2. Variable normale

2.1. Caractérisation de la loi normale

Définition 2.1-1 *On dit q'une variable aléatoire réelle X suit la loi normale ou la loi gaussienne d'espérance μ et une variance σ^2 si elle admet pour densité de probabilité la fonction $f(x)$ définie, pour tout nombre réel x, par :*

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

On note cette loi par : $\mathcal{N}(\mu, \sigma^2)$. et dans ce cas on définit sa fonction de répartition par :

$$F_X(x) = P(X \leq x) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{x} \exp\left(-\frac{(t-\mu)^2}{2\sigma^2}\right) dt$$

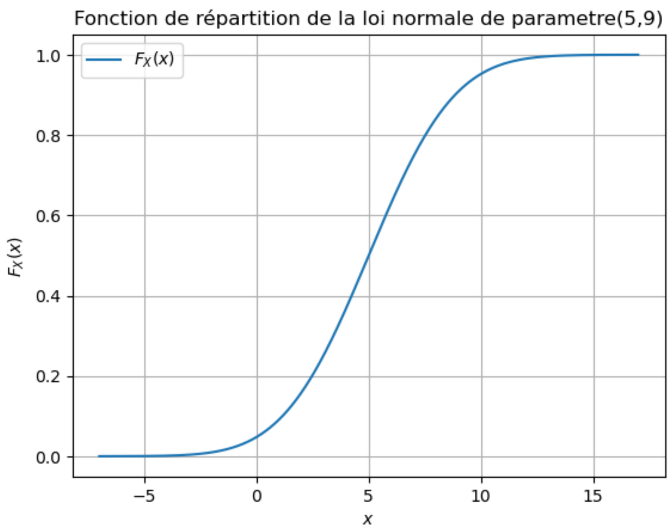

FIGURE 2.1 – Fonction de répartition de la loi normale de parametre(5,9)

Théorème 2.1-1 *Soit X une variable aléatoire réelle qui suit $\mathcal{N}(\mu, \sigma^2)$ alors :* $\mathbb{E}(X) = \mu$ *et* $\mathbb{V}(X) = \sigma^2$

Preuve : soit k un entier naturel, on considère la fonction définie par :

$$g(x) = x^k \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

g est continue sur $\mathbb{R}$. De plus,

$$\lim_{x\to\infty} x^2 g(x) = 0,$$

donc $g(x) = o\left(\frac{1}{x^2}\right)$ au voisinage de l'infini par croissance comparée. Donc g est intégrable sur $\mathbb{R}$ et par conséquent X admet une espérance et une variance.
L'espérance de X est :

$$\mathbb{E}(X) = \int_{-\infty}^{+\infty} x f(x)\, dx = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{+\infty} x \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx.$$

En effectuant le changement de variable $v = \frac{x-\mu}{\sigma}$, on obtient :

$$\mathbb{E}(X) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{+\infty} (\sigma v+\mu) \exp\left(-\frac{v^2}{2}\right) \sigma\, dv = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} (\sigma v+\mu) \exp\left(-\frac{v^2}{2}\right) dv.$$

Et on a donc :

$$\mathbb{E}(X) = \frac{1}{\sqrt{2\pi}} \left(\sigma \int_{-\infty}^{+\infty} v \exp\left(-\frac{v^2}{2}\right) dv + \mu \int_{-\infty}^{+\infty} \exp\left(-\frac{v^2}{2}\right) dv\right).$$

La première intégrale est nulle car elle est impaire, donc :

$$\mathbb{E}(X) = \frac{\mu}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \exp\left(-\frac{v^2}{2}\right) dv = \mu.$$

D'autre part : $\mathbb{E}(X^2)$, nous avons :

$$\mathbb{E}(X^2) = \int_{-\infty}^{+\infty} x^2 f(x)\, dx = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{+\infty} x^2 \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx.$$

En effectuant le même changement de variable $v = \frac{x-\mu}{\sigma}$, on obtient :

$$\mathbb{E}(X^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{+\infty} (\sigma v+\mu)^2 \exp\left(-\frac{v^2}{2}\right) \sigma\, dv = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} (\sigma^2 v^2+2\sigma\mu v+\mu^2) \exp\left(-\frac{v^2}{2}\right) dv.$$

En séparant les intégrales, on obtient :

$$\mathbb{E}(X^2) = \frac{1}{\sqrt{2\pi}} \left(\sigma^2 \int_{-\infty}^{+\infty} v^2 \exp\left(-\frac{v^2}{2}\right) dv + 2\sigma\mu \int_{-\infty}^{+\infty} v \exp\left(-\frac{v^2}{2}\right) dv + \mu^2 \int_{-\infty}^{+\infty} \exp\left(-\frac{v^2}{2}\right) dv\right)$$

La deuxième intégrale est nulle car elle est impaire, donc :

$$\mathbb{E}(X^2) = \sigma^2 \left(\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} v^2 \exp\left(-\frac{v^2}{2}\right) dv\right) + \mu^2.$$

La première intégrale est l'espérance d'une variable normale centrée réduite au carré, qui vaut 1, donc :

$$\mathbb{E}(X^2) = \sigma^2 + \mu^2.$$

Ainsi, la variance de X est :

$$\mathrm{Var}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2 = \sigma^2 + \mu^2 - \mu^2 = \sigma^2.$$

Proposition 2.1-1 *La fonction caractéristique de la variable aléatoire X qui suit la loi $\mathcal{N}(\mu, \sigma^2)$ est :*

$$\phi_X(t) = \exp(it\mu) \exp\left(-\frac{t^2\sigma^2}{2}\right).$$

Preuve : Par définition, la fonction caractéristique $\phi_X(t)$ est donnée par :

$$\phi_X(t) = \mathbb{E}(e^{itX}) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{+\infty} \exp(itx) \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx.$$

En effectuant le changement de variable $s = \frac{x-\mu}{\sqrt{2}\sigma}$, on obtient :

$$\begin{aligned}
\phi_X(t) &= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{+\infty} \exp\left(-\frac{1}{2}s^2\right) \exp\left(it(s\sigma+\mu)\right) \sigma \, ds \\
&= \exp(it\mu) \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \exp\left(-\frac{1}{2}(s-it\sigma)^2\right) \exp\left(-\frac{1}{2}(t\sigma)^2\right) ds \\
&= \exp(it\mu) \exp\left(-\frac{1}{2}(t\sigma)^2\right) \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \exp\left(-\frac{1}{2}(s-it\sigma)^2\right) ds \\
&= \exp(it\mu) \exp\left(-\frac{1}{2}(t\sigma)^2\right).
\end{aligned}$$

Proposition 2.1-2 *La fonction génératrice des moments de la variable aléatoire X qui suit la loi $\mathcal{N}(\mu, \sigma^2)$ est :*

$$g_X(t) = \exp\left(t\mu + \frac{t^2\sigma^2}{2}\right).$$

Preuve : Par définition, la fonction génératrice $g_X(t)$ est donnée par :

$$g_X(t) = \mathbb{E}(e^{tX}) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{+\infty} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \exp(tx) \, dx$$

$$\begin{aligned}
&= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{+\infty} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2} + tx\right) dx \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{+\infty} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2} + t(x-\mu+\mu)\right) dx \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{+\infty} \exp(t\mu) \exp\left(-\frac{(x-\mu)^2}{2\sigma^2} + t(x-\mu)\right) dx \\
&= \exp(t\mu) \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{+\infty} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2} + t(x-\mu)\right) dx.
\end{aligned}$$

Avec un simple changement de variable $s = x - \mu$, on aura :

$$g_X(t) = \exp(t\mu) \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{+\infty} \exp\left(-\frac{s^2}{2\sigma^2} + ts\right) ds$$

$$= \exp(t\mu) \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{+\infty} \exp\left(-\frac{1}{2\sigma^2}\left(s^2 - 2\sigma^2 ts + (t\sigma^2)^2 - (t\sigma^2)^2\right)\right) ds$$

$$= \exp(t\mu) \exp\left(\frac{t^2\sigma^2}{2}\right) \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{+\infty} \exp\left(-\frac{1}{2\sigma^2}\left(s - t\sigma^2\right)^2\right) ds$$

$$= \exp(t\mu)\exp\left(\frac{t^2\sigma^2}{2}\right) = \exp\left(t\mu + \frac{t^2\sigma^2}{2}\right).$$

Corollaire 2.1-1 *Soient $X_1, \ldots, X_n$ des variables aléatoires normales mutuellement indépendantes, avec des espérances respectives μ_i et des variances σ_i^2. Alors la somme $X = \sum_{i=1}^n X_i$ suit une loi normale avec une espérance égale à la somme des espérances $\sum_{i=1}^n \mu_i$ et une variance égale à la somme des variances $\sum_{i=1}^n \sigma_i^2$.*

Théorème 2.1-2 *Soient X et Y deux vecteurs gaussiens avec $X \sim \mathcal{N}(m_X, \sigma_X^2)$ et $Y \sim \mathcal{N}(m_Y, \sigma_Y^2)$. Alors, X est indépendant de Y si et seulement si cov$(X, Y) = 0$.*

2.2. Variable aléatoire normale standard

Définition 2.2-1 *On dit qu'une variable aléatoire X suit la loi normale standard si elle admet une densité de probabilité, la fonction f, définie sur $\mathbb{R}$, par :*

$$f(x) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{x^2}{2}\right)$$

.

et dans ce cas on définit sa fonction de répartition par :

$$F_X(x) = P(X \leq x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}\exp\left(-\frac{t^2}{2}\right) dt$$

Proposition 2.2-1 *Soit X qui suit la loi normale de paramètres μ et σ^2. Alors la variable*

$$Z = \frac{X-\mu}{\sigma}$$

suit la loi normale standard

Preuve : Pour $Z = \frac{X-\mu}{\sigma}$, effectuons le changement de variable $x = \frac{t-\mu}{\sigma}$. Ainsi, la fonction de répartition de Z est donnée par :

$$F_Z(z) = P(Z \leq z) = P\left(\frac{X-\mu}{\sigma} \leq z\right) = P(X \leq z\sigma + \mu) = F_X(z\sigma + \mu)$$

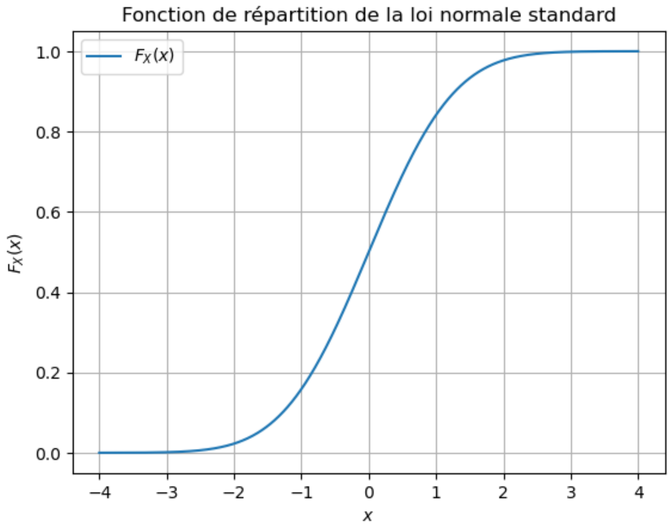

FIGURE 2.2 – Fonction de répartition de la loi normale standard

En utilisant le changement de variable $x = \frac{t-\mu}{\sigma}$, nous avons $dt = \sigma\, dx$. Donc,

$$F_Z(z) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{z\sigma+\mu} \exp\left(-\frac{(t-\mu)^2}{2\sigma^2}\right) dt$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} \exp\left(-\frac{x^2}{2}\right) dx$$

Ainsi, $Z = \frac{X-\mu}{\sigma}$ suit la loi normale standard $\mathcal{N}(0,1)$, car sa fonction de répartition est celle de $\mathcal{N}(0,1)$.

Proposition 2.2-2 *soit X une variable normale standard alors :*

1. ***Espérance :***

$$\mathbb{E}(X) = 0$$

2. ***Variance :***

$$Var(X) = 1$$

3. ***Fonction caractéristique :***

$$\phi_X(t) = \mathbb{E}\left(e^{itX}\right) = \exp\left(-\frac{t^2}{2}\right)$$

4. ***Fonction génératrice :***

$$g_X(t) = \mathbb{E}\left(e^{tX}\right) = \exp\left(\frac{t^2}{2}\right)$$

Théorème 2.2-1 ***Théorème de Box-Muller***
Soient U_1 et U_2 deux variables aléatoires indépendantes uniformément distribuées sur l'intervalle $[0,1]$. On définit les variables aléatoires Z_0 et Z_1 par :

$$Z_0 = \sqrt{-2\ln(U_1)}\cos(2\pi U_2)$$

$$Z_1 = \sqrt{-2\ln(U_1)}\sin(2\pi U_2)$$

Alors Z_0 et Z_1 sont deux variables aléatoires indépendantes suivant une loi normale centrée réduite, c'est-à-dire $Z_0, Z_1 \sim \mathcal{N}(0,1)$.

Définition 2.2-2 *Une variable aléatoire X suit la loi Gamma avec les paramètres a et b, noté $X \sim$ Gamma(a,b), si sa densité de probabilité est donnée par :*

$$f_X(x) = \frac{b^a}{\Gamma(a)} x^{a-1} e^{-bx}, \quad x > 0,$$

où $a > 0$ est le paramètre de forme, $b > 0$ est le paramètre de taux, et $\Gamma(a)$ est la fonction Gamma définie par :

$$\Gamma(a) = \int_0^\infty t^{a-1} e^{-t}\, dt.$$

Lemme 2.2-1 *Si $X \sim$ Gamma(a,b) et $Y \sim$ Gamma(c,b) avec X et Y sont indépendantes alors $X + Y \sim$ Gamma$(a+c,b)$*

Corollaire 2.2-1 *Soit $X_1, X_2, \ldots, X_k$ un ensemble de variables aléatoires indépendantes et identiquement distribuées, où chaque X_i suit une loi gamma de paramètres (a,b) :*

$$X_i \sim \Gamma(a,b) \quad \text{pour} \quad i = 1, 2, \ldots, k$$

$$X = \sum_{i=1}^{k} X_i$$

Alors

$$X \sim \Gamma(ka, b)$$

Lemme 2.2-2 *Si X suit la loi normale standard alors $X^2 \sim Gamma(\frac{1}{2}, \frac{1}{2})$*

Preuve : Soit $X \sim \mathcal{N}(0, 1)$, La densité de probabilité de X est donnée par :

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

Nous voulons calculer $\mathbb{E}(h(X^2))$. Par définition :

$$\mathbb{E}(h(X^2)) = \int_{-\infty}^{+\infty} h(x^2) f_X(x)\, dx$$

Nous faisons le changement de variable $z = x^2$. Donc $x = \sqrt{z}$ ou $x = -\sqrt{z}$, et la transformation jacobienne nous donne :

$$dx = \frac{1}{2\sqrt{z}}\, dz$$

Nous devons maintenant réécrire l'intégrale en termes de z. En utilisant le changement de variable $z = x^2$, l'intégrale devient :

$$\mathbb{E}(h(X^2)) = \int_{-\infty}^{+\infty} h(x^2) \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\, dx$$

En transformant la variable x en z, nous avons deux contributions (pour $x = \sqrt{z}$ et $x = -\sqrt{z}$), donc :

$$\mathbb{E}(h(X^2)) = \int_{0}^{+\infty} h(z) \left(\frac{1}{\sqrt{2\pi}} e^{-\frac{z}{2}} \frac{1}{\sqrt{z}} + \frac{1}{\sqrt{2\pi}} e^{-\frac{z}{2}} \frac{1}{\sqrt{z}} \right) dz$$

Simplifions cette expression :

$$\mathbb{E}(h(X^2)) = \int_{0}^{+\infty} h(z) \cdot 2 \cdot \frac{1}{\sqrt{2\pi z}} e^{-\frac{z}{2}}\, dz$$

Nous pouvons identifier la densité de probabilité de $Y = X^2$. En effet, la densité de probabilité $f_Y(y)$ de Y est donnée par :

$$f_Y(y) = \frac{1}{\sqrt{2\pi y}} e^{-\frac{y}{2}} \quad \text{pour} \quad y \geq 0$$

avec

$$\Gamma(\frac{1}{2}) = \int_0^\infty t^{-\frac{1}{2}} e^{-t} \, dt = \sqrt{\pi}$$

Donc $X^2 \sim \text{Gamma}(\frac{1}{2}, \frac{1}{2})$

Définition 2.2-3 *On dit qu'une variable aléatoire X suit la loi Khi-deux à k degrés de liberté, notée $\chi^2(k)$si $X \sim Gamma(\frac{k}{2}, \frac{1}{2})$*

Théorème 2.2-2 *Soient $Z_1, Z_2, \ldots, Z_k$ une suite de variables aléatoires indépendantes de même loi $\mathcal{N}(0,1)$. Alors la variable aléatoire $\sum_{i=1}^{k} Z_i^2$ suit une loi du khi-deux $\chi^2(k)$.*

Preuve : Soit $\{Z_i\}_{i=1}^k$ un ensemble de variables aléatoires indépendantes et identiquement distribuées, où chaque Z_i suit une loi normale standard :

$$Z_i \sim \mathcal{N}(0, 1) \quad \text{pour} \quad i = 1, 2, \ldots, k$$

D'après le lemme 2.2-2 Pour tout i entre 1 et k, on a :

$$Z_i^2 \sim \Gamma\left(\frac{1}{2}, \frac{1}{2}\right)$$

On pose

$$Y = \sum_{i=1}^{k} Z_i^2$$

Alors selon le corollaire 2.2-1 :

$$Y \sim \Gamma\left(\frac{k}{2}, \frac{1}{2}\right)$$

. Donc, Y suit une loi de chi-deux avec k degrés de liberté :

$$Y \sim \chi^2(k)$$

Théorème 2.2-3 *Soient Z et Qdeux variables aléatoires indépendantes telles que Z suit une loi normale $\mathcal{N}(0,1)$ et Q suit une loi du $\chi^2(\nu)$ avec ν degrés de liberté : Alors la variable aléatoire $T = \frac{Z}{\sqrt{\frac{Q}{\nu}}}$ suit une loi appelée loi de Student à ν degrés de liberté, notée $St(\nu)$; sa densité est donnée par :*

$$f(x) = \frac{1}{\sqrt{\nu\pi}} \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \left(1 + \frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}}.$$

Théorème 2.2-4 *Si on a deux variables indépendantes Q_1 et Q_2 telles que Q_1 suit une loi du $\chi^2(\nu_1)$ avec ν_1 degrés de liberté et Q_2 suit une loi du $\chi^2(\nu_2)$ avec ν_2 degrés de liberté, alors la variable aléatoire*

$$F = \frac{Q_1/\nu_1}{Q_2/\nu_2}$$

suit une loi de Fisher-Snedecor à ν_1 et ν_2 degrés de liberté, notée $F(\nu_1, \nu_2)$, telle que sa densité est donnée par :

$$f(x) = \begin{cases} \frac{\left(\frac{\nu_1}{\nu_2}\right)^{\frac{\nu_1}{2}} \left(\frac{\nu_1+\nu_2}{2}\right)\Gamma\left(\frac{\nu_1+\nu_2}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right)\Gamma\left(\frac{\nu_2}{2}\right)} x^{\frac{\nu_1}{2}-1} \left(1 + \frac{\nu_1}{\nu_2}x\right)^{-\frac{\nu_1+\nu_2}{2}}, & si\ x > 0, \\ 0, & sinon. \end{cases}$$

Théorème 2.2-5 ***Théorème centrale limite***

Soit $\{X_i\}_{i=1}^n$ une suite de variables aléatoires indépendantes et identiquement distribuées (i.i.d.) avec espérance $\mu = \mathbb{E}[X_i]$ et variance finie $\sigma^2 = \mathrm{Var}(X_i)$.Soit $S_n = \sum_{i=1}^n X_i$. Alors la variable aléatoire

$$Z_n = \frac{S_n - n\mu}{\sigma\sqrt{n}}$$

converge vers une loi normale standard $\mathcal{N}(0,1)$ lorsque $n \to \infty$:

$$\frac{S_n - n\mu}{\sigma\sqrt{n}} \to \mathcal{N}(0,1) \quad quand\ n \to \infty,$$

Preuve :L'idée de la preuve consiste à montrer que la fonction génératrice des moments de Z_n tend vers la fonction génératrice des moments d'une variable aléatoire suivant une loi $\mathcal{N}(0,1)$. Nous avons vu dans le chapitre précédent que la fonction génératrice des moments d'une variable aléatoire suivant une loi $\mathcal{N}(0,1)$ correspond à $\exp\left(\frac{t^2}{2}\right)$.

Notons $Y_i = \frac{X_i-\mu}{\sigma}$; les Y_i sont donc des variables aléatoires indépendantes et identiquement distribuées, d'espérance nulle et de variance 1. En notant $g(t)$ leur fonction génératrice des moments commune, nous avons :

$$g_{Z_n}(t) = \left(g\left(t\sqrt{\frac{1}{n}}\right)\right)^n$$

Pour calculer la limite de $g_{Z_n}(t)$, nous utilisons un développement de série de Taylor à l'ordre 2 de $g\left(t\sqrt{\frac{1}{n}}\right)$. Nous avons : $g(0)=1$, $g'(0)=0$ et $g''(0)=1$. Ceci nous donne :

$$g_{Z_n}(t) = \left(g(0) + t\sqrt{\frac{1}{n}} \cdot \frac{g'(0)}{1!} + \frac{t^2}{2n} \cdot \frac{g''(0)}{2!} + R\left(t\sqrt{\frac{1}{n}}\right)\right)^n$$

où $R\left(t\sqrt{\frac{1}{n}}\right)$ désigne le reste du développement avec $\lim_{n\to\infty} nR\left(t\sqrt{\frac{1}{n}}\right) = 0$. Notons $r_n = R\left(t\sqrt{\frac{1}{n}}\right)$. En remplaçant les valeurs de $g(0)$, $g'(0)$ et $g''(0)$, nous obtenons :

$$g_{Z_n}(t) = \left(1 + \frac{t^2}{2n} + r_n\right)^n$$

$$= \left(1 + \frac{t^2 + 2nr_n}{2n}\right)^n$$

$$= \left(1 + \frac{1}{n} \cdot \frac{t^2}{2} + nr_n\right)^n$$

Comme $\lim_{n\to\infty}\left(1+\frac{1}{x}\right)^x = e$, nous avons :

$$g_{Z_n}(t) = \left(1 + \frac{1}{n}\left(\frac{t^2}{2} + nr_n\right)\right)^n$$

$$\lim_{n\to\infty}\left(1 + \frac{1}{n}\left(\frac{t^2}{2} + nr_n\right)\right)^n = \exp\left(\frac{t^2}{2}\right)$$

car $nr_n \to 0$. Finalement, on voit que $g_{Z_n}(t)$ tend vers la fonction génératrice des moments d'une variable aléatoire suivant la loi $\mathcal{N}(0,1)$. Alors on peut conclure que :

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(\frac{X_i-\mu}{\sigma}\right)\xrightarrow{L}\mathcal{N}(0,1)$$