Reccueil d'exercices - Quantification des incertitudes pour la simulation

G. Perrin

Table des matières

1 Test statistique pour la détection d'une maladie bovine

— un laboratoire propose un test de dépistage de la maladie de la vache folle.
— La notice précise la qualité du test : si le test est appliqué sur une vache malade, le test est positif dans 99.8% des cas. Si le test est appliqué sur une vache saine, le test est négatif dans 99.6% des cas.
— On sait d'autre part qu'il y a une vache malade sur 100 000.

Peut-on avoir confiance dans ce test ? Pour cela, on cherche la réponse à la question : si le test est positif, quelle est la probabilité que la vache soit malade ?

On note S pour "sain", M pour "malade", $+$ pour test positif, $-$ pour test négatif.

$$\begin{aligned}
\mathbb{P}(M|+) &= \frac{\mathbb{P}(+|M)\mathbb{P}(M)}{\mathbb{P}(+)} \\
&\frac{\mathbb{P}(+|M)\mathbb{P}(M)}{\mathbb{P}(+|M)\mathbb{P}(M) + \mathbb{P}(+|S)\mathbb{P}(S)} \\
&\frac{\mathbb{P}(+|M)\mathbb{P}(M)}{\mathbb{P}(+|M)\mathbb{P}(M) + (1-\mathbb{P}(-|S))(1-\mathbb{P}(M))} \\
&\frac{0.998 \times 10^{-5}}{0.998 \times 10^{-5} + (1-0.996)(1-10^{-5})} \\
&\approx \frac{10^{-5}}{4 \times 10^{-3}} \approx 0.25\%.
\end{aligned}$$

2 Quantification robuste des incertitudes

1. On dispose d'une règle équilibrée de longueur 2L, centrée en $x = 0$, ainsi que d'un kilogramme de sable. Comment disposer le sable sur la règle afin d'avoir le maximum de sable sur son bord droit, à une distance supérieure à t de son centre ?

   On place m en t et $1-m$ en $-L$. Par équilibre des moments, on déduit :

$$tm = L(1-m) \;\Rightarrow\; m = L/(L+t).$$

2. Soit X une v.a. positive ($X \geq 0$) et de PDF p_X. La seule information sur cette PDF est la fait que la moyenne de X soit égale à m. Calculer alors :

$$\max_{p_X,\ \mathbb{E}[X]=m} \mathbb{P}(X \geq t).$$

On distinguera les cas $t \leq m$ et $t \geq m$.

Si $m > t$, $\max_{p_X,\ \mathbb{E}[X]=m} \mathbb{P}(X \geq t) = 1$ (on met une distribution dont le support est au dessus de t.

Pour $m \leq t$, par analogie avec la précédente question, on déduit que la solution est donnée par (δ_x étant la loi Dirac en x) :

$$p_X(x) = \alpha\delta_t(x) + (1-\alpha)\delta_0(x).$$

On trouve alors α en respectant la contrainte sur la moyenne :

$$m = \alpha \times t$$

On conclut : $\max_{p_X,\ \mathbb{E}[X]=m} \mathbb{P}(X \geq t) = \alpha = m/t$.

3 Positionnement par mesures sonores

En mars 1918, Paris est frappée par des projectiles qui semblent venir de nulle part. Le canon responsable de cette attaque est un des plus puissants et des plus mystérieux canons de l'époque, capable d'attendre des cibles à plus de 100km. Après la surprise, on s'organise pour mettre en place des méthodes de localisation de ce canon. Et pour de telles distances, c'est par le son qu'il est le plus sÃ»r et le plus efficace de repérer ce canon.

On assimile le monde au plan $(\boldsymbol{O}, \boldsymbol{e}_x, \boldsymbol{e}_y)$, tel que tout point $\boldsymbol{p}$ du plan s'écrit $= p_x\boldsymbol{e}_x + p_y\boldsymbol{e}_y$, ou plus simplement $\boldsymbol{p} = (p_x, p_y)$. On note alors :

- $\boldsymbol{s} = (s_x, s_y)$ la position **inconnue** du canon,
- $\tau > 0$ l'instant auquel le canon tire son projectile
- $\{\boldsymbol{z}_n = (z_x, z_y),\ 1 \leq n \leq N\}$ les positions de N observateurs, dont la mission est de localiser le canon,
- v la vitesse du son.

Les observateurs sont positionnés sur un cercle de rayon $R = 1000m$, centré en $\boldsymbol{O} = (0,0)$. Chaque observateur lance son chronomètre à $t = 0$, et on note t_n le temps auquel l'observateur $\boldsymbol{z}_n$ arrête son chronomètre car il a entendu le coup de canon.

1. Justifier le fait que la position du canon et le temps τ soient solutions du système d'équations suivant :

$$\begin{cases} ||\boldsymbol{z}_1 - \boldsymbol{s}||^2 = (v(t_1 - \tau))^2, \\ \vdots \\ ||\boldsymbol{z}_N - \boldsymbol{s}||^2 = (v(t_N - \tau))^2. \end{cases} \tag{1}$$

direct par équations des distances

2. En déduire que le vecteur $\boldsymbol{u} = (s_x, s_y, \tau)$ est solution du système linéaire suivant :

$$[m]\boldsymbol{u} = \boldsymbol{f}. \tag{2}$$

Déterminer la matrice $[m]$ et le vecteur $\boldsymbol{f}$.

$$[m] = 2\times \begin{bmatrix} x_1 - x_N & y_1 - y_N & v^2(\tau_N - \tau_1) \\ \vdots & \vdots & \vdots \\ x_{N-1} - x_N & y_{N-1} - y_N & v^2(\tau_N - \tau_{N-1}) \end{bmatrix}, \quad \boldsymbol{f} = \begin{pmatrix} x_1^2 - x_N^2 + y_1^2 - y_N^2 + v^2(\tau_N^2 - \tau_1^2) \\ \vdots \\ x_{N-1}^2 - x_N^2 + y_{N-1}^2 - y_N^2 + v^2(\tau_N^2 - \tau_{N-1}^2) \end{pmatrix} \tag{3}$$

3. Sans incertitudes sur les données du problème, indiquer le nombre minimal d'observateurs pour que le système défini par l'Eq. (1) admette une unique solution (on pourra effectuer un dessin pour justifier ce nombre).

Il faut au moins 3 observateurs en 2D.

4*. Lister les potentielles sources d'incertitudes pouvant affecter cette localisation, en les classant, selon vous et en expliquant pourquoi, des plus influentes aux moins influentes. Décrire ensuite succintement comment en déduire la position moyenne du canon.

- en 1 : le temps d'arrivée - en 2 : la vitesse du son - en 3 : ... - en ... : les positions des observateurs

4 Mise à la puissance et distribution associée

Soit X une variable uniformément distribuée sur $[0, 1]$. Calculer et représenter sur un même graphique les PDF de X^k pour $k = 1, 2, 3, 4$.

— $\mathbb{P}(X^k \leq z) = \mathbb{P}(X \leq z^{1/k}) = z^{1/k} \Rightarrow f_{X^k}(z) = z^{1/k-1}/k$.
— On trace alors pour $k = 1 : 4$ en cherchant les passages par 1.

5 Jeu statistique

On considère l'expérience aléatoire suivante :
— On tire une valeur, $X(\theta)$, d'une variable aléatoire X uniformément distribuée sur $[0, 1]$.
— On écrit alors cette valeur sur le verso, et son carré, $X(\theta)^2$, sur le recto d'une même feuille.
— Seule une des faces (choisie de manière purement aléatoire) de cette feuille est ensuite rendue visible. On nomme alors Y cette valeur visible.

Deux choix sont alors proposés : soit accepter le montant observé, soit préférer le montant inscrit de l'autre côté de la feuille. Comment définir une stratégie pour maximiser son gain ? En particulier, que faire face aux valeurs $\{0.1\ ;\ 0.24\ ;\ 0.29\ ;\ 0.35\ ;\ 0.5\ ;\ 0.75\}$?

— Intuition :
— $\mathbb{P}(X = Y) = 1$, $\mathbb{P}(X^2 = Y) = 1/(2\sqrt{Y})$
— $G(Y|Y = X) = Y - Y^2$, $G(Y|Y = X^2) = \sqrt{Y} - Y$.

— On retourne si $G(Y|Y = X^2)\mathbb{P}(X^2 = Y) \geq G(Y|Y = X)\mathbb{P}(X = Y) \Rightarrow (\sqrt{Y} - Y)/(2\sqrt{Y}) \geq Y - Y^2 \Rightarrow Y \geq Y \leq 0.3195$.

6 Loi d'un échantillon trié

Soit X une variable aléatoire de PDF f_X et de CDF F_X. Soient $X_1, \ldots, X_N$ N copies indépendantes et de mêmes lois que X. On note alors $Y_1, \ldots, Y_N$ les variables aléatoires associées au réarrangement par ordre croissant de $X_1, \ldots, X_N$:

$$Y_1 \leq Y_2 \leq \ldots \leq Y_N.$$

1. Calculer la loi de $Y_1 = \min_{1 \leq n \leq N} Y_n$ en fonction de f_X et F_X.

$$\begin{aligned}
\mathbb{P}(Y_1 \leq y) &= 1 - \mathbb{P}(Y_1 \geq y) \\
&= 1 - \mathbb{P}(X_1 \geq y, \ldots, X_N \geq y) \\
&= 1 - \mathbb{P}(X \geq y)^N \\
&= 1 - (1 - F_X(y))^n.
\end{aligned}$$

On déduit : $f_{Y_1}(y) = n f_X(y)(1 - F_X(y))^{n-1}$.

2. Calculer de même les lois de $Y_2, \ldots, Y_N$.

$$\begin{aligned}
\mathbb{P}(Y_j \leq y) &= \mathbb{P}(\text{"au moins j valeurs au dessous de y"}) \\
&= \sum_{k=j}^{N} \mathbb{P}(\text{"exactement k valeurs au dessous de y"}) \\
&= \sum_{k=j}^{N} \binom{N}{k} F_X(y)^k (1 - F_X(y))^{N-k} \\
&= \sum_{k=j}^{N-1} \binom{N}{k} F_X(y)^k (1 - F_X(y))^{N-k} + F_X(y)^N.
\end{aligned}$$

On déduit :

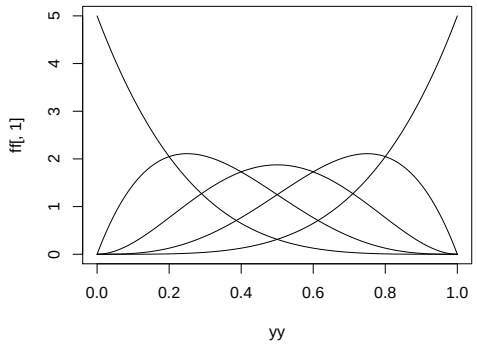
$$f_{Y_j}(y) = f_X(y) \left(N F_X(y)^{N-1} + \sum_{k=j}^{N-1} \binom{N}{k} \left(k F_X(y)^{k-1} (1 - F_X(y))^{N-k} - (N-k) F_X(y)^k (1 - F_X(y))^{N-k-1} \right) \right)$$

3. Calculer ces lois lorsque X est la loi uniforme.

Pour X uniforme, $F_X(y) = y$ et $f_X(y) = 1$. On déduit :

$$\begin{aligned}\mathbb{P}(Y_j \leq y) &= \sum_{k=j}^{N} \binom{N}{k} F_X(y)^k (1 - F_X(y))^{N-k} \\ &= \sum_{k=j}^{N} \binom{N}{k} y^k (1-y)^{N-k}.\end{aligned}$$

$$f_{Y_j}(y) = N y^{N-1} + \sum_{k=j}^{N-1} \binom{N}{k} \left(k y^{k-1}(1-y)^{N-k} + (N-k) y^k (1-y)^{N-k-1} \right).$$



7 Analyse du processus sinus

On définit ϕ une variable aléatoire uniformément distribuée sur $[0, 2\pi]$, et $\{X(t),\ t \in \mathbb{R}\}$ le processus aléatoire associé, tel que pour tout t :

$$X(t) = \sin(\omega t + \phi).$$

1. Calculer la moyenne $\mu(t) = \mathbb{E}[X(t)]$ de ce processus.

$\mu(t) = 0$

2. Calculer la fonction d'autocorrélation, $R(t, t') = \mathbb{E}\left[X(t)X(t')\right]$ de ce processus.

$$\begin{aligned}\mathbb{E}\left[X(t)X(t')\right] &= \frac{1}{2\pi} \int_0^{2\pi} \sin(\omega t + \phi) \sin(\omega t' + \phi) d\phi \\ &= \frac{1}{4\pi} \int_0^{2\pi} \left(\cos(\omega(t + t') + 2\phi) + \cos(\omega(t - t')) \right) d\phi \\ &= \frac{1}{2} \cos(\omega(t - t'))\end{aligned}$$

3. Que peut-on dire sur la stationnarité de X ?

Stationnaire d'ordre 2 : moyenne constante et covariance ne dépendant que de $|t - t'|$.

4. Pour t fixé, calculer la distribution de $X(t)$. (Pour cela, on passera par la fonction de répartition $F_{X(t)}(x) = \mathbb{P}(X(t) \leq x)$).

$$\begin{aligned}
\mathbb{P}(X(t) \leq x) &= \mathbb{P}(sin(\omega t + \phi) \leq x) \\
&= \mathbb{P}(sin(\phi) \leq x) \\
&= \mathbb{P}(sin(\Phi) \leq x), \text{ avec } \Phi \sim U[-\pi/2, \pi/2] \\
&= \mathbb{P}(\Phi \leq \text{asin}(x)) \\
&= \frac{1}{\pi}(\frac{\pi}{2} + \text{asin}(x))
\end{aligned}$$

On déduit : $f_{X(t)}(x) = \frac{1}{\pi\sqrt{1-x^2}}$.

5. Que peut-on dire sur le caractère gaussien de ce processus X ?

La distribution $f_{X(t)}$ n'est clairement pas gaussienne, donc le processus n'est pas gaussien.

8 Analyse du copule de Farlie-Gumbel-Morgenstern

On considère deux variables aléatoires X_1 et X_2 telles que pour $0 \leq x_1, x_2 \leq 1$, $0 \leq \theta \leq 1$:

$$\mathbb{P}(X_1 \leq x_1, X_2 \leq x_2) = x_1 x_2(1 + \theta(1 - x_1)(1 - x_2)). \tag{4}$$

1. Trouver les CDFs de X_1 et X_2, ainsi que la fonction de copule $C(x_1, x_2; \theta)$.

$$F_{X_1}(x_1) = \mathbb{P}(X_1 \leq x_1) = \mathbb{P}(X_1 \leq x_1, X_2 \leq 1) = x_1.$$

$$F_{X_2}(x_2) = \mathbb{P}(X_2 \leq x_2) = \mathbb{P}(X_1 \leq 1, X_2 \leq x_2) = x_2.$$

$$\mathbb{P}(X_1 \leq x_1, X_2 \leq x_2) = C(F_{X_1}(x_1), F_{X_2}(x_2); \theta).$$

On déduit $C(x_1, x_2; \theta) = x_1 x_2(1 + \theta(1 - x_1)(1 - x_2))$.

2. Calculer la PDF de $\boldsymbol{X} = (X_1, X_2)$.

$$f_{\boldsymbol{X}}(x_1, x_2) = \frac{\partial^2 F_{\boldsymbol{X}}(x_1, x_2)}{\partial x_1 \partial x_2} = 1 + \theta(1 - 2x_1 - 2x_2 + 4x_1 x_2).$$

3. En déduire les marginales f_{X_1} et f_{X_2}. Vérifier la cohérence avec les résultats précédents.

On retrouve bien $f_{X_1}(x_1) = f_{X_2}(x_2) = 1$, ce qui est cohérent avec les CDFs de 1.

4. Chercher le mode de $\boldsymbol{X}$.

Par symétrie, le maximum est obtenu en $x_1 = x_2$. On obtient :

$$f_{\boldsymbol{X}}(x, x) = 1 + \theta(1 - 4x(1 - x)),$$

qui attend son maximum en $x = 0$ ou $x = 1$. Attention, ici, la fonction est convexe, la valeur de x où la dérivée est nulle correspond à un minimum.

5. Calculer la moyenne et la covariance de $\boldsymbol{X}$.

$$\mathbb{E}[\boldsymbol{X}] = \int_0^1 \int_0^1 \boldsymbol{x} f_{\boldsymbol{X}}(\boldsymbol{x}) d\boldsymbol{x} = \frac{1}{2}.$$

En posant $\boldsymbol{u} = \boldsymbol{x} - \mathbb{E}[\boldsymbol{X}] = (x_1 - 1/2, x_2 - 1/2)$, on montre que :

$$f_{\boldsymbol{X}}(\boldsymbol{x}) = 1 + 4\theta u_1 u_2.$$

On déduit :

$$\begin{aligned}
\mathbb{E}[(\boldsymbol{X} - \mathbb{E}[\boldsymbol{X}]) \otimes (\boldsymbol{X} - \mathbb{E}[\boldsymbol{X}])] &= \int_0^1 \int_0^1 (\boldsymbol{x} - \mathbb{E}[\boldsymbol{X}]) \otimes (\boldsymbol{x} - \mathbb{E}[\boldsymbol{X}]) f_{\boldsymbol{X}}(\boldsymbol{x}) d\boldsymbol{x} \\
&= \int_{-1/2}^{1/2} \int_{-1/2}^{1/2} \boldsymbol{u} \otimes \boldsymbol{u} (1 + 4\theta u_1 u_2) du_1 du_2 \\
&= \frac{1}{12} \begin{bmatrix} 1 & \theta/3 \\ \theta/3 & 1 \end{bmatrix}.
\end{aligned}$$

6. Interpréter la valeur de θ.

9 Influence de la corrélation sur la fiabilité d'une poutre encastrée

Soient X_1 et X_2 deux efforts corrélés ($E[X_1] = E[X_2] = m$, $\sigma_{X_1} = \sigma_{X_2} = \sigma$, $\mathrm{Cov}(X_1, X_2) = \rho\sigma^2$) s'exerÃ§ant sur une poutre simplement appuyée.

On admet que la valeur maximale du moment, $M_{\max}$ dans la poutre, est donné par :

$$M^{\max}(X_1, X_2) = \max_x M(x) = \max\left(\frac{2X_1 + X_2}{3}a, \frac{2X_2 + X_1}{3}a\right),$$

et on suppose que la poutre casse lorsque $M^{\max}(X_1, X_2) \geq M_r$.

1. Donner l'expression de la probabilité de défaillance $P_f = \mathbb{P}(M^{\max}(X_1, X_2) \geq M_r)$.

$$\boldsymbol{C} = \sigma^2 \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$$

$$\begin{aligned}
P_f &= \mathbb{P}(M^{\max}(X_1, X_2) \geq M_r) \\
&= \int_{M^{\max}(X_1, X_2) \geq M_r} \frac{1}{2\pi\sqrt{det(\boldsymbol{C})}} \exp\left(-\frac{1}{2}(x_1 - m, x_2, m)^T \boldsymbol{C}^{-1} (x_1 - m, x_2, m)\right) dx_1 dx_2
\end{aligned}$$

2. Calculer les variables aléatoires centrées et non-corrélées Z_1 et Z_2 associées à X_1 et X_2.

On cherche le vecteur $\boldsymbol{Z}$ centré, de covariance identité, associé à $\boldsymbol{X}$. Pour cela, on définit d'abord le vecteur centré :

$$\boldsymbol{Y} = \begin{pmatrix} X_1 - m \\ X_2 - m \end{pmatrix},$$

$$[R_{\boldsymbol{Y}}] = \mathbb{E}(\boldsymbol{Y} \otimes \boldsymbol{Y}) = \sigma^2 \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} = [D][\lambda][D]^T, \text{ avec } [D]^T[D] = [I],$$

$$[D] = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \quad [\lambda] = \sigma^2 \begin{bmatrix} 1+\rho & 0 \\ 0 & 1-\rho \end{bmatrix},$$

On déduit :

$$\boldsymbol{Z} = [\lambda]^{-1/2}[D]^T\boldsymbol{Y} = \begin{pmatrix} \frac{X_1+X_2-2m}{\sigma\sqrt{2(1+\rho)}} \\ \frac{X_1-X_2-2m}{\sigma\sqrt{2(1-\rho)}} \end{pmatrix}.$$

3. Exprimer le moment de flexion $M^{\max}(X_1, X_2)$ en fonction de Z_1 et Z_2.

On calcule :

$$\begin{cases} \dfrac{2X_1 + X_2}{3} a = a \left[m + Z_1 \sigma \sqrt{\dfrac{1+\rho}{2}} + \dfrac{Z_2 \sigma}{3} \sqrt{\dfrac{1-\rho}{2}} \right] \\ \dfrac{2X_2 + X_1}{3} a = a \left[m + Z_1 \sigma \sqrt{\dfrac{1+\rho}{2}} - \dfrac{Z_2 \sigma}{3} \sqrt{\dfrac{1-\rho}{2}} \right] \end{cases}$$

Ainsi :

$$\frac{M^{\max}(Z_1, Z_2)}{a} - m = \max\left(bZ_1 + cZ_2, bZ_1 - cZ_2\right),$$

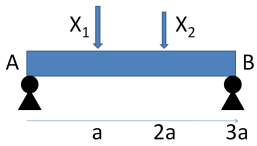
$$\begin{cases} b = \sigma \sqrt{\dfrac{1+\rho}{2}} \\ c = \dfrac{\sigma}{3} \sqrt{\dfrac{1-\rho}{2}} \end{cases}$$

4. Récrire P_f en fonction de la loi de (Z_1, Z_2).

$$P_f = \int_{\max(bZ_1+cZ_2, bZ_1-cZ_2) \geq \frac{M_r}{a} - m} \frac{1}{2\pi} \exp\left(-\frac{(z_1^2 + z_2^2)}{2}\right) dz_1 dz_2$$

5. Représenter graphiquement la zone de défaillance, et justifier l'inégalité suivante :

$$P_f \leq 2 \int_{\beta(\rho)}^{+\infty} \frac{1}{\sqrt{2\pi}} \exp \frac{u^2}{2} du,$$



avec $\beta(\rho)$ la distance minimale entre la zone de défaillance et le point $(z_1 = 0, z_2 = 0)$ à calculer.

La zone de défaillance est constituée de l'union de deux demi-plans, d'équations respectives $bZ_1 + cZ_2 = \frac{M_r}{a} - m$ et $bZ_1 - cZ_2 = \frac{M_r}{a} - m$. Pour calculer la probabilité de défaillance, il faut alors intégrer $\exp\left(-\frac{(z_1^2+z_2^2)}{2}\right)$, qui décroit très rapidement quand on s'éloigne du centre du repère sur cette zone de défaillance. P_f est alors clairement inférieur à 2 fois l'intégration sur un seul des deux demi-plans, l'intégration sur l'intersection étant comptée deux fois. Par symétrie de rotation de la loi de (Z_1, Z_2), on obtient l'équation fournie.

Par égalité des aires des triangles, on trouve alors :

$$\beta(\rho)\sqrt{\frac{1}{c^2} + \frac{1}{b^2}}\left(\frac{M_r}{a} - m\right) = \left(\frac{M_r}{a} - m\right)^2 \frac{1}{cb},$$

$$\beta(\rho) = \frac{3\left(\frac{M_r}{a} - m\right)}{\sigma\sqrt{5 + 4\rho}}.$$

6. Comparer la résistance de la poutre pour les valeurs de ρ égales à 0, -1 et 1.

Ainsi, on montre :

$$\begin{cases} \beta(0) = \beta_0, \\ \beta(1) = \beta_0 \dfrac{\sqrt{5}}{3} \approx 0.75\beta_0 \\ \beta(-1) = \beta_0\sqrt{5} \approx 2.24\beta_0 \end{cases}$$

10 Collection de figurines

N figurines différentes sont disposées de manière équirépartie dans des paquets de céréales (on suppose le nombre de paquets très grand si bien que les probabilités d'obtenir une figurine ne changent pas dans le temps). On cherche à évaluer le nombre moyen de paquets de céréales que l'on doit acheter pour avoir 99% de chances d'obtenir toutes les figurines.

Pour cela, on définit $\{X_n,\ n \geq 1\}$ le processus aléatoire à temps discret tel que X_n correspond au nombre (aléatoire) de figures différentes que l'on a obtenu en achetant n paquets.

1. Evaluer $\mathbb{P}(X_{n+1} = p)$ en fonction de la valeur de X_n, pour $p \geq 1$.

$$\mathbb{P}(X_{n+1}=p)=\begin{cases}\dfrac{p}{N} \text{ si } X_n=p\\ \dfrac{N-p+1}{N} \text{ si } X_n=p-1\\ 0 \text{ sinon.}\end{cases}$$

2. On définit par ailleurs $\boldsymbol{P}_n$ le vecteur tel que :

$$\boldsymbol{P}_n=\left(\mathbb{P}(X_n=1),\dots,\mathbb{P}(X_n=N)\right).$$

Exprimer alors $\boldsymbol{P}_{n+1}$ en fonction de $\boldsymbol{P}_n$.

On déduit :

$$\boldsymbol{P}_{n+1}=\begin{pmatrix}\mathbb{P}(X_{n+1}=1)\\ \mathbb{P}(X_{n+1}=2)\\ \vdots\\ \mathbb{P}(X_{n+1}=N)\end{pmatrix}=\begin{pmatrix}\frac{1}{N} & 0 & \cdots & \cdots & 0\\ \frac{N-1}{N} & \frac{2}{N} & \ddots & \ddots & \vdots\\ 0 & \frac{N-2}{N} & \frac{3}{N} & \ddots & \vdots\\ \vdots & \ddots & \ddots & \ddots & 0\\ 0 & \dots & 0 & \frac{1}{N} & 1\end{pmatrix}\begin{pmatrix}\mathbb{P}(X_n=1)\\ \mathbb{P}(X_n=2)\\ \vdots\\ \mathbb{P}(X_n=N)\end{pmatrix}=[A]\boldsymbol{P}_n.$$

3. En déduire la valeur de $\boldsymbol{P}_n$ pour tout $n\geq 1$.

Il vient :

$$\Rightarrow\ \boldsymbol{P}_{n+1}=[A]^n\boldsymbol{P}_1,\quad \boldsymbol{P}_1=(1,0,\dots,0).$$

4. Expliquer comment en déduire le nombre moyen de paquets à acheter pour avoir 99% de chances d'avoir toutes les figurines.

Par construction, $(\boldsymbol{P}_{n+1})_N=\mathbb{P}(X_{n+1}=N)$. Avec l'expression précédente, on a ainsi une expression analytique de cette probabilité. On cherche alors $n^\star$ tel que $\mathbb{P}(X_{n^\star}=N)=0.99$.

5. Sur la figure 1 est tracée l'évolution de $\mathbb{P}(X_{n^*}=N)=0.99$ en fonction de N. Commenter cette évolution en fonction de N.

L'évolution, de manière un peu surprenante semble relativement linéaire en fonction de N.

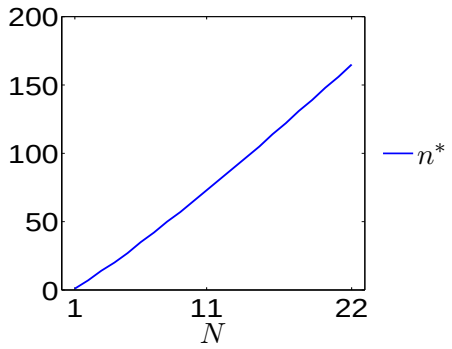


FIGURE 1 – Evolution de $\mathbb{P}(X_{n^*} = N) = 0.99$ en fonction de N.

11 Loi gaussienne et erreur expérimentale

On dispose de N mesures indépendantes, $X_1, \ldots, X_N$, d'une même quantité X. Le bruit de mesure permet de considérer que ces mesures sont aléatoires. On suppose que l'erreur de mesure est caractérisée par une unique densité f inconnue. On suppose néanmoins que les erreurs de mesure sont centrées sur X. On pose $\Phi = \log(f)$.

1. Ecrire la vraisemblance $\mathcal{L}(X)$ d'obtenir $X_1 = x_1, \ldots, X_N = x_n$ en fonction de f et X.

$$\mathcal{L}(X) = f(x_1 - X) \times \ldots \times f(x_N - X).$$

2. On suppose que cette fonction de vraisemblance est maximale lorsque $X = \frac{x_1 + \ldots + x_N}{N} =: \widehat{x}$, pour tout $N \geq 1$. En déduire que la quantité $\Phi(x_1 - X) + \ldots + \Phi(x_N - X)$ est également maximale en $X = \widehat{x}$.

Le résulat est direct par concavité de la fonction log.

3. En déduire que :

$$\Phi'(x_1 - \widehat{x}) + \ldots + \Phi'(x_N - \widehat{x}) = 0.$$

La fonction étant maximale, sa dérivée est nulle.

4. Déduire du cas $N = 2$ que Φ' est une fonction impaire.

On calcule :

$$0 = \Phi'(x_1 - \frac{x_1 + x_2}{2}) + \Phi'(x_2 - \frac{x_1 + x_2}{2}) = \Phi'(\frac{x_1 - x_2}{2}) + \Phi'(\frac{x_2 - x_1}{2}).$$

Ceci étant valable pour tout x_1, x_2, on déduit directement le résultat.

5. Déduire du cas $N = 3$ que Φ' vérifie, pour tout u, v : $\Phi'(u + v) = \Phi'(u) + \Phi'(v)$.

On a :

$$\Phi'(x_1 - \frac{x_1+x_2+x_3}{3}) + \Phi'(x_2 - \frac{x_1+x_2+x_3}{3}) + \Phi'(x_3 - \frac{x_1+x_2+x_3}{3}) = 0.$$

En posant : $u = \frac{2x_1-x_2-x_3}{3}$ et $v = \frac{2x_2-x_1-x_3}{3}$, on déduit que :

$$\Phi'(u) + \Phi'(v) + \Phi'(-(u+v)) = 0,$$

c'est à dire, par parité de Φ', que :

$$\Phi'(u+v) = \Phi'(u) + \Phi'(v).$$

6. En déduire que $\Phi(X) = \alpha\frac{X^2}{2} + \beta$.

De la question 5, on déduit que Φ' est linéaire : $\Phi' = \alpha X$. On intègre, et il vient : $\Phi(X) = \alpha\frac{X^2}{2} + \beta$.

7. En déduire que la fonction f est une fonction gaussienne.

En prenant l'exponentielle de l'expression précédente, on retrouve directement ce que l'on cherche.

12 Inférence de la loi d'une quantité d'intérêt à partir de mesures indirectes

On s'intéresse à la loi statistique de la variable aléatoire Y, dont on suppose une loi a priori gaussienne de moyenne μ et de variance σ^2. Pour cela, on dispose de la mesure d'une quantité Z, telle que :

$$Z = \alpha Y + \varepsilon,$$

où α est une constante connue, reliant la mesure à la variable aléatoire d'intérêt, et ε est une erreur de mesure aléatoire, de loi gaussienne de moyenne nulle et de variance connue σ^2_{mes}. On suppose que les grandeurs Y et ε sont indépendantes statistiquement.

1. Calculer la moyenne et la covariance de Z.

$\mathbb{E}[Z] = \alpha\mu$.
$\text{Cov}(Z) = \alpha^2\sigma^2 + \sigma^2_{\text{mes}}$

2. Calculer la covariance croisée entre Z et Y, $\mathbb{E}\left[(Y - \mathbb{E}[Y])(Z - \mathbb{E}[Z])\right]$.

$\mathbb{E}\left[(Y - \mathbb{Y})(Z - \mathbb{Z})\right] = \alpha\sigma^2$.

3. En déduire que le vecteur (Y, Z) est gaussien, de moyenne $(\mu, \alpha\mu)$, et de matrice de covariance $\begin{bmatrix} \sigma^2 & \alpha\sigma^2 \\ \alpha\sigma^2 & \alpha^2\sigma^2 + \sigma^2_{\text{mes}} \end{bmatrix}$.

4. A partir des formules de conditionnement gaussien, en déduire la loi de $(Y|Z)$.

$(Y|Z)$ est gaussien, de moyenne $\mu + \alpha\sigma^2(\alpha^2\sigma^2 + \sigma^2_{\text{mes}})^{-1}(Z - \alpha\mu)$, et de variance $\sigma^2 - \alpha^2\sigma^4(\alpha^2\sigma^2 + \sigma^2_{\text{mes}})^{-1}$.

13 Mesures redondantes et réduction des incertitudes

On cherche à mesurer deux quantités Y_1 et Y_2. On dispose pour cela de processus expérimentaux bruités permettant de mesurer directement Y_1 et Y_2, mais également $Y_1 + Y_2$. On note alors respectivement X_1^{obs}, X_2^{obs} et X_3^{obs} les résultats de mesure des grandeurs Y_1, Y_2 et $Y_1 + Y_2$, que l'on suppose indépendants statistiquement, et dont on peut supposer un comportement gaussien :

$$X_1^{\text{obs}} \sim \mathcal{N}(Y_1, V_1)\,,\ X_2^{\text{obs}} \sim \mathcal{N}(Y_2, V_2)\,,\ X_3^{\text{obs}} \sim \mathcal{N}(Y_1 + Y_2, V_3)\,,$$

où les incertitudes de mesure V_1, V_2 et V_3 sont connues.

1. Calculer la probabilité de l'évènement $(x_1 \leq X_1^{\text{obs}} \leq x_1 + dx_1, x_2 \leq X_2^{\text{obs}} \leq x_2 + dx_2, x_3 \leq X_3^{\text{obs}} \leq x_3 + dx_3)$.

$$\mathbb{P}(X_1^{\text{obs}} = x_1, X_2^{\text{obs}} = x_2, X_3^{\text{obs}} = x_3) \propto \exp\left(-\frac{1}{2}\left(\frac{(Y_1 - x_1)^2}{V_1} + \frac{(Y_2 - x_2)^2}{V_2} + \frac{(Y_1 + Y_2 - x_3)^2}{V_3}\right)\right).$$

2. Montrer que le logarithme de cette probabilité est proportionnel, à une constante additive près, à :

$$Q := Y_1^2\left(\frac{1}{V_1} + \frac{1}{V_3}\right) + Y_2^2\left(\frac{1}{V_2} + \frac{1}{V_3}\right) + \frac{2Y_1Y_2}{V_3} - 2Y_1\left(\frac{x_1}{V_1} + \frac{x_3}{V_3}\right) - 2Y_2\left(\frac{x_2}{V_2} + \frac{x_3}{V_3}\right).$$

Il suffit de développer.

3. En posant $\boldsymbol{Y} = (Y_1, Y_2)$, montrer que Q est égal, à une constante additive près, à :

$$(\boldsymbol{Y} - \boldsymbol{\mu})^T[R]^{-1}(\boldsymbol{Y} - \boldsymbol{\mu}),$$

où les grandeurs $\boldsymbol{\mu}$ et $[R]^{-1}$ sont à expliciter.

$$(\boldsymbol{Y}-\boldsymbol{\mu})^T[R]^{-1}(\boldsymbol{Y}-\boldsymbol{\mu}) \propto [R]^{-1}_{11}Y_1^2 + [R]^{-1}_{22}Y_2^2 + 2[R]^{-1}_{12}Y_1Y_2 - 2Y_1([R]^{-1}_{11}\mu_1 + [R]^{-1}_{12}\mu_2) - 2Y_2([R]^{-1}_{22}\mu_2 + [R]^{-1}_{12}\mu_1).$$

En posant :

$$[R]^{-1}_{11} = \frac{1}{V_1} + \frac{1}{V_3},\ [R]^{-1}_{12} = \frac{1}{V_3},\ [R]^{-1}_{22} = \frac{1}{V_1} + \frac{1}{V_3},\ \boldsymbol{\mu} = [R]\begin{pmatrix} \frac{x_1}{V_1} + \frac{x_3}{V_3} \\ \frac{x_2}{V_2} + \frac{x_3}{V_3} \end{pmatrix},$$

on retrouve l'expression recherchée.

4. Dans quelle mesure peut on en déduire que l'incertitude sur le couple (Y_1, Y_2) est gaussienne ? Commenter l'apport de la mesure de $Y_1 + Y_2$ sur cette incertitude.

Dans le plan (Y_1, Y_2), $\mathbb{P}(X_1^{\text{obs}} = x_1, X_2^{\text{obs}} = x_2, X_3^{\text{obs}} = x_3)$ correspond à une loi gaussienne de moyenne $\boldsymbol{\mu}$ et de covariance $[R]$, ce qui jusitifie le terme d'incertitudes gaussiennes sur Y_1 et

Y_2. On remarquera que les incertitudes sur ces deux grandeurs sont corrélées par la présence de la troisième mesure.

5. Sans refaire les calculs, expliquer comment intégrer une mesure indépendante de $Y_1 - Y_2$, une seconde mesure de Y_1 ou Y_2.

Il faut repartir de l'expression de la probabilité $\mathbb{P}(X_1^{\text{obs}} = x_1, X_2^{\text{obs}} = x_2, X_3^{\text{obs}} = x_3, X_4^{\text{obs}} = x_4)$ puis l'exprimer en fonction de (Y_1, Y_2) et identifier la moyenne et la covariance.

14 Estimateur du maximum de vraisemblance

On suppose disposer de N réalisations iid de $X \sim \mathcal{N}(\mu, \sigma^2)$, que l'on nomme $X_1, \ldots, X_N$.

1. Calculer la vraisemblance de $X_1, \ldots, X_N$ sachant μ et σ, que l'on note $\pi[\mathbb{X}|\mu, \sigma]$.

$$\pi[\mathbb{X}|\mu, \sigma] = \prod_{n=1}^{N} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(X_n - \mu)^2}{2\sigma^2}\right).$$

2. Montrer que :

$$\log(\pi[\mathbb{X}|\mu, \sigma]) \propto -L(\mu, \sigma) = -N \log(\sigma^2) - \frac{1}{\sigma^2} \sum_{n=1}^{N} (X_n - \mu)^2.$$

3. En déduire les estimateurs du maximum de vraisemblance de μ et σ.

Par monotonie de la fonction log, $\pi[\mathbb{X}|\mu, \sigma]$ est maximum quand $L(\mu, \sigma)$ est minimum. On en déduit que le MLE de μ, μ^{MLE}, vérifie :

$$\mu^{MLE} = \arg\min_{\mu} \sum_{n=1}^{N} (X_n - \mu)^2 = \frac{1}{N} \sum_{n=1}^{N} X_n.$$

On note alors $D = \sum_{n=1}^{N} (X_n - \mu^{MLE})^2 = \sum_{n=1}^{N} X_n^2 - N(\mu^{MLE})^2$, si bien que :

$$L(\mu^{MLE}, \sigma) = N \log(\sigma^2) + \frac{D}{\sigma^2}.$$

On calcule :

$$\frac{\partial L(\mu^{MLE}, \sigma)}{\partial \sigma^2} = \frac{1}{N\sigma^2} - \frac{D}{\sigma^4}.$$

Ainsi :

$$\left(\sigma^{MLE}\right)^2 = \frac{D}{N}.$$

4. Montrer que l'estimateur du maximum de vraisemblance de σ est biaisé.

$$\begin{aligned}
\mathbb{E}\left[\sigma^{MLE}\right] &= \frac{1}{N}\mathbb{E}\left[\sum_{n=1}^{N} X_n^2 - N(\mu^{MLE})^2\right] \\
&= \frac{1}{N}\left(N(\sigma^2+\mu^2) - \frac{1}{N}\mathbb{E}\left[\sum_{n,m=1}^{N} X_n X_m\right]\right) \\
&= \sigma^2 + \mu^2 - \frac{1}{N^2}\sum_{n,m=1}^{N}\mathbb{E}\left[X_n X_m\right] \\
&= \sigma^2 + \mu^2 - \frac{1}{N}(\sigma^2+\mu^2) - \frac{N(N-1)}{N^2}\mu^2 \\
&= \frac{(N-1)}{N}\sigma^2.
\end{aligned}$$

15 Principe du maximum d'entropie

On rappelle que l'entropie statistique $S(X)$ d'une variable aléatoire X de densité f_X définie sur $K \subset \mathbb{R}$ s'écrit :

$$S(X) = -\int_{\mathbb{R}} f_X(x)\log(f_X)(x)dx.$$

On admet que la loi maximisant l'entropie sous les contraintes $\int_{\mathbb{R}} g_m(x)f_X(x)dx = h_m,\ 1 \leq m \leq M$ s'écrit sous la forme :

$$f_X(x) = \exp\left(-\lambda_0 + \sum_{m=1}^{M}\lambda_m g_m(x)\right).$$

1. On suppose que l'on connaît uniquement la moyenne μ et la variance σ^2 de X. Montrer comment cela se traduit en terme de notations g_m, f_m.

$g_1(x) = x,\ f_1 = \mu,\ g_2(x) = (x-\mu)^2,\ f_2 = \sigma^2.$

2. En déduire que f_X peut s'écrire sous la forme :

$$f_X(x) = \exp\left(-\widehat{\lambda}_0 - \widehat{\lambda}_1 x - \widehat{\lambda}_2 x^2\right).$$

C'est direct, en enlevant le μ.

3. En déduire cette fois :

$$f_X(x) = c_0\exp\left(-\frac{(x-c_1)^2}{2c_2}\right),$$

où c_0, c_1 et c_2 doivent être exprimées en fonction de $\widehat{\lambda}_0, \widehat{\lambda}_1, \widehat{\lambda}_2$.

$\widehat{\lambda}_0 = -\log(c_0) + \frac{c_1^2}{2c_2}$

$\widehat{\lambda}_1 = -\frac{c_1}{c_2}$

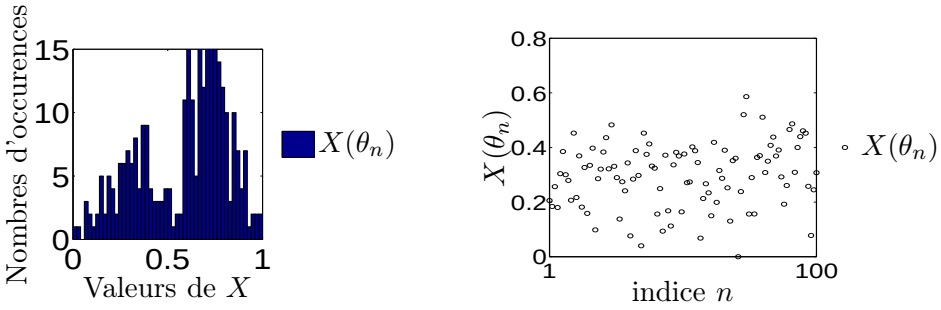


FIGURE 2 – Information disponible pour la caractérisation de a loi de X.

$\widehat{\lambda}_2 = \frac{1}{2c_2}$

4. En déduire l'expression de f_X.

On conclut que $c_0 = \frac{1}{\sqrt{2\pi}\sigma}$, $c_1 = \mu$, $c_2 = \sigma^2$.

16 Identification d'une loi bimodale

On dispose de $N = 100$ réalisations indépendantes d'une variable aléatoire X, nommées $\{X(\theta_n),\ 1 \leq n \leq N\}$, représentées sur la figure **??**. A partir de l'histogramme fourni, il semble naturel de décrire la PDF f_X de X sous forme hiérarchique : (1) une loi (marginale) pour L et (2) une loi (conditionnelle) pour X, dépendant de la valeur de L. Plus précisément :

— L suit une loi de Bernoulli :

$$\mathbb{P}(L = \ell) = \begin{cases} p \text{ si } \ell = 1, \\ 1 - p \text{ si } \ell = 2, \end{cases}$$

— $X|L$ suit une loi gaussienne dont les paramètres dépendent de la valeur de L :

$$\begin{cases} f_{X|L}(x|L=1) = \dfrac{\sqrt{\gamma_1}}{\sqrt{2\pi}} \exp\left(-\gamma_1 \left(x-\mu_1\right)^2/2\right) = \mathcal{N}(x\ ;\ \mu_1, \gamma_1), \\ f_{X|L}(x|L=2) = \dfrac{\sqrt{\gamma_2}}{\sqrt{2\pi}} \exp\left(-\gamma_2 \left(x-\mu_2\right)^2/2\right) = \mathcal{N}(x\ ;\ \mu_2, \gamma_2), \end{cases}$$

On nomme alors $\boldsymbol{\theta} = (\mu_1, \gamma_1, \mu_2, \gamma_2, p)$ le vecteur regroupant les paramètres à déterminer pour caractériser f_X. On note alors $f_{X|\boldsymbol{\theta}}$ l'approximation paramétrique de f_X.

1. Calculer la PDF $f_{X|\boldsymbol{\theta}}$ de $X|\boldsymbol{\theta}$ en fonction de $(\mu_1, \gamma_1, \mu_2, \gamma_2, p)$.

Direct par propriété des probabilités totales :

$$f_{X|\boldsymbol{\theta}}(x) = p\mathcal{N}(x\ ;\ \mu_1, \gamma_1) + (1-p)\mathcal{N}(x\ ;\ \mu_2, \gamma_2).$$

2. En supposant que μ_1, μ_2, p soient connues et que $\gamma_1 = \gamma_2 = \gamma$, quelle condition doit vérifier γ pour que l'on puisse distinguer les deux modes de $f_{X|\boldsymbol{\theta}}$?

Il faut que l'écart type, $1/\sqrt{\gamma}$ soit petit devant $|\mu_1 - \mu_2|$.

3. Calculer la vraisemblance $\pi[\mathbb{X}|\boldsymbol{\theta}]$ de $\boldsymbol{\theta}$ sachant les observations $\mathbb{X} = \{X(\theta_1), \ldots, X(\theta_N)\}$.

$$\pi[\mathbb{X}|\boldsymbol{\theta}] = \prod_{n=1}^{N} f_{X|\boldsymbol{\theta}}(X(\theta_n)).$$

4. On suppose que le vecteur des paramètres est également aléatoire, de distribution *a priori* $\pi_{\boldsymbol{\theta}}(\boldsymbol{\theta}) = \pi_{M_1}(\mu_1)\pi_{G_1}(\gamma_1)\pi_{M_2}(\mu_2)\pi_{G_2}(\gamma_2)\pi_P(p)$, où :

— π_{M_1} et π_{M_2} sont des gaussiennes de moyennes μ_0 et de variance $1/\gamma_0$,

$$\pi_{M_i}(\mu_i) \propto \exp(-\gamma_0(\mu_i - \mu_0)^2/2), \quad i \in \{1,2\}\,.$$

— π_{G_1} et π_{G_2} sont des lois gamma de paramètres α_0 et β_0,

$$\pi_{G_i}(\gamma_i) \propto \gamma_i^{\alpha_0 - 1} \exp(-\beta_0\gamma_i) 1_{\mathbb{R}^+}(\gamma_i), \quad i \in \{1,2\}\,.$$

En déduire la distribution a posteriori de $\boldsymbol{\theta}$, $\pi[\boldsymbol{\theta}|\mathbb{X}]$, à une constante près.

$$\pi[\boldsymbol{\theta}|\mathbb{X}] \propto \pi[\mathbb{X}|\boldsymbol{\theta}]\pi_{\boldsymbol{\theta}}(\boldsymbol{\theta}),$$

$$\propto (\gamma_1\gamma_2)^{\alpha_0-1} 1_{\mathbb{R}^+}(\gamma_1) 1_{\mathbb{R}^+}(\gamma_2) \exp\left\{-\gamma_0\left((\mu_1-\mu_0)^2 + (\mu_2-\mu_0)^2\right)/2 - \beta_0(\gamma_1+\gamma_2)\right\} \prod_{n=1}^{N} f_{X|\boldsymbol{\theta}}(X(\theta_n))$$

5. Dans le cas où $p = 1$, montrer que $\pi[\boldsymbol{\theta}|\mathbb{X}]$ est proportionnelle à :

$$\gamma_1^{N/2+\alpha_0-1} 1_{\mathbb{R}^+}(\gamma_1) \exp\left\{-\gamma_0(\mu_1-\mu_0)^2/2 - \beta_0\gamma_1 - \gamma_1\left(\sum_{n=1}^{N}(X(\theta_n)-\widehat{\mu})^2 + N(\widehat{\mu}-\mu_1)^2\right)/2\right\},$$

où $\widehat{\mu} = \frac{1}{N}\sum_{n=1}^{N} X(\theta_n)$.

Par construction, on a :

$$\pi[\boldsymbol{\theta}|\mathbb{X}] \propto \gamma_1^{N/2+\alpha_0-1} 1_{\mathbb{R}^+}(\gamma_1) \exp\left\{-\gamma_0(\mu_1-\mu_0)^2/2 - \beta_0\gamma_1 - \gamma_1\sum_{n=1}^{N}(X(\theta_n)-\mu_1)^2/2\right\},$$

$$\begin{aligned}
\sum_{n=1}^{N}(X(\theta_n)-\mu_1)^2 &= \sum_{n=1}^{N}(X(\theta_n)-\widehat{\mu}+\widehat{\mu}-\mu_1)^2 \\
&= \sum_{n=1}^{N}(X(\theta_n)-\widehat{\mu})^2 + \sum_{n=1}^{N}(\widehat{\mu}-\mu_1)^2 + 2(\widehat{\mu}-\mu_1)^2\sum_{n=1}^{N}(X(\theta_n)-\widehat{\mu}) \\
&= \sum_{n=1}^{N}(X(\theta_n)-\widehat{\mu})^2 + N(\widehat{\mu}-\mu_1)^2.
\end{aligned}$$

6. En déduire que les lois a posteriori conditionnelles de $\mu_1|\gamma_1$ et de $\gamma_1|\mu_1$ restent des lois gaussienne et gamma respectivement, de paramètres à définir.

De l'expression précédente, il vient que :
$\pi_{\mu_1|\gamma_1}(\mu_1|\gamma_1) \propto \exp\left\{-\gamma_0(\mu_1-\mu_0)^2/2 - \gamma_1 N(\widehat{\mu}-\mu_1)^2/2\right\}$,
$\pi_{\gamma_1|\mu_1}(\gamma_1|\mu_1) \propto \gamma_1^{N/2+\alpha_0-1} 1_{\mathbb{R}^+}(\gamma_1) \exp\left\{-\gamma_1(\beta_0 + 1/2\sum_{n=1}^N (X(\theta_n)-\widehat{\mu})^2 + N/2(\widehat{\mu}-\mu_1)^2)\right\}$
On déduit que $\mu_1 \sim \mathcal{N}(\mu_1|m = \frac{\gamma_0\mu_0+\gamma_1 N\widehat{\mu}}{\gamma_0+\gamma_1 N}, \gamma = \gamma_0+\gamma_1 N)$, et que γ_1 suit une loi gamma de paramètres $\alpha = N/2+\alpha_0$ et $\beta = \beta_0 + 1/2\sum_{n=1}^N (X(\theta_n)-\widehat{\mu})^2 + N(\widehat{\mu}-\mu_1)^2/2$.

7. En déduire que quand $N \to +\infty$, les lois conditionnées de μ_1 et γ_1 sont de plus en plus piquées autour de $\widehat{\mu}$ et $1/\widehat{\gamma} = \frac{1}{N}\sum_{n=1}^N (X(\theta_n)-\widehat{\mu})^2$.
direct en faisant tendre N vers l'infini.

17 Avantages et limites du cumul quadratique

Soient X_1 et X_2 deux variables indépendantes qui sont uniformément distribuées sur $[-1,1]$.

1. Rappeler la densité jointe, f_{X_1,X_2}, associée à ce couple de variables aléatoires.

$f_{X_1,X_2} = 1/4$.

2. On considère le modèle linéaire suivant :

$$Y(X_1,X_2) = aX_1 + bX_2 + cX_1X_2,$$

où a,b,c sont trois constantes positives. Calculer la moyenne et la variance de Y.

$$\mathbb{E}[Y] = 0,$$

$$\begin{aligned}\mathbb{E}[Y^2] &= \mathbb{E}[a^2X_1^2 + b^2X_2^2 + c^2X_1^2X_2^2 + 2abX_1X_2 + 2acX_1^2X_2 + 2bcX_2^2X_1],\\ &= \frac{a^2}{3} + \frac{b^2}{3} + \frac{c^2}{9}.\end{aligned}$$

3. Calculer les approximations par cumul quadratique de la moyenne et de la variance de Y.
$\mathbb{E}[Y] \simeq Y(\mathbb{E}[X_1],\mathbb{E}[X_2]) + \frac{1}{2}\sum_{i=1}^2 \frac{\partial^2 Y}{\partial X_i^2}_{(X_1,X_2)=\mathbb{E}[X_1],\mathbb{E}[X_2]} \sigma_i^2 = 0.$
$\mathrm{var}(Y) \simeq \sum_{i=1}^2 \left(\frac{\partial Y}{\partial X_i}_{(X_1,X_2)=\mathbb{E}[X_1],\mathbb{E}[X_2]}\right)^2 \sigma_i^2 = \frac{a^2}{3} + \frac{b^2}{3}.$

4. En déduire les conditions sur a,b,c qui doivent être vérifiées pour que l'approximation par cumul quadratique soit correct.
Il faut que c soit petit devant a et b.

18 Estimation de quantités déterministe et aléatoire

On dispose de mesures indépendantes d'une quantité $P_{\max}$ présentant une double source d'incertitudes :
— incertitudes de mesure dues au processus expérimental (incertitude "épistémique" ou "réductible"),

— variabilité "naturelle" entre échantillon (incertitude "aléatoire" ou "irréductible").

1. Cas 1 : la variabilité "naturelle" est négligeable ⇒ il existe une unique valeur de $P_{\max}$. En se basant sur l'égalité suivante, avec $\varepsilon_1^{\text{mes}}, \dots, \varepsilon_N^{\text{mes}}$ des copies indépendantes et identiquement distribuées (iid) d'une même variable aléatoire ε^{mes}, de PDF gaussienne centrée d'écart type σ, exprimer la vraisemblance des mesures.

$$P_{\max,n}^{\text{mes}} = P_{\max} + \varepsilon_n^{\text{mes}}, \quad \boldsymbol{y} = (P_{\max,1}^{\text{mes}}, \dots, P_{\max,N}^{\text{mes}}).$$

$$f(\boldsymbol{y}|P_{\max}) = \prod_{n=1}^{N} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(P_{\max,n}^{\text{mes}} - P_{\max})^2}{2\sigma^2}\right).$$

2. Commenter la limite de ce résultat quand N tend vers l'infini.

L'incertitude tend vers 0, on tend vers une connaissance parfaite de $P_{\max}$.

3. Cas 2 : la variabilité "naturelle" est majoritaire ⇒ il n'existe pas de valeur unique de $P_{\max}$. Dans ce cas, $P_{\max}$ est modélisée par une v.a. de loi $\pi(\cdot;\boldsymbol{\alpha})$, de paramètres $\boldsymbol{\alpha}$ (moyenne, écart-type,...) inconnus. En se basant sur l'égalité suivante, en déduire la nouvelle expression de la vraisemblance.

$$\pi(P_{\max}|\boldsymbol{y}) = \int_{\mathbb{A}} \pi(P_{\max}|\boldsymbol{\alpha})\pi(\boldsymbol{\alpha}|\boldsymbol{y})d\boldsymbol{\alpha}, \quad \pi(\boldsymbol{\alpha}|\boldsymbol{y}) \propto f(\boldsymbol{y}|\boldsymbol{\alpha})\pi(\boldsymbol{\alpha}).$$

$$f(\boldsymbol{y}|\boldsymbol{\alpha}) = \prod_{n=1}^{N} \int_{-\infty}^{\infty} \frac{\pi(P_{\max};\boldsymbol{\alpha})}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(P_{\max,n}^{\text{mes}} - P_{\max})^2}{2\sigma^2}\right) dP_{\max}.$$

4. Analyser la convergence de cette expression quand N tend vers l'infini.

Cette fois, on tend vers une PDF particulière. L'incertitude est non réductible.

19 Statistiques sur les processus gaussiens stationnaires

Soit $\{X(t),\ t \in \mathbb{R}\}$ un processus aléatoire gaussien stationnaire centré et de fonction de covariance $(t,t') \mapsto C(t,t') = \mathbb{E}[X(t)X(t')] = R(|t-t'|)$. On suppose que ce processus est dérivable, au sens où pour tout t, la limite

$$\lim_{dt \to 0} \frac{X(t+dt) - X(t)}{dt}$$

existe, et est notée $\dot{X}(t)$.

1. Montrer que $\dot{X}$ est un processus gaussien.

La somme de deux processus gaussiens étant gaussien, pour toute valeur de $dt > 0$, $\frac{X(t+dt)-X(t)}{dt}$ est donc gaussien, ce qui reste le cas pour sa limite.

2. Calculer les fonctions moyenne et de covariance du processus dérivée $\dot{X}$.

X étant stationnaire, on déduit $\mathbb{E}[X(t+dt)] = \mathbb{E}[X(t)]$, donc $\dot{X}$ est de moyenne nulle. Par ailleurs :

$$\begin{aligned} R_{\dot{X}\dot{X}}(\tau) &= \lim_{\epsilon_1,\epsilon_2\to 0} E\left[\frac{X(t+\epsilon_1)-X(t)}{\epsilon_1}\frac{\dot{X}(t+\epsilon_2+\tau)-\dot{X}(t+\tau)}{\epsilon_2}\right] \\ &= \lim_{\epsilon_1,\epsilon_2\to 0} \frac{1}{\epsilon_1\epsilon_2}\{(R(\tau+\epsilon_2-\epsilon_1)-R(\tau-\epsilon_1))-(R(\tau+\epsilon_2)-R(\tau))\} \\ &= \lim_{\epsilon_1\to 0}\frac{1}{\epsilon_1}\left\{\frac{dR}{d\tau}(\tau-\epsilon_1)-\frac{dR}{d\tau}(\tau)\right\} = (-1)\times\frac{d^2R}{d\tau^2}(\tau) \end{aligned}$$

3. Montrer que X et $\dot{X}$ sont décorrélés. En déduire que dans ce cas gaussien, ils sont indépendants.

$$\begin{aligned} E\left[X(t)\dot{X}(t)\right] &= \lim_{\epsilon\to 0}\frac{1}{\epsilon}\{E[X(t)X(t+\epsilon)]-E[X(t)X(t)]\} \\ &= \lim_{\epsilon\to 0}\frac{1}{\epsilon}\{R_X(\epsilon)-R_X(0)\} = \frac{dR_X}{\epsilon}(0). \end{aligned}$$

Comme la fonction d'autocorrélation d'un processus stationnaire dérivable est **paire**, alors $\frac{dR_X}{d\tau}(0) = 0$, ce qui prouve la propriété voulue.

4. Montrer que la probabilité pour X de dépasser un seuil ξ, avec pente positive, entre t et $t+dt$ s'écrit :

$$P_\xi^+ = \int_0^{+\infty} p(\xi,\dot{x})\dot{x}d\dot{x}dt,$$

La grandeur $\nu_\xi^+ = \int_0^{+\infty} p(\xi,\dot{x})\dot{x}d\dot{x}$ correspond à la fréquence attendue de franchissement du seuil ξ avec pente positive, ou encore probabilité de franchissement de ξ avec pente > 0 par unité de temps.

Au premier ordre en dt, on peut écrire : $X(t+dt) \approx X(t) + dt\dot{X}(t)$. Ainsi, dans le plan $(X,\dot{X})$, l'espace $\mathcal{D}_\xi^+$ des couples vérifiant la condition précédente, s'écrit :

$$\mathcal{D}_\xi^+ = \left\{\left(X(t),\dot{X}(t)\right),\quad \dot{X}(t) > 0\ ,\ \xi - \dot{X}(t)dt \le X(t) \le \xi\right\}.$$

Si p est la distribution jointe de $(X(t),\dot{X}(t))$, alors la probabilité P_ξ^+ d'être dans $\mathcal{D}_\xi^+$ se déduit par :

$$\begin{aligned} P_\xi^+ &= \int_{\mathcal{D}_\xi^+} p(x,\dot{x})dxd\dot{x} = \int_0^{+\infty}\left\{\int_{\xi-\dot{X}(t)dt}^{\xi} p(x,\dot{x})dx\right\}d\dot{x} \\ &\approx \int_0^{+\infty} p(\xi,\dot{x})\dot{x}d\dot{x}dt \ = \ \nu_\xi^+\times dt. \end{aligned}$$

5. Exprimer ν_ξ^+ en fonction de ξ et des écart-types σ_X et $\sigma_{\dot{X}}$ de X et $\dot{X}$ respectivement.

Si $X(t)$ est gaussien centré de variance σ_X^2, alors $\dot{X}(t)$ est également gaussien centré, de variance $\sigma_{\dot{X}}^2$, et on a :

$$p_X(x) = \frac{1}{\sqrt{2\pi}\sigma_X} \exp\left(-\frac{1}{2}\frac{x^2}{\sigma_X^2}\right), \quad p_{\dot{X}}(\dot{x}) = \frac{1}{\sqrt{2\pi}\sigma_{\dot{X}}} \exp\left(-\frac{1}{2}\frac{\dot{x}^2}{\sigma_{\dot{X}}^2}\right).$$

Comme $X(t)$ et $\dot{X}(t)$ sont décorrélés et gaussiens, ils sont **indépendants**, et $p(x, \dot{x}) = p_X(x)p_{\dot{X}}(\dot{x})$. On déduit :

$$\begin{aligned}\nu_\xi^+ &= \int_0^{+\infty} p(\xi, \dot{x})\dot{x}d\dot{x} = p_X(\xi)\int_0^{+\infty} \dot{x}p_{\dot{X}}(\dot{x})d\dot{x} \\ &= \frac{1}{2\pi}\frac{\sigma_{\dot{X}}}{\sigma_X}\exp\left(-\frac{\xi^2}{2\sigma_X^2}\right), \quad \text{avec} \int_0^{+\infty} \dot{x}\exp\left(-\frac{1}{2}\frac{\dot{x}^2}{\sigma_{\dot{X}}^2}\right)d\dot{x} = \sigma_{\dot{X}}^2.\end{aligned}$$

$$\nu_\xi^+ = \frac{1}{2\pi}\frac{\sigma_{\dot{X}}}{\sigma_X}\exp\left(-\frac{\xi^2}{2\sigma_X^2}\right).$$

6. En particulier, pour $\xi = 0$, montrer que

$$\nu_0^+ = \frac{\omega_0^+}{2\pi}, \quad \omega_0^+ = \frac{\sigma_{\dot{X}}}{\sigma_X}.$$

Interpréter cette valeur.

— Plus $\sigma_{\dot{X}}$ est grand, et plus on a d'oscillations, et donc ν_0^+ est grand.

— Plus σ_X est grand, et plus les amplitudes de X sont grandes, et donc plus longtemps on peut être éloigné de 0, et donc plus ν_0^+ est petit.

7. On note maintenant $P_m(\xi)$ la probabilité que X admette un maximum de niveau ξ entre t et $t + dt$. Montrer que $P_m(\xi)$ peut s'écrire sous la forme :

$$P_m(\xi) = \int_{V_m(\xi)} p_{X(t),\dot{X}(t),\ddot{X}(t)}(x, \dot{x}, \ddot{x})dxd\dot{x}d\ddot{x},$$

où $p_{X(t),\dot{X}(t),\ddot{X}(t)}$ est la loi jointe de $\left(X(t), \dot{X}(t), \ddot{X}(t)\right)$ et où $V_m(\xi)$ est un volume à définir.

Pour qu'un maximum ait lieu entre t et $t+dt$, il faut que les deux conditions suivantes soient vérifiées :

$$(1)\ \dot{X}(t) > 0, \quad (2)\ \dot{X}(t + dt) < 0.$$

Au premier ordre : $\dot{X}(t + dt) = \dot{X}(t) + \ddot{X}(t)dt$, si bien que la condition précédente s'écrit :

$$0 < \dot{X}(t) < -\ddot{X}(t)dt.$$

Comme $dt > 0$, cette inégalité implique en particulier la condition suivante sur $\ddot{X}(t)$:

$$\ddot{X}(t) < 0.$$

Soit $P_m(\xi)$ la probabilité pour que $X(t)$ admette un maximum de valeur ξ entre t et $t + dt$. On peut alors définir $V_m(\xi)$ le volume tel que :

$$P_m(\xi) = \int_{V_m(\xi)} p_{X(t),\dot{X}(t),\ddot{X}(t)}(x, \dot{x}, \ddot{x})dxd\dot{x}d\ddot{x}.$$

D'après les conditions précédentes, il vient :

$$V_m(\xi) = \{(x, \dot{x}, \ddot{x}),\ |\ \xi \leq x \leq \xi + dx,\ 0 < \dot{x} < -\ddot{x}dt,\ \ddot{x} < 0\}\,.$$

8. On note $p_m(\xi)$ la densité des maxima par unité de temps. On admet que pour des champs gaussiens stationnaires centrés,

$$p_m(\xi) = \frac{1}{\sqrt{2\pi m_0}} \left\{ \begin{array}{l} \epsilon \exp\left(-\dfrac{\xi^2}{2m_0\epsilon^2}\right) + \sqrt{1-\epsilon^2} \times \dfrac{\xi}{\sqrt{m_0}} \times \\ \exp\left(-\dfrac{\xi^2}{2m_0}\right) \displaystyle\int_{-\infty}^{\frac{\xi}{\sqrt{m_0}}\frac{\sqrt{1-\epsilon^2}}{\epsilon}} \exp\left(-\dfrac{y^2}{2}\right) dy \end{array} \right\},$$

$$\epsilon^2 = \frac{m_0 m_4 - m_2^2}{m_0 m_4} = 1 - \frac{m_2^2}{m_0 m_4} = 1 - \left(\frac{\nu_0^+}{\mu}\right)^2,$$

$$\mathbb{E}\left[(X(t), \dot{X}(t), \ddot{X}(t))(X(t), \dot{X}(t), \ddot{X}(t))^T\right] = \begin{bmatrix} m_0 & 0 & -m_2 \\ 0 & m_2 & 0 \\ -m_2 & 0 & m_4 \end{bmatrix}.$$

avec ν_0^+ la fréquence des passages par zéro et μ la fréquence des maxima. Commenter alors les distributions asymptotiques quand $\epsilon \to 0$ et $\epsilon \to 1$. Interpréter graphiquement ces résultats.

— Si X est à bande étroite à fréquence centrale f_0, alors $\mu \approx \nu_0^+ \approx f_0$, et $\epsilon \approx 0$.
— Si X est à bande large, alors il existe un nombre très important de maxima entre deux passages par 0, et $\mu \gg \nu_0^+$. Il vient : $\epsilon \to 1$.

Si $\epsilon \to 1$, alors :

$$p_m(\xi) \to \frac{1}{\sqrt{2\pi m_0}} \exp\left(-\frac{\xi^2}{2m_0}\right) \text{ (GAuSS).}$$

Si $\epsilon \to 0$, alors :

$$p_m(\xi) \to \frac{\xi}{m_0} \exp\left(-\frac{\xi^2}{2m_0}\right) \text{ (RAYLEIGH).}$$

20 Méthodes de réduction de variance

L'idée de cet exercice est d'illustrer les principales méthodes de réduction de variance, afin d'accélérer les convergences des extimateurs Monte-Carlo. On cherche ainsi à évaluer numériquement la valeur de π, à partir de tirages aléatoires de deux variables indépendantes X_1 et X_2, qui sont toutes deux uniformément réparties sur $[0, 1]$.

On note $\Omega = \left\{(x_1, x_2) \in [0,1] \times [0,1] \mid x_1^2 + x_2^2 \leq 1\right\}$, et on définit 1_Ω la fonction indicatrice définie sur $[0, 1] \times [0, 1]$ telle que $1_\Omega(x_1, x_2) = 1$ si $(x_1, x_2) \in \Omega$ et 0 sinon.

Approche Monte-Carlo classique.

1. Exprimer π en fonction de 1_Ω.

2. Définir l'estimateur Monte-Carlo, $\widehat{I}_N$, de π associé. Calculer la moyenne et la variance de cet extimateur, $\mathrm{Var}(\widehat{I}_N)$.

3. Déduire de cette variance le nombre moyen de tirages de X_1 et X_2 nécessaires à une estimation de π à 10^{-4} près avec une confiance de 95%.

Monte-Carlo conditionnel

4. En utilisant le fait que $x_1^2+x_2^2 \leq 1 \Leftrightarrow x_2 \leq \sqrt{1-x_1^2}$, récrire π en fonction de X_1 seulement. En déduire un nouvel estimateur de π, nommé $\widehat{I}_N^{\mathrm{cond}}$.

5. Evaluer la variance de ce nouvel estimateur, puis comparer au cas Monte-Carlo précédent.

6. On propose de combiner l'estimateur précédent à une **stratification** des tirages. Pour cela, on propose de concentrer αN tirages à l'intervalle $[0, 1/2]$ et seulement $(1-\alpha)N$ tirages à l'intervalles $[1/2, 1]$. On nomme alors Y et Z les variables aléatoires uniformément distribuées sur $[0, 1/2]$ et $[1/2, 1]$ respectivement. Proposer alors un nouvel estimateur de π, que l'on nomme $\widehat{I}^{\mathrm{ST}}$.

7. Montrer que :

$$\mathrm{Var}(\widehat{I}^{\mathrm{ST}}) = \frac{1}{4N}\left\{\frac{1}{\alpha}\mathrm{Var}(\sqrt{1-Y^2}) + \frac{1}{1-\alpha}\mathrm{Var}(\sqrt{1-Z^2})\right\}.$$

8. Calculer $C_Y = \mathrm{Var}(\sqrt{1-Y^2})$ et $C_Z = \mathrm{Var}(\sqrt{1-Z^2})$.

9. Calculer le rapport optimal $\alpha^\star$ minimisant $\mathrm{Var}(\widehat{I}^{\mathrm{ST}})$ en fonction de C_Y et C_Z.

10. Evaluer la valeur de la variance associée à cette valeur de $\alpha^\star$ et comparer aux cas précédents.

Variable de contrôle et tirage d'importance

11. On se donne une fonction de contrôle g_r, ainsi que la constante associée $I_r = \mathbb{E}\left[g_r(X_1)\right]$ que l'on suppose connue. On nomme alors $\widehat{I}_N^{VC} = bI_r + \frac{1}{N}\sum_{n=1}^N \left(\sqrt{1-X_1^2(\theta_n)} - bg_r(X_1)\right)$. Exprimer la variance de cet estimateur en fonction de b, $\mathrm{Var}(\widehat{I}_N^{\mathrm{cond}})$, $\mathrm{Cov}\left(\sqrt{1-X_1^2}, g_r(X_1)\right)$ et $\mathrm{Var}(g_r(X_1))$.

12. Evaluer la valeur optimale $b^\star$ permettant de minimiser la variance de $\widehat{I}_N^{VC}$.

13. Pour $g_r(x) = (1-x)^2$, on trouve :

$$\rho^2 = \frac{\mathrm{Cov}\left(\sqrt{1-X_1^2}, g_r(X)\right)^2}{\mathrm{Var}(g_r(X_1))\mathrm{Var}(\sqrt{1-X_1^2})} \approx 0.642.$$

En déduire la variance de l'estimateur associé à cette fonction de contrôle.

14. On récrit :

$$\frac{\pi}{4} = \int_0^1 \sqrt{1-x^2}dx = \int_0^1 \sqrt{1-x^2}\frac{f(x)}{f(x)}dx = \mathbb{E}_f\left[\frac{\sqrt{1-X_1^2}}{f(X_1)}\right].$$

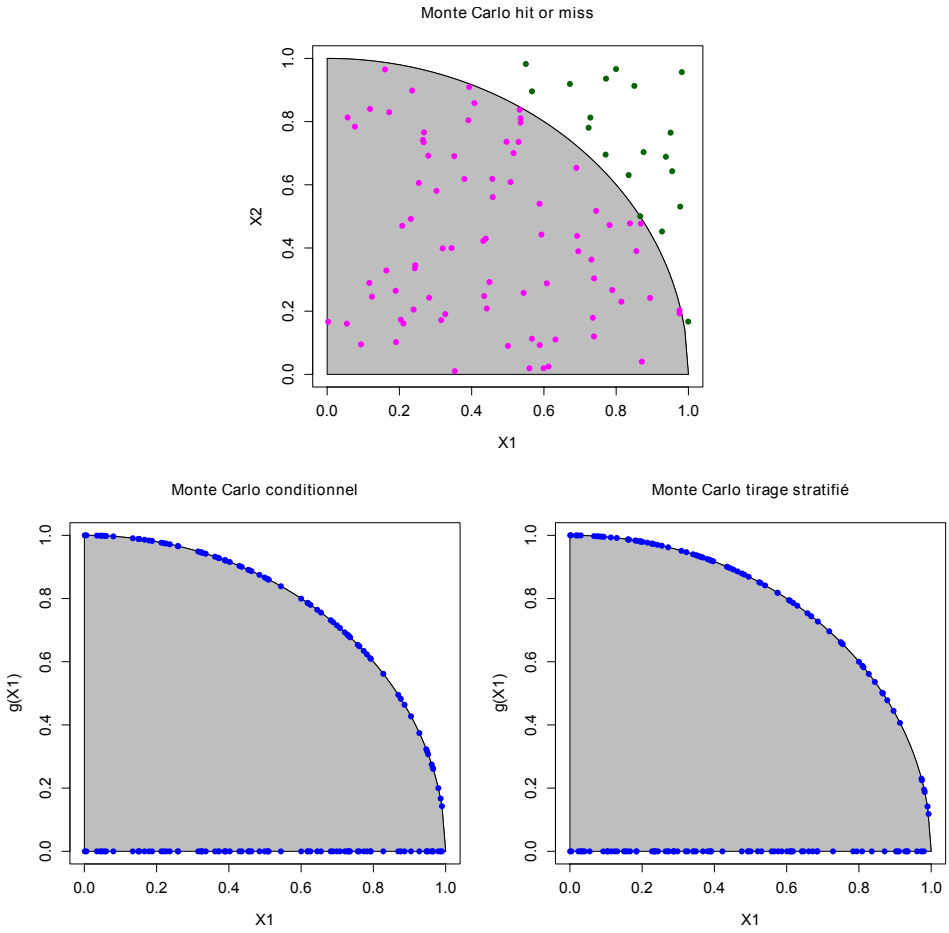
On définit alors l'estimateur de tirage d'importance :

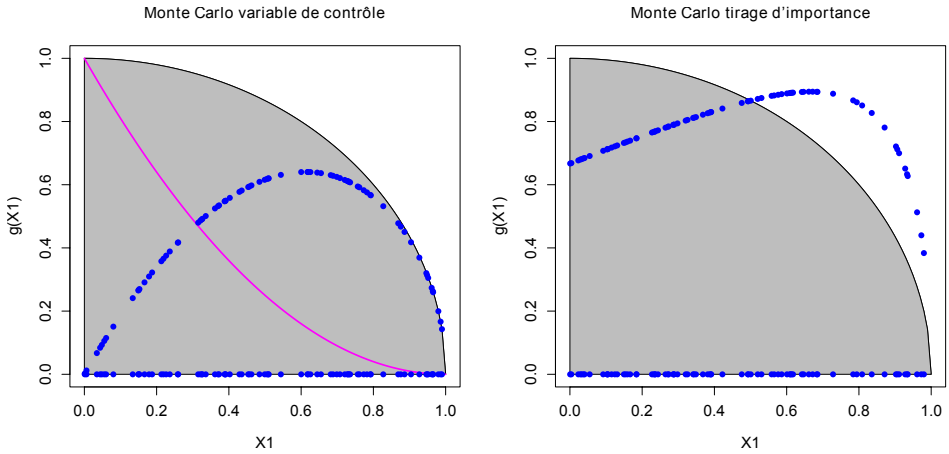
$$\widehat{I}_N^{TI} = \frac{1}{N}\sum_{i=1}^{N}\frac{\sqrt{1-\widetilde{X}^2}}{f(\widetilde{X})},$$

avec $\widetilde{X}$ une variable aléatoire de PDF f. Exprimer la variance de ce nouvel estimateur.

15. Pour $f(x) = 1.5 - x$, on admet que $\mathrm{Var}(\widehat{I}_N^{TI}) \approx \frac{0.0099}{N}$. Expliquer dans quelle mesure $f(x)$ permet de réduire cette variance.

16. Commenter l'ensemble des méthodes de réduction de variance proposées, en terme d'efficacité numérique.





Correction

1. $\frac{\pi}{4} = I = \int_0^1 \int_0^1 1_\Omega(x_1, x_2) dx_1 d_2 \approx 0.7854$.
2. $\widehat{I}_N = \frac{1}{N} \sum_{i=1}^N 1_\Omega(X_1^i, X_2^i)$. On déduit :

$$\mathbb{E}\left[\widehat{I}_N\right] = I,$$
$$\mathrm{Var}(\widehat{I}_N) = \mathbb{E}\left[\left(\widehat{I}_N - I\right)^2\right] = \frac{I(1-I)}{N}.$$

3. $\mathrm{Var}(\widehat{I}_N) \approx 0.1685/N \Rightarrow$ un intervalle contenant 95% de chance d'avoir $\pi/4$ est donné par $\Delta I = 2 \times 1.96 \times \sqrt{\frac{I(1-I)}{N}}$. Il faut alors environ $N = 2.590 \times 10^8$ pour $\Delta I = 10^{-4}$.

4. On remarque que :

$$I = \int_0^1 \int_0^1 1_{x_2 \leq \sqrt{1-x_1^2}}(x_1, x_2) dx_1 dx_2 = \int_0^1 \sqrt{1-x_1^2} dx_1.$$

On déduit alors le nouvel estimateur :

$$\widehat{I}_N^{\mathrm{cond}} = \frac{1}{N} \sum_{i=1}^N \sqrt{1-(X_1^i)^2}.$$

5. On calcule :

$$\mathrm{Var}(\widehat{I}_N^{\mathrm{cond}}) = \frac{\mathrm{Var}\left(\sqrt{1-X_1^2}\right)}{N},$$

$$\begin{aligned}
\mathrm{Var}\left(\sqrt{1-X_1^2}\right) &= \int_0^1 \left(\sqrt{1-x_1^2}\right)^2 dx_1 - \left(\int_0^1 \sqrt{1-x_1^2} dx_1\right)^2 \\
&= \left(1 - \frac{1}{3}\right) - \left(\int_0^{\pi/2} \sqrt{1-\sin(u)^2}\cos(u) du\right)^2 \\
&= \frac{2}{3} - \left(\frac{\pi}{4}\right)^2 \approx 0.0498.
\end{aligned}$$

On observe ainsi une nette réduction de la variance.

6. On stratifie. Pour cela, on définit $Y \sim u[0,1/2]$ et $Z \sim u[1/2,1]$, Y et Z indépendants, $N_Y = \alpha N$ et $N_Z = (1-\alpha)N$, et on déduit :

$$\int_0^1 \sqrt{1-x_1^2}dx_1 = 2\times\frac{1}{2}\int_0^{0.5}\sqrt{1-x_1^2}dx_1 + 2\times\frac{1}{2}\int_{0.5}^1\sqrt{1-x_1^2}dx_1 = \frac{1}{2}\mathbb{E}\left[\sqrt{1-Y^2}\right] + \frac{1}{2}\mathbb{E}\left[\sqrt{1-Z^2}\right],$$

$$\widehat{I}_N^{\text{strat}} = \frac{1}{2N_Y}\sum_{i=1}^{N_Y}\sqrt{1-Y_i^2} + \frac{1}{2N_Z}\sum_{j=1}^{N_Z}\sqrt{1-Z_j^2}.$$

7. On calcule :

$$\begin{aligned}\text{Var}\left(\widehat{I}_N^{\text{strat}}\right) &= \frac{1}{4}\left\{\frac{\text{Var}\left(\sqrt{1-Y^2}\right)}{N_Y} + \frac{\text{Var}\left(\sqrt{1-Z^2}\right)}{N_Z}\right\} \\ &= \frac{1}{4N}\left\{\frac{1}{\alpha}\text{Var}(\sqrt{1-Y^2}) + \frac{1}{1-\alpha}\text{Var}(\sqrt{1-Z^2})\right\}.\end{aligned}$$

8. On calcule :

$$\begin{aligned}\text{Var}\left(\sqrt{1-Y^2}\right) &= 2\int_0^{0.5}\left(\sqrt{1-y^2}\right)^2 dy - \left(2\int_0^{0.5}\sqrt{1-y^2}dy\right)^2 \\ &= 2\left(\frac{1}{2} - \frac{0.5^3}{3}\right) - 4\left(\int_0^{\pi/6}\sqrt{1-\sin(u)^2}\cos(u)du\right)^2 \\ &= \left(1-\frac{1}{12}\right) - 4\left(\int_0^{\pi/6}\cos^2(u)du\right)^2 \\ &= \left(1-\frac{1}{12}\right) - 4\left(\int_0^{\pi/6}\frac{1+\cos(2u)}{2}du\right)^2 \\ &= \left(1-\frac{1}{12}\right) - 4\left(\frac{\pi}{12} + \frac{\sqrt{3}}{8}\right)^2 \approx 0.0016.\end{aligned}$$

$$\begin{aligned}\text{Var}\left(\sqrt{1-Z^2}\right) &= 2\int_{0.5}^1\left(\sqrt{1-z^2}\right)^2 dz - \left(2\int_{0.5}^1\sqrt{1-z^2}dz\right)^2 \\ &= 2\left(\frac{1}{2} - \frac{1}{3} + \frac{0.5^3}{3}\right) - 4\left(\int_{\pi/6}^{\pi/2}\frac{1+\cos(2u)}{2}du\right)^2 \\ &= \left(1+\frac{1}{12}-\frac{2}{3}\right) - 4\left(\frac{\pi}{4} - \frac{\pi}{12} - \frac{\sqrt{3}}{8}\right)^2 \approx 0.0387.\end{aligned}$$

9. On calcule :

$$\frac{\partial \mathrm{Var}\left(\widehat{I}_N^{\mathrm{strat}}\right)}{\partial \alpha} = \frac{1}{4N}\left\{-\frac{1}{\alpha^2}C_Y + \frac{1}{(1-\alpha)^2 C_Z}\right\}.$$

Ainsi :

$$\begin{aligned}\frac{\partial \mathrm{Var}\left(\widehat{I}_N^{\mathrm{strat}}\right)}{\partial \alpha} = 0 &\Leftrightarrow \sqrt{C_Y}(1-\alpha) = \sqrt{C_Z}\alpha \\ &\Leftrightarrow \alpha = \frac{\sqrt{C_Y}}{\sqrt{C_Y}+\sqrt{C_Z}}\end{aligned}$$

On déduit : $\alpha^\star \approx 0.1663$.

10. On trouve $\mathrm{Var}\left(\widehat{I}_N^{\mathrm{strat}}\right) \approx \frac{0.0141}{N}$. On a encore beaucoup réduit la variance avec cette stratification.

11. On calcule :

$$\mathrm{Var}(\widehat{I}_N^{\mathrm{VC}}) = \mathrm{Var}(\widehat{I}_N^{\mathrm{cond}}) - \frac{2b}{N}\mathrm{Cov}\left(\sqrt{1-X_1^2}, g_r(X)\right) + \frac{b^2}{N}\mathrm{Var}(g_r(X)).$$

12. On calcule :

$$\frac{\partial \mathrm{Var}\left(\widehat{I}_N^{VC}\right)}{\partial b} = \frac{2}{N}\left\{b\mathrm{Var}(g_r(X)) - \mathrm{Cov}\left(\sqrt{1-X_1^2}, g_r(X)\right)\right\}.$$

Ainsi :

$$\frac{\partial \mathrm{Var}\left(\widehat{I}_N^{VC}\right)}{\partial b} = 0 \;\Leftrightarrow\; b = b^\star = \frac{\mathrm{Cov}\left(\sqrt{1-X_1^2}, g_r(X)\right)}{\mathrm{Var}(g_r(X))}.$$

13. En remplacant b par $b^\star$, on déduit :

$$\mathrm{Var}\left(\widehat{I}_N^{VC}\right) = (1-\rho^2)\mathrm{Var}\left(\widehat{I}_N^{\mathrm{cond}}\right) \approx \frac{0.0178}{N}.$$

14. Par construction, on a directement :

$$\mathrm{Var}(\widehat{I}_N^{TI}) = \frac{1}{N}\mathrm{Var}\left(\frac{\sqrt{1-\widetilde{X}^2}}{f(\widetilde{X})}\right).$$

15. On a choisi une PDF d'importance qui se rapproche de $\sqrt{1-x^2}$, ce qui permet de réduire fortement la variance de l'estimateur.

16. A chaque étape, en ajoutant de l'information, on peut réussir à réduire la variance de l'estimateur. Attention toutefois, ces réductions ne sont pas systématiques.

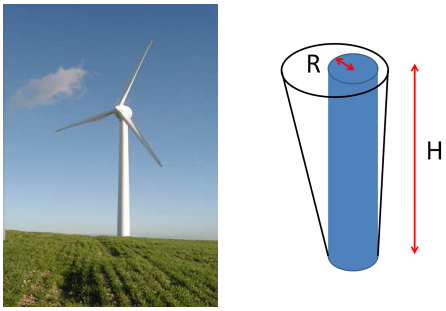


FIGURE 3 – Modélisation approchée d'une éolienne.

21 Analyse statique d'une éolienne et fiabilité

On définit les deux repères suivants :

$$\begin{array}{l}\text{repère cartésien} : (\boldsymbol{e}_x, \boldsymbol{e}_y, \boldsymbol{e}_z)\,, \\ \text{repère cylindrique} : (\boldsymbol{e}_r(\theta), \boldsymbol{e}_\theta(\theta), \boldsymbol{e}_z)\end{array}.$$

On modélise l'éolienne comme un cylindre plein de rayon R, de section S et de hauteur H, et l'action du vent sous la forme d'une force de pression non uniforme (cf figure **??**) :

$$p(\theta, z) = cz^2 \sin^2\left(\frac{\theta}{2}\right).$$

1. Justifier, d'après l'expression de la pression, que le déplacement maximal de l'éolienne se situe en $z = H$ et selon $\boldsymbol{e}_x$. On admet que l'amplitude maximale de ce déplacement (appelée flèche) vaut :

$$v^{\max} = \frac{13cH^6}{90ER^3},$$

avec E la rigidité de l'éolienne.

2. On observe que les dimensions R et H sont connues de manière incertaine, et prennent respectivement des valeurs uniformément réparties dans $[R_0, R_0 + \Delta R]$, $[H_0, H_0 + \Delta H]$. On définit alors P_f la probabilité que la flèche soit supérieure à un seuil V. Montrer que :

$$P_f = \mathbb{P}(R^{-3}H^6 > S),$$

avec S un seuil à déterminer.

$\mathbb{P}(v^{\max} > V) = \mathbb{P}\left(H^6R^{-3} > S\right)$, avec $S = \frac{90EV}{13c}$.

3. Représenter sur un graphique la zone de défaillance, dans le cas où :

$$\begin{cases} R_0 \leq (H_0 + \Delta H)^2 S^{-1/3} \leq R_0 + \Delta R, \\ H_0 \leq R_0^{1/2} S^{1/6} \leq H_0 + \Delta H. \end{cases}$$

Expliquer alors comment calculer P_f (sans développer tous les calculs).

La limite est donnée par $H^6 = SR^3 \Leftrightarrow H = R^{1/2}S^{1/6}$. La zone limite est dans la partie supérieure gauche dans le plan (R,H). Il faut alors intégrer $\frac{1}{\Delta R \Delta H}$ sur cette zone.

22 Marges de sécurité

On s'intéresse à la garantie d'un système $\mathcal{S}$, pour lequel on aimerait s'assurer que la valeur "réelle" d'une quantité d'intérêt pour la sécurité, notée $y^{\text{réel}}$ soit inférieure à un seuil S. Pour cela, on suppose que l'on dispose d'un simulateur de cette quantité d'intérêt, notée y^{sim}. Cette quantité simulée est aléatoire, du fait qu'elle intègre un certain nombre d'incertitudes (numériques, paramétriques, de modèle). Pour assurer le bon fonctionnement du système, on se donne alors un niveau de risque acceptable α (par exemple $\alpha = 5\%$), et on cherche à vérifier que :

$$p = \mathbb{P}(y^{\text{sim}} > S) < \alpha.$$

Afin d'estimer cette probabilité p, on suppose disposer de N tirages indépendants de y^{sim}, notés $y_1, \dots, y_N$.

1. Rappeler l'estimateur Monte Carlo de p, noté $\widehat{p}$, ainsi que sa loi asymptotique lorsque $N \to +\infty$.

$\widehat{p} = \frac{1}{N}\sum_{n=1}^{N} 1_{y_n > S} \to \mathcal{N}(p, \frac{p(1-p)}{N})$.

2. Si $p = \alpha$, en déduire que $\mathbb{P}(\widehat{p} < \alpha) = 1/2$.

Direct de part le caractère symétrique et centré sur α de la loi asymptotique.

3. Afin de minimiser les chances de garantir un système ne devant pas être garanti, on se donne le critère suivant :

"On accepte de garantir le système $\mathcal{S}$ si $\widehat{p} + a < \alpha$".

Quel doit être le signe de a pour que l'introduction de la marge permette de réduire le nombre de garanties à tord ? Commenter l'appellation "marge de sécurité" pour a, ainsi que l'appellation "risque de garantie à tord" pour $r = \max_{p \geq \alpha} \mathbb{P}(\widehat{p} + a < \alpha)$.

Il faut que $a > 0$. Par construction, a est bien une marge de sécurité entre α et $\widehat{p}$ au sens où plus a est grand, moins le risque de garantie à tord est élevé. La notion de risque de garantie à tord est directe par la définition, car c'est bien la probabilité d'une garantie du système, sachant qu'il ne devrait pas être garanti car $p \geq \alpha$.

4. En supposant que $\widehat{p}$ suit sa loi asymptotique, montrer que $\mathbb{P}(\widehat{p}+a < \alpha) = \Phi\left(\frac{(\alpha-p-a)\sqrt{N}}{\sqrt{p(1-p)}}\right)$, où Φ est la fonction de répartition de la loi normale centrée réduite.

On calcule :

$$\begin{aligned}
\mathbb{P}(\widehat{p} + a < \alpha) &= \mathbb{P}(\widehat{p} < \alpha - a) \\
&= \int_{-\infty}^{\alpha-a} \frac{1}{\sqrt{2\pi}} \sqrt{\frac{N}{p(1-p)}} \exp\left(-\frac{N}{2p(1-p)}(p-x)^2\right) dx \\
&= \Phi\left(\frac{(\alpha-p-a)\sqrt{N}}{\sqrt{p(1-p)}}\right)
\end{aligned}$$

5. On se place dans la configuration où $p < \frac{1}{2}$. Montrer alors que $p \mapsto \mathbb{P}(\widehat{p} + a < \alpha)$ est une fonction décroissante vis-à-vis de p.

$p \mapsto -\alpha - p - a$ est une fonction décroissante avec p, tandis que pour $p < 0.5$, $p \mapsto \sqrt{p(1-p)}$ est croissante. Ainsi, $p \mapsto \frac{(\alpha - p - a)\sqrt{N}}{\sqrt{p(1-p)}}$ est décroissante, et donc $p \mapsto \Phi\left(\frac{(\alpha - p - a)\sqrt{N}}{\sqrt{p(1-p)}}\right)$ également.

6. En déduire la plus petite valeur de a permettant d'assurer que pour tout $p \geq \alpha$, $\mathbb{P}(\widehat{p} + a < \alpha) \leq 1 - \gamma$, c'est à dire la valeur de a permettant d'assurer que $r \leq 1 - \gamma$, avec γ un niveau de confiance donné (par exemple $\gamma = 95\%$).

Comme $\mathbb{P}(\widehat{p} + a < \alpha)$ est décroissante avec p, la plus grande valeur de $\mathbb{P}(\widehat{p} + a < \alpha)$ est donnée par p le plus petit possible, c'est à dire choisir $p = \alpha$:

$$\max_{p \geq \alpha} \mathbb{P}(\widehat{p} + a < \alpha) = \Phi\left(\frac{-a\sqrt{N}}{\sqrt{\alpha(1-\alpha)}}\right).$$

La valeur de a seuil est alors donnée par $\Phi\left(\frac{-a\sqrt{N}}{\sqrt{\alpha(1-\alpha)}}\right) = 1 - \gamma$, c'est à dire $\frac{-a\sqrt{N}}{\sqrt{\alpha(1-\alpha)}} = q^{\mathcal{N}}_{1-\gamma}$, soit $a = -q^{\mathcal{N}}_{1-\gamma}\sqrt{\frac{\alpha(1-\alpha)}{N}}$.

23 Fiabilité et processus ponctuels

On s'intéresse à une quantité Y de PDF f_Y. On cherche à calculer $P_f = \mathbb{P}(Y > S)$, où S est un seuil donné.

1. Soient $Y_1, \ldots, Y_N$ N réalisations indépendantes de Y. En déduire l'estimateur Monte Carlo de P_f, que l'on nomme $\widehat{P}_f$.

$\widehat{P}_f = \frac{1}{N}\sum_{n=1}^{N} 1_{Y_n > S}$.

2. Calculer la moyenne et la variance de cet estimateur $\widehat{P}_f$.

$\mathbb{E}\left[\widehat{P}_f\right] = P_f$, $\mathrm{Var}(\widehat{P}_f) = \frac{P_f(1-P_f)}{N}$.

3. On note $\delta = \sqrt{\mathrm{Var}(\widehat{P}_f)}/\mathbb{E}\left[\widehat{P}_f\right]$ le coefficient de variation de cet estimateur. En déduire la valeur de N pour que δ soit inférieur ou égal à 0.1. Donner les valeurs de N associées aux valeurs $P_f \in \left\{10^{-1}, 10^{-3}, 10^{-5}\right\}$. Commenter.

$0.1^2 = 0.01 = \frac{1-P_f}{P_f N} \Leftrightarrow N = \frac{1-P_f}{P_f 0.01}$. On trouve alors $N = 900, 99900, 9999900$.

4. Pour accélérer cette estimation de P_f, on introduit les deux probabilités suivantes, pour $S_1 < S$:

$$P_1 = \mathbb{P}(Y > S_1), \quad P_2 = \mathbb{P}(Y > S \mid Y > S_1), \tag{5}$$

et on note $\widehat{P}_1$ et $\widehat{P}_2$ leurs estimateurs Monte Carlo associés à $N_1 \leq N$ tirages indépendants de Y et $N - N_1$ tirages indépendants de $(Y > S \mid Y > S_1)$. En déduire un nouvel estimateur de P_f en fonction de $\widehat{P}_1$ et $\widehat{P}_2$, que l'on nomme cette fois $\widehat{P}_f$.

$\widetilde{P}_f = \widehat{P}_1 \times \widehat{P}_2$.

5. Montrer que $\mathbb{E}\left[\widetilde{P}_f\right] = P_f$ et que :

$$\mathrm{Var}(\widetilde{P}_f) = \frac{P_1(1-P_1)P_2(1-P_2)}{N_1(N-N_1)} + \frac{P_1(1-P_1)P_2^2}{N_1} + \frac{P_2(1-P_2)P_1^2}{N-N_1}. \tag{6}$$

Comme $\widehat{P}_1$ et $\widehat{P}_2$ sont des estimateurs non biaisés et indépendants statistiquement, on déduit directement que $\mathbb{E}\left[\widetilde{P}_f\right] = \mathbb{E}\left[\widehat{P}_1\right]\mathbb{E}\left[\widehat{P}_2\right] = P_1P_2 = P_f$. Par ailleurs, on a :

$$\begin{aligned}\mathrm{Var}(\widetilde{P}_f) &= \mathrm{Var}(\widehat{P}_1\widehat{P}_2)\\ &= \mathrm{Var}(\widehat{P}_1)\mathrm{Var}(\widehat{P}_2) + \mathrm{Var}(\widehat{P}_1)\mathbb{E}\left[\widehat{P}_2\right]^2 + \mathrm{Var}(\widehat{P}_2)\mathbb{E}\left[\widehat{P}_1\right]^2\\ &= \frac{P_1(1-P_1)}{N_1}\frac{P_2(1-P_2)}{N-N_1} + \frac{P_1(1-P_1)}{N_1}P_2^2 + \frac{P_2(1-P_2)}{N-N_1}P_1^2.\end{aligned} \tag{7}$$

6. Dans le cas où N est fixe et $P_1 = 1$, en déduire la valeur de N_1 permettant de minimiser $\mathrm{Var}(\widetilde{P}_f)$.

Si $P_1 = 1$, alors $\mathrm{Var}(\widetilde{P}_f) = \frac{P_2(1-P_2)}{N-N_1}$, et il faut prendre $N_1 = 0$.

7. De même, dans le cas où N est fixe et $P_2 = 1$, en déduire la valeur de N_1 permettant de minimiser $\mathrm{Var}(\widetilde{P}_f)$.

Si $P_2 = 1$, alors $\mathrm{Var}(\widetilde{P}_f) = \frac{P_1(1-P_1)}{N_1}$, et il faut prendre $N_1 = N$.

8. Dans le cas où $P_1 = P_2$, trouver la valeur de N_1 permettant de minimiser $\mathrm{Var}(\widetilde{P}_f)$.

Si $P_1 = P_2 = P$, alors

$$\begin{aligned}\mathrm{Var}(\widetilde{P}_f) &= \frac{P^2(1-P)^2}{N_1(N-N_1)} + P^3(1-P)\left(\frac{1}{N_1} + \frac{1}{N-N_1}\right),\\ &= \frac{P^2(1-P)^2 + P^3(1-P)N}{N_1(N-N_1)},\end{aligned} \tag{8}$$

qui est minimal en $N_1 = N/2$.

9. A valeur de N fixée, en déduire un critère sur la valeur de P_f à partir de laquelle une telle décomposition est intéressante, au sens où $\mathrm{Var}(\widetilde{P}_f) \leq \mathrm{Var}(\widehat{P}_f)$.

On sait que $P_1P_2 = P^2 = P_f$. On déduit : $\mathrm{Var}(\widetilde{P}_f) = \frac{4\left(P_f(1-\sqrt{P_f})^2 + P_f^{3/2}(1-\sqrt{P_f})N\right)}{N^2}$. Il vient :

$$\mathrm{Var}(\widetilde{P}_f) \leq \mathrm{Var}(\widehat{P}_f) \Leftrightarrow \frac{4\left(P_f(1-\sqrt{P_f})^2 + P_f^{3/2}(1-\sqrt{P_f})N\right)}{N^2} \leq \frac{(1-P_f)P_f}{N}. \tag{9}$$

10. Application numérique. Pour $N = 10^4$, et $P_f \in \{0.5, 0.1, 0.01\}$, indiquer s'il vaut mieux considérer $\widetilde{P}_f$ ou $\widehat{P}_f$.

On trouve $\mathrm{Var}(\widetilde{P}_f) - \mathrm{Var}(\widehat{P}_f) = 1.642307e-05, -3.490192e-07, -6.296760e-07$. Ce qui permet de conclure.

24 Indices de Sobol de la fonction Ishigami

On considère le modèle suivant :

$$y(\boldsymbol{X}) = \sin(X_1) + 7\sin(X_2)^2 + 0.1X_3^4 \sin(X_1), \quad \boldsymbol{X} = (X_1, X_2, X_3, X_4) \sim \sqcap\left([-\pi, \pi]^4\right).$$

1. Rappeler l'expression des indices de Sobol d'ordre 1, $S_1^{(i)}$, et totaux, $S_T^{(i)}$, caractérisant l'influence des entrées X_i sur la variance de y.

$S_1^{(i)} = \frac{\mathrm{Var}(\mathbb{E}[y(\boldsymbol{X})|X_i])}{\mathrm{Var}(y(\boldsymbol{X}))}, \quad S_T^{(i)} = 1 - \frac{\mathrm{Var}(\mathbb{E}[y(\boldsymbol{X})|X_{-i}])}{\mathrm{Var}(y(\boldsymbol{X}))}$

2. Donner la PDF de $\boldsymbol{X}$, $f_{\boldsymbol{X}}$.

$f_{\boldsymbol{X}}(\boldsymbol{x}) = \frac{1}{(2\pi)^4}$.

3. Calculer $\mathbb{E}\left[\sin(X_i)\right]$, $\mathbb{E}\left[\sin^2(X_i)\right]$, $\mathbb{E}\left[\sin^4(X_i)\right]$, $\mathbb{E}\left[X_i^4\right]$ et $\mathbb{E}\left[X_i^8\right]$.

$\mathbb{E}\left[\sin(X_i)\right] = 0$

$\mathbb{E}\left[\sin^2(X_i)\right] = \frac{1}{2\pi}\int_{-\pi}^{\pi} sin^2(X_i)dX_i = \frac{1}{2}$.

$$\begin{aligned}
\mathbb{E}\left[\sin^4(X_i)\right] &= \frac{1}{2\pi}\int_{-\pi}^{\pi} sin^4(X_i)dX_i \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{(1-\cos(2X_i))}{2}\frac{(1-\cos(2X_i))}{2}dX_i \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{1}{4}\left(1 + \cos^2(2X_i) - 2\cos(2X_i)\right)dX_i \\
&= \frac{1}{2\pi}\left(\frac{2\pi}{4} + \frac{\pi}{4}\right) \\
&= \frac{3}{8}
\end{aligned}$$

$$\begin{aligned}
\mathbb{E}\left[X_i^4\right] = &= \frac{1}{2\pi}\int_{-\pi}^{\pi} X_i^4 dX_i \\
&= \frac{1}{2\pi}\left(\frac{2\pi^5}{5}\right) \\
&= \frac{\pi^4}{5}
\end{aligned}$$

$$\begin{aligned}\mathbb{E}\left[X_i^8\right] = &= \frac{1}{2\pi}\int_{-\pi}^{\pi} X_i^8 dX_i \\ &= \frac{1}{2\pi}\left(\frac{2\pi^9}{9}\right) \\ &= \frac{\pi^8}{9}\end{aligned}$$

4. Calculer la moyenne de y.

$\mathbb{E}[y] = \frac{7}{2}$

5. Calculer la variance de y.

$$\begin{aligned}\mathbb{E}\left[y^2\right] &= \mathbb{E}\left[\sin^2(X_1)\right] + 49\mathbb{E}\left[\sin^4(X_2)\right] + 0.01\mathbb{E}\left[X_3^8\right]\mathbb{E}\left[\sin^2(X_1)\right] + 0.2\mathbb{E}\left[\sin^2(X_1)\right]\mathbb{E}\left[X_3^4\right] + 0 + 0 \\ &= \frac{1}{2} + \frac{49\times 3}{8} + 0.01\frac{\pi^8}{18} + 0.2\frac{\pi^4}{10}\end{aligned}$$

$$\mathrm{Var}(y) = \mathbb{E}\left[y^2\right] - \mathbb{E}[y]^2 \approx 13.84$$

6. Calculer les indices $S_1^{(i)}$ et $S_T^{(i)}$.

Pour X_4 tout est toujours nul.

$\mathbb{E}[y|X_1] = \frac{7}{2} + \sin(X_1)\left(1 + 0.1\pi^4/5\right)$
$\mathbb{E}[y|X_2] = 7\sin^2(X_2)$
$\mathbb{E}[y|X_3] = \frac{7}{2}$
$\mathbb{E}[y|X_2, X_3] = 7\sin^2(X_2)$
$\mathbb{E}[y|X_1, X_3] = \frac{7}{2} + \sin(X_1)\left(1 + 0.1X_3^4\right)$
$\mathbb{E}[y|X_1, X_2] = 7\sin^2(X_2) + \sin(X_1)\left(1 + 0.1\pi^4/5\right)$

On déduit :
$\mathrm{Var}\left(\mathbb{E}[y|X_1]\right) = \frac{1}{2}\left(1 + 0.1\pi^4/5\right)^2 \approx 4.35 \;\Rightarrow S_1^{(1)} \approx 0.31$
$\mathrm{Var}\left(\mathbb{E}[y|X_2]\right) = \frac{49}{8} = 6.125 \;\Rightarrow S_1^{(2)} \approx 0.44$
$\mathrm{Var}\left(\mathbb{E}[y|X_3]\right) = 0 \;\Rightarrow S_1^{(3)} = 0$
$\mathrm{Var}\left(\mathbb{E}[y|X_2, X_3]\right) = \frac{49}{8} = 6.125 \;\Rightarrow S_T^{(1)} \approx 0.55$
$\mathrm{Var}\left(\mathbb{E}[y|X_1, X_3]\right) = \frac{1}{2}\left(1 + 0.2\pi^4/5 + 0.01\pi^8/9\right) \approx 7.72 \;\Rightarrow S_T^{(2)} \approx 0.44$
$\mathrm{Var}\left(\mathbb{E}[y|X_1, X_2]\right) = \frac{49}{8} + \frac{1}{2}\left(1 + 0.1\pi^4/5\right)^2 = 10.47 \;\Rightarrow S_T^{(3)} \approx 0.24$

7. Commenter les résultats précédents.

X_3 ne joue pas seul sur la moyenne de y, mais joue beaucoup conjointement à X_1.

8. Sans refaire tous les calculs, commenter l'influence d'un augmentation du domaine de définition de $\boldsymbol{X}$ à $[0, 2\pi]^4$.

Le seul impacté sera X_3 dont l'influence devrait fortement augmenter (les autres sont 2π périodiques).

25 Indices de Sobol pour variables corrélées

On considère la fonction :

$$y : \begin{cases} \mathbb{R}^3 \ \to \ \mathbb{R} \\ \boldsymbol{x} = (x_1, x_2, x_3) \ \mapsto \ y(\boldsymbol{x}) = \sqrt{2}x_1 + x_2 + \dfrac{x_3}{\sqrt{2}}. \end{cases}$$

On suppose que $\boldsymbol{x}$ est un vecteur gaussien **centré** et de matrice de covariance :

$$\boldsymbol{R} := \begin{bmatrix} 1 & 0 & \rho \\ 0 & 1 & 0 \\ \rho & 0 & 1 \end{bmatrix}, \quad -1 \leq \rho \leq 1.$$

1. Rappeler la densité $f_{\boldsymbol{x}}$ de $\boldsymbol{x}$ en fonction de $\boldsymbol{R}$. Peut-on dire que les variables x_1, x_2, x_3 sont indépendantes statistiquement ?

$$f_{\boldsymbol{x}}(\boldsymbol{x}) = \frac{1}{(2\pi)^{3/2}\sqrt{\det(\boldsymbol{R})}} \exp\left(-\frac{1}{2}\boldsymbol{x}^T \boldsymbol{R}^{-1}\boldsymbol{x}\right). \tag{10}$$

Les composantes x_1 et x_3 sont dépendantes statistiquement.

2. Calculer la moyenne et la variance de $y(\boldsymbol{x})$.

$\mathbb{E}[y(\boldsymbol{x})] = 0,$
$\mathrm{Var}(y(\boldsymbol{x})) = \mathbb{E}\left[(y(\boldsymbol{x}))^2\right] = 2 + 1 + 1/2 + 2\rho = 7/2 + 2\rho.$

On rappelle que si $\boldsymbol{y} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{R}_y)$ et $\boldsymbol{z} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{R}_z)$ sont deux vecteurs gaussiens centrés corrélés, tels que $\mathrm{Cov}(\boldsymbol{y}, \boldsymbol{z}) = \boldsymbol{R}_{yz}$, alors :

$$\mathbb{E}\left[\boldsymbol{y}|\boldsymbol{z}\right] = \boldsymbol{R}_{yz}\boldsymbol{R}_z^{-1}\boldsymbol{z}.$$

3. Calculer $\mathbb{E}[y(\boldsymbol{x})|x_1]$, $\mathbb{E}[y(\boldsymbol{x})|x_2]$, $\mathbb{E}[y(\boldsymbol{x})|x_3]$.

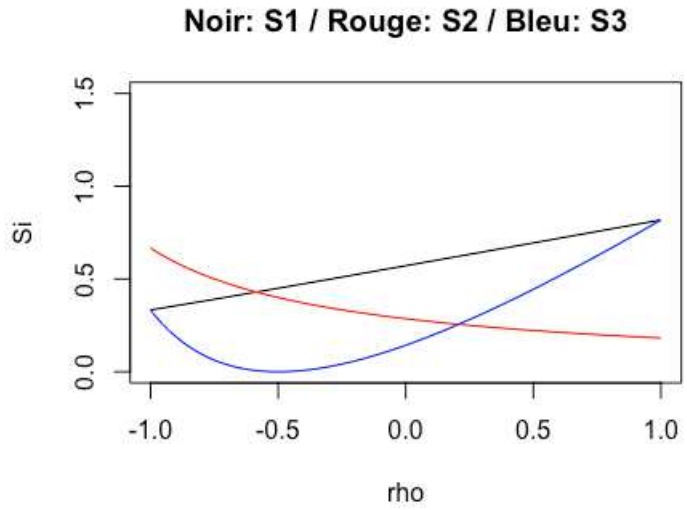
$\mathrm{Cov}(y(\boldsymbol{x}), x_1) = \sqrt{2} + \rho/\sqrt{2}.$
$\mathrm{Cov}(y(\boldsymbol{x}), x_2) = 1.$
$\mathrm{Cov}(y(\boldsymbol{x}), x_3) = \sqrt{2}\rho + 1/\sqrt{2}.$
On déduit :
$\mathbb{E}[y(\boldsymbol{x})|x_1] = (\sqrt{2} + \rho/\sqrt{2})x_1$
$\mathbb{E}[y(\boldsymbol{x})|x_2] = x_2$
$\mathbb{E}[y(\boldsymbol{x})|x_3] = (\sqrt{2}\rho + 1/\sqrt{2})x_3$

4. En déduire les indices de Sobol d'ordre 1 associés à x_1, x_2, x_3, respectivement notés S_1, S_2, S_3.

$\mathrm{Var}(\mathbb{E}[y(\boldsymbol{x})|x_1]) = (\sqrt{2} + \rho/\sqrt{2})^2$
$\mathrm{Var}(\mathbb{E}[y(\boldsymbol{x})|x_2]) = 1$
$\mathrm{Var}(\mathbb{E}[y(\boldsymbol{x})|x_1]) = (\sqrt{2}\rho + 1/\sqrt{2})^2$

$S_1 = (\sqrt{2} + \rho/\sqrt{2})^2/(7/2 + 2\rho)$
$S_2 = 1/(7/2 + 2\rho)$
$S_3 = (\sqrt{2}\rho + 1/\sqrt{2})^2/(7/2 + 2\rho)$

5. Tracer sur un même graphique l'évolution de ces indices en fonction de $\rho \in [-1, 1]$. Commenter le résultat (en particulier les cas $\rho \in \{-1, 0, 1\}$).



Pour $\rho = -1$, $x_1 = -x_3$, et c'est x_2 qui a le plus d'influence, car $\sqrt{2} - 1/\sqrt{2} < 1$. Pour $\rho = 0$, on est dans le cas indépendant, et on retrouve $x_1 > x_2 > x_3$ (tri par ordre décroissant des coefficients devant les x_i). Par contre, pour $\rho = 1$, $x_1 = x_3$ et ce sont maintenant x_1 et x_3 qui jouent le plus. Pour les valeurs intermédiaires, on vérifie que l'on a toujours bien $x_1 > x_3$ en terme d'influence, et x_1 devient rapidement le plus influent.

6. Calculer la somme des indices d'ordre 1. Que constatez vous comme différence (surprenante) avec le cas des variables indépendantes (on pourra à nouveau se concentrer sur les valeurs $\rho \in \{-1, 0, 1\}$) ? Comment l'expliquez vous ?

$S_1 + S_2 + S_3 = 1 + \frac{5\rho^2/2+2\rho}{7/2+2\rho}$
On obtient des valeurs pouvant être plus grandes que 1 (notamment en $\rho = -1$ ou 1, alors que pour le cas indépendant ($\rho = 0$), la somme vaut bien 1. Cela s'explique par la corrélation, qui fait contribuer x_1 pour S_1 et S_3, et x_3 pour S_1 et S_3 également => des contributions sont comptées deux fois, d'où la somme plus grande que 1.

26 Approches spectrales et indices de Sobol

On s'intéresse à l'influence des entrées $\boldsymbol{X} = (X_1, \ldots, X_d)$, que l'on suppose indépendantes statistiquement et de même loi (après renormalisation par exemple), sur la variance de $y(\boldsymbol{X})$. On propose alors de passer par une approche spectrale. Pour cela, à partir d'un nombre limité d'évaluations de y, on suppose avoir identifié une approximation de y sous la forme :

$$y(\boldsymbol{X}) \approx \sum_{\boldsymbol{\alpha}=(\alpha_1,\ldots,\alpha_d)\in\mathcal{A}} c_{\boldsymbol{\alpha}} \psi_{\alpha_1}(X_1) \times \cdots \times \psi_{\alpha_d}(X_d),$$

où $\{\psi_0, \psi_1, \psi_2, \ldots\}$ est une base orthonormée par rapport à la loi des X_i, telle que :

$$\psi_0(X_i) = 1, \quad \mathbb{E}\left[\psi_m(X_i)\psi_n(X_i)\right] = \delta_{nm}, \quad n, m \geq 0, \quad 1 \leq i \leq d,$$

et où $\mathcal{A}$ est un sous ensemble fini de $\mathbb{N}^d$.

1. Calculer la moyenne de y.

$\mathbb{E}[y] = c_{0,0,\ldots,0}$.

2. Calculer la variance de y.

$$\mathbb{E}\left[y(\boldsymbol{X})^2\right] = \sum_{\boldsymbol{\alpha}\in\mathcal{A}} \sum_{\boldsymbol{\alpha}'\in\mathcal{A}} c_{\boldsymbol{\alpha}} c_{\boldsymbol{\alpha}'} \prod_{i=1}^{d} \mathbb{E}\left[\psi_{\alpha_i}(X_i)\psi_{\alpha_i'}(X_i)\right] = \sum_{\boldsymbol{\alpha}\in\mathcal{A}} c_{\boldsymbol{\alpha}}^2$$

donc $\mathrm{Var}\left(y(\boldsymbol{X})\right) = \sum_{g\alpha\in\mathcal{A}\setminus\{0,\ldots,0\}} c_{\boldsymbol{\alpha}}^2$.

3. Calculer $\mathbb{E}\left[y(\boldsymbol{X})|X_i\right]$.

$$\begin{aligned}\mathbb{E}\left[y(\boldsymbol{X})|X_i\right] &= \sum_{\boldsymbol{\alpha}\in\mathcal{A}} c_{\boldsymbol{\alpha}} c_{\boldsymbol{\alpha}'} \psi_{\alpha_i}(X_i) \prod_{j\neq i} \mathbb{E}\left[\psi_{\alpha_j}(X_j)\right] \\ &= \sum_{\boldsymbol{\alpha}\in\{\boldsymbol{\alpha}\in\mathcal{A}|\alpha_j=0,\ j\neq i\}} c_{\boldsymbol{\alpha}}\psi_{\alpha_i}(X_i)\end{aligned}$$

4. En déduire $S_1^{(i)} = \mathrm{Var}(\mathbb{E}\left[y(\boldsymbol{X})|X_i\right])/\mathrm{Var}(y(\boldsymbol{X}))$.

$$\begin{aligned}\mathrm{Var}(\mathbb{E}\left[y(\boldsymbol{X})|X_i\right]) &= -c_{0,\ldots,0}^2 + \sum_{\alpha_i} \sum_{\alpha_i'} c_{0,\ldots,\alpha_i,\ldots,0} c_{0,\ldots,\alpha_i',\ldots,0} \mathbb{E}\left[\psi_{\alpha_i}(X_i)\psi_{\alpha_i'}(X_i)\right] \\ &= \sum_{\alpha_i\neq 0} c_{0,\ldots,\alpha_i,\ldots,0}^2\end{aligned}$$

On ne prend alors que les coefficients ne dépendant que de i uniquement et on divise par la variance !

5. Expliquer alors comment calculer $S_2^{(i,j)}, \ldots, S_d^{(1,\ldots,d)}$ et $S_T^{(i)}$.

— Pour $S_2^{(i,j)}$, on ne prend que ceux ne dépendant que de i et j uniquement,
— ...
— Pour $S_d^{(1,\ldots,d)}$, on ne prend que ceux ne dépendant que de $1, \ldots, d$ uniquement,
— Pour $S_T^{(i)}$, on prend tout ceux dans lesquels i intervient.

27 Planification d'expériences

On s'intéresse au phénomène linéaire $y(x) = ax + b$, $x \in [0, 1]$. Pour l'estimation des paramètres a et b, on peut effectuer des mesures aux points x_1, x_2, x_3. Ces mesures sont bruitées, et notées $y^{\mathrm{mes}}(x_n) = y(x_n) + \varepsilon_n$, $1 \leq n \leq 3$, où ε_n sont des variables aléatoires centrées indépendantes de mêmes lois de probabilité (on note σ_{mes}^2 leur variance).

1. Montrer que $y(x)$ peut se mettre sous la forme $\boldsymbol{f}(x)^T\boldsymbol{\beta}$, où le vecteur de fonctions $\boldsymbol{f}$ et le vecteur $\boldsymbol{\beta}$ sont à expliciter.

$\boldsymbol{f} = (1, x)$ et le vecteur $\boldsymbol{\beta} = (b, a)$.

2. En supposant que $\boldsymbol{\beta}$ et les ϵ_n sont indépendants, rappeler l'expression du meilleur estimateur de $\boldsymbol{\beta}$ (au sens des moindres carrés), que l'on nomme $\widehat{\boldsymbol{\beta}}$.

$\widehat{\boldsymbol{\beta}} = ([F]^T[F])^{-1}[F]^T\boldsymbol{Y}$, $\boldsymbol{Y} = (y^{\text{mes}}(x_1), y^{\text{mes}}(x_2), y^{\text{mes}}(x_3))$, $[F] = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ 1 & x_3 \end{bmatrix}$.

3. Calculer le vecteur moyenne et la matrice de covariance de $\widehat{\boldsymbol{\beta}}$.

$\mathbb{E}\left[\widehat{\boldsymbol{\beta}}\right] = \boldsymbol{\beta}$, $\text{Cov}(\widehat{\boldsymbol{\beta}}) = \sigma^2_{\text{mes}}([F]^T[F])^{-1}$.

4. Afin de réduire l'incertitude sur $\widehat{\boldsymbol{\beta}}$, on cherche à minimiser le déterminant de sa matrice de covariance. Montrer que cela revient à maximiser la fonction coÃ»t $\mathcal{C}$:

$$\mathcal{C}(x_1, x_2, x_3) = (x_1 - x_2)^2 + (x_2 - x_3)^2 + (x_1 - x_3)^2.$$

$\det(\text{Cov}(\widehat{\boldsymbol{\beta}})) \propto \frac{1}{\det([F]^T[F])}$. Ainsi, minimiser $\det(\text{Cov}(\widehat{\boldsymbol{\beta}}))$ revient à maximiser $\det([F]^T[F])$. On calcule :
$\det([F]^T[F]) = 3(x_1^2 + x_2^2 + x_3^2) - (x_1 + x_2 + x_3)^2 = 2\left(x_1^2 + x_2^2 + x_3^2 - x_1x_2 - x_1x_3 - x_2x_3\right) = (x_1 - x_2)^2 + (x_2 - x_3)^2 + (x_1 - x_3)^2$.

5. Dans le cas où seulement deux mesures sont effectivement possibles, en déduire les positions x_1 et x_2 optimales vis à vis de ce critère sur le déterminant de la matrice de covariance.

Dans le cas à deux expériences, la fonction coÃ»t s'écrit : $\mathcal{C}(x_1, x_2) = (x_1 - x_2)^2$, qui est maximum sur $[0, 1]^2$ en $x_1 - x_2 = \pm 1$, c'est à dire un point en 0 et un point en 1.

6. Dans le cas où 3 mesures sont possibles, en déduire les positions optimales de x_1, x_2 et x_3 optimales vis à vis de ce critère sur le déterminant de la matrice de covariance.

Soit $x_1 \mapsto g(x_1) = \mathcal{C}(x_1, x_2, x_3)$. On calcule $\frac{dg}{dx_1}(x_1) = 4\left(x_1 - \frac{x_2+x_3}{2}\right)$. Pour toute valeur de x_2 et x_3, la fonction g est ainsi décroissante de 0 à $\frac{x_2+x_3}{2}$ puis croissante de $\frac{x_2+x_3}{2}$ à 1. Le maximum est ainsi obtenu sur les bords. Par symétrie du problème, on pose alors $x_1 = 0$. On note cette fois $x_2 \mapsto h(x_2) = \mathcal{C}(x_1 = 0, x_2, x_3) = x_2^2 + x_3^2 + (x_2 - x_3)^2$. De même que précédemment $\frac{dh}{dx_2}(x_2) = 4\left(x_2 - \frac{x_3}{2}\right)$. De même que précédemment, par symétrie en x_2 et x_3, on en déduit que la fonction coÃ»t est maximale quand x_2 et x_3 sont sur les bords du domaine, l'un en 1 et l'autre en 0.

7. Commenter l'efficacité d'un tel modèle s'il s'avère que y n'est plus linéaire mais parabolique.

Si y est parabolique, le plan optimal pour le cas linéaire n'est plus aussi performant. Illustration de la différence entre plan space filling "robuste" et plan optimisé donc non robuste.

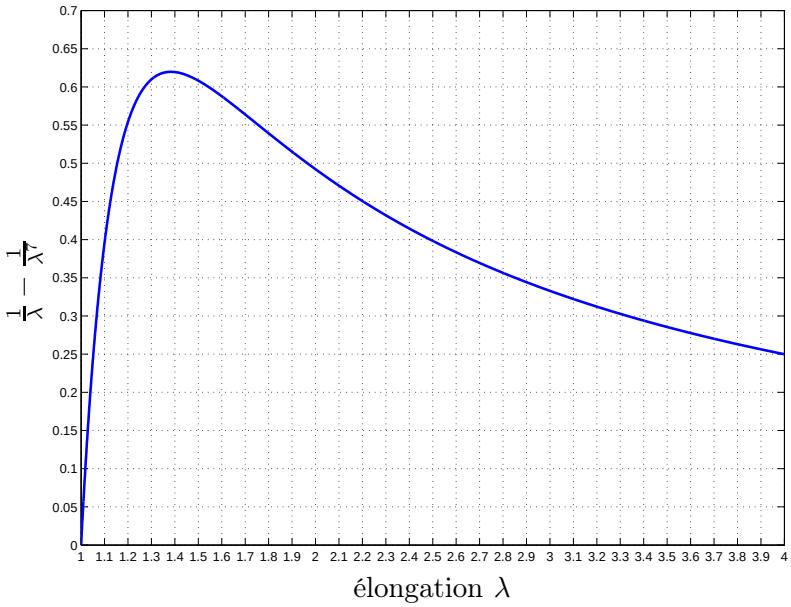


FIGURE 4 – Evolution de la pression adimensionnée en fonction de l'élongation

28 Optimisation économico-fiabiliste du gonflement d'un ballon de baudruche

L'objectif de cet exercice est d'optimiser la valeur maximale de pression d'une pompe de gonflement de ballon de baudruche. Soit $\mathcal{B}$ un ballon de baudruche sphérique, de rayons intérieur $r_i(t)$ et extérieur $r_e(t)$ et d'épaisseur $e(t) = r_e(t) - r_i(t)$. On note R_i, R_e les rayons interne et externe de ce ballon à l'état initial $t = 0$ et $E = R_e - R_i = e(0)$ l'épaisseur correspondante. On définit enfin $P(t)$ la pression dans le ballon au cours de son gonflement. On suppose la pression nulle en dehors du ballon.

1. On suppose l'épaisseur du ballon très faible devant son rayon, $e(t) = r_e(t) - r_i(t) \ll r_i(t)$, $E = R_e - R_i \ll R_i$, et on pose $\lambda = \frac{r_i(t)}{R_i}$ l'élongation interne du ballon. On admet alors que la pression interne en fonction de l'élongation s'écrit :

$$P(t) = \frac{4C_0E}{R_i}\left(\frac{1}{\lambda} - \frac{1}{\lambda^7}\right), \tag{11}$$

où C_0 est une constante matériau. La courbe 4 représente l'évolution de la pression adimensionnée en fonction de l'élongation λ. Commenter l'évolution de la pression interne au cours du gonflement.

La pression croît puis décroît en fonction de l'élongation : le gonflement est plus dur au début qu'à la fin, il faut juste dépasser un pression maximale pour que ce soit de plus en plus facile.

2. On considère maintenant deux ballons $\mathcal{B}^{(1)}$ et $\mathcal{B}^{(2)}$ initialement identiques (mêmes dimensions $R_i = R_i^{(1)} = R_i^{(2)}$, $E = E^{(1)} = E^{(2)}$ et mêmes matériaux $C_0 = C_0^{(1)} = C_0^{(2)}$). Le ballon $\mathcal{B}^{(2)}$

est davantage gonflé que le ballon $\mathcal{B}^{(1)}$, correspondant à des élongations $\lambda^{(2)} = 2.2 > \lambda^{(1)} = 1.2$. Lire sur la Figure 4 les pressions adimensionnées dans les deux ballons avant leur mise en communication. En déduire le sens du flux d'air d'un ballon dans l'autre.

En lisant le graphique, on remarque que le ballon 1, qui est le plus petit ($\lambda^{(1)} < \lambda^{(2)}$) est gonflé à une pression supérieure : $P^{(1)} = 0.55 > P^{(2)} = 0.45$. Ainsi, l'air est communiqué du petit ballon au plus grand pour atteindre une pression égale.

3. Calculer la pression minimale $P^*(E, C_0)$ à imposer afin de permettre le gonflement d'un ballon d'une élongation λ d'environ 4, en fonction de l'épaisseur E, du rayon interne initial R_i, et des caractéristiques matériaux C_0 du ballon.

On cherche le maximum de la pression :

$$\frac{\partial P}{\partial \lambda}(\lambda) = \frac{4C_0E}{R_i}\left(\frac{7}{\lambda^8} - \frac{1}{\lambda^2}\right),$$

$$\frac{\partial P}{\partial \lambda}(\lambda^*) = 0 \;\Rightarrow\; \lambda^* = 7^{1/6}, \quad P^* = \frac{4C_0E}{R_i}\left(\frac{1}{7^{1/6}} - \frac{1}{7^{7/6}}\right).$$

4. On considère maintenant le gonflement d'une série de ballons, dont les propriétés géométriques et les caractéristiques matériaux sont variables. Pour effectuer ces gonflements, on réfléchit à l'acquisition d'une pompe pouvant délivrer une pression maximale $P^{\max}$. Les constructeurs de ballon garantissent par ailleurs que l'épaisseur des ballons est comprise entre $E_{\min}$ et $E_{\min} + \Delta E$, tandis que les propriétés de l'élastomère les constituant ont une caractéristique matériau comprise entre $C_{\min}$ et $C_{\min} + \Delta C$. En supposant que l'épaisseur du ballon, E, ainsi que la constante matériau, C_0, sont uniformément distribués sur leurs domaines de définition, calculer la probabilité que $P^*(E, C_0)$ soit inférieure à $P^{\max}$, $\mathbb{P}(P^*(E, C_0) \leq P^{\max})$.

On calcule :

$$\mathbb{P}(P^* \leq P^{\max}) = \mathbb{P}(C_0E \leq S^{\max}),\; S^{\max} = \frac{P^{\max}R_i}{4\left(\frac{1}{7^{1/6}} - \frac{1}{7^{7/6}}\right)},$$
$$= \frac{1}{\Delta E \Delta C}\int_{C_{\min}}^{\min(C_{\min}+\Delta C, S^{\max}/E)} \left\{ \int_{E}^{\min(E+\Delta E, S^{\max}/\min)} dE \right\} dC.$$

En définissant $C^* = \max\left(C_{\min}, S^{\max}/(E + \Delta E)\right)$, on déduit :

$$\mathbb{P}(P^* \leq P^{\max})\Delta E \Delta C = \int_{C_{\min}}^{C^*} \Delta E dC + \int_{C^*}^{\min(C_{\min}+\Delta C, S^{\max}/E)} \left(\frac{S^{\max}}{C} - E\right) dC,$$
$$= \Delta E\left(C^* - C_{\min}\right) + S^{\max} \ln\left(\frac{\min\left(C_{\min} + \Delta C, S^{\max}/E\right)}{C^*}\right) - E\left(\min\left(C_{\min} + \Delta C, S^{\max}/E\right) - C^*\right).$$

5. On évalue le gain, G, relatif au gonflement de la série de ballons, en fonction du gain relatif à la vente des ballons, C_v, et des dépenses relatives à l'achat et l'entretien de la pompe, par la formule suivante :

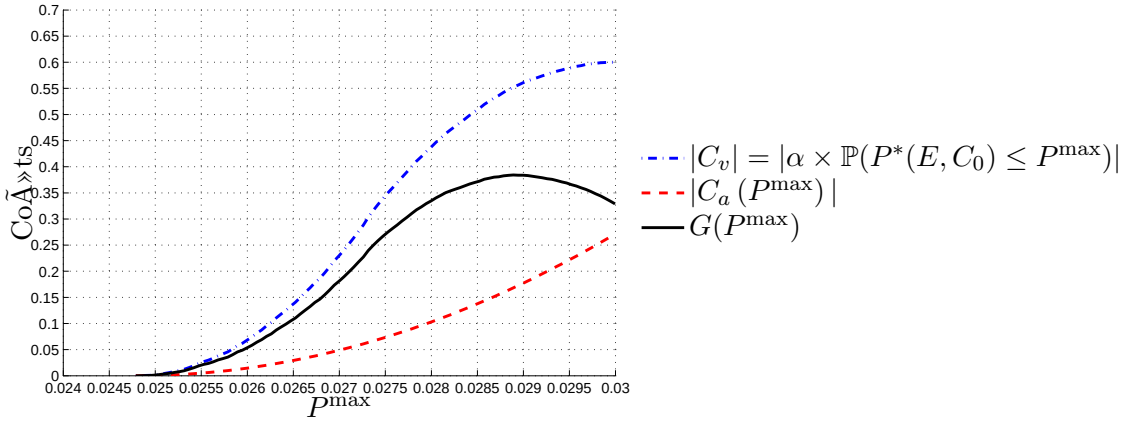


FIGURE 5

$$G(P^{\max}) = C_v\left\{\mathbb{P}(P^*(E, C_0) \leq P^{\max})\right\} + C_a\left\{P^{\max}\right\}.$$

Commenter qualitativement le caractère croissant ou décroissant, positif ou négatif, de ces deux fonctions C_v et C_a, en fonction de $P^{\max}$.

C_v est positif tandis que C_a est négatif. C_v croît avec $P^{\max}$ car davantage de ballons peuvent être gonflés. Néanmoins, plus $P^{\max}$ est grand et plus les coÃ»ts sont élevés, donc fonction croissante en absolu pour C_a, mais comme elle est négative, fonction décroissante.

6. L'évolution de G est représentée sur la Figure 5, dans le cas où C_v est linéaire en $|\mathbb{P}(P^*(E, C_0) \leq P^{\max})|$, c'est à dire dans le cas où il existe α tel que $C_v = |\alpha \times \mathbb{P}(P^*(E, C_0) \leq P^{\max})|$. évaluer numériquement (par lecture de graphique) la pression $P^{\max}$ permettant de maximiser le gain G, ainsi que la valeur de α (on admettra que $\mathbb{P}(P^*(E, C_0) \leq P^{\max}) = 1$ quand $P^{\max}$ attend sa valeur maximale).

On trouve par lecture graphique que le maximum de gain est atteint pour $P^{\max} = 0.0288$. Le rapport α est donné par 0.6 en regardant la valeur de C_v au niveau de la valeur maximale de $P^{\max}$.

7. En déduire le pourcentage moyen de ballons gonflés correspondant à la configuration de gain maximal.

En $P^{\max} = 0.0288$, on a $C_V = 0.55$. Ça correspond à une probabilité :

$$\mathbb{P}(P^* \leq P^{\max}) = \frac{0.55}{0.6} \approx 92\%.$$

Ainsi, 8% des ballons ne seront pas gonflés en moyenne.

29 Optimisation technico-économique de la fiabilité

On suppose disposer d'une méthode (supposée précise) d'évaluation de la probabilité de ruine, notée P_r, d'un système physique (éolienne, pont, plateforme pétrolière...), potentiellement

sollicité par une ou plusieurs sources aléatoires.

Le coût total C_T de conception du système se décompose souvent en deux termes. Le coût initial, noté $C_i(P_r)$, comprend l'ensemble des coûts liés directement au système et à son fonctionnement. De manière empirique, une loi logarithmique est souvent observée :

$$C_i(P_r) = C_0 - C_1 \ln(P_r).$$

Par ailleurs, le coût attendu de la ruine, noté $C_r(P_r)$, est égal au coût de ruine C_R, qui dépend des coûts induits par la ruine, pondéré par la probabilité de ruine de la structure :

$$C_T(P_r) = C_i(P_r) + C_r(P_r), \quad C_r(P_r) = C_R \times P_r.$$

1. Attribuer les coûts suivants au coût initial C_i ou au coût de ruine C_R en le justifiant :
— coûts liés aux pertes de vies humaines,
— coûts de conception,
— coûts liés aux pertes de production,
— coûts de construction,
— coûts liés à l'inspection,
— coûts liés à la maintenance,
— coûts liés aux dommages sur l'environnement,
— coûts liés à une dégradation de réputation ou d'image,
— coûts d'exploitation.

Le coût initial :
— conception
— construction
— équipement
— exploitation
— inspection
— maintenance...
Coût de ruine :
— perte de vies humaines,
— dommage sur l'environnement,
— perte de production,
— réputation, image,...

2. Indiquer le caractère croissant ou décroissant des coûts C_i et C_r en fonction de P_r. Cette évolution est-elle intuitive ou contre-intuitive selon vous ? Pourquoi ?

Plus la probabilité de ruine est grande, plus il y a de chance que le système casse, et donc en moyenne, plus le coût de ruine est grand (fonction linéaire), c'est tout à fait naturel. Par contre, plus le système coute initialement chère, et plus il y a de chances que la probabilité de ruine soit petite, donc C_i doit être décroissant en P_r.

3. A coefficients C_0, C_1 et C_R fixés, calculer la probabilité de ruine permettant de minimiser le coût total du système.

Un optimum peut ainsi être obtenu pour :

$$\frac{\partial C_T}{\partial P_r}(P_r^*) = 0 \;\Rightarrow\; P_r^* = \frac{C_1}{C_R}.$$

4. En notant P_{seuil} une probabilité de ruine maximale autorisée et C_{seuil} un coût total maximal, deux visions pour l'optimisation de ce système sous contrainte fiabiliste sous souvent introduites :
   — minimiser C_T sous contrainte $P_r(C_T) \leq P_{\text{seuil}}$,
   — minimiser P_r sous contrainte $C_T(P_r) \leq C_{\text{seuil}}$.

Commenter les avantages et les désavantages de ces deux approches, ainsi que les différences auxquelles elles pourraient conduire.

30 Révisions

Dans cet exercice, toutes les questions sont indépendantes.

1. On s'intéresse à la fonction $y(x_1, x_2, x_3, x_4) = x_2 + 0.1x_1 + 0.01x_3$ pour x_1, x_2, x_3, x_4 des variables aléatoires indépendantes et de même loi. Sans calcul mais en le justifiant, hiérarchiser les variables x_1 à x_4 de la plus influente à la moins influente sur y.

   Direct : $x_2 > x_1 > x_3 > x_4$

2. On s'intéresse à la fonction $y(x_1, x_2, x_3) = x_1 + x_2 + x_3$ pour x_1, x_2, x_3 trois variables aléatoires indépendantes, uniformément distribuées sur $[0, 1]$, $[0, 0.1]$, $[0, 0.01]$ respectivement. Sans calcul mais en le justifiant, hiérarchiser les variables x_1 à x_3 de la plus influente à la moins influente sur y.

   Direct : $x_1 > x_2 > x_3$.

3. Si $X_1, \ldots, X_N$ sont N réalisations indépendantes d'une même variable aléatoire gaussienne X de loi $\mathcal{N}(\mu, \sigma^2)$ (μ et σ étant inconnus), donner un intervalle de confiance I_1 pour μ ne dépendant que de $X_1, \ldots, X_N$, tel que :

$$\mathbb{P}(\mu \in I_1) = 0.95.$$

   $I_1 = \widehat{\mu} \pm q_{t(N-1),0.975}\widehat{\sigma}/\sqrt{N}$, $\widehat{\mu} = \frac{1}{N}\sum_{n=1}^N X_n$, $\widehat{\sigma}^2 = \frac{1}{N-1}\sum_{n=1}^N (X_n - \widehat{\mu})^2$.

4. Si $X_1, \ldots, X_N$ sont N réalisations indépendantes d'une même variable aléatoire gaussienne X de loi $\mathcal{N}(\mu, \sigma^2)$ (μ et σ étant inconnus), donner un intervalle de prédiction I_2 pour X ne dépendant que de $X_1, \ldots, X_N$, tel que :

$$\mathbb{P}(X \in I_2) = 0.95.$$

   $I_2 = \widehat{\mu} \pm q_{t(N-1),0.975}\widehat{\sigma}$, $\widehat{\mu} = \frac{1}{N}\sum_{n=1}^N X_n$, $\widehat{\sigma}^2 = \frac{1}{N-1}\sum_{n=1}^N (X_n - \widehat{\mu})^2$.

5. Soit x une variable aléatoire exponentielle de paramètre 1. Exprimer le plus petit intervalle I tel que $\mathbb{P}(x \in I) = \alpha$.

   Par antisymétrie de la loi de x, $I = [0, a]$, avec $\int_0^a \exp(-x)dx = \alpha$, soit $a = -\log(1 - \alpha)$.

6. Soient $X_1, \ldots, X_N$ N variables aléatoires indépendantes de même fonction de répartition F_X. Calculer $\mathbb{P}(\max_{1 \leq n \leq N} X_n \leq x)$ en fonction de F_X.

   $\mathbb{P}(\max_{1 \leq n \leq N} X_n \leq x) = \mathbb{P}(X_1 \leq x)^N = F_X(x)^N$.

7. Soit X une variable aléatoire uniformément répartie sur $[0,1]$. En déduire le domaine de définition et l'expression de la fonction densité de probabilité de $Y = \log\left(1+\sqrt{X}\right)$.

$$\begin{aligned}\mathbb{P}(Y \leq y) &= \mathbb{P}(\log\left(1+\sqrt{X}\right) \leq y)\\ &= \mathbb{P}(X \leq (\exp(y)-1)^2)\\ &= (\exp(y)-1)^2.\end{aligned}$$

On déduit : $f_Y(y) = 2\exp(y)(\exp(y)-1)$, pour tout $y \in [0, \log(2)]$.

8. Pour $\theta \geq 1$, peut on dire que la fonction $C(u,v) = \exp\left[-\left((-\log(u))^\theta + (-\log(v))^\theta\right)^{1/\theta}\right]$ est une fonction copule ? Justifier.

log est croissant, donc $-\log$ est décroissant, donc $u \mapsto (-\log(u))^\theta$ est décroissant, et donc $u \mapsto \exp\left[-\left((-\log(u))^\theta + (-\log(v))^\theta\right)^{1/\theta}\right]$ est croissant. Pareil pour v.
Quand u tend vers 0, C(u,v) tend vers 0. Quand u=1, C(u=1,v)=v. Pareil pour v.
Quand u, v tendent vers 1, C(u,v) tend vers 1.
Donc fonction copule.

9. Soit Y une quantité aléatoire, s un seuil et $p = \mathbb{P}(Y > s)$ une probabilité de dépassement de seuil. En notant $y_1, \ldots, y_N$ N réalisations indépendantes de Y, rappeler l'estimateur Monte Carlo de p, ainsi que sa loi asymptotique lorsque $N \to +\infty$.

$\widehat{p} = \frac{1}{N}\sum_{n=1}^{N} 1_{y_n > S} \to \mathcal{N}(p, \frac{p(1-p)}{N})$.

10. Afin de minimiser les chances de garantir un système ne devant pas être garanti, on se donne le critère suivant :

"On accepte de garantir le système si $\widehat{p} + a < \alpha$",

avec $\widehat{p}$ un estimateur de la vraie probabilité p de problème sur le système, et α une constante entre 0 et 1. Quel doit être le signe de a pour que l'introduction de la marge permette de réduire le nombre de garanties à tord ? Commenter l'appellation "marge de sécurité" pour a, ainsi que l'appellation "risque de garantie à tord" pour $r = \max_{p \geq \alpha} \mathbb{P}(\widehat{p}+a < \alpha)$.

Il faut que $a > 0$. Par construction, a est bien une marge de sécurité entre α et $\widehat{p}$ au sens où plus a est grand, moins le risque de garantie à tord est élevé. La notion de risque de garantie à tord est directe par la définition, car c'est bien la probabilité d'une garantie du système, sachant qu'il ne devrait pas être garanti car $p \geq \alpha$.

11. On s'intéresse au développement d'un service aux propriétaires intéressés pour louer leur voiture à des particuliers, qui leur permette de maximiser leur gain moyen. On note p le prix de location et $\boldsymbol{x}$ le vecteur des caractéristiques du véhicule (gamme, marque, ancienneté, note moyenne attribuée par les précédents loueurs...). On propose de modéliser la probabilité qu'un client intéressé loue une voiture de caractéristique $\boldsymbol{x}$ au prix p sous la forme :

$$f \circ h(\boldsymbol{\beta}, \boldsymbol{x}, p), \quad h(\boldsymbol{\beta}, \boldsymbol{x}, p) = \beta_0 + \boldsymbol{\beta}_1 \cdot \boldsymbol{x} + \beta_2 p, \quad f(u) = \frac{1}{1+\exp(-u)}.$$

Expliquer, au regard des fonctions f et h, dans quelle mesure ce choix de modélisation peut être pertinent. Selon vous, quel est le signe attendu de β_2 ?

f est une fonction croissante de 0 à 1 : cela définit bien une fonction de répartition, avec un seuil paramétré par h permettant d'adapter la probabilité aux données. A priori, il serait cohérent d'avoir $\beta_2 < 0$, car la probabilité d'avoir une location est vraisemblablement décroissante en fonction du prix.

12. On s'intéresse à la prédiction d'une quantité y. On dispose d'une mesure $y^{\text{mes}} = y + \varepsilon^{\text{mes}}$, et d'une simulation $y = y^{\text{sim}}(\beta) + \varepsilon^{\text{sim}}$, où β est un vecteur à identifier. $\varepsilon^{\text{mes}} \sim \mathcal{N}(0, \sigma^2)$, $\varepsilon^{\text{sim}} \sim \mathcal{N}(0, \tau^2)$ sont des erreurs de mesure et de simulation gaussiennes, supposés indépendantes statistiquement. En déduire la fonction de vraisemblance associée au paramètre β.

$L(\beta) = \frac{1}{\sqrt{2\pi}\sqrt{\sigma^2+\tau^2}} \exp\left(-\frac{1}{2}\frac{(y^{\text{mes}}-y^{\text{sim}}(\beta))^2}{\sigma^2+\tau^2}\right)$

13. On cherche à minimiser la fonction $x \mapsto \sin(6\pi\sqrt{x}))$ sur $[0, 1]$. Identifier les potentielles valeurs optimales de x, et les classer de la plus à la moins robuste vis-à-vis d'incertitudes sur x.

Les minima sont pour $6\pi\sqrt{x} = -\pi/2 + 2k\pi$, $k > 1$, c'est à dire :
$x = \left(\frac{k}{3} - \frac{1}{12}\right)^2$, $k = 1...3$. Plus k est grand, plus le minimum est "robuste".

14. Trouver $x^\star \geq 0$ solution du problème de maximisation sous incertitudes et sous contrainte fiabiliste suivant :

$$\max_{x\geq 0,\ \mathbb{P}(x+\tau\geq 2)\leq\alpha} \mathbb{E}\left[(x+\tau)^2\right],\ \tau \sim \mathcal{N}(0,1),\ 0 < \alpha < 1.$$

On calcule d'abord : $\mathbb{E}\left[(x+\tau)^2\right] = x^2 + 1$, donc on cherche la valeur de x la plus grande respectant les contraintes.
L'optimum sera ainsi atteint pour des contraintes activées, c'est à dire quand $\mathbb{P}(x + \tau \geq 2) = \alpha$, c'est à dire $x = 2 - q_{1-\alpha}$, avec $q_{1-\alpha}$ le quantile de niveau $1-\alpha$ de la loi gaussienne centrée réduite.

15. Soit y une fonction telle que pour tout x, $y(x) = h(x)\beta + \varepsilon$, avec $\beta \sim \mathcal{N}(\widetilde{\beta}, \sigma_\beta^2)$ et $\varepsilon \sim \mathcal{N}(0, \tau^2)$ deux quantités aléatoires indépendantes. Donner la loi du couple $(y(x), \beta)$, puis en déduire la loi de $\beta|y(x)$.

$$(y(x), \beta) \sim \mathcal{N}\left(\begin{pmatrix} h(x)\widetilde{\beta} \\ \widetilde{\beta} \end{pmatrix}, \begin{bmatrix} \sigma_\beta^2 h(x)^2 + \tau^2 & \sigma_\beta^2 h(x) \\ \sigma_\beta^2 h(x) & \sigma_\beta^2 \end{bmatrix}\right),$$
$$\beta|y(x) \sim \mathcal{N}(\widetilde{\beta} + h(x)\sigma_\beta^2/(\sigma_\beta^2 h(x)^2 + \tau^2)(y(x) - h(x)\widetilde{\beta}), \sigma_\beta^2 - h(x)^2\sigma_\beta^4/(\sigma_\beta^2 h(x)^2 + \tau^2)).$$