Donc un intervalle de confiance de niveau 0,95 pour θ est donné par $\left[T - \frac{1}{\sqrt{0{,}2n}}; T + \frac{1}{\sqrt{0{,}2n}}\right]$. Mais cette inégalité est une approximation grossière. En fait, si n est grand, alors $(T-\theta)\sqrt{n/\theta(1-\theta)}$ suit approximativement la loi $\mathcal{N}(0,1)$ sous $\mathbb{P}_\theta$. Un intervalle de confiance de niveau 0,95 est donné par $[T - 0,98/\sqrt{n}, T + 0,98/\sqrt{n}]$, qui est plus petit donc meilleur que le précédent lorsque n est grand (en fait $n \geq 30$).

2. Soit un n-échantillon de loi $\mathcal{N}(\theta, \sigma^2)$ avec σ^2 connu. La loi de $\bar{X}_n - \theta$ sous $\mathbb{P}_\theta$ est la loi $\mathcal{N}(0, \sigma^2/n)$, donc $\sqrt{n}(\bar{X}_n - \theta)/\sigma$ suit la loi $\mathcal{N}(0,1)$. Fixons un niveau $\alpha = 0,95$. On lit sur la table de la loi gaussienne que

$$\mathbb{P}_\theta(|\sqrt{n}(\bar{X}_n - \theta)/\sigma| > 1,96) = 0,05$$

et donc un intervalle de confiance de niveau 0,95 pour la moyenne est $\left[\bar{X}_n - 1,96\frac{\sigma}{\sqrt{n}}, \bar{X}_n + 1,96\frac{\sigma}{\sqrt{n}}\right]$.

Rappelons que l'estimateur sans biais de la variance d'un $n-$échantillon $X_1, \ldots, X_n$ est définie par

$$s_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2. \tag{13.1}$$

Proposition 13.2.1. *Soit $(X_1, \ldots, X_n)$ un n-échantillon de loi $\mathcal{N}(\mu, \sigma^2)$. Alors $\bar{X}_n$ et s_n^2 sont indépendantes. La variable $\bar{X}_n$ suit la loi $\mathcal{N}(\mu, \sigma^2/n)$ et $(n-1)s_n^2/\sigma^2$ suit la loi χ^2_{n-1}.*

Démonstration. On sait déjà que $\bar{X}_n$ suit la loi $\mathcal{N}(\mu, \sigma^2/n)$. Il suffit de montrer le résultat lorsque $\mu = 0$ et $\sigma^2 = 1$. Notons $f = \frac{1}{\sqrt{n}}(1, \ldots, 1)$ et F le sous-espace vectoriel de $\mathbb{R}^n$ engendré par f. D'après le théorème de Cochran, $p_F X$ et $p_{F^\perp} X$ sont indépendantes, avec $\|p_F X\|_2^2$ suit la loi χ^2_1 et $\|p_{F^\perp} X\|_2^2$ suit la loi χ^2_{n-1}. On a $p_F X = \langle X, f\rangle f = \bar{X}_n(1, \ldots, 1)$ et

$$\|p_{F^\perp} X\|_2^2 = \|X - p_F X\|_2^2 = \sum_{k=1}^{n} (X_k - \bar{X}_n)^2 = (n-1)s_n^2,$$

ce qui donne les résultats voulus.

□

Définition. Soient X et Y deux variables indépendantes. Si X suit la loi $\mathcal{N}(0,1)$ et Y suit la loi χ^2_n, alors on dit que $T = \frac{X}{\sqrt{Y/n}} = X\sqrt{n/Y}$ suit la loi de Student à n degrés de liberté et on la note t_n.

Proposition 13.2.2. *Soit $(X_1, \ldots, X_n)$ un n-échantillon de loi $\mathcal{N}(\mu, \sigma^2)$, où μ et σ^2 sont inconnues. Alors $\sqrt{n}(\bar{X}_n - \mu)/S$ suit la loi de Student t_{n-1} à $n-1$ degrés de liberté.*