# INTRODUCTION AUX VARIABLES ALEATOIRES

Ibrahima Dione (Ph.D.) & Van Son Lai (Ph.D., CFA)

Simulations Stochastiques et Applications en Finance

- Variables et Vecteurs de Variables Aléatoires

- Transformation de Variables et de Vecteurs Aléatoires

- Approximation de la Distribution Cumulée de la Loi Normale

# Variables et Vecteurs de Variables Aléatoires

### Définition

Une variable aléatoire $X$ est une fonction définie sur l'ensemble des événements simples $S$ d'une expérience aléatoire à valeur réelle

$$\begin{aligned} X : S &\longrightarrow \mathbb{R} \\ w &\longrightarrow X(w) \end{aligned}$$

tel que pour tout $x \in \mathbb{R}$, l'événement $A = \{w \in S \text{ tel que } X(w) \leq x\}$ appartient à l'algèbre des événements $\zeta$ de l'expérience aléatoire.

**Note:** Dans ce cas, la variable aléatoire $X$ est dite mesurable.

### Définition

Soit $X$ une variable aléatoire. Sa fonction de distribution de probabilité notée $F_X(x)$ est définition par

$$F_X(x) = P(A) \text{ où } A = \{w \in S \text{ tel que } X(w) \leq x\} \tag{1}$$

**Propriété(s) :** On déduit de la définition (1), les propriétés suivantes:

- ⋆ $0 \leq F_X(x) \leq 1$.
- ⋆ $F_X(x)$ est monotone, non décroissante, c'est à dire si $x_1 < x_2$ alors on a $F_X(x_1) \leq F_X(x_2)$.
- ⋆ $F_X(-\infty) = 0$ et $F_X(+\infty) = 1$.

▷ Quand $F_X(x)$ est continue, $X$ est qualifiée de variable aléatoire continue.

▷ Si $F_X(x)$ est une fonction étagère, $X$ est dite variable aléatoire discrète.

▷ Il se peut que $X$ soit un mélange de nature continue et discrète.

▷ Considérons $X$ une variable aléatoire continue. La fonction de densité de probabilité notée $f_X(x)$ est:

$$f_X(x) = \frac{dF_X(x)}{dx} \tag{2}$$

**Propriété(s) :** On a les propriétés suivantes:

- ⋆ $f_X(\cdot)$ est non négative.
- ⋆ $F_X(x) = \int_{-\infty}^{x} f_X(s)ds$.

Pour une variable discrète $X$ , on parle de la masse de probabilité c'est-à-dire la probabilité que $X$ prenne une valeur précise: $P(X = k)$.

**Exemple:**

1. On lance un dé et la variable aléatoire discrète $X$ est définie par la face qui se montre. Alors $X$ prend les valeurs $1, 2, 3, 4, 5$, et $6$, où la probabilité que $X$ prenne une de ces valeurs est

$$P(X = k) = \frac{1}{6}, \forall k \in \{1, 2, 3, 4, 5, 6\}$$

2. On définit la variable aléatoire discrète $X$ comme étant le nombre d'appels reçus à une centrale téléphonique. Alors $X$ prend les valeurs $0, 1, 2, 3, \cdots, \infty$ et obéit à la loi de Poisson dont la masse de probabilité est définie par

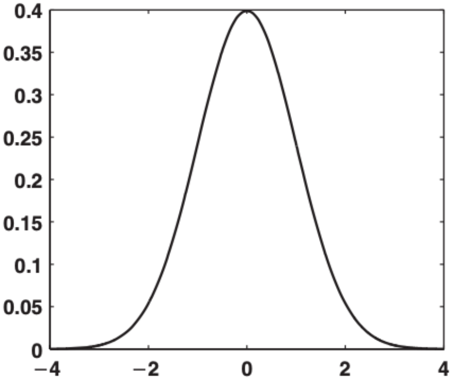
$$P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}, \; k = 0, 1, 2, 3, \cdots, \infty$$

**Exemple:**

Les fonctions de densité de probabilité continues les plus connues sont:

1. La loi normale ou gaussienne dont la fonction de densité est

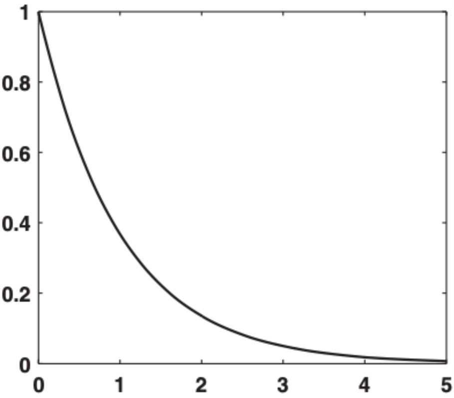
$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}},$$




Courbe de la loi normale de moyenne $\mu = 0$ et de variance $\sigma^2 = 1$.

2. La densité exponentielle dont la fonction de densité est donnée par

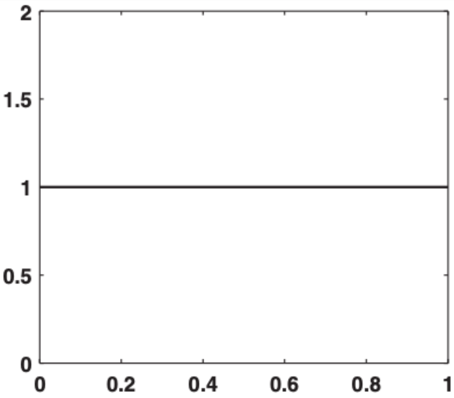
$$f_X(x) = \alpha e^{-\alpha x}, \ \forall \, x \in [0, +\infty],$$




Courbe de la loi exponentielle.

3. La densité uniforme dont la fonction de densité est donnée par

$$f_X(x) = \begin{cases} \frac{1}{b-a}, & \forall\, x \in [a, b] \\ 0 & \text{sinon} \end{cases}$$




Courbe de la loi uniforme où $a = 0$ et $b = 1$.

- Physiquement quand on veut déterminer la valeur moyenne d'une expérience aléatoire, on effectue $N$ expériences et on note les résultats obtenus $x_i, i \in \{1, 2, \cdots, N\}$.

- Si parmi les $N$ expériences, le résultat $x_i$ se produit $n_i$ fois alors la valeur moyenne des résultats est

$$\overline{X} = \sum_{i=1}^{N} \left(\frac{n_i}{N}\right) x_i$$

**Note:** La quantité $\frac{n_i}{N}$ mesure la chance ou la probabilité d'obtenir le résultat $x_i$, et la moyenne expérimentale dans ce cas s'écrit

$$\overline{X} = \sum_{i=1}^{+\infty} x_i P\left(X = x_i\right)$$

- Quand on fait augmenter $N$ indéfiniment dans le cas où $X$ est une variable aléatoire continue, alors le rapport $\frac{n_i}{N}$ tend vers la quantité infinitésimale représentée par $f_X(x_i)dx$ au point $x_i$.

- On introduit ainsi la quantité $m_X$ appelée moyenne statistique et définie

$$m_X = E[X] = \int_{-\infty}^{+\infty} x f_X(x) dx$$

**Note:** La notation $E[X]$ désigne l'opérateur valeur moyenne communément appelé espérance mathématique. Cet opérateur est linéaire, c'est à dire il vérifie l'égalité suivante

$$E[aX + bY] = aE[X] + bE[Y]$$

où $a$ et $b$ sont déterministes; $X$ et $Y$ sont des variables aléatoires.

- Si on s'intéresse maintenant à une autre variable aléatoire $Y$ liée à la variable $X$ par la transformation $Y = g(X)$, la valeur moyenne de $Y$ est

$$\overline{Y} = \sum_{i=1}^{N} \left(\frac{n_i}{N}\right) y_i = \sum_{i=1}^{N} \left(\frac{n_i}{N}\right) g(x_i)$$

- Quand $N$ tend vers l'infini, on obtient la moyenne statistique de $Y$

$$m_Y = E[Y] = \int_{-\infty}^{+\infty} g(x) f_X(x) dx$$

▷ En particulier quand $Y = X^2$, on obtient le moment de second ordre de $X$

$$E[Y] = E[X^2] = \int_{-\infty}^{+\infty} x^2 f_X(x) dx$$

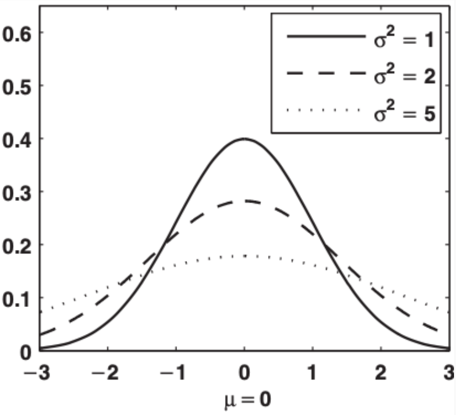
▷ On peut centrer la variable $X$ autour de sa valeur moyenne $Y = X - E[X]$.

▷ Dans ce cas $E[Y] = 0$ et le moment de second ordre de la variable $Y$ est

$$E[Y^2] = E\left[(X - E[X])^2\right] = E[X^2] - (E[X])^2$$

**Note:** Ce moment de second ordre de la variable centrée $Y$ est par définition la variance de la variable aléatoire $X$ souvent noté $\sigma_X^2$. $\sigma_X$ est appelé l'écart type ou la déviation standard de $X$.

▷ Physiquement $\sigma_X$ mesure la dispersion des résultats obtenus autour de la valeur moyenne. Quand $\sigma_X$ est faible, le graphe de la densité de probabilité de $X$ est une courbe qui se concentre autour de la moyenne. Par contre quand $\sigma_X$ est grand, cette courbe s'aplatit tout en s'élargissant.




Courbes de la distribution normale de moyenne $\mu = 0$ et pour différentes variances.

- ▷ De la même façon on introduit les moments d'ordre *n* de la variable *X*

$$M_X^n = E\left[X^n\right] = \int_{-\infty}^{+\infty} x^n f_X(x) dx$$

- ▷ Ainsi que les moments centrés d'ordre *n* notés $C_X^n$ et définis par

$$C_X^n = E\left[(X - m_X)^n\right] = \int_{-\infty}^{+\infty} (x - m_X)^n f_X(x) dx$$

**Exemple:**

1. Soit $X$ une variable binaire telle que $P(X=1)=p$ et $P(X=0)=1-p$. Alors, sa moyenne est donnée par

$$m_X = p \times 1 + (1-p) \times 0 = p$$

et son moment de second ordre est

$$M_X^2 = p \times 1^2 + (1-p) \times 0^2 = p$$

Sa variance devient

$$\begin{aligned} \sigma_X^2 &= M_X^2 - (m_X)^2 \\ &= p - p^2 \\ &= p(1-p) \end{aligned}$$

2. Soit $X$ une variable aléatoire gaussienne de densité de probabilité

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

La valeur moyenne de la variable gaussienne $X$ est donnée par

$$\begin{aligned} E[X] &= \int_{-\infty}^{+\infty} x f_X(x) dx \\ &= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{+\infty} x e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \\ &= \mu, \quad \text{(Pour les détailts de calcul, voir le livre du cours)} \end{aligned}$$

En effectuant les mêmes transformations que le calcul de la moyenne, on obtient la variance

$$\begin{aligned} E[X] &= E\left[(X - m_X)^2\right] \\ &= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{+\infty} (x-\mu)^2 e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \\ &= \sigma^2, \quad \text{(Voir le livre pour les détailts de ce calcul)} \end{aligned}$$

## Définition

Si $X$ une variable aléatoire ayant la fonction de densité $f_X(x)$. La fonction caractéristique de $X$, notée $\phi_X(\cdot)$, est définie par la transformée de Fourier de sa fonction de densité, c'est-à-dire

$$\phi_X(w) = E[e^{iwx}] = \int_{-\infty}^{+\infty} e^{iwx} f_X(x)dx$$

- La fonction caractéristique est un outil analytique important permettant d'analyser une somme de variables indépendantes.
- En prenant le développement de Taylor de la fonction caractéristique, on établit qu'elle contient tous les moments de $X$

$$\begin{aligned}\phi_X(w) &= \int_{-\infty}^{+\infty} \left[\sum_{k=0}^{\infty} \frac{(iw)^k}{k!} x^k f_X(x)dx\right] \\ &= \sum_{k=0}^{\infty} \frac{(iw)^k}{k!} \underbrace{\int_{-\infty}^{+\infty} x^k f_X(x)dx}_{M_X^k(w)}\end{aligned}$$

▷ On en déduit ainsi la formule suivante du moment d'ordre $k$ de $X$

$$M_X^k(w) = E[X^k] = \left[\frac{1}{i^k}\frac{d^k\phi_X(w)}{dw^k}\right]_{w=0}$$

**Exemple:**

Soit une variable aléatoire gaussienne de moyenne nulle et dont la fonction de densité est donnée par

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{x^2}{2\sigma^2}}$$

On peut établir que sa fonction caractéristique est

$$\phi_X(w) = e^{-\frac{\sigma^2}{2}w^2}$$

En développant la fonction caractéristique de cette loi gaussienne en série entière, on en arrive à l'expression

$$\phi_X(w) = \sum_{k=0}^{\infty}\left(\frac{\sigma^2}{2}\right)^k\frac{\left(-w^2\right)^k}{k!}$$

- Pour simplifier les écritures, nous nous bornerons à des variables de deux dimensions. Les formules de $n$ dimensions peuvent être déduites de façon similaire.

**Définition**

Soit $X$ un vecteur aléatoire à 2 dimensions, e.i. $X = (X_1, X_2)$. La fonction de distribution conjointe des variables aléatoires $X_1$ et $X_2$ est

$$F_{X_1,X_2}(x_1, x_2) = P(E), \quad \text{où } E = \{X_1 \leq x_1 \ \text{ et } \ X_2 \leq x_2\}$$

**Propriété(s) :** La fonction de distribution $F_{X_1,X_2}(x_1, x_2)$ possède les propriétés:

- $0 \leq F_{X_1,X_2}(x_1, x_2) \leq 1$,
- $F_{X_1,X_2}(+\infty, +\infty) = 1$,
- $F_{X_1,X_2}(x_1, -\infty) = 0$, $F_{X_1,X_2}(-\infty, x_2) = 0$
- Si $(x_1, x_2)$ et $(x_1', x_2')$ sont tels que $x_1 \leq x_1'$ et $x_2 \leq x_2'$, alors

$$F_{X_1,X_2}(x_1', x_2') + F_{X_1,X_2}(x_1, x_2) - F_{X_1,X_2}(x_1, x_2') - F_{X_1,X_2}(x_1', x_2) \geq 0$$

- Comme dans le cas d'une variable aléatoire, la fonction de densité conjointe $f_{X_1,X_2}(\cdot,\cdot)$ des variables $X_1$ et $X_2$ est donnée par

$$P(E) = \int\int_E f_{X_1,X_2}(u,v)dudv$$

- La fonction de densité conjointe des variables aléatoires $X_1$ et $X_2$ est

$$f_{X_1,X_2}(x_1,x_2) = \frac{\partial^2 F_{X_1,X_2}(x_1,x_2)}{\partial x_1 \partial x_2} \tag{3}$$

- Quand les variables aléatoires $X_1$ et $X_2$ sont de type discret, nous préférons parler de masse de probabilité définie par

$$P(X_1 = x_1 \text{ et } X_2 = x_2) \tag{4}$$

où $x_1$ et $x_2$ sont des valeurs discrètes que prennent $X_1$ et $X_2$.

- La fonction de répartition marginale de X est

$$F_X(x) = F_{X,Y}(x,+\infty) = \int_{-\infty}^{x}\int_{-\infty}^{+\infty} f_{X,Y}(u,v)dudv \tag{5}$$

▷ La fonction de densité de probabilité marginale de X est donnée par

$$f_X(x) = \frac{dF_X(x)}{dx} = \int_{-\infty}^{+\infty} f_{X,Y}(x,v)dv$$

**Exemple:**

Considérons la fonction de densité conjointe des deux variables aléatoires gaussiennes $X$ et $Y$

$$f_{X,Y}(x,y) = \frac{1}{2\pi\sqrt{1-\rho^2}\sigma_X\sigma_Y} e^{\frac{-1}{2\left(1-\rho^2\right)}\left(\frac{(x-m_X)^2}{\sigma_X^2}+\frac{(y-m_Y)^2}{\sigma_Y^2}-\frac{2\rho}{\sigma_X\sigma_Y}(x-m_X)(y-m_Y)\right)}$$

Si on s'intéresse à la densité marginale de $X$, on intègre la fonction de densité conjointe par rapport à $y$ sur $[-\infty, +\infty]$

$$f_X(x) = \int_{-\infty}^{+\infty} f_{X,Y}(x,y)dy = \frac{1}{\sqrt{2\pi}\sigma_X} e^{-\frac{(x-m_X)^2}{2\sigma_X^2}}$$

Ainsi la densité marginale de $X$ est la loi gaussienne de moyenne $m_X$ et de variance $\sigma_X^2$.

▷ On définit les moments conjoints d'ordre $n, m$ de $X$ et $Y$ par

$$M_{X,Y}^{n,m} = E[X^n Y^m] = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x^n y^m f_{X,Y}(x,y) dx dy \tag{6}$$

▷ Et les moments centrés d'ordre $n, m$ par

$$\begin{aligned} C_{X,Y}^{n,m} &= E\left[(X - m_X)^n (Y - m_Y)^m\right] \\ &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x - m_X)^n (y - m_Y)^m f_{X,Y}(x,y) dx dy \end{aligned} \tag{7}$$

▷ $C_{X,Y}^{1,1}$ est appelé la covariance entre $X$ et $Y$. La matrice $\Lambda$ est appelée la matrice de variance-covariance du vecteur $(X, Y)$

$$\Lambda = \begin{bmatrix} C_{X,X}^{1,1} & C_{X,Y}^{1,1} \\ C_{Y,X}^{1,1} & C_{Y,Y}^{1,1} \end{bmatrix} = E\left[\begin{pmatrix} X - m_X \\ Y - m_Y \end{pmatrix} (X - m_X, Y - m_Y)\right] \tag{8}$$

▷ Le coefficient de corrélation des variables aléatoires $X$ et $Y$ est défini

$$\rho = \frac{E\left[(X - m_X)(Y - m_Y)\right]}{\rho_X \rho_Y} = \frac{C_{X,Y}^{1,1}}{\rho_X \rho_Y}$$

- La fonction caractéristique du vecteur aléatoire $(X, Y)$ est la transformée de Fourier à 2 dimensions du vecteur $(X, Y)$. Elle est définie par

$$\phi_{X,Y}(w_1, w_2) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{i(xw_1 + yw_2)} f_{X,Y}(x, y) dxdy \tag{9}$$

- Si $X$ et $Y$ sont indépendantes alors

$$\phi_{X,Y}(w_1, w_2) = \phi_X(w_1)\phi_Y(w_2) \tag{10}$$

- La fonction caractéristique contient les moments des variables $X$ et $Y$

$$E\left[X^n Y^m\right] = \left[\frac{1}{i^{n+m}} \frac{\partial^{n+m} \phi_{X,Y}(w_1, w_2)}{\partial w_1^n \partial w_2^m}\right]_{w_1 = w_2 = 0} \tag{11}$$

# Transformation de Variables et de Vecteurs Aléatoires

- Soit $X$ une variable aléatoire de fonction de densité $f_X(x)$.
- $Y$ la variable aléatoire obtenue par la transformation $Y = g(X)$.
- La fonction de densité de probabilité $f_Y(\cdot)$ de la variable $Y$ est

$$f_Y(y) = \sum_{k=1}^{n} f_X(x) \left| \frac{dx_k}{dy} \right| = \sum_{k=1}^{n} \frac{f_X(x)}{\left| \frac{dy}{dx} \right|_{x=x_k}} \tag{12}$$

les $x_k$ étant les solutions de $y = g(x)$, $x_k = g^{-1}(y)$, $k = 1, 2, \cdots, n$.

**Exemple:**

1. Soit la variable aléatoire uniforme de fonction de densité $f_X(x)$ définie par

$$f_X(x) = \begin{cases} \frac{1}{b-a} & a \leq x \leq b \\ 0 & \text{ailleurs} \end{cases}$$

On considère la transformation $y = \alpha x + \beta$ qui n'admet qu'une seule solution donnée par $x = \frac{y-\beta}{\alpha}$. La densité de probabilité de la variable $Y$ est obtenue à partir de la formule (12) par

$$f_Y(y) = \frac{f_X(x)}{\left|\frac{dy}{dx}\right|} = \frac{f_X(\frac{y-\beta}{\alpha})}{|\alpha|} = \begin{cases} \frac{1}{|\alpha|(b-a)} & \alpha a + \beta \leq y \leq \alpha b + \beta \\ 0 & \text{ailleurs} \end{cases}$$

**Note:** On observe donc qu'une transformation affine transforme une variable aléatoire uniforme en une autre variable aléatoire uniforme.

2. Soit une variable aléatoire gaussienne de fonction de densité

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma_X} e^{-\frac{(x-m_X)^2}{2\sigma_X^2}}$$

On considère la transformation $y = \alpha x + \beta$ qui n'admet qu'une seule solution donnée par $x = \frac{y-\beta}{\alpha}$. La densité de probabilité de $Y$ est

$$f_Y(y) = \frac{f_X(x)}{\left|\frac{dy}{dx}\right|} = \frac{f_X(\frac{y-\beta}{\alpha})}{|\alpha|} = \frac{1}{\sqrt{2\pi}|\alpha|\sigma_X} e^{-\frac{(y-\beta-\alpha m_X)^2}{2\alpha^2\sigma_X^2}}$$

**Note:** La transformation affine transforme une variable gaussienne en une nouvelle variable aléatoire gaussienne dont la nouvelle moyenne et la nouvelle variance se calcule de façon directe

$$E[Y] = m_Y = \alpha m_X + \beta, \text{ et } \sigma_Y^2 = \alpha^2 \sigma_X^2$$

- Soit $\underline{X} = (X_1, \cdots, X_n)^T$ un vecteur aléatoire de dimension $n$ et de densité $F_{\underline{X}}(\underline{x})$, et $\underline{Y}$ le vecteur aléatoire obtenue par la transformation $\underline{Y} = \underline{g}(\underline{X})$

$$Y_k = g_k(\underline{X}) = g_k(X_1, X_2, \cdots, X_n), k = 1, 2, \cdots, n \tag{13}$$

- Si $\underline{Y} = \underline{g}(\underline{X})$ admet comme image inverse $m$ vecteurs $\underline{x}^1, \cdots, \underline{x}^m$ avec $\underline{y} = \underline{g}(\underline{x}^i)$, $i = 1, 2, \cdots, m$, alors le Jacobien $J$ de la transformation est défini par

$$J = det \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \cdots & \frac{\partial y_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial y_n}{\partial x_1} & \cdots & \frac{\partial y_n}{\partial x_n} \end{bmatrix} \tag{14}$$

- La fonction de densité de probabilité conjointe de $Y$ est ainsi donnée

$$f_{\underline{Y}}(\underline{y}) = \sum_{i=1}^{m} \frac{f_{\underline{X}}(\underline{x}^i)}{|J(\underline{x}^i)|} \tag{15}$$

Exemple:

Soit un vecteur aléatoire gaussien $\underline{X}$ de moyenne $m_{\underline{X}}$ et de matrice de covariance $\Lambda_{\underline{X}}$

$$\underline{X} \sim N(m_{\underline{X}}, \sqrt{\Lambda_{\underline{X}}})$$

En définissant la transformation affine suivante

$$\underline{Y} = \Sigma \underline{X} + \mu$$

où $\Sigma$ est une matrice de dimension appropriée, alors le vecteur gaussien $\underline{X}$ est transformé en un nouveau vecteur aléatoire gaussien $\underline{Y}$ dont la moyenne est donnée par

$$E[\underline{Y}] = m_{\underline{Y}} = \Sigma m_{\underline{X}} + \mu$$

et la matrice de variance-covariance est donnée par

$$\Lambda_{\underline{Y}} = \Sigma \Lambda_{\underline{X}} \Sigma^T$$

# Approximation de la Distribution Cumulée de la Loi Normale

- Soit $N(x)$ la fonction de distribution cumulée de la loi normale.

- Une approximation exacte à six décimales près de la fonction $N(x)$ est

$$N(x) = \begin{cases} 1 - n(x)\left(a_1 k + a_2 k^2 + a_3 k^3 + a_4 k^4 + a_5 k^5 +\right) & \text{si } x > 0, \\ 1 - N(-x) & \text{si } x \leq 0 \end{cases}$$

où $k = \frac{1}{1+\lambda x}$, $\lambda = 0.2316419$ et où les paramètres $a_1, a_2, a_3, a_4, a_5$ donnés

$$a_1 = 0.319381530, a_2 = -0.356563782,$$
$$a_3 = 1.781477937, a_4 = -1.821255978,$$
$$a_5 = 1.330274429$$

- La fonction de densité est obtenue à partir de la dérivée de $N$ suivante

$$n(x) = N'(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \tag{16}$$

- Notes et lectures complémentaires, voir [2, 1, 3].

[1] R. V. Hogg and A. T. Craig.
*Introduction to mathematical statistics.*
Prentice Hall, 5th edition, 1995.

[2] S. M. Ross.
*A first course in probability.*
Prentice Hall, 6th edition, 2002.

[3] W. M. Wackerly, D. D. and R. L. Scheaffer.
*Mathematical statistics with applications.*
Duxbury Press, 5th edition, 1996.