Scribe Séance N° 04

Mathématiques pour l'Intelligence Artificielle

Probabilités

Etudiants :
Ibrahim Riazi
Victor Tancrez

Enseignant :
Thomas Schatz

Master 2 Intelligence Artificielle & Apprentissage Automatique
Aix-Marseille Université

September 19, 2023

1 Univers et mesure de probabilité

1.1 Univers (Espace d'échantillonnage)

L'univers noté Ω, également appelé espace d'échantillonnage, est l'ensemble de tous les résultats possibles d'une expérience aléatoire ou d'un phénomène aléatoire.

1.2 Événement

Un événement A est un sous-ensemble de l'univers, c'est-à-dire un ensemble de résultats possibles. Les événements peuvent être simples (un seul résultat) ou composés (plusieurs résultats).

1.3 l'Ensemble des Événements (Tribu)

L'ensemble des événements, souvent noté $\mathcal{F}$, est une collection d'ensembles d'événements qui représente un espace des probabilités. Cette collection d'ensembles doit être stable par union dénombrable et par passage au complémentaire.

1.4 Mesure de Probabilité

- La mesure de probabilité est une fonction mathématique qui attribue des probabilités à chaque événement dans l'univers.
- Elle quantifie la chance qu'un événement se produise. La probabilité d'un événement est généralement exprimée comme un nombre réel compris entre 0 (impossible) et 1 (certain).
- La somme de toutes les probabilités des événements dans l'univers est égale à 1, ce qui signifie que l'un de ces événements se produira certainement.

$$P : \mathcal{F} \to \mathbb{R}$$
$$\forall A \in \mathcal{F}, P(A) \geq 0$$
$$P(\Omega) = 1$$

Pour tout ensemble dénombrable d'événements disjoints $A_1, A_2, \ldots$, c'est-à-dire $\forall i, j \in \mathbb{N}$, et $i \neq j$, On a, $A_i \cap A_j = \emptyset$, alors :

$$P\left(\bigcup_i A_i\right) = \sum_i P(A_i)$$

2 Variable aléatoire

Une variable aléatoire X est une fonction qui associe chaque élément de l'univers Ω à un élément de l'espace d'arrivée E.

$$X : \Omega \to E$$

2.1 Le tribu sur l'espace d'arrivée

Le tribu $\mathcal{T}$ est un ensemble de sous-ensembles de l'espace d'arrivée E, qui doit être stable par union dénombrable et par passage au complémentaire.

$$\mathcal{T} \subset \mathcal{P}(E)$$

$$\forall T \in \mathcal{T}, X^{-1}(T) \in \mathcal{F}$$

2.2 La loi de X

La loi de X (mesure de probabilité induite sur E) est définie comme suit :

$$\forall T \in \mathcal{T}, \quad P_X(T) = P(X^{-1}(T))$$

2.3 Conditionnemene et l'indépendance

Indépendance mutuelle d'événements

La formule indique que les événements A, B et C sont mutuellement indépendants si et seulement si la probabilité de l'intersection de tous les événements est égale au produit des probabilités individuelles de chaque événement.

$$P(A \cap B \cap C) = P(A)P(B|A)P(C|A \cap B)$$

Loi de probabilité conditionnelle

La loi de probabilité conditionnelle donne la probabilité d'un événement A sachant que l'événement B est survenu.

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Dans le cas de A et B deux événements indépendants
Si A et B sont deux événements indépendants, alors :

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \Leftrightarrow P(A|B) = \frac{P(A)P(B)}{P(B)} \Leftrightarrow P(A|B) = P(A)$$

Variables aléatoires indépendantes

Les variables aléatoires indépendantes sont des variables aléatoires qui ne sont pas liées ou corrélées entre elles.
Soient les variables aléatoires X et Y:
On dit que X et Y sont des variables aléatoires indépendantes si et seulement si leur fonction de distribution conjointe est égale au produit de leurs fonctions de distribution marginales.

$$P(X = x, Y = y) = P(X = x)P(Y = y)$$

2.4 Variables aléatoires réelles

Fonction de masse

La fonction de masse $P(X = x)$ donne la probabilité que la variable aléatoire X prenne la valeur x.

$$P(X = x)$$

Fonction de répartition

La fonction de répartition $F(x)$ donne la probabilité que la variable aléatoire X soit inférieure ou égale à x.

$$F(x) = P(X \leq x)$$

Densité de probabilité

La fonction de densité de probabilité $f(x)$ décrit comment la probabilité est répartie sur l'ensemble des valeurs possibles de la variable aléatoire continue. Cette fonction est la dérivée de la fonction de répartition $F(x)$.

$$f(x) = \frac{d}{dx}F(x)$$

L'intégrale de la densité de probabilité sur l'ensemble de toutes les valeurs possibles de x est égale à 1 (**L'aire sous la courbe de densité doit être égale à 1**) :

$$\int_{-\infty}^{+\infty} f(x)\,dx = 1$$

Espérance (intégrale de Lebesgue)

L'espérance représente la "moyenne pondérée" de la variable aléatoire continue X, où les valeurs sont pondérées par leurs probabilités. Cette espérance est calculée à partir de l'intégrale de Lebesgue de x multiplié par $f(x)$ sur l'ensemble de l'espace d'échantillonnage.

$$E[X] = \int xf(x)dx$$

Cas continu

Dans le cas continu, les fonctions de densité et de distribution sont définies sur un intervalle continu.

$$P(a \leq X \leq b) = \int_a^b f(x)dx$$

Cas discret

Dans le cas discret, les fonctions de masse et de distribution sont définies pour des valeurs discrètes de X.

$$P(X = x) = p_x$$

2.5 Plusieurs variables aléatoires réelles

Définition de la variable aléatoire vectorielle

Une variable aléatoire vectorielle $X = (X_1, X_2, \ldots, X_n)$ est une fonction de l'espace d'échantillon Ω vers $\mathbb{R}^n$.

$$X : \Omega \to \mathbb{R}^n$$

Fonction de masse de probabilité jointe

La fonction de masse de probabilité jointe p_X est donnée par

$$p_X(x_1, x_2, \ldots, x_n) = P_X(\{x_1, x_2, \ldots, x_n\})$$

Fonction de masse marginale

La fonction de masse marginale de X, notée $P(X = x)$, est calculée comme suit à partir de la fonction de masse conjointe $P(X = x, Y = y)$:

$$P(X = x) = \sum_{\text{toutes les valeurs de } y} P(X = x, Y = y)$$

Fonction de répartition (jointe)

La fonction de répartition jointe F_X est donnée par

$$F_X(x_1, x_2, \ldots, x_n) = P_X(\{y_1, y_2, \ldots, y_n | y_1 \leq x_1, y_2 \leq x_2, \ldots, y_n \leq x_n\})$$

Densité de probabilité (jointe)

La densité de probabilité (jointe) est donnée par

$$pdX : x_1, x_2, \ldots, x_n \mapsto \frac{\partial^n F_X}{\partial x_1 \partial x_2 \ldots \partial x_n}(x_1, x_2, \ldots, x_n)$$

Densité de probabilité marginale

La fonction de densité marginale permet de décrire la distribution de probabilité individuelle de chaque variable aléatoire, en ignorant les autres variables aléatoires. Pour calculer la densité de probabilité marginale de X, notée $f_X(x)$, à partir de la densité de probabilité conjointe de X et Y, on peut utiliser la formule suivante :

$$f_X(x) = \int_{-\infty}^{+\infty} f(x, y)\, dy$$

Espérance avec plusieurs variables aléatoires (cas continu)

Dans le cas continu, l'espérance de $f(X)$ est donnée par

$$E[f(X)] = \int \dots \int f(x_1, x_2, \dots, x_n) pdX(x_1, x_2, \dots, x_n) dx_1 dx_2 \dots dx_n$$

Espérance avec plusieurs variables aléatoires (cas discret)

Dans le cas discret, l'espérance de $f(X)$ est donnée par

$$E[f(X)] = \sum_{x_1} \sum_{x_2} \dots \sum_{x_n} f(x_1, x_2, \dots, x_n) p_X(x_1, x_2, \dots, x_n)$$

2.6 Modèle graphique

Un modèle graphique est une représentation d'objets probabilistes. C'est un graphe qui représente les dépendances de variables aléatoires.

$$p(X_1, X_2, \dots, X_n) = \prod_{i=1}^{n} P(X_i | X_{\sim i})$$

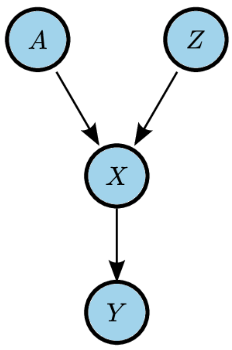

Figure 1: exemple de modèle graphique

Dans cette exemple

2.7 Moments

Variance

$$\text{Var}[X] = E\left[(X - E[X])^2\right]$$

Cette formule représente la variance de la variable aléatoire X, qui mesure la dispersion des valeurs de X par rapport à sa moyenne $E[X]$.

Moment

$$E[X^k]$$

Les moments d'une variable aléatoire sont des indicateurs de la dispersion d'une variable. Le moment d'ordre 1 avec k=1 équivaut à l'esperance de la variable X, le moment d'ordre 2 avec k=2 correspond à la variance X.

Moment centré

$$E\left[(X - E[X])^k\right]$$

Le premier moment centré est égale 0

Covariance

$$\mathrm{Cov}(X,Y) = E\left[(X - E[X])\,(Y - E[Y])\right]$$

Cette formule représente la covariance entre les variables aléatoires X et Y, qui mesure le degré de variation simultanée de X et Y. elle peut être positive si l'écart entre les variables et leurs moyennes a le même signe. Négative sinon X

Linéarité de l'espérance

$$E\left[\sum_{i=1}^{n} a_i X_i\right] = \sum_{i=1}^{n} a_i E[X_i]$$

Cette formule indique que l'espérance d'une combinaison linéaire de variables aléatoires est égale à la combinaison linéaire de leurs espérances respectives.

Bilinéarité de la variance

$$\mathrm{Var}\left[\sum_{i=1}^{n} a_i X_i\right] = \sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j \mathrm{Cov}(X_i, X_j)$$

2.8 moments conditionnels

Espérance conditionnelle

$$f(y) = E[X|Y=y] = \int x P_{X|Y=y}(dx)$$

Cette formule définit l'espérance de X conditionnelle à $Y = y$, celle-ci donne la valeur moyenne de cette variable quand un certain événement est réalisé

Variance conditionnelle

$$g(y) = \mathrm{Var}[X|Y=y] = E\left[(X - E[X])^2\,|Y=y\right]$$

Cette formule donne la variance de X conditionnelle à $Y = y$, qui mesure la dispersion des valeurs de X autour de leur moyenne conditionnelle, pour une valeur donné de Y.

Loi de l'espérance totale

$$E[X] = E[E[X|Y]]$$

L'espérance de l'espérance conditionnelle de X sachant Y est la même que l'espérance de X.

Loi de la variance totale

$$\mathrm{Var}[X] = \mathrm{Var}[E[X|Y]] + E[\mathrm{Var}[X|Y]]$$