Faculté de Droit & des Sciences Économiques — Université de Limoges

**ANNÉE UNIVERSITAIRE 2023-2024**

**Session 1** **Semestre 4**

## Licence Economie-Gestion – 2e année

**Matière :** Statistiques et probabilités – Éléments de corrections **Durée :** 2 heures
**Enseignant :** Vincent Jalby

### Problème

Selon l'observatoire communal de l'habitat de la ville de Limoges, le prix moyen du m² pour la vente d'une maison dans la ville de Limoges s'élèvait à 1945 € **en 2022** (prix moyen basé sur les 744 transactions de l'année 2022). L'écart-type est estimé à 415 €. (*Source : Observatoire communal de l'habitat de la Ville de Limoges. Marché de l'année 2022*)

Un site internet immobilier souhaite actualiser ces données pour le **premier trimestre 2024**. Pour cela, il collecte auprès de son réseau d'agences immobilières partenaires le prix du m² sur un échantillon de transactions effectuées en début d'année 2024. (*Les résultats suivants sont basés sur les données du site* `www.meilleursagents.com`)

**Partie I** (20 min, 3 points)

Les données (prix du m²) récoltées sur l'échantillon sont résumées dans la sortie SPSS suivante :

**Descriptive Statistics**

| | N | Minimum | Maximum | Mean | | Std. Deviation |
|---|---|---|---|---|---|---|
| | Statistic | Statistic | Statistic | Statistic | Std. Error | Statistic |
| m2 | 71 | 889 | 2601 | 1848.72 | 47.469 | 399.979 |
| Valid N (listwise) | 71 | | | | | |

**1)** On étudie sur la population ($\Omega$) des transaction immobilières sur la ville de Limoges la variable $X$ représentant le prix de vente au m² (en 2024). On note $\mu = \mathrm{E}(X)$ la prix moyen au m² et $\sigma = \sigma_X$ son écart-type, les deux étant inconnus (en 2024). Pour cela, on utilise un 71-échantillon $(X_1, \dots, X_{71})$ correspondant à 71 variables aléatoires indépendantes et de même loi que $X$.

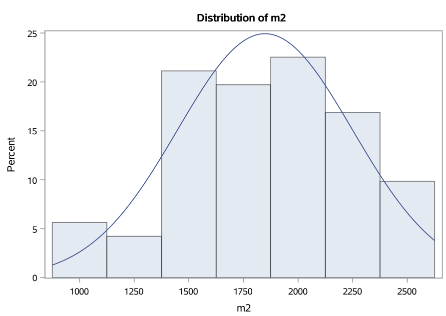
**2)** Sur les $n = 71$ transactions effectuées, le prix moyen du m² est $\bar{x} = 1848.72$ € avec un écart-type de $s = 399.98$ € ; le prix variant de 889 € à 2601 €. L'estimation du prix moyen du m² pour T1-2024 est donc de 1848.72 € avec une précision donnée par l'erreur standard se = 47.47 €.

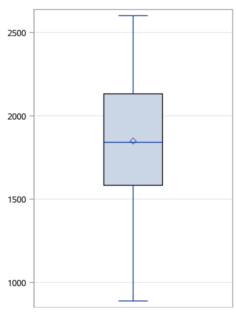
**3)** La taille de l'échantillon étant suffisamment grande ($n = 71 > 50$), l'intervalle de confiance à 90 % pour $\mu$ au premier trimestre 2024 est donné par la formule :

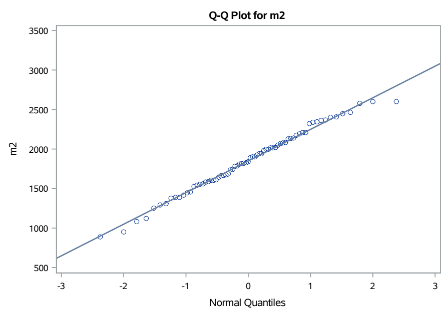
$$\mathrm{ic}_{0.90}(\mu) = \left[\bar{x} \pm t_{0.95}\mathrm{se}\right] = [1848.72 \pm 1.667 \times 47.47] = [1769.59, 1927.85]$$

**Partie II** (10 min, 2 points)

A l'aide des sorties SAS suivantes, vérifier que le prix du m² suit bien une loi normale.






**Tests for Normality**

| Test | Statistic | | p Value | |
|---|---|---|---|---|
| Shapiro-Wilk | W | 0.987287 | Pr < W | 0.6945 |
| Kolmogorov-Smirnov | D | 0.049669 | Pr > D | >0.1500 |
| Cramer-von Mises | W-Sq | 0.018338 | Pr > W-Sq | >0.2500 |
| Anderson-Darling | A-Sq | 0.15785 | Pr > A-Sq | >0.2500 |

Le premier graphique (histogramme) n'est pas vraiment convaincant quant à la normalité de l'échantillon. La boite à moustache par contre souligne bien un échantillon symétrique avec des moustaches équilibrées signes de normalité. Cela est confirmé par le diagramme Q-Q où les points suivent bien la droite définissant la normalité. Finalement, le test de Shapiro-Wilk, avec une probabilité critique d'environ $0.7 > 0.05$, confirme qu'on ne peut pas rejeter l'hypothèse de normalité de $X$.

**Partie III** (25 min, 4 points)

On souhaite vérifier que les variations de prix du m² sont restées stables par rapport à 2022.

**1)** Il s'agit d'un test sur une variance (ou écart-type) sur une population normale (hypothèse vérifiée dans la partie précédente). En 2022, l'écart-type du prix au m² était de $\sigma_0 = 415$. On souhaite donc tester $H_0 : \sigma = 415$ contre $H_1 : \sigma \neq 415$.

**2)** La variable de décision est

$$K^2 = \frac{(n-1)S^2}{\sigma_0^2} \hookrightarrow \chi^2(n-1) = \chi^2(70) \implies k^2 = \frac{70 \times 399.98^2}{415^2} = 65.02$$

Avec un risque de première espèce de 5 %, la région critique (bilatérale) est

$$W = [0, k^2_{0.025}[ \cup ]k^2_{0.975}, +\infty[ = [0, 48.76[ \cup ]95.02, +\infty[$$

Comme $k^2 \notin W$, on ne peut pas rejeter l'hypothèse $H_0$.

**3)** Les données collectées ne démontrent pas de différence des variations du prix du m² en 2024 (par rapport à 2022). Le risque pris est le risque de seconde espèce $\beta = P(H_0|H_1)$ qui est inconnu (mais calculable).

**4)** Commenter la sortie STATA suivante :

```
Stata. sdtest m2 == 415

One-sample test of variance
------------------------------------------------------------------------------
Variable |     Obs        Mean    Std. err.   Std. dev.   [95% conf. interval]
---------+--------------------------------------------------------------------
      m2 |      71    1848.718    47.46872    399.9785    1754.045    1943.392
------------------------------------------------------------------------------
    sd = sd(m2)                                             c = chi2 =  65.0242
H0: sd = 415                                       Degrees of freedom =       70

    Ha: sd < 415              Ha: sd != 415               Ha: sd > 415
  Pr(C < c) = 0.3541         2*Pr(C < c) = 0.7083        Pr(C > c) = 0.6459
```

La sortie stata correspond au test précédent. On retrouve la valeur de la variable de décision `chi2=65.0242`. La probabilité critique associée au test bilatéral est 0.7083, nettement supérieure au risque de première espèce de 5 % confirmant qu'on ne peut pas rejeter $H_0$.

**Partie IV** (30 min, 5 points)

On souhaite à présent tester si le prix moyen du m² a baissé en 2024 (par rapport à 2022).

**1)** Il s'agit d'un test sur une moyenne. En 2022, le prix moyen du m² était $\mu_0 = 1945$ €. L'hypothèse alternative, correspondant à une baisse est donc $H_1 : \mu < 1945$. L'hypothèse nulle, correspondant à une stabilité du prix est $H_0 : \mu = 1945$. Il s'agit donc d'un test unilatéral à gauche.
La variable de décision est

$$T = \frac{\overline{X} - \mu_0}{S/\sqrt{n}} \hookrightarrow \mathrm{St}(n-1) = \mathrm{St}(70) \implies t = \frac{1848.72 - 1945}{47.47} = -2.03$$

Avec un risque de première espèce de 5 %, la région critique (à gauche) est $W = ]-\infty, t_{0.05}[ = ]-\infty, -1.667[$. Comme $t \in W$, on rejette l'hypothèse $H_0$ et on accepte l'hypothèse $H_1$.

**2)** Les données collectées démontrent une baisse du prix moyen du m² au premier trimestre 2024 (par rapport à 2022). Le risque pris est le risque de première espèce $\alpha = P(H_1|H_0) = 5$ %.

**3)** La probabilité critique associée à ce test est

$$p\text{-}value = P(T < -2.03) = 1 - P(T < 2.03) \approx 1 - 0.98 = 0.02 < 5\ \%$$

La probabilité critique étant inférieure au risque de première espèce ($\alpha = 5$ %), cela confirme le rejet de l'hypothèse nulle.

**4)** Commenter la sortie R suivante :

```
R> t.test(data$m2, mu=1945)
        One Sample t-test
data:  data$m2
t = -2.0283, df = 70, p-value = 0.04633
alternative hypothesis: true mean is not equal to 1945
95 percent confidence interval:
 1754.045 1943.392
sample estimates:
mean of x
 1848.718
```

La sortie R correspond au test bilatéral. On retrouve la valeur de la variable de décision `t = -2.0283` et la probabilité critique bilatérale `p-value = 0.04633`. La probabilité critique unilatérale correspondant au test précédent est donc *p-value* $= 0.04633/2 \approx 0.023$ proche de la valeur trouvée à la question précédente.

**Partie V** (20 min, 3 points)

La sortie JAMOVI suivante précise le prix du m² des maisons de l'échantillon selon leur localisation (Centre-Ville ou Périphérie).

Descriptives

| | Localisation | N | Mean | SE | SD | Minimum | Maximum |
|---|---|---|---|---|---|---|---|
| m2 | Périphérie | 44 | 1776 | 62.89 | 417.1 | 950 | 2601 |
| | Centre-Ville | 27 | 1967 | 66.51 | 345.6 | 889 | 2449 |

**1)** La population statistiques $\Omega$ (transactions immobilières de maisons) est à présent divisée en deux : $\Omega_1$ pour les transactions de maisons en périphérie et $\Omega_2$ pour celles en centre-ville. Le prix (au m²) des transactions en périphérie est noté $X$ et $Y$ pour celui en centre-ville. On note respectivement $\mu_X, \bar{x}, \sigma_X, s_X$ et $\mu_Y, \bar{y}, \sigma_Y, s_Y$ la moyenne, moyenne empirique, écart-type, écart-type empirique correspondants.
La sortie JAMOVI montre un prix moyen au m² ($\bar{x} = 1776$) sur l'échantillon de la périphérie (de taille $n_X = 44$) inférieur à celui ($\bar{y} = 1967$) sur l'échantillon du centre-ville (de taille $n_Y = 27$). Les variations (sur chaque échantillon) ne semblent pas très différentes ($s_X = 417.1$ vs $s_Y = 345.6$).

**2)** On souhaite donc effectuer le test $H_0 : \mu_X = \mu_Y$ contre $H_1 : \mu_X < \mu_Y$. Avant de pouvoir effectuer ce test, on vérifie l'homogénéité des variances (test $H_0 : \sigma_X = \sigma_Y$) correspondant à la première sortie JAMOVI.
Le test (bilatéral) de Fisher (ligne Variance ratio) a une probabilité critique *p-value* = 31 % très supérieure au risque habituel (5 ou 10 %). On ne peut donc pas rejeter l'hypothèse d'égalité des variances.
On peut alors procéder à un test de comparaison des moyennes de Student. La sortie JAMOVI donne la probabilité critique unilatérale : *p-value* = 2.5 % < 5 %. On conclut donc que le prix moyen du m² des maisons de périphérie est significativement inférieur à celui des maison de centre-ville. En outre, la taille de l'effet (Cohen's $d$) égal à $0.487 \approx 0.5$ démontre une différence relativement forte.

Homogeneity of Variances Tests

| | | F | df | df2 | p |
|---|---|---|---|---|---|
| m2 | Levene's | 2.118 | 1 | 69 | 0.150 |
| | Variance ratio | 1.457 | 43 | 26 | 0.310 |

Independent Samples T-Test

| | | Statistic | df | p | | Effect Size |
|---|---|---|---|---|---|---|
| m2 | Student's t | -1.995 | 69.00 | 0.025 | Cohen's d | -0.4877 |

*Note.* $H_a$ $\mu_{\text{Périphérie}} < \mu_{\text{Centre-Ville}}$