Probabilités et Statistiques

Dans cette partie, on présente quelques résultats en probabilités et statistiques dans le cadre de ce cours. Pour plus d'information, vous pouvez vous référer au cours STT-1000, à Wasserman (2010) (en anglais) et à Delmas (2013) (en français).

Modéliser le hasard

Beaucoup de phénomènes réels ne sont pas prévisibles et généralement, leurs résultats contiennent une certaine variabilité. Cette variabilité est prise en compte grâce à une mesure de l'incertitude que l'on appelle **mesure de probabilités**.

Définition

L'**espace d'évènements** S est l'ensemble de tous les résultats possibles d'un phénomène. Un **évènement** est un sous-ensemble de l'espace d'évènements S.

Exemples

1. Si l'expérience consiste à lancer un pièce, $S = \{0, 1\}$. Le résultat de cette expérience ne peut pas être connu à l'avance. Par exemple, $E = \{1\}$ est un évènement de S.
2. Si on s'intéresse à la durée de vie d'un téléphone, $S = \mathbb{R}_+$. On peut aussi choisir $S = [0, M]$, car cette durée de vie n'est probablement pas infini ! L'évènement $E = [10, \infty)$ représente l'évènement "une durée de vie de plus de 10 unités de temps".
3. Pour le nombre de jours sans neige à Québec dans l'année, on peut choisir $S = \mathbb{N}$. L'évènement $E = (0, 5]$ représente l'évènement "moins de 5 jours sans neige à Québec dans l'année".

Définition

Une **mesure de probabilités** $\mathbb{P}$ sur S est une application (fonction) définie sur l'espace d'évènements et satisfaisant les propriétés suivantes :

1. Pour chaque évènement E, $\mathbb{P}(E) \in [0, 1]$.
2. $\mathbb{P}(S) = 1$.
3. Soient $E_1, E_2, \ldots$, une séquence d'évènements (finie ou infinie) mutuellement exclusive, i.e. $\forall i \neq j, E_i \cap E_j = \emptyset$. On a

$$\mathbb{P}(\bigcup_{n=1}^{\infty} E_n) = \sum_{n=1}^{\infty} \mathbb{P}(E_n).$$

On appelle $\mathbb{P}(E)$, la probabilité de l'évènement E.

La définition de mesures de probabilités peut être subjective et lié à l'expérience du statisticien. En reprenant l'exemple 3 sur le nombre de jours sans neige à Québec dans l'année. Une personne venant d'arriver au Canada peut vouloir donner la même probabilité à chacun des jours, alors qu'un Québécois aura plus d'information et pourras faire varier les probabilités en fonction de cette connaissance.

Définition

Deux évènements E et F sont dits **indépendants** si $\mathbb{P}(E \cap F) = \mathbb{P}(E) \times \mathbb{P}(F)$.

Définition

Soient E et F, deux évènements, la **probabilité conditionelle** que E se réalise sachant que F s'est réalisé est définie par :

$$\mathbb{P}(E \mid F) = \frac{\mathbb{P}(E \cap F)}{\mathbb{P}(F)}.$$

De façon intuitive, deux évènements sont indépendants si la connaissance de l'un ne donne aucune information sur la réalisation de l'autre. On a aussi $\mathbb{P}(E \mid F) = \mathbb{P}(E)$.

Variables aléatoires

En probabilité, la convention est d'exprimer le résultat d'expériences comme la valeur d'une fonction appelé **variable aléatoire**. Cette caractérisation est toujours possible.

Soit $X = (X_1, \ldots, X_p)^\top \in \mathbb{R}^p$, un vecteur aléatoire de taille p.

Définition

Espérance :

$$\mathbb{E}(X) = \begin{pmatrix} \mathbb{E}(X_1) \\ \vdots \\ \mathbb{E}(X_p) \end{pmatrix} := \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_p \end{pmatrix} := \mu.$$

Matrice de variance/covariance :

$$\mathrm{Var}(X) = \begin{pmatrix} \mathrm{Var}(X_1) & \cdots & \mathrm{Cov}(X_1, X_p) \\ \vdots & \ddots & \vdots \\ \mathrm{Cov}(X_1, X_p) & \cdots & \mathrm{Var}(X_p) \end{pmatrix} := \begin{pmatrix} \sigma_1^2 & \cdots & \sigma_{1,p} \\ \vdots & \ddots & \vdots \\ \sigma_{1,p} & \cdots & \sigma_p \end{pmatrix} := \Sigma.$$

Matrice de corrélation :

$$\mathrm{Cor}(X) = \begin{pmatrix} 1 & \cdots & \mathrm{Corr}(X_1, X_p) \\ \vdots & \ddots & \vdots \\ \mathrm{Corr}(X_1, X_p) & \cdots & 1 \end{pmatrix} := \begin{pmatrix} 1 & \cdots & \rho_{1,p} \\ \vdots & \ddots & \vdots \\ \rho_{1,p} & \cdots & 1 \end{pmatrix} := R.$$

Propriété des moments :

Properties

Soit X un vecteur aléatoire de moyenne $\mathbb{E}(X) = \mu$ et de variance $\mathrm{Var}(X) = \Sigma$, et soit M une matrice de constantes et v un vecteur de constantes.

1. Σ est définie non-négative et symétrique.
2. $\Sigma = \mathbb{E}\left[(X - \mu)(X - \mu)^\top\right] = \mathbb{E}(XX^\top) - \mu\mu^\top$.
3. $\mathbb{E}(MX + v) = \mathbb{E}(X) + v$.
4. $\mathrm{Var}(MX + v) = \mathrm{Var}(MX) = M\Sigma M^\top$.
5. $\Sigma = \Delta R \Delta \iff R = \Delta^{-1}\Sigma\Delta^{-1}$.

Ajouter qqc sur l'indépendance de variables aléatoires !

Loi normale multivariée : On dit qu'un vecteur aléatoire X de dimension p suit une loi normale multidimensionnelle de moyenne μ et de variance $\Sigma \sim \mathcal{N}_p(\mu, \sigma^2)$, si sa densité est données par

$$f_X(x) = \frac{1}{(2\pi)^{p/2}} \cdot \frac{1}{(\det\Sigma)^{1/2}} \cdot \exp\left\{-\frac{1}{2}\,(x - \mu)^\top\,\Sigma^{-1}\,(x - \mu)\right\}, \quad x \in \mathbb{R}^p.$$

Estimation avec un échantillon : En practique, nous ne connaissons pas les valeurs de μ et de Σ et nous voulons les estimer à partir d'un échantillon. Soit $X_1, \dots, X_n$, n réalisations indépendantes d'un vecteur aléatoire X de moyenne μ et de variance Σ. On estime μ et Σ par :

$$\hat{\mu} = \overline{X} := \frac{1}{n}\sum_{i=1}^{n} X_i$$

$$\widehat{\Sigma} = S^2 := \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})(X_i - \overline{X})^\top.$$

Notons $D = \{\mathrm{diag}(S^2)\}^{1/2}$, la matrice des écarts-types calculés sur l'échantillon. On peut calculer la matrice des corrélations sur l'échantillon par :

$$\widehat{R} = D^{-1}S^2D^{-1} \iff S^2 = D\widehat{R}D.$$

Delmas, Jean-François. 2013. *Introduction au calcul des probabilités et à la statistique : exercices, problèmes et corrections (2e édition)*. Les Presses de l'ENSTA.

Wasserman, Larry. 2010. *All of Statistics : A Concise Course in Statistical Inference*. Springer Publishing Company, Incorporated.