Fiche de Statistiques

Contents

Part I
Rappels et formules utiles

Ceci est plus une liste des choses à réviser qu'un rappel en soit, tout le reste se trouve dans la pougne de proba. Il faut aussi savoir faire des calculs d'espérance. Toutes les formules ici sont à connaître, elles sont toutes utiles et utilisées.

1 Rappels généraux

Définition d'un espace probabilisé $(\Omega, \mathcal{A}, P)$ est un espace probabilisé où Ω un ensemble, $\mathcal{A} \in \mathcal{P}(\Omega)$ est une tribu et P probabilité

Tribu borélienne $\mathcal{B}(\mathbb{R}) = \{]-\infty, a[, a \in \mathbb{R}\}$

Probabilités conditionnelles

Formule des probabilités totales Soit C_n famille de $\mathcal{A}$ où aucun des termes n'a une probabilité nulle,

$$\forall A \in \mathcal{A}, P(A) = \sum_{n=1}^{+\infty} P(A|C_n)P(C_n)$$

Formule de Bayes

$$\text{Si } P(A) \neq 0, P(C_i|A) = \frac{P(A|C_i)P(C_i)}{\sum_{n=1}^{+\infty} P(A|C_n)P(C_n)}$$

2 Espérance mathématique

Calcul v.a discrète $E[X] = \sum_{\omega \in \Omega} X(\omega)P(\omega) = \sum_{k \in (1, \dots n)} a_n P(X = a_n)$

v.a continue $E[X] = \int_{\omega \in \Omega} X(\omega)dP(\omega) = \int_{\mathbb{R}} x dP_X(x)$

Propriétés

- $E[X+Y] = E[X] + E[Y]$
- Linéarité : $E[aX+b] = aE[X] + b$, $a, b \in \mathbb{R}$ ($\Rightarrow E[XE[X]] = E[X]^2$ car $E[X]$ est un scalaire)
- Formule à retenir : $E[\mathbf{1}_C] = P(C)$
- En pratique si les X_i sont n v.a.i.i.d., $E[\sum X_i] = nE[X_1]$

3 Variance

Calcul $Var(X) \begin{cases} = E[(X - E[X])^2] & \textit{(Définition)} \\ = \int_{\mathbb{R}} (x - E[X])^2 f_X(x)dx & \textit{(Théorème de transfert)} \\ = E[X^2] - E[X]^2 & \textit{(Pour les calculs. ATTENTION à vérifier que } X^2 \textit{ est intégrable)} \end{cases}$

Covariance Définition : $Cov(X,Y) = E[(X - E[X])(Y - E[Y])] = E[XY] - E[X]E[Y]$

Propriétés

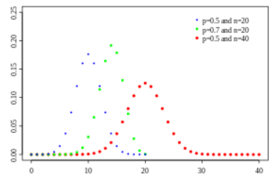
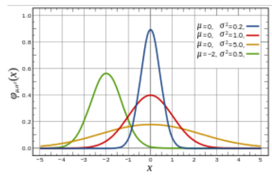
- $Var(aX + b) = a^2 Var(X)$
- $Var(X + Y) = Var(X) + Var(Y) + 2Cov(X, Y)$
- Si X et Y sont indépendants alors $E[XY] = E[X]E[Y]$ et $Cov(X, Y) = 0$ donc $Var(X + Y) = Var(X) + Var(Y)$

Ecart type $\sigma_X = \sqrt{Var(X)}$

Relations entre les convergences

$$\begin{array}{ccccc} & & X_n \xrightarrow{L^\alpha} X & & \\ & & \Downarrow & & \\ X_n \xrightarrow{p.s.} X & \Rightarrow & X_n \xrightarrow{P} X & \Rightarrow & X_n \xrightarrow{\mathcal{L}} X \end{array}$$

Les principales lois à connaître (minimum)

Nom	Formule	Espérance	Variance	Graph
Uniforme $[a, b]$	$f_X(x) = \frac{1}{b-a} 1_{[a,b]}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$	
Binomiale $\mathcal{B}(n, p)$	$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$	np	$np(1-p)$	
Normale $\mathcal{N}(\mu, \sigma)$	$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[\frac{-(x-\mu)^2}{2\sigma^2}\right]$	μ	σ^2	

Part II
Estimateurs

1 Définition

Introduction Lorsque l'on a un problème complexe avec beaucoup de données, cela devient vite intraitable. On décide donc de prendre un échantillon, ce qui implique ne plus avoir des valeures excates mais des échantillons. Une estimation est une valeur calculée sur un échantillon, que l'on espère proche de la valeur d'un paramètre et qui permet de caractériser la population totale. En général, la problématique ici est de trouver le "meilleur" estimateur.

Contexte Soient $(X_1, ...X_n)$ v.a. indépendantes de loi commune appartenant à la famille $(P_\theta)_{\theta\in\Theta}$, la famille de modèle probabiliste $(A \subset \mathbb{R}, \mathcal{A}_n \otimes P_\theta)_{\theta\in\Theta}$ est dite modèle statistique

2 Qualité d'un estimateur

Minimisation de l'erreur Si $\hat{T}_n$ estimateur de t, on minimise les critères suivants :

- biais($\hat{T}_n$) $= |E[t - \hat{T}_n]| = |t - E[\hat{T}_n]|$
- Var($\hat{T}_n$) $= E[(\hat{T}_n - E[\hat{T}_n])^2]$ (stabilité autour de la valeur)
- Erreur quadratique : EQM($\hat{T}_n$) $= E[(t - \hat{T}_n)^2] = Var(\hat{T}_n) +$ biais$(\hat{T}_n)^2$

Remarque : l'erreur quadratique est UNE façon de mesurer l'erreur, mais il y en a d'autre..

Convergence en probabilité si $\forall \epsilon > 0, \lim \xrightarrow{n \to +\infty} P(|t - \hat{\mathrm{T}}_n| > 0) = 0$

Estimateur sans biais si $E[\hat{\mathrm{T}}_n] = t$

Construction des estimateurs

- Sans biais (on s'arrange pour qu'après calcul, le biais soit nul)
- Maximum de vraissemblance ($\widehat{\theta}(x_1, ..., x_n) = arg \max_\theta \log f_\theta^n(x_1, ..., x_n)$)
- Méthode des moments (exemple $E[X] = k\theta \Rightarrow \theta = kE[X] \Rightarrow \widehat{\theta}_n = \frac{k}{n} \sum X_i$
- Moindre carrés (p37)

En pratique : Il faut savoir faire des calculs d'espérance...

3 Stratégie bayésienne

Notation poly p59

- θ paramètre
- $\hat{s}$ estimateur
- ξ v.a. de densité f
- $\prod = (\prod_1, ..., \prod_k)$ loi de ξ, probabilité "a priori"
- $\prod_i^x = \frac{\prod_i f_i(x)}{\sum_j \prod_j f_j(x)}$ probabilité "a posteriori"

Fonction de perte $L = \{ \begin{array}{ccc} 0 & si & \theta_i = \theta_j \\ \lambda_{i,j} & si & \theta_i \neq \theta_j \end{array}$ où $\lambda_{i,j}$ modélise la gravité de l'erreur "on a choisi θ_i au lieu de θ_j

Risque $R(\hat{s}, \theta) = E[L(\hat{s}(X), \theta)] = \int_{\mathbb{R}} L(\hat{s}(x), \theta) dP_\theta(x)$

Stratégie bayésienne ou estimateur bayésien $\hat{s}_B$permet de prendre des décisions et vérifie $E[L(\hat{s}_B(X), \xi)] = \min_{\hat{s}} E[L(\hat{s}(X, \xi)]$
mais surtout (2.10) $[\hat{s}_B = \theta_m] \leftrightarrow [\forall 1 \leq i \leq k, \sum_j \lambda_{mj} \prod_j^x \leq \sum_j \lambda_{ij} \prod_j^x$

Probabilité d'erreur se calcule souvent grâce aux probabilités conditionnelles

Coût d'une erreur (perte moyenne) formule 2.8 $E[L(\hat{s}(X), \xi)]$

PS : Le poly est sympa, y'a plein plein d'annales à la fin...