

## conformite\_grand\_echantillon

October 1, 2024

```
[23]: import pandas as pd
import numpy as np
import random
random.seed(10)

moy_cible = 58.00
n = 100
alpha = 0.05

data = np.random.normal(loc=58.2, scale=0.2, size=n)
serie = pd.Series(data)

moy_obs = serie.mean()
display(moy_obs)
```

58.207151045410285

```
[3]: # échantillon sup à 30 obs. d'après tcl, le résultat du test sera valide.
# on peut vérifier avec test adéquatation tq Shapiro-Wilk.
from scipy.stats import shapiro
shapiro(serie)
```

```
[3]: ShapiroResult(statistic=0.9929735660552979, pvalue=0.8864582180976868)
```

Le test de Shapiro-Wilk confirme la normalité de l'échantillon.

```
[4]: # statistique de test.
std = np.std(serie, ddof = 1)
t = (moy_obs - moy_cible) / (std / np.sqrt(n))
display(t)
```

9.036274695663673

```
[5]: # quantile loi de student.
from scipy import stats
quantile = stats.t.ppf(1-alpha/2, n-1)
quantile
```

[5]: 1.9842169515086827

Avec un risque d'erreur de 5%, on peut rejeter l'hypothèse nulle d'égalité des moyennes. En effet, d'après les relevés effectués, nous avons suffisamment d'évidence pour conclure à la défaillance machine.

```
[6]: # pvalue
from scipy import stats
pvalue = 2*(1-stats.t.cdf(t, n-1))
pvalue
```

[6]: 1.3766765505351941e-14

La probabilité d'observer notre statistique de test est largement inférieur au seuil de singificativité. On peut donc conclure à la défaillance machine.

```
[7]: # IC_95
ic_sup = moy_obs + quantile * std/np.sqrt(n)
ic_inf = moy_obs - quantile * std/np.sqrt(n)

display(ic_inf, moy_cible, ic_sup)
```

58.171463255736086

58.0

58.26795148964892

Effectivement, la moyenne cible n'appartient pas à l'intervalle de confiance de notre moyenne observée. Il est donc peu probable qu'on conclut à tort à la défaillance machine.

```
[15]: # test de student à un échantillon
from scipy.stats import ttest_1samp as test_conformite

results = test_conformite(serie, moy_cible, alternative='two-sided')
display(results.statistic, results.pvalue)
```

9.036274695663673

1.3814898377141245e-14

```
[18]: results.confidence_interval(confidence_level=0.95)
```

[18]: ConfidenceInterval(low=58.171463255736086, high=58.26795148964892)

Nos observations précédentes étaient bonnes. Avec un risque d'erreur de 5%, on peut rejeter l'hypothèse nulle. On conclut donc à la défaillance machine.