

PROBLÈME DE CALIBRATION (G10, J)

(18 / 10 / 2019, © Monfort, Dicostat2005, 2005-2019)

(i) Un **problème de calibration** consiste en la donnée :

(a) d'une **expérience de calibration** (eg, selon le contexte, une **expérience** de contrôle ou d'apprentissage) dans laquelle on observe N réponses Y_n à N_1 niveaux prédéterminés X_n ($n = 1, \dots, N_1$) d'un stimulus ξ , la relation entre le **stimulus** et la **réponse** η (cf **variable endogène**) étant supposée être une **régression non linéaire** :

$$(1) \quad \eta = f(\xi, b) + \varepsilon,$$

où f est (analytiquement) donnée (et, en général, monotone pr à ξ) et où $b \in \mathbf{R}^Q$ est un **paramètre** à déterminer. On pose :

$$(2) \quad Y_n = f(X_n, b) + u_n, \quad \forall n = 1, \dots, N_1,$$

et l'on suppose souvent que les u_n forment une **suite iid** selon $\mathcal{N}_1(0, \sigma^2)$ (**loi normale** centrée) ;

(b) d'une **expérience** « prévisionnelle » dans laquelle on observe N_2 réponses Y_n ($n = N_1 + 1, \dots, N_1 + N_2 = N$) relatives à un même niveau X_0 , inconnu cette fois, du stimulus ξ . On note :

$$(3) \quad Y_n = f(X_0, b) + u_n, \quad \forall n = N_1 + 1, \dots, N,$$

où l'on suppose toujours que les u_n sont iid selon $\mathcal{N}_1(0, \sigma^2)$.

(ii) L'objet d'un problème de calibration porte principalement sur X_0 (**estimation, tests**) et, le cas échéant, sur (b, σ^2) .

Une approche bayésienne (cf **théorie bayésienne**) consiste, dans ce contexte, et moyennant une hypothèse d'**indépendance** entre les deux types d'expériences, à doter le triplet (X_0, b, σ^2) d'une **loi a priori** $\mathcal{L}(X_0, b, \sigma^2)$. Cette loi est donc le produit des lois $\mathcal{L}(X_0)$ de X_0 et $\mathcal{L}(b, \sigma^2)$ de (b, σ^2) .

La méthode conduit à la **loi a posteriori** $\mathcal{L}((X_0, b, \sigma^2) / (X(1), Y))$, conditionnelle aux observations $X(1) = (X_1, \dots, X_{N(1)})$ et $Y = (Y_1, \dots, Y_N)$. Par suite, la loi a posteriori (marginale) $\mathcal{L}(X_0 / (X(1), Y))$ de X_0 permet d'étudier ce niveau du stimulus ξ (**mode**, etc).

Dans les sciences expérimentales, où les expériences peuvent être planifiées (cf **plan d'expérience**), on peut renouveler plusieurs fois l'**expérience de calibration** et l'expérience prévisionnelle. L'**inférence statistique** peut donc en tenir compte. La

courbe définie par le graphe de f est dite **courbe de calibration**, ou **courbe stimulus-réponse**.

(iii) Dans une **expérience biologique** (en anglais : « *bioassay* »), par exemple, ξ représente le taux de concentration d'un **traitement** (le stimulus) et η le taux de réponse (ou taux de réaction) des **unités statistiques** (**unités expérimentales**) repérées par $n = 1, \dots, N$.

La méthode est applicable dans de nombreux contextes : physique (technométrie), biologie (pharmacologie), écologie (agronomie), etc. Ainsi, elle peut notamment contribuer à « remettre en état » un instrument de **mesure** (ou un **système d'observation**) qui peut s'altérer à l'usage (déformations, **biais**, etc) (cf aussi **dispositif expérimental**).