Principes et méthodes statistiques

William SCHMITT

2018-2019

Table des matières

1 Estimation

Grâce aux statistiques descriptives, on propose un modèle.

On le valide en :
— utilisant un graphe de probabilités
— trouvant que c'est la seule possibilité
Le modèle dépend d'un (uniforme) ou plusieurs paramètres (exponentielle).

Le (ou les) paramètre(s) n'est pas connu dans la vraie vie.

Estimation valeur réelle caclculée à partir des données x_i dont on aimerait qu'elle soit "proche" du paramètre inconnnu θ qu'on cherche à estimer.

1.1 Critères pour construire une estimation

Idée : faire coller la théorie (modèle proposé, i.e. loi, paramètres) aux résultats obtenus.

1.1.1 Critères usuels

Moment d'ordre 1 l'espérance théorique représente la moyenne, l'échantillon a une moyenne. On approxime l'espérance théorique par la moyenne des résultats obtenus.

$$E(X) \sim \frac{1}{n}\sum_{i=1}^{n} x_i$$
$$f(\theta) \sim \frac{1}{n}\sum_{i=1}^{n} x_i$$
$$f(\theta_{\text{est}}) = \frac{1}{n}\sum_{i=1}^{n} x_i$$
$$\theta_{\text{est}} = f^{-1}(\frac{1}{n}\sum_{i=1}^{n} x_i)$$

Cette méthode ne permet pas d'évaluer la variance.

Moment d'ordre 2 Même démarche avec la variance

Maximum de vraisemblance

1.2 Qualité d'estimation

L'estimation devient la valeur prise par l'estimateur (i.e. la **réalisation** de l'estimateur).

Si on recommence l'expérience aléatoire qui consiste à collecter des données dans les mêmes conditions que celles dont on dispose (même nombre, même protocole de recueil), les variables aléatoires restent les mêmes (l'estimateur aussi), par contre les valeurs vont probablement être différentes, et donc l'estimation obtenue va très vraisemblablement être différente.

1.2.1 Biais

On souhaite que l'estimation tombe sur la **vraie valeur** θ du paramètre inconnu. On calcule donc l'**espérance de l'estimateur** :

— Si l'espérance permet de retrouver le paramètre inconnu : c'est un **estimateur sans biais**

— Sinon, il y a un biais (qu'on peut éventuellement corriger avec la linéarité de l'espérance).

Dans le cas de Jeannette, $\mathbb{E}(T_{\text{est}}) = \frac{n}{n+1}t$, on peut débiaiser en remplaçant l'estimateur de Jeannette par $\frac{n+1}{n} \times$ Max pour trouver un estimateur débiaisé. Cela fonctionne pour des facteurs multiplicatifs mais également pour des additions.

1.2.2 Convergence

Une fois l'estimateur débiaisé, on souhaite minimiser la dispersion de l'estimateur autour de la **vraie valeur** θ. On calcule donc la variance de l'estimateur : elle est meilleure si la variance est faible.

1.2.3 Appréciation de la qualité

- Théorique : quand c'est possible
- Simulations : sinon