Marcel Délèze
Edition 2017

Thème : Statistiques II, § 2 Variables aléatoires indépendantes
Lien vers les énoncés des exercices:
https://www.deleze.name/marcel/sec2/applmaths/csud/statistique_2/2-stat_II.pdf

Corrigé de l'exercice 2.3 - 3

$$V(X_1) = p(1-p)$$

$$V(S) = V(X_1) + ... + V(X_n) = p(1-p) + ... + p(1-p) = n p (1-p)$$

$$\sigma_S = \sqrt{n p (1-p)}$$

Corrigé de l'exercice 2-6-1

La distribution de la variable aléatoire "obtenir un six en lançant le dé" est

```
six = BernoulliDistribution[1/6];
```
distribution de Bernoulli

Son espérance mathématique est

```
μ = 1/6;
```

et son écart-type théorique est

```
σ = Sqrt[μ (1 - μ)]; N[σ]
```
valeur

```
0.372678
```

Organisons une simulation constituée de k échantillons de taille n:

```
n = 50; k = 40;
simul = Table[echant = RandomInteger[six, n];
    m = Apply[Plus, echant]/n;
    s = Sqrt[m (1 - m)];
    N[{m, s}], {k}]
```
table — entier aléatoire — valeur numérique

```
{{0.28, 0.448999}, {0.18, 0.384187}, {0.18, 0.384187}, {0.18, 0.384187},
 {0.18, 0.384187}, {0.2, 0.4}, {0.2, 0.4}, {0.2, 0.4}, {0.18, 0.384187}, {0.2, 0.4},
 {0.2, 0.4}, {0.2, 0.4}, {0.22, 0.414246}, {0.1, 0.3}, {0.18, 0.384187}, {0.04, 0.195959},
 {0.2, 0.4}, {0.12, 0.324962}, {0.12, 0.324962}, {0.2, 0.4}, {0.16, 0.366606},
 {0.12, 0.324962}, {0.18, 0.384187}, {0.2, 0.4}, {0.32, 0.466476}, {0.14, 0.346987},
 {0.12, 0.324962}, {0.14, 0.346987}, {0.2, 0.4}, {0.2, 0.4}, {0.06, 0.237487},
 {0.1, 0.3}, {0.16, 0.366606}, {0.1, 0.3}, {0.12, 0.324962}, {0.12, 0.324962},
 {0.22, 0.414246}, {0.18, 0.384187}, {0.12, 0.324962}, {0.18, 0.384187}}
```

Il faut maintenant interpréter m="moyenne de n tirages" comme une variable aléatoire qui, lors d'expériences successives, prend les valeurs empiriques suivantes:

```
{em, es} = Transpose[simul];
```
transposée

```
em
```

```
{0.28, 0.18, 0.18, 0.18, 0.18, 0.2, 0.2, 0.2, 0.18, 0.2, 0.2, 0.2, 0.22, 0.1,
 0.18, 0.04, 0.2, 0.12, 0.12, 0.2, 0.16, 0.12, 0.18, 0.2, 0.32, 0.14, 0.12,
 0.14, 0.2, 0.2, 0.06, 0.1, 0.16, 0.1, 0.12, 0.12, 0.22, 0.18, 0.12, 0.18}
```

L'espérance mathématique et l'écart-type de m="moyenne de n tirages" sont:

$$N\left[\left\{\mu, \frac{\sigma}{\sqrt{n}}\right\}\right]$$
valeur numérique

```
{0.166667, 0.0527046}
```

On observe un accord entre la théorie (théorème de la limite centrale) et l'expérience (simulation) dans le sens que, par exemple,
"m appartient à l'intervalle

$$N\left[\left\{\mu - \frac{\sigma}{\sqrt{n}}, \mu + \frac{\sigma}{\sqrt{n}}\right\}, 3\right]$$
valeur numérique

```
{0.114, 0.219}
```

avec une probabilité de 68.3%".

Par ailleurs, comparons les valeurs théoriques $\{\mu, \frac{\sigma}{\sqrt{n}}\}$ avec les valeurs empiriques correspondantes de la variable aléatoire m="moyenne de n tirages":

$$mm = \frac{\text{Apply}[\text{Plus}, em]}{k};$$

$$sm = \sqrt{\frac{\text{Apply}\left[\text{Plus}, (em - mm)^2\right]}{k}};$$

```
N[{mm, sm}]
```
valeur numérique

```
{0.1675, 0.053233}
```

En comparant les valeurs théoriques et expérimentales, il faut garder à l'esprit que l'on considère ici un échantillon de seulement k valeurs de m.

Organisons une simulation constituée de k échantillons d'une plus grande taille n:

```
n = 5000; k = 40;
simul = Table[echant = RandomInteger[six, n];
        ⌊table               ⌊entier aléatoire
          m = Apply[Plus, echant] / n;
          s = √(m (1 - m)) ;
          N[{m, s}], {k}]
          ⌊valeur numérique
```

```
{{0.1676, 0.373511}, {0.1702, 0.375808}, {0.166, 0.372081}, {0.165, 0.371181},
 {0.1658, 0.371901}, {0.1606, 0.367162}, {0.1658, 0.371901}, {0.1694, 0.375105},
 {0.17, 0.375633}, {0.1622, 0.368634}, {0.1684, 0.374221}, {0.165, 0.371181},
 {0.1706, 0.376159}, {0.1598, 0.36642}, {0.1752, 0.380138}, {0.174, 0.379109},
 {0.1698, 0.375457}, {0.1672, 0.373154}, {0.1714, 0.376858}, {0.1724, 0.377728},
 {0.1574, 0.364177}, {0.1698, 0.375457}, {0.1634, 0.36973}, {0.1634, 0.36973},
 {0.1616, 0.368083}, {0.1674, 0.373333}, {0.164, 0.370276}, {0.1626, 0.369},
 {0.1704, 0.375984}, {0.146, 0.353106}, {0.1654, 0.371541}, {0.1644, 0.370638},
 {0.1662, 0.37226}, {0.1634, 0.36973}, {0.1628, 0.369183}, {0.1648, 0.371},
 {0.169, 0.374752}, {0.1604, 0.366977}, {0.1732, 0.37842}, {0.1682, 0.374044}}
```

Il faut interpréter m="moyenne de n tirages" comme une variable aléatoire qui, lors d'expériences successives, prend les valeurs empiriques suivantes:

```
{em, es} = Transpose[simul];
           ⌊transposée
```

```
em
```

```
{0.1676, 0.1702, 0.166, 0.165, 0.1658, 0.1606, 0.1658, 0.1694, 0.17, 0.1622,
 0.1684, 0.165, 0.1706, 0.1598, 0.1752, 0.174, 0.1698, 0.1672, 0.1714, 0.1724,
 0.1574, 0.1698, 0.1634, 0.1634, 0.1616, 0.1674, 0.164, 0.1626, 0.1704, 0.146,
 0.1654, 0.1644, 0.1662, 0.1634, 0.1628, 0.1648, 0.169, 0.1604, 0.1732, 0.1682}
```

L'espérance mathématique et l'écart-type de m="moyenne de n tirages" sont:

$$N\left[\left\{\mu, \frac{\sigma}{\sqrt{n}}\right\}\right]$$

⌊valeur numérique

```
{0.166667, 0.00527046}
```

On observe un accord entre la théorie (théorème de la limite centrale) et l'expérience (simulation) dans le sens que, par exemple,
"m appartient à l'intervalle

$$N\left[\left\{\mu - \frac{\sigma}{\sqrt{n}}, \mu + \frac{\sigma}{\sqrt{n}}\right\}, 3\right]$$

⌊valeur numérique

```
{0.161, 0.172}
```

avec une probabilité de 68.3%".

Par ailleurs, comparons les valeurs théoriques $\{\mu, \frac{\sigma}{\sqrt{n}}\}$ avec les valeurs empiriques correspondantes de la variable aléatoire m="moyenne de n tirages":

```
mm = Apply[Plus, em] / k;
sm = Sqrt[Apply[Plus, (em - mm)^2] / k];
N[{mm, sm}]
```

valeur numérique

{0.166005, 0.00517213}

En comparant les valeurs théoriques et expérimentales, il faut garder à l'esprit que l'on considère ici un échantillon de seulement k valeurs de m.

Comparons maintenant les deux simulations. En passant d'un échantillon de taille 50 à 5000, l'écart-type de m est divisé par environ $\sqrt{100}$ =10.
On constatera que cette relation entre les deux valeurs de sm est vérifiée qualitativement; mais ici la faible taille k de l'échantillon des valeurs de m peut produire des écarts aléatoires.