ASE3 - Cheat Sheet 📚

Partie 1 : Couple de variables aléatoires discrètes et analyse des données

Loi conjointe d'un couple (X, Y)

Propriétés

Soient X et Y deux variables aléatoires avec :

- $X(\Omega) = \{x_i, i\epsilon\mathbb{N}\}$
- $Y(\Omega) = \{y_j, j\epsilon\mathbb{N}\}$

Concernant les probabilités nous avons la propriété suivante :

$$p_{i,j} = P((X = x_i) \cap (Y = y_j)) = P(X = x_i, Y = y_j)$$

Ces propriétés sont souvents données dans un tableau à double entrées.

⚠️ Attention à bien noter la propriété suivante pour les sommes :

$$\sum_{i,j\epsilon\mathbb{N}} p_{i,j} = 1$$

Exemple

Exemple du dé, soit X la variable qui vaut 1 si le résultat est impair et 2 si il est pair. Soit Y qui vaut 0 si le résultat est 1, 1 si 6 et 2 sinon.

On obtient alors le tableau à double entrée suivant :

X/Y	0	1	2	p_i
1	$\frac{1}{6}$	0	$\frac{2}{6}$	$\frac{1}{2}$
2	0	$\frac{1}{6}$	$\frac{2}{6}$	$\frac{1}{2}$
p_j	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{4}{6}$	1

Les probabilités ont été obtenues avec les calculs suivant :

- $P((X = 1) \cap (Y = 0)) = P(X = 1) = \frac{1}{6}$
- $P((X = 1) \cap (Y = 1)) = P(\emptyset) = 0$
- $P((X = 1) \cap (Y = 2)) = P(X = 3 \cup X = 5) = \frac{2}{6}$
- $P((X = 2) \cap (Y = 0)) = P(\emptyset) = 0$
- $P((X = 2) \cap (Y = 1)) = P(X = 6) = \frac{1}{6}$
- $P((X = 2) \cap (Y = 2)) = P(X = 2 \cap X = 4) = \frac{2}{6}$

Lois Marginales

Définition
On appelle la première / deuxième **loi marginale** la loi de la première composante X / deuxième composante Y. On les obtient de la façon suivante :
- $p_i = P(X = x_i) = \sum_{j \epsilon \mathbb{N}} P(X = x_i, Y = y_j) = \sum_{j \epsilon \mathbb{N}} p_{i,j}$
- $p_j = P(Y = y_j) = \sum_{i \epsilon \mathbb{N}} P(X = x_i, Y = y_j) = \sum_{i \epsilon \mathbb{N}} p_{i,j}$

Exemple

Dans l'exemple précedent les colonnes p_i et p_j sont les lois marginales respectivement de X et Y.

Lois Conditionnelles

Propriétés

Définition
La **loi conditionnelle** est la loi qui pour soit A un évènement tel que $P(A) \neq 0$ et soit X une variable aléatoire réelle sur $\{\Omega, C, P\}$ est la probabilité de X sachant A

Ainsi la loi conditionnelle de X sachant $A = (x_i, P_A(X = x_i))$ est :
$P_A(X = x_i) = \frac{P((X=x_i) \cap A)}{P(A)}$

Dans le cas où A est l'évènement $Y = y_j$:

$$P_{(Y=y_i)}(X = x_i) = \frac{P((X=x_i) \cap (Y=y_i))}{P(Y=y_i)}$$

Pour la loi d'un couple (X, Y) :

$$P[(X, Y) = (i, j)] = P(X = i)P_{X=i}(Y = j)$$

Example

Soit (X, Y) un couple dont la loi conjointe est :

X/Y	1	2	3	4	p_i
1	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{4}$
2	0	$\frac{2}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{4}$
3	0	0	$\frac{3}{16}$	$\frac{1}{16}$	$\frac{1}{4}$
4	0	0	0	$\frac{4}{16}$	$\frac{1}{4}$
p_j	$\frac{1}{16}$	$\frac{3}{16}$	$\frac{5}{16}$	$\frac{7}{16}$	1

Dont on veut obtenir la loi conditionnelle telle que $P_{(Y=3)}(X = x_i)$. On dresse alors le tableau suivant :

X_i	**1**	**2**	**3**	**4**
$P_{(Y=3)}(X = x_i)$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{3}{5}$	0

Les probabilités ont été calculées de la manière suivante :

- $P_{(Y=3)}(X = 1) = \frac{P((X=1) \cap (Y=3))}{P(Y=3)} = \frac{\frac{1}{16}}{\frac{5}{16}} = \frac{1}{5}$

- $P_{(Y=3)}(X=2) = \frac{P((X=2)\cap(Y=3))}{P(Y=3)} = \frac{\frac{1}{16}}{\frac{5}{16}} = \frac{1}{5}$
- $P_{(Y=3)}(X=3) = \frac{P((X=3)\cap(Y=3))}{P(Y=3)} = \frac{\frac{3}{16}}{\frac{5}{16}} = \frac{3}{5}$
- $P_{(Y=3)}(X=4) = \frac{P((X=4)\cap(Y=3))}{P(Y=3)} = \frac{0}{\frac{5}{16}} = 0$

Indépendance de variables aléatoires discrètes

X et Y sont dites indépendantes si et seulement si :
$P(X = x_i, Y = y_j) = P(X = x_i)P(Y = y_j)$
$\Leftrightarrow P_{i,j} = P_i P_j$

Calcul de probabilités d'une fonction

Propriétés

Soit g une fonction de $\mathbb{R}^2 \to \mathbb{R}$, définie sur l'ensemble des valeurs prises par (X, Y).
Soit $Z = g(X, Y), Z_n = g(x_i, y_i)\epsilon Z(\Omega)$, alors :
$(Z = z_k) = U_{(i,j)}((X = x_i) \cap (Y = y_i))$ avec $z_k = g(x_i, y_i)$
$\to P(Z = z_k) = \sum_{x,y} P((X = x_i) \cap (Y = y_j))$ avec $z_k = g(x_i, y_i)$

Dans les cas particuliers :
- $Z = X + Y = g(X, Y)$
  $P(Z = z) = \sum(x, y)P((X = x) \cap (T = y))$ avec $x + y = z$
- $Z = X.Y$
  $P(X.Y = z) = \sum(x, y)P((X = x) \cap (T = y))$ avec $x.y = z$

Exemple

X/Y	1	2	3	4	p_i
1	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{4}$
2	0	$\frac{2}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{4}$
3	0	0	$\frac{3}{16}$	$\frac{1}{16}$	$\frac{1}{4}$
4	0	0	0	$\frac{4}{16}$	$\frac{1}{4}$
p_j	$\frac{1}{16}$	$\frac{3}{16}$	$\frac{5}{16}$	$\frac{7}{16}$	1

Déterminer la loi de $Z = X + Y$

$Z = \{2, 3, 4, 5, 6, 7, 8\}$

Z_k	**2**	**3**	**4**	**5**	**6**	**7**	**8**
$P_{(Y=3)}(X = x_i)$	$\frac{1}{16}$	$\frac{3}{16}$	$\frac{3}{16}$	...			

Avec les calculs suivants :

- $P(Z = 2) = P(X + Y = 2) = P((X = 1) \cap (Y = 1)) = \frac{1}{16}$
- $P(Z = 3) = P(X + Y = 3) = P((X = 1) \cap (Y = 2)) + P((X = 2) \cap (Y = 1)) = \frac{1}{16} + \frac{2}{16} = \frac{3}{16}$
- $P(Z = 4) = P(X + Y = 4) = P((X = 1) \cap (Y = 3)) + P((X = 2) \cap (Y = 2)) + P((X = 3) \cap (Y = 1)) = \frac{1}{16} +$
- $P(Z = 5) = P(X + Y = 5) = P((X = 1) \cap (Y = 4)) + P((X = 2) \cap (Y = 3)) + P((X = 3) \cap (Y = 2)) + P(($

Déterminer la loi de $Z = X.Y$

$Z = \{1, 2, 3, 4, 6, 8, 9, 12, 16\}$

$P(Z=4) = P((X=1) \cap (Y=4)) + P((X=2) \cap (Y=2)) + P((X=4) \cap (Y=1)) = \frac{1}{16} + \frac{2}{16} + 0 = \frac{3}{16}$
...

Espérance d'une fonction de 2 variables discrètes

Propriétés

Soit X et Y tels que $Z = g(X,Y)$
$E(Z) = E(g(X,Y)) = \sum_{i,j} g(x_i, y_j)P((X=x_i) \cap (y=y_j))$
$\Leftrightarrow E(Z) = \sum_{i,j} g(x_i, y_j)p_{i,j}$

On note aussi :
- $E(X) = \sum_i iP(X=x_i)$
- $E(Y) = \sum_j jP(Y=y_j)$

Si X et Y sont deux V.A. indépendantes alors :
$E(X.Y) = E(X).E(Y)$
⚠️

La réciproque est fausse

Example

$g(X,Y) = X.Y$
$E(X.Y) = \sum_{i,j} g(x_i, y_j)P((X=x_i) \cap (y=y_j))$
$E(X.Y) = \sum_{i,j} x_i.y_jP_{ij}$

Covariance et coefficient de corrélation linéaire

Covariance

Définition
X et Y deux variables discrètes, on appelle la covariance le réel :
$Cov(X,Y) = E((X - E(X)(Y - E(y)))$

NB
Si X et Y admettent tous les 2 une variance, alors on aura :
$Cov(X,Y) = E(XY) - E(X)E(Y)$

⚠️ Si X et Y sont indépendantes alors $Cov(X,Y) = 0$ ⚠️

CCL

Définition
On suppose que X et Y admettent une variance non nulle, le coefficeient de corrélation linéaire de (X,Y), noté $C(x,y)$ / $p(x,y)$, est :
$p(X,Y) = \frac{Cov(X,Y)}{\sigma_x \sigma_y}$

On notera bien :
- $\sigma_x = \sqrt{V(X)}$
- $\sigma_y = \sqrt{V(Y)}$

Rappels utiles

Lois discrètes

Loi Binomiale $B(n,p)$

La loi binomiale modélise la fréquence du nombre de succès obtenus lors de la répétition de plusieurs expériences aléatoires identiques et indépendantes.

On considère une expérience aléatoire E et un événement A lié à E, de probabilité non nulle, avec P(A) = p.
On appelle succès, la réalisation de A et échec, celle de

- $P(X=k) = \binom{n}{k}p^k(1-p)^{n-k}$ avec $\binom{n}{k}$ = $\frac{n!}{k!(n-k)!}$
- $E(X) = np$
- $V(X) =$$npq$ avec $q = 1-p$

Loi de Bernouilli de paramètre p **:** $B(p)$

X ne peut prendre que les valeurs 1 ou 0, c'est une distribution discrète de probabilité, qui prend la valeur 1 avec la probabilité p et 0 avec la probabilité $q = 1-p$.

- $P(X=k) = p$ et $P(X=0) = 1-p = q$
- $E(X) = p$
- $V(X) = p(1-p)$

Loi de Poisson : $P(\lambda)$ **(λ paramètre)**

Décrit le comportement du nombre d'événements se produisant dans un intervalle de temps fixé, si ces événements se produisent avec une fréquence moyenne ou espérance connue, et indépendamment du temps écoulé depuis l'événement précédent.

- $P(X=k) = \frac{\lambda^k}{k!}e^{-\lambda}$
- $E(X) = V(X) = \lambda$

Suites géométriques

$\sum_{k=0}^{+\infty} X^n = \frac{1}{1-X}$

Les tricks

$e^X = \sum_{j=0}^{+\infty} \frac{x^j}{j!}$
Dans le cas de $X = 1$:
$e = \sum_{j=0}^{+\infty} \frac{1}{j!}$

$\sum x = \frac{x(x+1)}{2}$

Les sommes

$\sum_{i=1}^{n} \sum_{j=i}^{n} a_{ij} = \sum_{j=1}^{n} \sum_{i=1}^{j} a_{ij}$

Variance

- $E(X^2) - (E(X))^2 = V(X)$
- $V(X+Y) = V(X) + V(Y) + 2Cov(X,Y)$

Binôme de Newton

$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$

Partie 2 : Analyse en composantes principales

Rappels utiles :

1. *Multiplication de matrice*
   - La matrice résultante doit être : $A_{n*m} * B_{m*p} => C_{n*p}$
   - Trois propriétés importantes :
     - $AB \neq BA$
     - $I_n A = A I_n = A$
     - Si A et B sont carrées de taille n, alors $AB = I_n => BA = I_n$ et $B = A^{-1}$ (inverse de A)
2. *Calcul de la trace*
   - Somme des termes de la diagonale d'une matrice carrée
   - Notée sous la forme $Tr(A)$ ou $Trace(A)$
3. *Représentation par transposition*
   - Échange des lignes et des colonnes d'une matrice
   - Notée sous la forme A' ou A^T

Matrice des poids p_i

Nous associons à chaque individu un poids $p_i \geq 0$ qui correspond à la probabilité de choisir un individu. La somme des poids de la matrice est égale à 1.

Formule :

$$\sum_{i=1}^{n} p_i = p_1 + p_2 + p_3 + \ldots = 1$$

$$D = \begin{pmatrix} p_1 & 0 & \ldots & 0 \\ 0 & p_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & p_n \end{pmatrix}$$

Cas uniforme : Si tous les individus ont le même poids, alors $p_i = \frac{1}{n} \Rightarrow D = \frac{1}{n} I_n$, où I_n est la matrice identité.

Moyenne des variables $\bar{X}^{(j)}$ et centre de gravité g

Définition : La moyenne d'une colonne $\bar{X}^{(j)}$ (avec j le numéro de colonne et $j \in [1, p]$) s'obtient en additionnant chaque valeur de colonne et en multipliant l'ensemble par son poids p_i :

$$\bar{X}^{(j)} = \sum_{i=j}^{n} P_i X_i^{(j)}$$

Définition : Le centre de gravité, représenté par le vecteur g des moyennes arithmétiques de chaque variable $X^{(j)}$, est définit par $g = (\bar{X}^{(1)}, \bar{X}^{(2)}, \dots, \bar{X}^{(j)})$.

Matrice des données centrées Y

Définition : La matrice Y s'obtient en soustrayant chaque moyenne $\bar{X}^{(j)}$ de la matrice initiale X, c'est-à-dire :

$$Y_i^{(j)} = X_i^{(j)} - \bar{X}^{(j)} \quad , \forall j \in [1, p], \forall i \in [1, n]$$

Matrice de variance-covariance V

Définition : La matrice de variance-covariance V (ou *var-covariance*) est une matrice carrée de dimension *p* représentée sous la forme suivante :

$$V = \begin{pmatrix} \sigma_1^2 & \sigma_{1,2} & \dots & \sigma_{1,p} \\ \sigma_{2,1} & \sigma_{2,2} & \dots & \sigma_{2,p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{p,1} & \sigma_{p,2} & \dots & \sigma_p^2 \end{pmatrix}$$

Formule : Cette matrice s'obtient avec la formule

$$V = Y^T * D * Y$$

Symétrie : La matrice V est symétrique, donc $V^T = V$.

Diagonalisation d'une matrice

Rappel : Soit une matrice A, diagonaliser cette matrice revient à chercher une matrice diagonale D ainsi qu'une matrice inversible P telle que :

$$A = P * D * P^{-1}$$

Dans le cours d'ASE3, diagonaliser revient à calculer les valeurs propres de la matrice afin d'en déterminer par la suite ses composantes et facteurs principaux.

Valeurs propres

Définition (rappel) : Soit une matrice A, on appelle polynôme caractéristique de A, noté en général P_A, le polynôme défini par

$$P_A(\lambda) = det(A - \lambda I_n)$$

En calculant ce polynôme, nous pouvons trouver les **valeurs propres** de la matrice A.

Pourcentage d'inertie

Définition : L'inertie totale mesure l'étalement du nuage de points d'une matrice. L'inertie de l'axe α est calculée divisant sa valeur propre λ_α par la somme des valeurs propres des différents axes.

Formule :

$$\text{Inertie de l'axe } \alpha = \frac{\lambda_\alpha}{\sum_1^n \lambda_n}$$

⚠️ La résultat doit être présenté sous forme de pourcentage.

Facteurs principaux

Définition : Les facteurs principaux sont les vecteurs propres associés aux plus grandes valeurs propres.

Formule :
Pour trouver les vecteurs propres de V, nous posons $E(\lambda_\alpha) = Ker(V - \lambda_\alpha I_n)$

Méthode :

$$\forall u = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in E(\lambda_\alpha) <=> (V - \lambda_\alpha I_n) \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \vec{0}$$

Après calcul par intégration linéaire, on trouve :

$$E(\lambda_\alpha) = Vect(\begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix})$$

(où α, β et γ sont les solutions de l'équation linéaire trouvée pour i).

Pour calculer u, on pose :

$$u^{(i)} = \frac{1}{\sqrt{\alpha^2 + \beta^2 + \gamma^2}} \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix}$$

Remarque : $E(\lambda_\alpha)$ est une droite vectorielle et u est normé.

Composantes principales

La composante principale C est définie par :

$$C^{(i)} = Y * u^{(i)}$$

Remarque : Les composantes principales contiennent les projections d'individus sur les axes factoriels.

Coefficients de corrélation linéaire

Définition : La méthode la plus naturelle pour donner une signification à une composante principale $C^{(i)}$ est de la relier aux variables $X^{(j)}$ (variables intiales) en calculant les coefficients de corrélation linéaire :

$$p(X^{(j)}, C^{(i)})$$

Formule :

$$p(X^{(j)}, C^{(i)}) = \frac{Cov(X^{(j)}, C^{(i)})}{\sigma_{X^{(j)}}\sigma_{C^{(i)}}}$$

Remarque :

$$Cov(X^{(j)}, C^{(i)}) =< y^{(j)}, C^{(i)} >$$

où $y^{(j)}$ est une variable centrée.