même sens et $V = U$, alors $\mathrm{Cov}[X, Y] \geq 0$.
Même chose si f et g sont monotones en sens inverse et $V = 1 - U$,

Soient

$$\begin{aligned} X &= f(U) \text{ où } U \sim U(0,1) \quad \text{et} \\ Y &= g(V) \text{ où } V \sim U(0,1). \end{aligned}$$

Si f et g sont monotones dans le même sens et $V = U$, alors $\mathrm{Cov}[X,Y] \geq 0$.
Même chose si f et g sont monotones en sens inverse et $V = 1 - U$,

Et si on permute $V = U$ avec $V = 1 - U$, on obtient $\mathrm{Cov}[f(U), g(V)] \leq 0$.

Dépendance par quadrants

Conditions minimales de dépendance sur (U, V) pour assurer que $f(U)$ et $g(V)$ sont positivement (ou négativement) corrélées dès que f et g sont monotones dans le même sens?

Dépendance par quadrants

Conditions minimales de dépendance sur (U, V) pour assurer que $f(U)$ et $g(V)$ sont positivement (ou négativement) corrélées dès que f et g sont monotones dans le même sens?

Ici, U et V ne sont pas nécessairement uniformes.

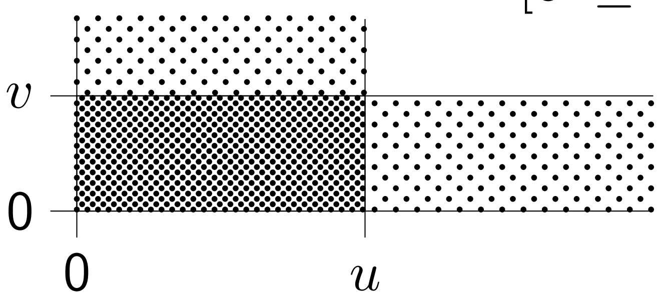
Dépendance par quadrants

Conditions minimales de dépendance sur (U, V) pour assurer que $f(U)$ et $g(V)$ sont positivement (ou négativement) corrélées dès que f et g sont monotones dans le même sens?

Ici, U et V ne sont pas nécessairement uniformes.

Définition. U et V sont positivement dépendantes par quadrant (PQD) si pour tous u et v,

$$\mathbb{P}[U \leq u, V \leq v] \geq \mathbb{P}[U \leq u] \cdot \mathbb{P}[V \leq v].$$



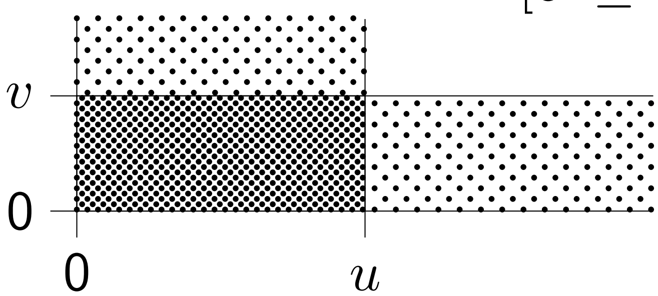
Dépendance par quadrants

Conditions minimales de dépendance sur (U, V) pour assurer que $f(U)$ et $g(V)$ sont positivement (ou négativement) corrélées dès que f et g sont monotones dans le même sens?

Ici, U et V ne sont pas nécessairement uniformes.

Définition. U et V sont positivement dépendantes par quadrant (PQD) si pour tous u et v,

$$\mathbb{P}[U \le u, V \le v] \ge \mathbb{P}[U \le u] \cdot \mathbb{P}[V \le v].$$



Avec $\le$, U et V sont négativement dépendantes par quadrant (NQD).

On note $\mathcal{P}^+$ et $\mathcal{P}^-$ les familles de paires de v.a. PQD et NQD.

Théorème.

(i) $(U, U) \in \mathcal{P}^+$;

Théorème.

(i) $(U, U) \in \mathcal{P}^+$;

(ii) $(U, V) \in \mathcal{P}^+$ ssi $(U, -V) \in \mathcal{P}^-$;

Théorème.

(i) $(U, U) \in \mathcal{P}^+$;

(ii) $(U, V) \in \mathcal{P}^+$ ssi $(U, -V) \in \mathcal{P}^-$;

(iii) Si $(U, V) \in \mathcal{P}^+$, et f et g sont monotones dans le même sens, alors $(f(U), g(V)) \in \mathcal{P}^+$, et de même pour $\mathcal{P}^-$;

Théorème.

(i) $(U, U) \in \mathcal{P}^+$;

(ii) $(U, V) \in \mathcal{P}^+$ ssi $(U, -V) \in \mathcal{P}^-$;

(iii) Si $(U, V) \in \mathcal{P}^+$, et f et g sont monotones dans le même sens, alors $(f(U), g(V)) \in \mathcal{P}^+$, et de même pour $\mathcal{P}^-$;

(iv) Si $(U, V) \in \mathcal{P}^+$ alors $\mathrm{Cov}[U, V] \geq 0$, et la covariance est nulle seulement si U et V sont indépendantes.

Théorème.

(i) $(U, U) \in \mathcal{P}^+$;

(ii) $(U, V) \in \mathcal{P}^+$ ssi $(U, -V) \in \mathcal{P}^-$;

(iii) Si $(U, V) \in \mathcal{P}^+$, et f et g sont monotones dans le même sens, alors $(f(U), g(V)) \in \mathcal{P}^+$, et de même pour $\mathcal{P}^-$;

(iv) Si $(U, V) \in \mathcal{P}^+$ alors $\mathrm{Cov}[U, V] \geq 0$, et la covariance est nulle seulement si U et V sont indépendantes.
De même, si $(U, V) \in \mathcal{P}^-$ alors $\mathrm{Cov}[U, V] \leq 0$.

Théorème.

(i) $(U, U) \in \mathcal{P}^+$;

(ii) $(U, V) \in \mathcal{P}^+$ ssi $(U, -V) \in \mathcal{P}^-$;

(iii) Si $(U, V) \in \mathcal{P}^+$, et f et g sont monotones dans le même sens, alors $(f(U), g(V)) \in \mathcal{P}^+$, et de même pour $\mathcal{P}^-$;

(iv) Si $(U, V) \in \mathcal{P}^+$ alors $\mathrm{Cov}[U, V] \geq 0$, et la covariance est nulle seulement si U et V sont indépendantes.
De même, si $(U, V) \in \mathcal{P}^-$ alors $\mathrm{Cov}[U, V] \leq 0$.

(v) $(U, V) \in \mathcal{P}^+$ ssi $\mathrm{Cov}[f(U), g(V)] \geq 0$ pour toutes fonctions non-décroissantes f, g telles que la covariance existe.

Théorème.

(i) $(U, U) \in \mathcal{P}^+$;

(ii) $(U, V) \in \mathcal{P}^+$ ssi $(U, -V) \in \mathcal{P}^-$;

(iii) Si $(U, V) \in \mathcal{P}^+$, et f et g sont monotones dans le même sens, alors $(f(U), g(V)) \in \mathcal{P}^+$, et de même pour $\mathcal{P}^-$;

(iv) Si $(U, V) \in \mathcal{P}^+$ alors $\mathrm{Cov}[U, V] \geq 0$, et la covariance est nulle seulement si U et V sont indépendantes.
De même, si $(U, V) \in \mathcal{P}^-$ alors $\mathrm{Cov}[U, V] \leq 0$.

(v) $(U, V) \in \mathcal{P}^+$ ssi $\mathrm{Cov}[f(U), g(V)] \geq 0$ pour toutes fonctions non-décroissantes f, g telles que la covariance existe.

Exemple. Si $X = F^{-1}(U)$ et $Y = G^{-1}(U)$ (inversion), alors $(X, Y) \in \mathcal{P}^+$.

Fonctions de plusieurs variables

Définition. Deux fonctions $f(U_1, U_2, \ldots)$ et $g(V_1, V_2, \ldots)$ sont comonotones [antimonotone] pour leur j-ième argument si elles sont monotones dans la même direction [dans des directions opposées] par rapport à cet argument.

Fonctions de plusieurs variables

Définition. Deux fonctions $f(U_1, U_2, \ldots)$ et $g(V_1, V_2, \ldots)$ sont comonotones [antimonotone] pour leur j-ième argument si elles sont monotones dans la même direction [dans des directions opposées] par rapport à cet argument.

Théorème. Soient $\{(U_j, V_j),\ j \geq 1\}$ des paires indépendantes de v.a., $Y_1 = f(U_1, U_2, \ldots)$ et $Y_2 = g(V_1, V_2, \ldots)$ deux v.a. de variance finie, et supposons que pour tout $j \geq 1$, avec probabilité 1, l'une des trois conditions suivantes tient:

Fonctions de plusieurs variables

Définition. Deux fonctions $f(U_1, U_2, \ldots)$ et $g(V_1, V_2, \ldots)$ sont comonotones [antimonotone] pour leur j-ième argument si elles sont monotones dans la même direction [dans des directions opposées] par rapport à cet argument.

Théorème. Soient $\{(U_j, V_j),\ j \geq 1\}$ des paires indépendantes de v.a., $Y_1 = f(U_1, U_2, \ldots)$ et $Y_2 = g(V_1, V_2, \ldots)$ deux v.a. de variance finie, et supposons que pour tout $j \geq 1$, avec probabilité 1, l'une des trois conditions suivantes tient:

$(U_j, V_j) \in \mathcal{P}^+$ et f et g sont comonotones pour leur argument j;

Fonctions de plusieurs variables

Définition. Deux fonctions $f(U_1, U_2, \ldots)$ et $g(V_1, V_2, \ldots)$ sont comonotones [antimonotone] pour leur j-ième argument si elles sont monotones dans la même direction [dans des directions opposées] par rapport à cet argument.

Théorème. Soient $\{(U_j, V_j),\ j \geq 1\}$ des paires indépendantes de v.a., $Y_1 = f(U_1, U_2, \ldots)$ et $Y_2 = g(V_1, V_2, \ldots)$ deux v.a. de variance finie, et supposons que pour tout $j \geq 1$, avec probabilité 1, l'une des trois conditions suivantes tient:

$(U_j, V_j) \in \mathcal{P}^+$ et f et g sont comonotones pour leur argument j;
$(U_j, V_j) \in \mathcal{P}^-$ et f et g sont antimonotones pour leur argument j;

Fonctions de plusieurs variables

Définition. Deux fonctions $f(U_1, U_2, \ldots)$ et $g(V_1, V_2, \ldots)$ sont comonotones [antimonotone] pour leur j-ième argument si elles sont monotones dans la même direction [dans des directions opposées] par rapport à cet argument.

Théorème. Soient $\{(U_j, V_j),\ j \geq 1\}$ des paires indépendantes de v.a., $Y_1 = f(U_1, U_2, \ldots)$ et $Y_2 = g(V_1, V_2, \ldots)$ deux v.a. de variance finie, et supposons que pour tout $j \geq 1$, avec probabilité 1, l'une des trois conditions suivantes tient:

$(U_j, V_j) \in \mathcal{P}^+$ et f et g sont comonotones pour leur argument j;
$(U_j, V_j) \in \mathcal{P}^-$ et f et g sont antimonotones pour leur argument j;
U_j et V_j sont indépendantes.

Fonctions de plusieurs variables

Définition. Deux fonctions $f(U_1, U_2, \ldots)$ et $g(V_1, V_2, \ldots)$ sont comonotones [antimonotone] pour leur j-ième argument si elles sont monotones dans la même direction [dans des directions opposées] par rapport à cet argument.

Théorème. Soient $\{(U_j, V_j),\ j \geq 1\}$ des paires indépendantes de v.a., $Y_1 = f(U_1, U_2, \ldots)$ et $Y_2 = g(V_1, V_2, \ldots)$ deux v.a. de variance finie, et supposons que pour tout $j \geq 1$, avec probabilité 1, l'une des trois conditions suivantes tient:

$(U_j, V_j) \in \mathcal{P}^+$ et f et g sont comonotones pour leur argument j;
$(U_j, V_j) \in \mathcal{P}^-$ et f et g sont antimonotones pour leur argument j;
U_j et V_j sont indépendantes.

Alors, $(Y_1, Y_2) \in \mathcal{P}^+$ et donc $\mathrm{Cov}[Y_1, Y_2] \geq 0$.

Fonctions de plusieurs variables

Définition. Deux fonctions $f(U_1, U_2, \ldots)$ et $g(V_1, V_2, \ldots)$ sont comonotones [antimonotone] pour leur j-ième argument si elles sont monotones dans la même direction [dans des directions opposées] par rapport à cet argument.

Théorème. Soient $\{(U_j, V_j),\ j \geq 1\}$ des paires indépendantes de v.a., $Y_1 = f(U_1, U_2, \ldots)$ et $Y_2 = g(V_1, V_2, \ldots)$ deux v.a. de variance finie, et supposons que pour tout $j \geq 1$, avec probabilité 1, l'une des trois conditions suivantes tient:

$(U_j, V_j) \in \mathcal{P}^+$ et f et g sont comonotones pour leur argument j;
$(U_j, V_j) \in \mathcal{P}^-$ et f et g sont antimonotones pour leur argument j;
U_j et V_j sont indépendantes.

Alors, $(Y_1, Y_2) \in \mathcal{P}^+$ et donc $\mathrm{Cov}[Y_1, Y_2] \geq 0$.

De même si les conditions tiennent en échangeant $\mathcal{P}^+$ et $\mathcal{P}^-$, alors $(Y_1, Y_2) \in \mathcal{P}^-$.

Fonctions de plusieurs variables

Définition. Deux fonctions $f(U_1, U_2, \ldots)$ et $g(V_1, V_2, \ldots)$ sont comonotones [antimonotone] pour leur j-ième argument si elles sont monotones dans la même direction [dans des directions opposées] par rapport à cet argument.

Théorème. Soient $\{(U_j, V_j),\ j \geq 1\}$ des paires indépendantes de v.a., $Y_1 = f(U_1, U_2, \ldots)$ et $Y_2 = g(V_1, V_2, \ldots)$ deux v.a. de variance finie, et supposons que pour tout $j \geq 1$, avec probabilité 1, l'une des trois conditions suivantes tient:

$(U_j, V_j) \in \mathcal{P}^+$ et f et g sont comonotones pour leur argument j;
$(U_j, V_j) \in \mathcal{P}^-$ et f et g sont antimonotones pour leur argument j;
U_j et V_j sont indépendantes.

Alors, $(Y_1, Y_2) \in \mathcal{P}^+$ et donc $\mathrm{Cov}[Y_1, Y_2] \geq 0$.

De même si les conditions tiennent en échangeant $\mathcal{P}^+$ et $\mathcal{P}^-$, alors $(Y_1, Y_2) \in \mathcal{P}^-$.

En pratique, les fonctions f et g ne dépendent habituellement que d'un nombre fini (souvent aléatoire) de leurs arguments.

Valeurs aléatoires communes (VAC)

On utilise $\Delta = X_2 - X_1$ pour estimer $\mu_2 - \mu_1 = \mathbb{E}[X_2] - \mathbb{E}[X_1]$. On a

$$\mathrm{Var}[\Delta] = \mathrm{Var}[X_1] + \mathrm{Var}[X_2] - 2\,\mathrm{Cov}[X_1, X_2].$$

Valeurs aléatoires communes (VAC)

On utilise $\Delta = X_2 - X_1$ pour estimer $\mu_2 - \mu_1 = \mathbb{E}[X_2] - \mathbb{E}[X_1]$. On a

$$\mathrm{Var}[\Delta] = \mathrm{Var}[X_1] + \mathrm{Var}[X_2] - 2\,\mathrm{Cov}[X_1, X_2].$$

Objectif: induire une corrélation positive entre X_1 et X_2 sans changer leurs lois individuelles.

Valeurs aléatoires communes (VAC)

On utilise $\Delta = X_2 - X_1$ pour estimer $\mu_2 - \mu_1 = \mathbb{E}[X_2] - \mathbb{E}[X_1]$. On a

$$\mathrm{Var}[\Delta] = \mathrm{Var}[X_1] + \mathrm{Var}[X_2] - 2\,\mathrm{Cov}[X_1, X_2].$$

Objectif: induire une corrélation positive entre X_1 et X_2 sans changer leurs lois individuelles.

Technique: utiliser les mêmes nombres aléatoires pour simuler les deux systèmes, en essayant de maintenir la synchronisation le mieux possible.

Valeurs aléatoires communes (VAC)

On utilise $\Delta = X_2 - X_1$ pour estimer $\mu_2 - \mu_1 = \mathbb{E}[X_2] - \mathbb{E}[X_1]$. On a

$$\mathrm{Var}[\Delta] = \mathrm{Var}[X_1] + \mathrm{Var}[X_2] - 2\,\mathrm{Cov}[X_1, X_2].$$

Objectif: induire une corrélation positive entre X_1 et X_2 sans changer leurs lois individuelles.

Technique: utiliser les mêmes nombres aléatoires pour simuler les deux systèmes, en essayant de maintenir la synchronisation le mieux possible.

Si $X_k = f_k(\mathbf{U}_k) = f_k(U_{k,1}, U_{k,2}, \ldots)$ pour $k = 1, 2$, utiliser des VAC partout veut dire prendre $\mathbf{U}_1 = \mathbf{U}_2$.

Valeurs aléatoires communes (VAC)

On utilise $\Delta = X_2 - X_1$ pour estimer $\mu_2 - \mu_1 = \mathbb{E}[X_2] - \mathbb{E}[X_1]$. On a

$$\mathrm{Var}[\Delta] = \mathrm{Var}[X_1] + \mathrm{Var}[X_2] - 2\,\mathrm{Cov}[X_1, X_2].$$

Objectif: induire une corrélation positive entre X_1 et X_2 sans changer leurs lois individuelles.

Technique: utiliser les mêmes nombres aléatoires pour simuler les deux systèmes, en essayant de maintenir la synchronisation le mieux possible.

Si $X_k = f_k(\mathbf{U}_k) = f_k(U_{k,1}, U_{k,2}, \ldots)$ pour $k = 1, 2$, utiliser des VAC partout veut dire prendre $\mathbf{U}_1 = \mathbf{U}_2$.

Au mieux, on a $\rho(X_1, X_2) = 1$ et $\mathrm{Var}[\Delta] = 0$;
au pire, $\rho(X_1, X_2) = -1$ et $\mathrm{Var}[\Delta] = 2(\mathrm{Var}[X_1] + \mathrm{Var}[X_2])$.

Pour maximiser la corrélation, il faudrait générer X_1 et X_2 directement par inversion! Habituellement trop difficile.

Conditions suffisantes de réduction de variance

Supposons que $\mathbb{E}[X_1^2] < \infty$ et $\mathbb{E}[X_2^2] < \infty$.

Conditions suffisantes de réduction de variance

Supposons que $\mathbb{E}[X_1^2] < \infty$ et $\mathbb{E}[X_2^2] < \infty$.
On construit $\mathbf{U}_2$ à partir de $\mathbf{U}_1$ comme suit:

$U_{2,j} = U_{1,j}$ pour $j \in \Psi^+$,
$U_{2,j} = 1 - U_{1,j}$ pour $j \in \Psi^-$,
et $U_{2,j}$ indépendant de $U_{1,j}$ pour les autres j.

Conditions suffisantes de réduction de variance

Supposons que $\mathbb{E}[X_1^2] < \infty$ et $\mathbb{E}[X_2^2] < \infty$.
On construit $\mathbf{U}_2$ à partir de $\mathbf{U}_1$ comme suit:

$U_{2,j} = U_{1,j}$ pour $j \in \Psi^+$,
$U_{2,j} = 1 - U_{1,j}$ pour $j \in \Psi^-$,
et $U_{2,j}$ indépendant de $U_{1,j}$ pour les autres j.

Cet échantillonnage est dit VAC-concordant si f_1 et f_2 sont comonotones par rapport à leur $j$ième argument $U_{k,j}$ si $j \in \Psi^+$ et antimonotones si $j \in \Psi^-$.

Conditions suffisantes de réduction de variance

Supposons que $\mathbb{E}[X_1^2] < \infty$ et $\mathbb{E}[X_2^2] < \infty$.
On construit $\mathbf{U}_2$ à partir de $\mathbf{U}_1$ comme suit:

$U_{2,j} = U_{1,j}$ pour $j \in \Psi^+$,
$U_{2,j} = 1 - U_{1,j}$ pour $j \in \Psi^-$,
et $U_{2,j}$ indépendant de $U_{1,j}$ pour les autres j.

Cet échantillonnage est dit VAC-concordant si f_1 et f_2 sont comonotones par rapport à leur $j$ième argument $U_{k,j}$ si $j \in \Psi^+$ et antimonotones si $j \in \Psi^-$.

Si on permute $U_{1,j}$ et $1 - U_{1,j}$ dans la définition, on obtient un échantillonnage VAC-discordant.

Conditions suffisantes de réduction de variance

Supposons que $\mathbb{E}[X_1^2] < \infty$ et $\mathbb{E}[X_2^2] < \infty$.
On construit $\mathbf{U}_2$ à partir de $\mathbf{U}_1$ comme suit:

$U_{2,j} = U_{1,j}$ pour $j \in \Psi^+$,
$U_{2,j} = 1 - U_{1,j}$ pour $j \in \Psi^-$,
et $U_{2,j}$ indépendant de $U_{1,j}$ pour les autres j.

Cet échantillonnage est dit VAC-concordant si f_1 et f_2 sont comonotones par rapport à leur $j$ième argument $U_{k,j}$ si $j \in \Psi^+$ et antimonotones si $j \in \Psi^-$.

Si on permute $U_{1,j}$ et $1 - U_{1,j}$ dans la définition, on obtient un échantillonnage VAC-discordant.

Théorème: Pour un échantillonnage VAC-concordant on a $\mathrm{Cov}[X_1, X_2] \geq 0$.
Pour un échantillonnage VAC-discordant on a $\mathrm{Cov}[X_1, X_2] \leq 0$.

Conditions suffisantes de réduction de variance

Supposons que $\mathbb{E}[X_1^2] < \infty$ et $\mathbb{E}[X_2^2] < \infty$.
On construit $\mathbf{U}_2$ à partir de $\mathbf{U}_1$ comme suit:

$U_{2,j} = U_{1,j}$ pour $j \in \Psi^+$,
$U_{2,j} = 1 - U_{1,j}$ pour $j \in \Psi^-$,
et $U_{2,j}$ indépendant de $U_{1,j}$ pour les autres j.

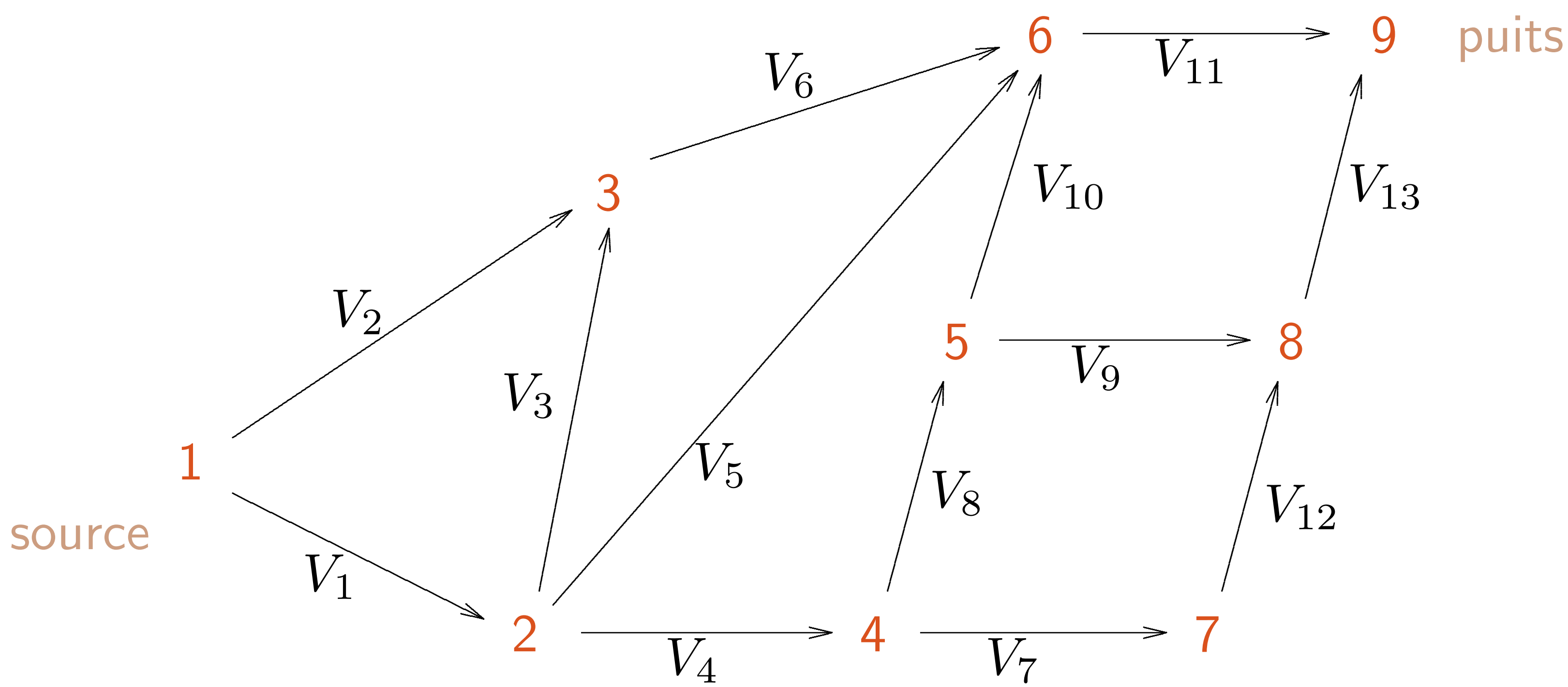
Cet échantillonnage est dit VAC-concordant si f_1 et f_2 sont comonotones par rapport à leur $j$ième argument $U_{k,j}$ si $j \in \Psi^+$ et antimonotones si $j \in \Psi^-$.

Si on permute $U_{1,j}$ et $1 - U_{1,j}$ dans la définition, on obtient un échantillonnage VAC-discordant.

Théorème: Pour un échantillonnage VAC-concordant on a $\mathrm{Cov}[X_1, X_2] \geq 0$.
Pour un échantillonnage VAC-discordant on a $\mathrm{Cov}[X_1, X_2] \leq 0$.

Preuve: Découle directement du théorème sur les fonctions comonotones de plusieurs v.a. PQD deux à deux.

Exemple: Réseau d'activités stochastique.



La durée V_j de l'activité j a la fonction de répart. $F_{1,j}$ dans la configuration 1 et $F_{2,j}$ dans l'autre. On génère toutes les durées par: $V_{k,j} = F_{k,j}^{-1}(U_j)$, puis $X_k = f_k(\mathbf{U}_k) = \mathbb{I}[T_k > x]$ pour $k = 1, 2$.
Le théorème nous assure que $\mathrm{Cov}[X_1, X_2] \geq 0$.

Exemple: Chaine de Markov stochastiquement monotone.

$\{\mathcal{S}_j,\, j \geq 0\}$; $\mathbb{P}[\mathcal{S}_{j+1} \geq y \mid \mathcal{S}_j = x]$ non-décroissant en x.

$C_j = c(\mathcal{S}_j)$ et $X = \sum_j C_j$.

Si toutes les transitions de la chaine sont generées par inversion et si c est non décroissante, alors X est une fonction non décroissante de chaque uniforme.

Exemple: Chaine de Markov stochastiquement monotone.

$\{\mathcal{S}_j,\, j \geq 0\}$; $\mathbb{P}[\mathcal{S}_{j+1} \geq y \mid \mathcal{S}_j = x]$ non-décroissant en x.

$C_j = c(\mathcal{S}_j)$ et $X = \sum_j C_j$.

Si toutes les transitions de la chaine sont generées par inversion et si c est non décroissante, alors X est une fonction non décroissante de chaque uniforme.

Exemple: processus de Lindley,

$W_{i+1} = \max(0,\, W_i + S_i - A_i)$, où $S_i - A_i$ est indépendant de W_i.

C'est une CM stochastiquement monotone.

Si X_1 et X_2 sont des fonctions non-décroissantes des W_i pour deux processus de Lindley simulés avec des VAC, alors $\mathrm{Cov}[X_1, X_2] \geq 0$.

Exemple: Chaine de Markov stochastiquement monotone.

$\{\mathcal{S}_j,\, j \ge 0\}$; $\mathbb{P}[\mathcal{S}_{j+1} \ge y \mid \mathcal{S}_j = x]$ non-décroissant en x.

$C_j = c(\mathcal{S}_j)$ et $X = \sum_j C_j$.

Si toutes les transitions de la chaine sont generées par inversion et si c est non décroissante, alors X est une fonction non décroissante de chaque uniforme.

Exemple: processus de Lindley,

$W_{i+1} = \max(0,\, W_i + S_i - A_i)$, où $S_i - A_i$ est indépendant de W_i.

C'est une CM stochastiquement monotone.

Si X_1 et X_2 sont des fonctions non-décroissantes des W_i pour deux processus de Lindley simulés avec des VAC, alors $\mathrm{Cov}[X_1, X_2] \ge 0$.

Par contre, s'il y a des abandons, les choses se compliquent.

Exemple: Chaine de Markov stochastiquement monotone.

$\{\mathcal{S}_j,\ j \ge 0\}$; $\mathbb{P}[\mathcal{S}_{j+1} \ge y \mid \mathcal{S}_j = x]$ non-décroissant en x.

$C_j = c(\mathcal{S}_j)$ et $X = \sum_j C_j$.

Si toutes les transitions de la chaine sont generées par inversion et si c est non décroissante, alors X est une fonction non décroissante de chaque uniforme.

Exemple: processus de Lindley,

$W_{i+1} = \max(0,\, W_i + S_i - A_i)$, où $S_i - A_i$ est indépendant de W_i.

C'est une CM stochastiquement monotone.

Si X_1 et X_2 sont des fonctions non-décroissantes des W_i pour deux processus de Lindley simulés avec des VAC, alors $\mathrm{Cov}[X_1, X_2] \ge 0$.

Par contre, s'il y a des abandons, les choses se compliquent.

Exemple: mouvement Brownien géométrique

$\{S(t),\, t \ge 0\}$ et fonction de revenu non-décroissante $g(S(t_1), \ldots, S(t_d))$.

Si on en simule deux comme cela avec des VAC (les même Z_j's),

$X_1 = g_1(S_1(t_1), \ldots, S_1(t_d))$ et $X_2 = g_2(S_2(t_1), \ldots, S_2(t_d))$,
alors $\mathrm{Cov}[X_1, X_2] \ge 0$.

Parfois il est très difficile de vérifier les conditions du théorème.
Par exemple, pour le centre d'appels, la monotonicité est difficile à vérifier à cause des possibilités d'abandon.

Les conditions du théorème sont suffisantes, mais pas nécessaires.

Parfois il est très difficile de vérifier les conditions du théorème.
Par exemple, pour le centre d'appels, la monotonicité est difficile à vérifier à cause des possibilités d'abandon.

Les conditions du théorème sont suffisantes, mais pas nécessaires.

Pour tester si c'est efficace en pratique: faire un expérience pilote avec les VAC et estimer $\mathrm{Cov}(X_1, X_2)$.

Pas besoin de faire de simulations sans les VAC.
Pour estimer la variance qu'on aurait sans les VAC, il suffit de prendre la version empirique de $\mathrm{Var}[X_1] + \mathrm{Var}[X_2]$.

VAC pour des différences très petites

Supposons que $X = f(\theta, \mathbf{U})$, où $\theta \in [a, b]$.

Soient $X_1 = f(\theta_1, \mathbf{U}_1)$, $X_2 = f(\theta_1 + \delta, \mathbf{U}_2)$, et $\Delta = X_2 - X_1$, où $\delta \geq 0$ est petit.
On veut estimer $\mathbb{E}[\Delta]$ ou encore $\mathbb{E}[\Delta/\delta]$.

VAC pour des différences très petites

Supposons que $X = f(\theta, \mathbf{U})$, où $\theta \in [a, b]$.

Soient $X_1 = f(\theta_1, \mathbf{U}_1)$, $X_2 = f(\theta_1 + \delta, \mathbf{U}_2)$, et $\Delta = X_2 - X_1$, où $\delta \geq 0$ est petit. On veut estimer $\mathbb{E}[\Delta]$ ou encore $\mathbb{E}[\Delta/\delta]$.

E.g., pour analyse de sensibilité ou estimation de gradient par différences finies.

θ peut-être un paramètre d'une loi de probabilité du modèle, ou n'importe quel paramètre continu du modèle (vitesse d'un serveur ou d'un convoyeur, paramètre d'une option, etc.).

VAC pour des différences très petites

Supposons que $X = f(\theta, \mathbf{U})$, où $\theta \in [a, b]$.

Soient $X_1 = f(\theta_1, \mathbf{U}_1)$, $X_2 = f(\theta_1 + \delta, \mathbf{U}_2)$, et $\Delta = X_2 - X_1$, où $\delta \geq 0$ est petit.
On veut estimer $\mathbb{E}[\Delta]$ ou encore $\mathbb{E}[\Delta/\delta]$.

E.g., pour analyse de sensibilité ou estimation de gradient par différences finies.

θ peut-être un paramètre d'une loi de probabilité du modèle, ou n'importe quel paramètre continu du modèle (vitesse d'un serveur ou d'un convoyeur, paramètre d'une option, etc.).

Avec des VAI:
$\mathrm{Var}[\Delta/\delta] = \frac{(\mathrm{Var}[X_1]+\mathrm{Var}[X_2])}{\delta^2} = \Theta(\delta^{-2}) \to \infty$ quand $\delta \to 0$.

Avec des VAC:
Peut-on borner cette variance? Ou la faire augmenter moins vite?

Example: Asian call option under GMB model.
The payoff at time T when $S(0) = s_0$ can be written as

$$Y(s_0) = e^{-rT} \max(0,\, s_0 W - K)$$

where

$$W \quad = \quad \frac{1}{d} \sum_{i=1}^{d} \exp\left[(r - \sigma^2/2) t_i + \sigma \sum_{j=1}^{i} \sqrt{t_j - t_{j-1}} \Phi^{-1}(U_j) \right]$$

and $U_1, \ldots, U_d$ are i.i.d. $U(0,1)$.

Let $v(s_0, T) = \mathbb{E}[Y(s_0)]$ be the expected discounted payoff.
Want to estimate $v'(s_0, T) = \partial v(s_0, T)/\partial s_0$, via finite differences.

With CRNs, for $\delta > 0$, we have

$$\Delta = Y(s_0+\delta) - Y(s_0) = \begin{cases} 0 & \text{if } Y(s_0+\delta) = Y(s_0) = 0; \\ e^{-rT}\delta W & \text{if } Y(s_0+\delta) > Y(s_0) > 0; \\ e^{-rT}[(s_0+\delta)W - K] & \text{if } Y(s_0+\delta) > Y(s_0) = 0, \end{cases}$$

and the finite difference estimator Δ/δ satisfies $0 \le \Delta/\delta \le e^{-rT}W$.

With CRNs, for $\delta > 0$, we have

$$\Delta = Y(s_0+\delta) - Y(s_0) = \begin{cases} 0 & \text{if } Y(s_0+\delta) = Y(s_0) = 0; \\ e^{-rT}\delta W & \text{if } Y(s_0+\delta) > Y(s_0) > 0; \\ e^{-rT}[(s_0+\delta)W - K] & \text{if } Y(s_0+\delta) > Y(s_0) = 0, \end{cases}$$

and the finite difference estimator Δ/δ satisfies $0 \le \Delta/\delta \le e^{-rT}W$.

Therefore, $\mathrm{Var}[\Delta/\delta] \le e^{-2rT}\mathrm{Var}[W]$, which is bounded uniformly in δ.

We can take $\delta > 0$ as small as we want and the variance remains bounded.

With CRNs, for $\delta > 0$, we have

$$\Delta = Y(s_0+\delta) - Y(s_0) = \begin{cases} 0 & \text{if } Y(s_0+\delta) = Y(s_0) = 0; \\ e^{-rT}\delta W & \text{if } Y(s_0+\delta) > Y(s_0) > 0; \\ e^{-rT}[(s_0+\delta)W - K] & \text{if } Y(s_0+\delta) > Y(s_0) = 0, \end{cases}$$

and the finite difference estimator Δ/δ satisfies $0 \le \Delta/\delta \le e^{-rT}W$.

Therefore, $\mathrm{Var}[\Delta/\delta] \le e^{-2rT}\mathrm{Var}[W]$, which is bounded uniformly in δ.

We can take $\delta > 0$ as small as we want and the variance remains bounded.

Best δ to use in practice?
With finite differences, we must pay attention to the limited accuracy of computers.

We can do much better by taking the estimator to the limit: use the sample derivative

$$Y'(s_0) = \lim_{\delta \to 0} \frac{Y(s_0 + \delta) - Y(s_0)}{\delta} = \begin{cases} 0 & \text{if } Y(s_0) = 0; \\ e^{-rT}W & \text{if } Y(s_0) > 0, \end{cases}$$

to estimate the derivative

$$v'(s_0, T) = \frac{\partial v(s_0, T)}{\partial s_0} = \frac{\partial \mathbb{E}[Y(s_0)]}{\partial s_0}.$$

Since Δ/δ is bounded by $e^{-rT}W$ uniformly in δ, the dominated convergence theorem guarantees that we can interchange the limit and expectation:

$$\mathbb{E}[Y'(s_0)] = \mathbb{E}\left[\lim_{\delta \to 0} \frac{Y(s_0 + \delta) - Y(s_0)}{\delta}\right] = \lim_{\delta \to 0} \mathbb{E}\left[\frac{Y(s_0 + \delta) - Y(s_0)}{\delta}\right] = v'(s_0, T).$$

That is, we have an unbiased estimator of the derivative, whose variance is bounded by $e^{-2rT}\mathrm{Var}[W]$.

For this model, we can develop a similar estimator for the derivative with respect to σ.

For this model, we can develop a similar estimator for the derivative with respect to σ.

But taking the sample derivative as a derivative estimator does not always work.

There are many situations where the variance of Δ/δ with CRNs increases to infinity when $\delta \to 0$, although often at a slower rate than $O(\delta^{-2})$.

For this model, we can develop a similar estimator for the derivative with respect to σ.

But taking the sample derivative as a derivative estimator does not always work.

There are many situations where the variance of Δ/δ with CRNs increases to infinity when $\delta \to 0$, although often at a slower rate than $O(\delta^{-2})$.

Example. Replace the payoff $Y(s_0)$ in the previous example by the indicator function:

$$\tilde{Y}(s_0) = \mathbb{I}[Y(s_0) > 0] = \mathbb{I}[s_0 W > K],$$

where $K > 0$.

For this model, we can develop a similar estimator for the derivative with respect to σ.

But taking the sample derivative as a derivative estimator does not always work.
There are many situations where the variance of Δ/δ with CRNs increases to infinity when $\delta \to 0$, although often at a slower rate than $O(\delta^{-2})$.

Example. Replace the payoff $Y(s_0)$ in the previous example by the indicator function:

$$\tilde{Y}(s_0) = \mathbb{I}[Y(s_0) > 0] = \mathbb{I}[s_0 W > K],$$

where $K > 0$. For a given realization of W, the payoff is 0 for $s_0 \leq K/W$, and 1 for $s_0 > K/W$, so the sample derivative $\tilde{Y}'(s_0)$ is 0 with probability 1.
On the other hand, the derivative of $\mathbb{E}[\tilde{Y}(s_0)] = \mathbb{P}[W > K/s_0]$ with respect to s_0 is strictly positive. Thus, we no longer have an unbiased estimator.

If we use the finite-difference estimator $\Delta = \mathbb{I}[Y(s_0+\delta) > 0] - \mathbb{I}[Y(s_0) > 0]$, we have

$$\Delta = \mathbb{I}\left[s_0 \le K/W < s_0+\delta\right] = \mathbb{I}\left[\frac{K}{s_0+\delta} < W < \frac{K}{s_0}\right].$$

Both W and K/W have a bounded and continuous density in a small neighborhood of s_0.

Thus, if κ is the density of K/W at s_0, then $\mathbb{P}[\Delta > 0] = \kappa\delta + o(\delta)$ when δ is small, and

$$\mathrm{Var}[\Delta/\delta] = \frac{\mathbb{P}[\Delta > 0] - \mathbb{P}^2[\Delta > 0]}{\delta^2} = \frac{\kappa\delta - (\kappa\delta)^2 + o(\delta)}{\delta^2} = \frac{\kappa}{\delta} + o(1/\delta).$$

Still blows up when $\delta \to 0$, but at rate $\Theta(1/\delta)$ instead of $\Theta(1/\delta^2)$ with IRNs.

If we use the finite-difference estimator $\Delta = \mathbb{I}[Y(s_0+\delta) > 0] - \mathbb{I}[Y(s_0) > 0]$, we have

$$\Delta = \mathbb{I}\left[s_0 \le K/W < s_0+\delta\right] = \mathbb{I}\left[\frac{K}{s_0+\delta} < W < \frac{K}{s_0}\right].$$

Both W and K/W have a bounded and continuous density in a small neighborhood of s_0.

Thus, if κ is the density of K/W at s_0, then $\mathbb{P}[\Delta > 0] = \kappa\delta + o(\delta)$ when δ is small, and

$$\mathrm{Var}[\Delta/\delta] = \frac{\mathbb{P}[\Delta > 0] - \mathbb{P}^2[\Delta > 0]}{\delta^2} = \frac{\kappa\delta - (\kappa\delta)^2 + o(\delta)}{\delta^2} = \frac{\kappa}{\delta} + o(1/\delta).$$

Still blows up when $\delta \to 0$, but at rate $\Theta(1/\delta)$ instead of $\Theta(1/\delta^2)$ with IRNs.

Here, the finite-difference estimator can only take the values 0 and $1/\delta$. This second value is unbounded when $\delta \to 0$ and this is the reason why we cannot apply the dominated convergence theorem.

But the probability of this second value is only $\mathcal{O}(\delta)$.

Le théorème qui suit couvre les exemples précédents et bien plus encore.

Notre premier exemple correspond à la partie (iii) avec $\alpha = 2$ (ou au corollaire).

Le second exemple correspond à (iv) avec α arbitrairement grand (ou $\alpha = \infty$) et $\beta = 1$ (puisque $\mathbb{P}[\Delta^2 > 0] = O(1/\delta)$ dans cet exemple).

Le théorème qui suit couvre les exemples précédents et bien plus encore.

Notre premier exemple correspond à la partie (iii) avec $\alpha = 2$ (ou au corollaire).

Le second exemple correspond à (iv) avec α arbitrairement grand (ou $\alpha = \infty$) et $\beta = 1$ (puisque $\mathbb{P}[\Delta^2 > 0] = O(1/\delta)$ dans cet exemple).

Théorème: Soit $\Delta = X_2 - X_1 = f(\theta_1 + \delta, \mathbf{U}_2) - f(\theta_1, \mathbf{U}_1)$.

(i) Si $\mathbf{U}_1$ et $\mathbf{U}_2$ sont indépendants et si

$$\sup_{\theta \in \Theta} \mathrm{Var}[f(\theta, \mathbf{U})] < \infty,$$

alors $\mathrm{Var}[\Delta] = \Theta(1)$ lorsque $\delta \to 0$.

(ii) Supposons que $\mathbf{U}_1 = \mathbf{U}_2 = \mathbf{U}$ et

$$K_z(\delta, q) \stackrel{\text{def}}{=} \sup_{\theta_1, \theta_2 \in \Theta} \mathbb{E}[\Delta^{2q}] \leq \delta^{\nu} K_0$$

pour $q > 1$, $\nu \geq 0$, et $K_0 < \infty$. S'il existe une v.a. Γ, indépendante de θ_1 et δ, et des constantes α, β, K_δ, et K_γ telles que $\mathbb{E}[\Gamma^2] \leq K_\gamma$ et

$$p(\delta, \alpha) \stackrel{\text{def}}{=} \sup_{\theta_1, \theta_2 \in \Theta} \mathbb{P}\{\Delta^2 > \Gamma^2 \delta^{\alpha}\} \leq K_\delta \delta^{\beta},$$

alors

$$\mathrm{Var}[\Delta] \leq K_\gamma \delta^{\alpha} + K_z(\delta, q)^{1/q} (K_\delta \delta^{\beta})^{1-1/q} = O(\delta^{\alpha} + \delta^{[\nu + \beta(q-1)]/q}).$$

(iii) Soient $\mathbf{U}_1 = \mathbf{U}_2 = \mathbf{U}$. S'il existe une v.a. Γ, indépendante de θ_1 et δ, une constante K_γ telle que $\mathbb{E}[\Gamma^2] \leq K_\gamma$, et une constante α telle que

$$\sup_{\theta_1, \theta_2 \in \Theta} \mathbb{P}\{\Delta^2 > \Gamma^2 \delta^\alpha\} = 0,$$

alors

$$\mathrm{Var}[\Delta] \leq K_\gamma \delta^\alpha = O(\delta^\alpha).$$

(iii) Soient $\mathbf{U}_1 = \mathbf{U}_2 = \mathbf{U}$. S'il existe une v.a. Γ, indépendante de θ_1 et δ, une constante K_γ telle que $\mathbb{E}[\Gamma^2] \leq K_\gamma$, et une constante α telle que

$$\sup_{\theta_1,\theta_2\in\Theta} \mathbb{P}\{\Delta^2 > \Gamma^2\delta^\alpha\} = 0,$$

alors

$$\mathrm{Var}[\Delta] \leq K_\gamma\delta^\alpha = O(\delta^\alpha).$$

(iv) Soient $\mathbf{U}_1 = \mathbf{U}_2 = \mathbf{U}$. S'il existe une constante K_h telle que

$$\mathbb{P}[|\Delta| \leq K_h \text{ pour tous } \theta_1, \theta_1 + \delta \in \Theta] = 1,$$

et si $p(\delta, \alpha) \leq K_\delta\delta^\beta$, alors

$$\mathrm{Var}[\Delta] \leq K_\gamma\delta^\alpha + K_h^2 K_\delta\delta^\beta = O(\delta^\alpha + \delta^\beta).$$

Corollaire. Supposons qu'avec probabilité 1, $f(\cdot, \mathbf{U})$ est continue dans Θ et différentiable dans $D(\mathbf{U}) \subseteq \Theta$, où $\Theta \setminus D(\mathbf{U})$ est dénombrable, et

$$\sup_{\theta \in D(\mathbf{U})} |f'(\theta, \mathbf{U})| \leq \Gamma(\mathbf{U}),$$

où $\Gamma = \Gamma(\mathbf{U})$ est une v.a. indépendante de θ telle que $\mathbb{E}[\Gamma^2] \leq K_\gamma < \infty$.

Corollaire. Supposons qu'avec probabilité 1, $f(\cdot, \mathbf{U})$ est continue dans Θ et différentiable dans $D(\mathbf{U}) \subseteq \Theta$, où $\Theta \setminus D(\mathbf{U})$ est dénombrable, et

$$\sup_{\theta \in D(\mathbf{U})} |f'(\theta, \mathbf{U})| \leq \Gamma(\mathbf{U}),$$

où $\Gamma = \Gamma(\mathbf{U})$ est une v.a. indépendante de θ telle que $\mathbb{E}[\Gamma^2] \leq K_\gamma < \infty$. Alors $\mathrm{Var}[\Delta/\delta] \leq K_\gamma$.

Corollaire. Supposons qu'avec probabilité 1, $f(\cdot, \mathbf{U})$ est continue dans Θ et différentiable dans $D(\mathbf{U}) \subseteq \Theta$, où $\Theta \setminus D(\mathbf{U})$ est dénombrable, et

$$\sup_{\theta \in D(\mathbf{U})} |f'(\theta, \mathbf{U})| \leq \Gamma(\mathbf{U}),$$

où $\Gamma = \Gamma(\mathbf{U})$ est une v.a. indépendante de θ telle que $\mathbb{E}[\Gamma^2] \leq K_\gamma < \infty$. Alors $\mathrm{Var}[\Delta/\delta] \leq K_\gamma$.

Preuve: Par le théorème de la valeur moyenne,

$$\frac{\Delta}{\delta} = \frac{|f(\theta_1 + \delta, \mathbf{U}) - f(\theta_1, \mathbf{U})|}{\delta} \leq \sup_{\theta \in [\theta_1, \theta_2] \cap D(\mathbf{U})} |f'(\theta, \mathbf{U})| \leq \Gamma(\mathbf{U}),$$

avec probabilité 1.

Corollaire. Supposons qu'avec probabilité 1, $f(\cdot, \mathbf{U})$ est continue dans Θ et différentiable dans $D(\mathbf{U}) \subseteq \Theta$, où $\Theta \setminus D(\mathbf{U})$ est dénombrable, et

$$\sup_{\theta \in D(\mathbf{U})} |f'(\theta, \mathbf{U})| \leq \Gamma(\mathbf{U}),$$

où $\Gamma = \Gamma(\mathbf{U})$ est une v.a. indépendante de θ telle que $\mathbb{E}[\Gamma^2] \leq K_\gamma < \infty$.
Alors $\mathrm{Var}[\Delta/\delta] \leq K_\gamma$.

Preuve: Par le théorème de la valeur moyenne,

$$\frac{\Delta}{\delta} = \frac{|f(\theta_1 + \delta, \mathbf{U}) - f(\theta_1, \mathbf{U})|}{\delta} \leq \sup_{\theta \in [\theta_1, \theta_2] \cap D(\mathbf{U})} |f'(\theta, \mathbf{U})| \leq \Gamma(\mathbf{U}),$$

avec probabilité 1. Il s'ensuit que

$$\mathrm{Var}[\Delta] \leq \mathbb{E}[\Delta^2] \leq \mathbb{E}[\Gamma^2(\mathbf{U})\delta^2]$$

Corollaire. Supposons qu'avec probabilité 1, $f(\cdot, \mathbf{U})$ est continue dans Θ et différentiable dans $D(\mathbf{U}) \subseteq \Theta$, où $\Theta \setminus D(\mathbf{U})$ est dénombrable, et

$$\sup_{\theta \in D(\mathbf{U})} |f'(\theta, \mathbf{U})| \leq \Gamma(\mathbf{U}),$$

où $\Gamma = \Gamma(\mathbf{U})$ est une v.a. indépendante de θ telle que $\mathbb{E}[\Gamma^2] \leq K_\gamma < \infty$. Alors $\mathrm{Var}[\Delta/\delta] \leq K_\gamma$.

Preuve: Par le théorème de la valeur moyenne,

$$\frac{\Delta}{\delta} = \frac{|f(\theta_1 + \delta, \mathbf{U}) - f(\theta_1, \mathbf{U})|}{\delta} \leq \sup_{\theta \in [\theta_1, \theta_2] \cap D(\mathbf{U})} |f'(\theta, \mathbf{U})| \leq \Gamma(\mathbf{U}),$$

avec probabilité 1. Il s'ensuit que

$$\mathrm{Var}[\Delta] \leq \mathbb{E}[\Delta^2] \leq \mathbb{E}[\Gamma^2(\mathbf{U})\delta^2] \leq K_\gamma \delta^2 = O(\delta^2). \quad \square$$

Corollaire. Supposons qu'avec probabilité 1, $f(\cdot, \mathbf{U})$ est continue dans Θ et différentiable dans $D(\mathbf{U}) \subseteq \Theta$, où $\Theta \setminus D(\mathbf{U})$ est dénombrable, et

$$\sup_{\theta \in D(\mathbf{U})} |f'(\theta, \mathbf{U})| \le \Gamma(\mathbf{U}),$$

où $\Gamma = \Gamma(\mathbf{U})$ est une v.a. indépendante de θ telle que $\mathbb{E}[\Gamma^2] \le K_\gamma < \infty$. Alors $\mathrm{Var}[\Delta/\delta] \le K_\gamma$.

Preuve: Par le théorème de la valeur moyenne,

$$\frac{\Delta}{\delta} = \frac{|f(\theta_1 + \delta, \mathbf{U}) - f(\theta_1, \mathbf{U})|}{\delta} \le \sup_{\theta \in [\theta_1, \theta_2] \cap D(\mathbf{U})} |f'(\theta, \mathbf{U})| \le \Gamma(\mathbf{U}),$$

avec probabilité 1. Il s'ensuit que

$$\mathrm{Var}[\Delta] \le \mathbb{E}[\Delta^2] \le \mathbb{E}[\Gamma^2(\mathbf{U})\delta^2] \le K_\gamma \delta^2 = O(\delta^2). \quad \square$$

Dans ce cas, $\mathrm{Var}[\Delta/\delta] = O(1)$ et on peut prendre la limite quand $\delta \to 0$ comme estimateur de la dérivée!

Example: File $GI/G/1$.
Dist. de durée de service G_θ dépend de $\theta \in \Theta$. Les temps d'attente obéissent à l'équation de Lindley: $W_1(\theta) = 0$ et

$$W_{j+1}(\theta) = \max(0,\ W_j(\theta) + S_j(\theta) - A_j),$$

Example: File $GI/G/1$.
Dist. de durée de service G_θ dépend de $\theta \in \Theta$. Les temps d'attente obéissent à l'équation de Lindley: $W_1(\theta) = 0$ et

$$W_{j+1}(\theta) = \max(0,\ W_j(\theta) + S_j(\theta) - A_j),$$

Soit $f(\theta, \mathbf{U}) = \bar{W}_t(\theta) = \frac{1}{t}\sum_{j=1}^{t} W_j(\theta)$ et $\mu(\theta) = w_t(\theta) = E_\theta[\bar{W}_t(\theta)]$.
On a $\Delta = \bar{W}_t(\theta_2) - \bar{W}_t(\theta_1)$.

Example: File $GI/G/1$.
Dist. de durée de service G_θ dépend de $\theta \in \Theta$. Les temps d'attente obéissent à l'équation de Lindley: $W_1(\theta) = 0$ et

$$W_{j+1}(\theta) = \max(0,\ W_j(\theta) + S_j(\theta) - A_j),$$

Soit $f(\theta, \mathbf{U}) = \bar{W}_t(\theta) = \frac{1}{t}\sum_{j=1}^{t} W_j(\theta)$ et $\mu(\theta) = w_t(\theta) = E_\theta[\bar{W}_t(\theta)]$.
On a $\Delta = \bar{W}_t(\theta_2) - \bar{W}_t(\theta_1)$.

Essayons d'appliquer le corollaire.
En supposant qu'on utilise l'inversion, $f(\theta, \mathbf{U})$ est continue en θ ssi $S_j(\theta) = G_\theta^{-1}(U_{2j})$ est une fonction continue de θ pour tout U_{2j}.

Example: File $GI/G/1$.
Dist. de durée de service G_θ dépend de $\theta \in \Theta$. Les temps d'attente obéissent à l'équation de Lindley: $W_1(\theta) = 0$ et

$$W_{j+1}(\theta) = \max(0,\ W_j(\theta) + S_j(\theta) - A_j),$$

Soit $f(\theta, \mathbf{U}) = \bar{W}_t(\theta) = \frac{1}{t}\sum_{j=1}^{t} W_j(\theta)$ et $\mu(\theta) = w_t(\theta) = E_\theta[\bar{W}_t(\theta)]$.
On a $\Delta = \bar{W}_t(\theta_2) - \bar{W}_t(\theta_1)$.

Essayons d'appliquer le corollaire.
En supposant qu'on utilise l'inversion, $f(\theta, \mathbf{U})$ est continue en θ ssi $S_j(\theta) = G_\theta^{-1}(U_{2j})$ est une fonction continue de θ pour tout U_{2j}.

Vrai si θ est un paramètre d'échelle ou de localisation, par exemple.
Paramètre d'échelle: $S_j(\theta) = \theta\, G_1^{-1}(U_{2j})$. Ex.: $S_j(\theta) = -\theta \ln(1 - U_{2j})$.

Example: File $GI/G/1$.
Dist. de durée de service G_θ dépend de $\theta \in \Theta$. Les temps d'attente obéissent à l'équation de Lindley: $W_1(\theta) = 0$ et

$$W_{j+1}(\theta) = \max(0,\ W_j(\theta) + S_j(\theta) - A_j),$$

Soit $f(\theta, \mathbf{U}) = \bar{W}_t(\theta) = \frac{1}{t}\sum_{j=1}^{t} W_j(\theta)$ et $\mu(\theta) = w_t(\theta) = E_\theta[\bar{W}_t(\theta)]$.
On a $\Delta = \bar{W}_t(\theta_2) - \bar{W}_t(\theta_1)$.

Essayons d'appliquer le corollaire.
En supposant qu'on utilise l'inversion, $f(\theta, \mathbf{U})$ est continue en θ ssi $S_j(\theta) = G_\theta^{-1}(U_{2j})$ est une fonction continue de θ pour tout U_{2j}.

Vrai si θ est un paramètre d'échelle ou de localisation, par exemple.
Paramètre d'échelle: $S_j(\theta) = \theta\, G_1^{-1}(U_{2j})$. Ex.: $S_j(\theta) = -\theta\, \ln(1 - U_{2j})$.

La dérivée de $f(\theta, \mathbf{U})$ est

$$f'(\theta, \mathbf{U}) = \frac{1}{t}\sum_{j=1}^{t} W_j'(\theta),$$

où $W_1'(\theta) = 0$ et

$$W_{j+1}'(\theta) = \begin{cases} W_j'(\theta) + S_j'(\theta) & \text{if } W_j(\theta) + S_j(\theta) - A_j > 0, \\ 0 & \text{if } W_j(\theta) + S_j(\theta) - A_j < 0 \end{cases}$$

où $W_1'(\theta) = 0$ et

$$W_{j+1}'(\theta) = \begin{cases} W_j'(\theta) + S_j'(\theta) & \text{if } W_j(\theta) + S_j(\theta) - A_j > 0, \\ 0 & \text{if } W_j(\theta) + S_j(\theta) - A_j < 0 \end{cases}$$

où $S_j'(\theta) = \partial G_\theta^{-1}(U)/\partial\theta$.

où $W_1'(\theta) = 0$ et

$$W'_{j+1}(\theta) = \begin{cases} W'_j(\theta) + S'_j(\theta) & \text{if } W_j(\theta) + S_j(\theta) - A_j > 0, \\ 0 & \text{if } W_j(\theta) + S_j(\theta) - A_j < 0 \end{cases}$$

où $S'_j(\theta) = \partial G_\theta^{-1}(U)/\partial\theta$.
La dérivée n'existe pas aux valeurs de θ où $W_j(\theta) + S_j(\theta) - A_j = 0$ pour un j (la fonction a un angle), mais ce n'est qu'un nombre fini de points.

où $W_1'(\theta) = 0$ et

$$W_{j+1}'(\theta) = \begin{cases} W_j'(\theta) + S_j'(\theta) & \text{if } W_j(\theta) + S_j(\theta) - A_j > 0, \\ 0 & \text{if } W_j(\theta) + S_j(\theta) - A_j < 0 \end{cases}$$

où $S_j'(\theta) = \partial G_\theta^{-1}(U)/\partial\theta$.
La dérivée n'existe pas aux valeurs de θ où $W_j(\theta) + S_j(\theta) - A_j = 0$ pour un j (la fonction a un angle), mais ce n'est qu'un nombre fini de points. On a

$$f'(\theta, \mathbf{U}) \leq \frac{1}{t} \sum_{j=1}^{t} \sum_{\ell=1}^{j-1} S_j'(\theta).$$

où $W_1'(\theta) = 0$ et

$$W_{j+1}'(\theta) = \begin{cases} W_j'(\theta) + S_j'(\theta) & \text{if } W_j(\theta) + S_j(\theta) - A_j > 0, \\ 0 & \text{if } W_j(\theta) + S_j(\theta) - A_j < 0 \end{cases}$$

où $S_j'(\theta) = \partial G_\theta^{-1}(U)/\partial\theta$.
La dérivée n'existe pas aux valeurs de θ où $W_j(\theta) + S_j(\theta) - A_j = 0$ pour un j (la fonction a un angle), mais ce n'est qu'un nombre fini de points. On a

$$f'(\theta, \mathbf{U}) \le \frac{1}{t} \sum_{j=1}^{t} \sum_{\ell=1}^{j-1} S_j'(\theta).$$

Il suffit donc qu'il y ait une v.a. $\Gamma = \Gamma(U)$ de variance finie telle que

$$\sup_{\theta\in\Theta} |\partial G_\theta^{-1}(U)/\partial\theta| \le \Gamma$$

et on aura alors $\mathrm{Var}[\Delta] = O(\delta^2)$.

Cas concret: file $M/M/1$, taux d'arrivée $= 1$,
durées de service exponentielles de moyenne θ.
On a $S_j(\theta) = -\theta \ln(1 - U_j)$ et $S'_j(\theta) = -\ln(1 - U_j) = \Gamma(U_j)$.

Cas concret: file $M/M/1$, taux d'arrivée $= 1$,
durées de service exponentielles de moyenne θ.
On a $S_j(\theta) = -\theta \ln(1 - U_j)$ et $S'_j(\theta) = -\ln(1 - U_j) = \Gamma(U_j)$.

Expérience numérique: $t = 50$, $\theta_1 = 0.5$, $\theta_2 = \theta_1 + \delta$, $n = 10000$.

Cas concret: file $M/M/1$, taux d'arrivée $= 1$,
durées de service exponentielles de moyenne θ.
On a $S_j(\theta) = -\theta \ln(1 - U_j)$ et $S'_j(\theta) = -\ln(1 - U_j) = \Gamma(U_j)$.

Expérience numérique: $t = 50$, $\theta_1 = 0.5$, $\theta_2 = \theta_1 + \delta$, $n = 10000$.

Méthode	$\delta = 0.01$		$\delta = 0.001$		$\delta = 0.0001$	
	$\bar{\Delta}_n$	$\widehat{\mathrm{Var}}[\Delta]$	$\bar{\Delta}_n$	$\widehat{\mathrm{Var}}[\Delta]$	$\bar{\Delta}_n$	$\widehat{\mathrm{Var}}[\Delta]$
VAI	1.9E-1	2.3E-1	-4.7E-3	2.2E-1	-7.1E-3	2.2E-1
VAC	2.6E-2	5.3E-4	2.6E-3	4.8E-6	2.6E-4	4.8E-8

Cas concret: file $M/M/1$, taux d'arrivée $= 1$,
durées de service exponentielles de moyenne θ.
On a $S_j(\theta) = -\theta \ln(1 - U_j)$ et $S'_j(\theta) = -\ln(1 - U_j) = \Gamma(U_j)$.

Expérience numérique: $t = 50$, $\theta_1 = 0.5$, $\theta_2 = \theta_1 + \delta$, $n = 10000$.

Méthode	$\delta = 0.01$		$\delta = 0.001$		$\delta = 0.0001$	
	$\bar{\Delta}_n$	$\widehat{\mathrm{Var}}[\Delta]$	$\bar{\Delta}_n$	$\widehat{\mathrm{Var}}[\Delta]$	$\bar{\Delta}_n$	$\widehat{\mathrm{Var}}[\Delta]$
VAI	1.9E-1	2.3E-1	-4.7E-3	2.2E-1	-7.1E-3	2.2E-1
VAC	2.6E-2	5.3E-4	2.6E-3	4.8E-6	2.6E-4	4.8E-8

On peut aussi calculer directement la dérivée échantillonnale

$$f'(\theta, \mathbf{U}) = \frac{1}{t} \sum_{j=1}^{t} W'_j(\theta),$$

comme estimateur de la dérivée $\partial \mathbb{E}[f(\theta, \mathbf{U})]/\partial\theta$.
Une seule simulation suffit pour cela!

Considérons maintenant plutôt le nombre de clients dont la durée de séjour dépasse une constante K donnée:

$$f(\theta, \mathbf{U}) = \sum_{j=1}^{t} \mathbb{I}[W_j(\theta) + S_j(\theta) > K].$$

Considérons maintenant plutôt le nombre de clients dont la durée de séjour dépasse une constante K donnée:

$$f(\theta, \mathbf{U}) = \sum_{j=1}^{t} \mathbb{I}[W_j(\theta) + S_j(\theta) > K].$$

Cette fonction n'est pas continue en θ et est constante par morceaux. Donc le corollaire ne s'applique pas.

Considérons maintenant plutôt le nombre de clients dont la durée de séjour dépasse une constante K donnée:

$$f(\theta, \mathbf{U}) = \sum_{j=1}^{t} \mathbb{I}[W_j(\theta) + S_j(\theta) > K].$$

Cette fonction n'est pas continue en θ et est constante par morceaux. Donc le corollaire ne s'applique pas.

Mais on peut appliquer la partie (iv) du théorème pour montrer que $\mathrm{Var}[\Delta] = O(\delta)$, si G_θ se comporte bien.
On peut montrer que $\mathbb{P}[\Delta > 0] = O(\delta)$ ($\alpha = \infty$ et $\beta = 1$ dans le théorème).

Considérons maintenant plutôt le nombre de clients dont la durée de séjour dépasse une constante K donnée:

$$f(\theta, \mathbf{U}) = \sum_{j=1}^{t} \mathbb{I}[W_j(\theta) + S_j(\theta) > K].$$

Cette fonction n'est pas continue en θ et est constante par morceaux. Donc le corollaire ne s'applique pas.

Mais on peut appliquer la partie (iv) du théorème pour montrer que $\mathrm{Var}[\Delta] = O(\delta)$, si G_θ se comporte bien.
On peut montrer que $\mathbb{P}[\Delta > 0] = O(\delta)$ ($\alpha = \infty$ et $\beta = 1$ dans le théorème).

Expérience pour le cas exponentiel, avec $K = 2$.

Considérons maintenant plutôt le nombre de clients dont la durée de séjour dépasse une constante K donnée:

$$f(\theta, \mathbf{U}) = \sum_{j=1}^{t} \mathbb{I}[W_j(\theta) + S_j(\theta) > K].$$

Cette fonction n'est pas continue en θ et est constante par morceaux. Donc le corollaire ne s'applique pas.

Mais on peut appliquer la partie (iv) du théorème pour montrer que $\mathrm{Var}[\Delta] = O(\delta)$, si G_θ se comporte bien.
On peut montrer que $\mathbb{P}[\Delta > 0] = O(\delta)$ ($\alpha = \infty$ et $\beta = 1$ dans le théorème).

Expérience pour le cas exponentiel, avec $K = 2$.

Méthode	$\delta = 0.01$		$\delta = 0.001$		$\delta = 0.0001$	
	$\bar{\Delta}_n$	$\widehat{\mathrm{Var}}[\Delta_i]$	$\bar{\Delta}_n$	$\widehat{\mathrm{Var}}[\Delta_i]$	$\bar{\Delta}_n$	$\widehat{\mathrm{Var}}[\Delta_i]$
VAI	0.40	66	-0.05	63	-0.09	63
VAC	4.9E-1	6.3E-1	4.9E-2	5.2E-2	5.4E-3	5.3E-3

Exemple du centre d'appels: On montre que $\mathrm{Var}[\Delta] = O(\delta^{1+\epsilon})$ via la partie (ii) du théorème.

Exemple du centre d'appels: On montre que $\mathrm{Var}[\Delta] = O(\delta^{1+\epsilon})$ via la partie (ii) du théorème.

Réseau d'activités stochastique:
Pour estimer un changement sur $\mathbb{E}[T]$: $\mathrm{Var}[\Delta] = O(\delta^2)$.
Pour estimer un changement sur $\mathbb{P}[T > x]$: $\mathrm{Var}[\Delta] = O(\delta)$.

Exemple du centre d'appels: On montre que $\mathrm{Var}[\Delta] = O(\delta^{1+\epsilon})$ via la partie (ii) du théorème.

Réseau d'activités stochastique:
Pour estimer un changement sur $\mathbb{E}[T]$: $\mathrm{Var}[\Delta] = O(\delta^2)$.
Pour estimer un changement sur $\mathbb{P}[T > x]$: $\mathrm{Var}[\Delta] = O(\delta)$.

Évaluation d'options: sans barrière, avec barrière.
Important pour l'évaluation de la sensibilité (les "greeks").

VAC pour estimer une fonction de plusieurs espérances.

On veut estimer $g(\boldsymbol{\mu}) = g(\mu_1, \ldots, \mu_d)$ par $g(\bar{\mathbf{X}}_n)$.

VAC pour estimer une fonction de plusieurs espérances.

On veut estimer $g(\boldsymbol{\mu}) = g(\mu_1, \ldots, \mu_d)$ par $g(\bar{\mathbf{X}}_n)$.

Exemples: $g(\mu_1, \mu_2) = \mu_2 - \mu_1$, $g(\mu_1, \mu_2) = \mu_1/\mu_2$, etc.

VAC pour estimer une fonction de plusieurs espérances.

On veut estimer $g(\boldsymbol{\mu}) = g(\mu_1, \ldots, \mu_d)$ par $g(\bar{\mathbf{X}}_n)$.

Exemples: $g(\mu_1, \mu_2) = \mu_2 - \mu_1$, $g(\mu_1, \mu_2) = \mu_1/\mu_2$, etc.

Si $\sqrt{n}(\bar{\mathbf{X}}_n - \boldsymbol{\mu}) \Rightarrow N(\mathbf{0}, \boldsymbol{\Sigma}_x)$ alors $\sqrt{n}(g(\bar{\mathbf{X}}_n) - g(\boldsymbol{\mu}))/\sigma_{\mathrm{g}} \Rightarrow N(0,1)$ où

$$\sigma_{\mathrm{g}}^2 = (\nabla g(\boldsymbol{\mu}))^{\mathsf{t}} \boldsymbol{\Sigma}_x \nabla g(\boldsymbol{\mu}).$$

VAC pour estimer une fonction de plusieurs espérances.

On veut estimer $g(\boldsymbol{\mu}) = g(\mu_1, \ldots, \mu_d)$ par $g(\bar{\mathbf{X}}_n)$.

Exemples: $g(\mu_1, \mu_2) = \mu_2 - \mu_1$, $g(\mu_1, \mu_2) = \mu_1/\mu_2$, etc.

Si $\sqrt{n}(\bar{\mathbf{X}}_n - \boldsymbol{\mu}) \Rightarrow N(\mathbf{0}, \boldsymbol{\Sigma}_x)$ alors $\sqrt{n}(g(\bar{\mathbf{X}}_n) - g(\boldsymbol{\mu}))/\sigma_{\mathrm{g}} \Rightarrow N(0, 1)$ où

$$\sigma_{\mathrm{g}}^2 = (\nabla g(\boldsymbol{\mu}))^{\mathsf{t}} \boldsymbol{\Sigma}_x \nabla g(\boldsymbol{\mu}).$$

Si les $\mathbf{X}_i$ ont leurs coordonnées indépendantes, alors $\boldsymbol{\Sigma}_x$ est une matrice diagonale qui contient les variances des coordonnées de $\mathbf{X}_i$.

VAC pour estimer une fonction de plusieurs espérances.

On veut estimer $g(\boldsymbol{\mu}) = g(\mu_1, \ldots, \mu_d)$ par $g(\bar{\mathbf{X}}_n)$.

Exemples: $g(\mu_1, \mu_2) = \mu_2 - \mu_1$, $g(\mu_1, \mu_2) = \mu_1/\mu_2$, etc.

Si $\sqrt{n}(\bar{\mathbf{X}}_n - \boldsymbol{\mu}) \Rightarrow N(\mathbf{0}, \boldsymbol{\Sigma}_x)$ alors $\sqrt{n}(g(\bar{\mathbf{X}}_n) - g(\boldsymbol{\mu}))/\sigma_{\mathrm{g}} \Rightarrow N(0,1)$ où

$$\sigma_{\mathrm{g}}^2 = (\nabla g(\boldsymbol{\mu}))^{\mathsf{t}} \boldsymbol{\Sigma}_x \nabla g(\boldsymbol{\mu}).$$

Si les $\mathbf{X}_i$ ont leurs coordonnées indépendantes, alors $\boldsymbol{\Sigma}_x$ est une matrice diagonale qui contient les variances des coordonnées de $\mathbf{X}_i$.

Les VAC réduisent la variance de $g(\bar{\mathbf{X}}_n)$ (asymptotiquement) ssi elles réduisent σ_{g}^2.

VAC pour estimer une fonction de plusieurs espérances.

On veut estimer $g(\boldsymbol{\mu}) = g(\mu_1, \ldots, \mu_d)$ par $g(\bar{\mathbf{X}}_n)$.

Exemples: $g(\mu_1, \mu_2) = \mu_2 - \mu_1$, $g(\mu_1, \mu_2) = \mu_1/\mu_2$, etc.

Si $\sqrt{n}(\bar{\mathbf{X}}_n - \boldsymbol{\mu}) \Rightarrow N(\mathbf{0}, \boldsymbol{\Sigma}_x)$ alors $\sqrt{n}(g(\bar{\mathbf{X}}_n) - g(\boldsymbol{\mu}))/\sigma_{\mathrm{g}} \Rightarrow N(0,1)$ où

$$\sigma_{\mathrm{g}}^2 = (\nabla g(\boldsymbol{\mu}))^{\mathsf{t}} \boldsymbol{\Sigma}_x \nabla g(\boldsymbol{\mu}).$$

Si les $\mathbf{X}_i$ ont leurs coordonnées indépendantes, alors $\boldsymbol{\Sigma}_x$ est une matrice diagonale qui contient les variances des coordonnées de $\mathbf{X}_i$.
Les VAC réduisent la variance de $g(\bar{\mathbf{X}}_n)$ (asymptotiquement) ssi elles réduisent σ_{g}^2.

Cas particulier: $g(\mu_1, \mu_2) = \mu_1/\mu_2 = \nu$. Estimateur: $\bar{X}_n/\bar{Y}_n$.

VAC pour estimer une fonction de plusieurs espérances.

On veut estimer $g(\boldsymbol{\mu}) = g(\mu_1, \ldots, \mu_d)$ par $g(\bar{\mathbf{X}}_n)$.

Exemples: $g(\mu_1, \mu_2) = \mu_2 - \mu_1$, $g(\mu_1, \mu_2) = \mu_1/\mu_2$, etc.

Si $\sqrt{n}(\bar{\mathbf{X}}_n - \boldsymbol{\mu}) \Rightarrow N(\mathbf{0}, \boldsymbol{\Sigma}_x)$ alors $\sqrt{n}(g(\bar{\mathbf{X}}_n) - g(\boldsymbol{\mu}))/\sigma_{\mathrm{g}} \Rightarrow N(0,1)$ où

$$\sigma_{\mathrm{g}}^2 = (\nabla g(\boldsymbol{\mu}))^{\mathsf{t}} \boldsymbol{\Sigma}_x \nabla g(\boldsymbol{\mu}).$$

Si les $\mathbf{X}_i$ ont leurs coordonnées indépendantes, alors $\boldsymbol{\Sigma}_x$ est une matrice diagonale qui contient les variances des coordonnées de $\mathbf{X}_i$.
Les VAC réduisent la variance de $g(\bar{\mathbf{X}}_n)$ (asymptotiquement) ssi elles réduisent σ_{g}^2.

Cas particulier: $g(\mu_1, \mu_2) = \mu_1/\mu_2 = \nu$. Estimateur: $\bar{X}_n/\bar{Y}_n$. La constante

$$\sigma_{\mathrm{g}}^2 = \frac{\mathrm{Var}[X_i] + \nu^2 \mathrm{Var}[Y_i] - 2\nu \mathrm{Cov}(X_i, Y_i)}{\mu_2^2}$$

est réduite ssi $\mathrm{Cov}(X_i, Y_i) > 0$.