MACS - Quantification des incertitudes pour la simulation

TD 6 - Année 2022-2023

1 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.

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.

3. Calculer les approximations par cumul quadratique de la moyenne et de la variance de Y.

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

2 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 $\Rightarrow$ 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.

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

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

3. Cas 2 : la variabilité "naturelle" est majoritaire $\Rightarrow$ 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.

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

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

3 Statistiques sur les processus gaussiens stationnaires

Soit $\{X(t),\ t \in \mathbb{R}\}$ un processus aléatoire gaussien stationnaire centré et de fonction de covariance $(t,t') \mapsto C(t,t') = \mathbb{E}\left[X(t)X(t')\right] = R(|t-t'|)$. 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.

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

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

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.

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

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.

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.

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.