Licence AES 3ème année

Université de Lille II

Notes du cours d'introduction à l'économétrie

P. De Vreyer

Partie I: Introduction au calcul des probabilités.

1. Variables aléatoires et distributions de probabilité.

Définition 1: Variable aléatoire.
Une variable aléatoire est une variable dont la valeur est indéterminée *a priori*.

Par exemple: si je sors dans la rue et que je demande aux gens combien ils gagnent par mois, je ne sais par ce qu'ils vont me répondre. De mon point de vue (qui est en général celui de l'économètre) leur réponse est aléatoire.

Dans ce qui suit, les variables aléatoires seront repérées par des lettres majuscules (ex. X) et les valeurs qu'elles peuvent prendre par des lettres minuscules (ex. x).
Une variable aléatoire peut être discrète ou continue. Une variable aléatoire discrète prend ses valeurs dans un ensemble de taille finie ou infini et dénombrable. Une variable aléatoire continue prend ses valeurs dans un ensemble infiniment divisible et non dénombrable.

Par exemple: Si X est le résultat du lancement d'un dé à 6 faces, X est une variable aléatoire discrète. Elle peut prendre ses valeurs dans l'ensemble {1,2,3,4,5,6}. Si X est la moyenne des températures au cours du mois de janvier, X est une variable aléatoire continue: elle prend ses valeurs dans un intervalle, par exemple [-12,+10].

L'ensemble des valeurs possibles que peut prendre une variable aléatoire est appelé univers.

Définition 2: Evènement.
On appelle événement le fait qu'une variable aléatoire prend une valeur particulière. On notera {X=x} ou (X=x) l'évènement « X est égal à x ».

Définition 3: Probabilité et densité.
Une probabilité est une fonction, f(x), associée aux valeurs que peut prendre une variable aléatoire X et qui vérifie les propriétés suivantes:

- $0 \leqslant f(x) \leqslant 1$ pour tout x.
- $\sum_x f(x) = 1$ dans le cas discret ou $\int_x f(x)dx = 1$ dans le cas continu.

Dans le cas d'une variable aléatoire discrète, f(x) est la probabilité que l'évènement {X=x} se réalise, notée P(X=x).

Exemple: On lance un dé à 6 faces. Soit X le résultat du lancement. X peut prendre les valeurs suivantes: {1,2,3,4,5,6}. L'évènement {X=5} a une probabilité égale à P(X=5) = f(5) = 1/6.

Dans le cas continu, la fonction f(x) est appelée densité de X. Pour une variable aléatoire continue, la probabilité d'une valeur particulière de X est nulle (pensez-vous par exemple qu'il y ait une probabilité différente de zéro que la moyenne des températures du mois de mars 2006 soit égale à 12,9823456726351 ?) Dans le cas d'une variable aléatoire continue seuls les intervalles ont des probabilités non nulles *a priori*. La probabilité pour qu'une réalisation de la variable aléatoire X soit comprise entre α et β est égale à:

$$P(\alpha \leq x \leq \beta) = P(\alpha < x \leq \beta) = P(\alpha \leq x < \beta) = P(\alpha < x < \beta) = \int_{\alpha}^{\beta} f(x)\,dx$$

Les trois premières égalités découlent du fait que la probabilité d'une valeur particulière est nulle.

<u>Définition 4: Fonction de répartition.</u>
La fonction de répartition d'une variable aléatoire X est la fonction F définie par:
$F(t) = P(X \leq t)$

Dans le cas discret: $F(t) = \sum_{x \leq t} P(X = x)$

Dans le cas continu: $F(t) = \int_{-\infty}^{t} f(x)\,dx$

Propriétés de la fonction de répartition:

$Si\, t_1 > t_{2,}\ F(t_1) \geq F(t_2)$

$0 \leq F(t) \leq 1$ pour toute valeur de t

$F(+\infty) = 1$

$F(-\infty) = 0$

$P(\alpha < X \leq \beta) = F(\beta) - F(\alpha)$

2. Espérance et variance d'une variable aléatoire.

<u>Définition 5: Espérance d'une variable aléatoire.</u>

1. Cas discret.
Soit X une variable aléatoire de loi de probabilité P. L'espérance de X est égale à:
$\sum_x x.P(X = x)$

2. Cas continu.
Soit X une variable aléatoire de densité de probabilité f. L'espérance de X est égale à:
$\int_x x.f(x).dx$

Interprétation: l'espérance d'une variable aléatoire est une moyenne pondérée par leur probabilité de réalisation des valeurs que peut prendre cette variable aléatoire.

Propriétés de l'espérance.
Soit g une fonction de la variable aléatoire X. g est elle même une variable aléatoire et son espérance est égale à

$E(g(X)) = \sum_x g(x).P(X = x)$ dans le cas discret

$E(g(X)) = \int_x g(x).f(x).dx$ dans le cas continu.

L'espérance est un opérateur linéaire. C'est à dire que si X et Y sont deux variables aléatoires

et si a et b sont deux constantes:
$E(a.X+b.Y)=a.E(X)+b.E(Y)$
L'espérance d'une constante est égale à la constante elle-même: $E(a)=a$.

Définition 6: Variance et écart-type d'une variable aléatoire.
La variance d'une variable aléatoire X est égale à l'espérance de $(X-E(X))^2$:
$Var(X)=V(X)=E[(X-E(X))^2]$
En d'autres termes:
$V(X)=\sum_x (x-E(X))^{2.} P(X=x)$ dans le cas discret
$V(X)=\int_x (x-E(X))^{2.} f(x).dx$ dans le cas continu.

L'écart-type d'une variable aléatoire est la racine carrée de sa variance: $\sigma_X=\sqrt{(V(X))}$.

Propriétés de la variance:
Si X est une variable aléatoire et a une constante:
$V(X)=E(X^2)-E(X)^2$
$V(aX)=a^2V(X)$
$V(a)=0$

3. Couples de variables aléatoires.

Définition 7: Loi de probabilité et densité jointes de deux variables aléatoires.
Soient X et Y deux variables aléatoires prenant des valeurs notées x et y, respectivement. La loi de probabilité jointe de X et de Y est une fonction f(x,y) telle que:
$0\leq f(x,y)\leq 1$ pour tout couple (x,y)
$\sum_x \sum_y f(x,y)=1$ dans le cas discret
$\int_x \int_y f(x,y)dxdy=1$ dans le cas continu.
Dans le cas discret, $f(x,y)=P(X=x\, et\, Y=y)$ que l'on notera souvent $P(X=x, Y=y)$.
Dans le cas continu, la probabilité d'un couple de valeurs particulières est nulle, comme dans le cas d'une seule variable aléatoire. Seule la probabilité d'un pavé (croisement de deux intervalles) est *a priori* non nulle:
$$P(\alpha\leq X\leq\beta, \gamma\leq Y\leq\delta)=\int_\alpha^\beta \int_\gamma^\delta f(x,y)dxdy$$
Dans le cas continu, la fonction f est appelée densité jointe de X et de Y.

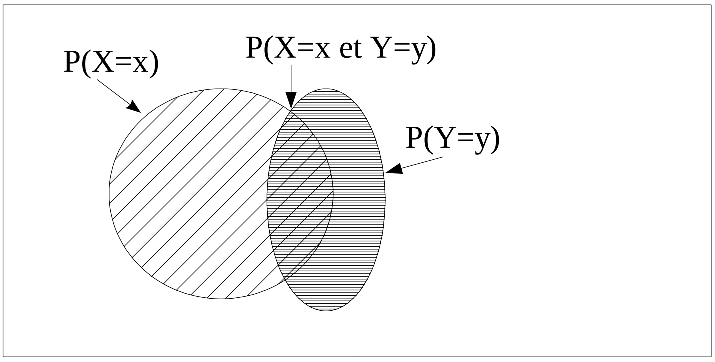
Propriétés:
Soient X et Y deux variables aléatoires. Soient {X=x} et {Y=y} deux événements. La probabilité que l'un ou l'autre de ces évènements se réalise est égale à:

$$P(X=x \text{ ou } Y=y)=P(X=x)+P(Y=y)-P(X=x \text{ et } Y=y)$$

Notez bien la différence entre l'évènement $\{X=x \text{ ou } Y=y\}$ et l'évènement $\{X=x \text{ et } Y=y\}$.

Si {X=x} et {Y=y} ne peuvent survenir simultanément (on dit que ce sont deux évènements incompatibles), alors $P(X=x \text{ et } Y=y)=0$ et dans ce cas:
$P(X=x \text{ ou } Y=y)=P(X=x)+P(Y=y)$

On peut donner une représentation graphique intuitive de ces relations.



La surface du rectangle représente la somme des probabilités de l'ensemble des événements qui peuvent se produire. Sa valeur est donc égale à l'unité. La probabilité de l'évènement {X=x} est une portion de cette surface. Celle de l'évènement {Y=y} également. On suppose dans la représentation graphique que les deux événements ne sont pas incompatibles: de sorte que la probabilité que {X=x} se réalise en même temps que {Y=y} n'est pas nulle. Dans le graphique ci-dessus c'est la zone doublement hachurée qui représente cette probabilité. La probabilité que l'un ou l'autre des évènements {X=x} et {Y=y} se réalise est égale à la somme des probabilités P(X=x) et P(Y=y), lorsque ces deux évènements sont incompatibles (ou, sur le graphique, lorsque les surfaces qui correspondent à P(X=x) et P (Y=y) sont disjointes). Lorsque les évènements ne sont pas incompatibles, pour calculer P (X=x ou Y=y), il faut additionner les probabilités P(X=x) et P(Y=y) et en retirer la probabilité que {X=x} et {Y=y} se réalisent simultanément, car sinon cette probabilité est comptée deux fois dans le total.

<u>Définition 8: fonction de répartition d'un couple de variables aléatoires.</u>
Soient X et Y deux variables aléatoires. La fonction de répartition de la loi jointe de X et Y est la fonction F définie par:

$$F(t_{1,} t_2) = P(X \leq t_{1,} \; Y \leq t_2)$$

4. Distributions marginales et indépendance de variables aléatoires.

<u>Définition 9: Distribution marginale.</u>
Soient X et Y deux variables aléatoires de loi de probabilité jointe f(x,y). La densité marginale de X est donnée par:

$$f_X(x) = \sum_y f(x,y) = \sum_y P(X=x, Y=y) = P(X=x) \text{ dans le cas discret.}$$

$$f_X(x) = \int_y f(x,y)\,dy = f(x) \text{ dans le cas continu.}$$

La distribution marginale de X est identique à celle de X lorsque seules les réalisations de X sont observées.

<u>Définition 10: indépendance de deux variables aléatoires.</u>
Soient X et Y deux variables aléatoires. X et Y sont dites indépendantes lorsque la loi jointe de X et de Y est égale au produit de leurs lois marginales:

$$f(x,y) = f_X(x) . f_Y(y)$$

Dans le cas discret ceci s'exprime:

$$P(X=x, Y=y) = P(X=x \text{ et } Y=y) = P(X=x) . P(Y=y)$$

L'indépendance signifie que la réalisation de X ne dépend pas de celle de Y et réciproquement.

<u>Définition 11: Espérance dans une distribution jointe.</u>

La moyenne et la variance des variables dans une distribution jointe sont calculées à partir des distributions marginales. Par exemple pour l'espérance de X:

$$E(X)=\sum_x x.P(X=x)=\sum_x x.\sum_y P(X=x,Y=y) \text{ dans le cas discret}.$$

$$E(X)=\int_x x.f_X(x)dx=\int_x x.(\int_y f(x,y)dy)dx \text{ dans le cas continu}.$$

Définition 12: Espérance d'une fonction de deux variables aléatoires.
Soient X et Y deux variables aléatoires et soit g une fonction de X et de Y. g(X,Y) est une variable aléatoire, dont l'espérance est donnée par:

$$E(g(X,Y))=\sum_x\sum_y g(x,y).P(X=x,Y=y) \text{ dans le cas discret}.$$

$$E(g(X,Y))=\int_x\int_y g(x,y).f(x,y).dx.dy \text{ dans le cas continu}.$$

5. Covariance et corrélation.

Définition 13: covariance de deux variables aléatoires.
Soient X et Y deux variables aléatoires. La covariance de X et de Y est égale à l'espérance de la fonction de X et de Y égale à $(X-E(X)).(Y-E(Y))$. Autrement dit:

$$cov(X,Y)=E((X-E(X))(Y-E(Y))).$$

Ainsi:

$$cov(X,Y)=\sum_x\sum_y (x-E(X))(y-E(Y))P(X=x,Y=y) \text{ dans le cas discret}.$$

$$cov(X,Y)=\int_x\int_y (x-E(X))(y-E(Y))f(x,y)dxdy \text{ dans le cas continu}.$$

Propriétés: on peut exprimer la covariance de X et Y de la façon suivante:

$$cov(X,Y)=E(X.Y)-E(X).E(Y).$$

« Covarier » signifie « varier ensemble ». La covariance est positive lorsque des valeurs de X plus petites que la moyenne (E(X)), tendent à survenir en même temps que des valeurs de Y également plus petites que la moyenne (E(Y)). La covariance est négative lorsque, au contraire, des valeurs de X plus petites que la moyenne tendent à survenir en même temps que des valeurs de Y plus grandes que la moyenne. Cependant la covariance est sensible à l'ordre de grandeur des valeurs prises par les variables aléatoires X et Y. Pour cette raison on préfère parfois utiliser le coefficient de corrélation pour évaluer l'association entre deux variables aléatoires.

Définition 14: coefficient de corrélation.
Soient X et Y deux variables aléatoires. Le coefficient de corrélation entre X et Y est égal à:

$$corr(X,Y)=\frac{cov(X,Y)}{\sigma_X\sigma_Y} \text{ avec } \sigma_X=\sqrt{(V(X))} \text{ et } \sigma_Y=\sqrt{(V(Y))}.$$

Propriétés:

$$corr(X,Y)\in[-1,+1].$$

$$corr(X,Y)=+1 \text{ si, et seulement si, il existe } a>0 \text{ et } b \text{ tels que } X=aY+b.$$

$$corr(X,Y)=-1 \text{ si, et seulement si, il existe } a<0 \text{ et } b \text{ tels que } X=aY+b.$$

Le coefficient de corrélation mesure la corrélation entre deux variables aléatoires. Son signe s'interprête de la même façon que celui de la covariance.

Quelques identités importantes:

$E(aX+bY+c)=aE(X)+bE(Y)+c$
$V(aX+bY+c)=a^2V(X)+b^2V(Y)+2abcov(X,Y)$
$cov(aX+bY,cX+dY)=acV(X)+bdV(Y)+(bc+ad)cov(X,Y)$
Si $cov(X,Y)=0$ alors $V(X+Y)=V(X-Y)=V(X)+V(Y)$
Soient X et Y deux variables aléatoires indépendantes et g_1 et g_2 deux fonctions de X et Y. Alors: $E(g_1(X).g_2(Y))=E(g_1(X)).E(g_2(Y))$.

6. Densité et espérance conditionnelles.

Définition 15: Densité conditionnelle.
Si deux variables aléatoires X et Y ne sont pas indépendantes alors lorsque l'une des deux variables, disons X, prend une valeur donnée, disons x_0, les probabilités assorties aux valeurs que peut prendre la variable Y ne sont pas les mêmes que lorsque X prend une autre valeur, disons x'_0. La distribution des valeurs de Y est donc conditionnée à celle de X. On parle de la distribution conditionnelle de Y sachant la valeur de X. Cette distribution conditionnelle a une densité conditionnelle, notée $f(y|x)$ définie par:

$$f(y|x)=\frac{f(x,y)}{f_X(x)}$$

La densité conditionnelle de Y sachant X est donc définie comme le rapport entre la densité jointe de X et de Y et la densité marginale de X.

Propriétés:
Si X et Y sont deux variables indépendantes alors $f(y|x)=f_Y(y)$ et $f(x|y)=f_X(x)$.
Etant données les définitions des densités marginales et conditionnelles il est toujours vrai que: $f(x,y)=f(y|x)f_X(x)=f(x|y)f_Y(y)$.

Définition 16: Espérance conditionnelle.
L'espérance conditionnelle est une espérance calculée en utilisant la distribution conditionnelle. Ainsi, dans le cas continu:

$$E(Y|X=x)=\int_y y.f(y|x).dy$$

Dans le cas discret cette relation devient:

$$E(Y|X=x)=\sum_y y.P(Y=y|X=x) \text{ où } P(Y=y|X=x)=\frac{P(X=x,Y=y)}{P(X=x)}.$$

Lorsque la valeur de X est connue (X=x). L'espérance conditionnelle de Y sachant X=x, $E(Y|X=x)$, est un nombre qui peut être calculé. Lorsque X est inconnu (la réalisation de X n'est pas encore intervenue ou n'a pas été observée), alors l'espérance conditionnelle de Y sachant X, notée $E(Y|X)$, est une variable aléatoire qui admet une moyenne et une variance.

Propriété: loi des espérances itérées.
Puisque $E(Y|X)$ est une variable aléatoire, on peut en calculer l'espérance. Celle-ci est donnée par:

$$E_X(E(Y|X))=E(Y)$$

Ici l'espérance est calculée en tenant compte de la loi de X (ce qui est indiqué par l'indice X), ce qui est logique puisque la valeur de $E(Y|X)$ dépend de celle prise par X. La loi des espérances itérées établit que cette espérance est égale à celle de Y. Ainsi, dans le cas continu:

$$E_X(E(Y|X)) = \int_x E(Y|X=x)\, f(x)\, dx$$

et, dans le cas discret:

$$E_X(E(Y|X)) = \sum_x E(Y|X=x)\, P(X=x)$$

La loi des espérances itérées est alors facile à démontrer:

$$\begin{aligned}
E_X(E(Y|X)) &= \int_x \left(\int_y y f(y|x)\, dy\right) f(x)\, dx \\
&= \int_x \left(\int_y y \frac{f(x,y)}{f(x)}\, dy\right) f(x)\, dx \\
&= \int_x \int_y y f(x,y)\, dx dy \\
&= \int_y y \left(\int_x f(x,y)\, dx\right) dy \\
&= \int_y y f_Y(y)\, dy \\
&= E(Y).
\end{aligned}$$

Une démonstration analogue peut être faite dans le cas discret.

Définition 17: Variance conditionnelle.

La variance conditionnelle est calculée en utilisant la distribution conditionnelle. Ainsi la variance conditionnelle de Y sachant que X=x est donnée par:

$$V(Y|X=x) = E(((Y - E(Y|X=x))^2 | X=x)$$

Comme pour l'espérance conditionnelle, la variance conditionnelle de Y sachant X=x est un nombre conditionné à la valeur de X. La variance conditionnelle de Y sachant X est donc une variable aléatoire, notée $V(Y|X)$, et à ce titre a une espérance et une variance.

Propriété: loi de décomposition de la variance.

Soient X et Y deux variables aléatoires. On peut écrire:

$$V(Y) = V_X(E(Y|X)) + E_X(V(Y|X)).$$

7. Vecteurs aléatoires.

Définition 18: Vecteur aléatoire.

Un vecteur aléatoire est un vecteur dont tous les éléments sont aléatoires.

Exemple: $X = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}$ est un vecteur aléatoire de dimension n. Chacun de ses éléments, X_i, est une variable aléatoire.

Définition 19: Espérance d'un vecteur aléatoire.

L'espérance d'un vecteur aléatoire est le vecteur des espérances de ses composantes:

$$E(X) = \begin{pmatrix} E(X_1) \\ \vdots \\ E(X_n) \end{pmatrix}$$

En revanche la variance d'un vecteur aléatoire n'est pas le vecteur des variances de ses composantes, mais une matrice carrée. Pour comprendre comment l'on obtient cette matrice nous devons au préalable rappeler quelques notions d'algèbre linéaire.

Rappels d'algèbre linéaire:

i. transposé d'une matrice.

Soit $A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{np} \end{pmatrix}$.

La matrice transposée de A, notée A' est égale à: $A'=\begin{pmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ \vdots & \vdots & \vdots & \vdots \\ a_{1p} & a_{2p} & \cdots & a_{np} \end{pmatrix}$. Transposer une matrice équivaut donc à inverser lignes et colonnes.

Il est facile d'en déduire comment transposer un vecteur. Soit $X=\begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}$ un vecteur « colonne ». Son transposé est un vecteur « ligne »: $X'=(X_1,\cdots,X_n)$.

ii. Dimension d'une matrice.
Une matrice comporte des lignes et des colonnes. La dimension d'une matrice est un couple dont le premier élément est le nombre de lignes et le second le nombre de colonnes. Ainsi une matrice de dimension (n,p) comporte n lignes et p colonnes.

iii. Produit de deux matrices.
Soients $A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{np} \end{pmatrix}$ et $B=\begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1q} \\ b_{21} & b_{22} & \cdots & b_{2q} \\ \vdots & \vdots & \vdots & \vdots \\ b_{p1} & b_{p2} & \cdots & b_{pq} \end{pmatrix}$ deux matrices de dimensions (n,p) et (p,q) respectivement. Le produit des deux matrices A*B est une matrice de dimension (n,q) égale à:

$$A*B=\begin{pmatrix} \sum_{i=1}^{p} a_{1i}b_{i1} & \sum_{i=1}^{p} a_{1i}b_{i2} & \cdots & \sum_{i=1}^{p} a_{1i}b_{iq} \\ \sum_{i=1}^{p} a_{2i}b_{i1} & \sum_{i=1}^{p} a_{2i}b_{i2} & \cdots & \sum_{i=1}^{p} a_{2i}b_{iq} \\ \vdots & \vdots & \vdots & \vdots \\ \sum_{i=1}^{p} a_{ni}b_{i1} & \sum_{i=1}^{p} a_{ni}b_{i2} & & \sum_{i=1}^{p} a_{ni}b_{iq} \end{pmatrix}$$

iv. Règles de multiplication des matrices.
Pour que deux matrices puissent être multipliées entre elles il faut qu'elles soient **compatibles**. La règle est la suivante:

> *Soient A et B deux matrices, le produit A*B existe si, et seulement si, le nombre de colonnes de A est égal au nombre de lignes de B.*

Le produit d'une matrice de dimension (n,p), par une matrice de dimension (p,q) est une matrice de dimension (n,q). On pourra se souvenir de ce résultat de la façon suivante:

(n,p)*(p,q) = (n,q)

La conséquence de ces régles de multiplication est que le produit de deux matrices n'est en général par commutatif, c'est à dire que, dans l'exemple précédent, le produit B*A n'existe pas.

v. Produit de deux vecteurs.
Un vecteur est une matrice qui ne contient qu'une seule ligne (on parle alors de vecteur ligne) ou une seule colonne (vecteur colonne). Deux vecteurs peuvent donc être multipliés, au même titre que deux matrices. Par exemple:

Soient $X=(X_{1,}\dots,X_n)'\, et\, Y=(Y_{1,}\dots,Y_p)'$ deux vecteurs colonnes (notez bien l'opérateur « transposé »). X est de dimension (n,1) et Y de dimension (p,1). Les produits suivants existent et valent:

$$X*Y'=\begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix}(Y_{1,}\dots,Y_p)=\begin{pmatrix} X_1Y_1 & X_1Y_2 & \cdots & X_1Y_p \\ X_2Y_1 & X_2Y_2 & \cdots & X_2Y_p \\ \vdots & \vdots & \vdots & \vdots \\ X_nY_1 & X_nY_2 & \cdots & X_nY_p \end{pmatrix}$$

$$X'*X=(X_{1,}\dots,X_n)\begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix}=\sum_{i=1}^{n} X_i^2$$

Le premier produit est une matrice de dimension (n,p). Ce résultat est facile à prévoir puisque X est de dimension (n,1) et Y' de dimension (1,p). Par conséquent le produit XY' est de dimension (n,1)*(1,p) = (n,p). Pour la même raison le second produit est de dimension (1,1) (autrement dit c'est un nombre), puisque (1,n)*(n,1) = (1,1).
En revanche, le produit X'Y n'existe pas puisque les dimensions (1,n) et (p,1) sont incompatibles.

vi. Transposé d'un produit de matrices.
Soient A, B deux matrices compatibles. La transposée du produit A*B est égale à:

(A*B)'=B'*A'

<u>Définition 20: Variance d'un vecteur aléatoire.</u>
Nous pouvons maintenant donner l'expression de la variance d'un vecteur aléatoire. Soit $X=(X_{1,}\dots,X_n)'$ un vecteur aléatoire. Soit $\mu=E(X)=(\mu_{1,}\dots,\mu_n)'$ l'espérance de ce vecteur. La variance de X, appelée matrice de variances-covariances de X est égale à:

$$V(X)=E[(X-\mu)(X-\mu)']$$

$$=E\begin{pmatrix} (X_1-\mu_1)^2 & (X_2-\mu_2)(X_1-\mu_1) & \dots & (X_n-\mu_n)(X_1-\mu_1) \\ (X_1-\mu_1)(X_2-\mu_2) & (X_2-\mu_2)^2 & \cdots & (X_n-\mu_n)(X_2-\mu_2) \\ \vdots & \vdots & \vdots & \vdots \\ (X_1-\mu_1)(X_n-\mu_n) & (X_2-\mu_2)(X_n-\mu_n) & \cdots & (X_n-\mu_n)^2 \end{pmatrix}$$

$$=\begin{pmatrix} \sigma_1^2 & \sigma_{12} & \dots & \sigma_{n1} \\ \sigma_{12} & \sigma_2^2 & \cdots & \sigma_{n2} \\ \vdots & \vdots & \vdots & \vdots \\ \sigma_{1n} & \sigma_{2n} & \cdots & \sigma_n^2 \end{pmatrix} \text{ où } \sigma_i^2=V(X_i) \text{ et } \sigma_{ij}=cov(X_i,X_j).$$

Propriétés: Espérance et variance d'une fonction linéaire de variables aléatoires.
Soit $X=(X_{1,}\dots,X_n)'$ un vecteur aléatoire et a un vecteur de constantes de même

dimension. Soient μ et Σ la moyenne et la matrice de variances covariances de X. On a:

$$E(a'x)=E(a_1 X_1+...+a_n X_n)=a_1 E(X_1)+...+a_n X_n=a'\mu$$

$$V(a'X)=a'V(X)a=a'\Sigma a$$

De même si A est une matrice constante de dimension compatible avec X:

$$E(AX)=AE(X)=A\mu$$

$$V(AX)=AV(X)A'=A\Sigma A'$$