



Intitulé du Module : Estimation

Code Module : MAT4055

Date : 15/07/2013

Professeur : Myriam Maumy-Bertrand

Durée : 1h30

Nombre de pages : 4

Examen : ☒

Contrôle: ☐

Classe : SYS1, INFO1, INFO2

Documents autorisés : Oui ☒

Non ☐

Calculatrice autorisée : Oui ☒

Non ☐

Ordinateur portable autorisé : Oui ☐

Non ☒

Précision sur le barème si QCM :

Commentaires :

Nom & Prénom de l'étudiant :

Code étudiant :

Classe :

## Sujet numéro 2

*Le sujet comporte trois exercices indépendants. Il vous est demandé de ne traiter que deux exercices parmi les trois. La rédaction de trois exercices entraîne automatiquement un rejet de correction.*

- Les calculatrices sont autorisées.
- Le cours, les exercices de travaux dirigés, leurs corrigés ainsi que les notes de cours sont autorisés. Tout autre document est interdit
- Afin de pouvoir traiter les questions, plusieurs résultats numériques et graphiques ont été intégrés au document.
- Vous prendrez un soin particulier à préciser quelles sont les hypothèses testées.
- Tous les tests seront effectués au seuil de signification  $\alpha = 5\%$ .

### Exercice 1. Video killed the radio star.

Une radio s'inquiète de sa perte d'audience. Une enquête faite auprès de 1 286 Français choisis au hasard dans l'annuaire et ayant accepté de répondre montre que 339 d'entre eux ont déclaré écouter au moins de temps en temps la station. Pour assurer son avenir, elle voudrait améliorer sa connaissance des habitudes de ses plus jeunes auditeurs : parmi les sondés, 202 étaient étudiants, et 40 l'écoutaient au moins de temps en temps. La direction de la prospective va diligenter un nouveau sondage destiné uniquement aux étudiants.

- a) Modéliser le problème (pour chacun des deux sondages), en précisant la population, les données et de le modèle. Indiquer les paramètres d'intérêt respectifs.
- b) Dédire du premier sondage une précision et un intervalle de confiance pour les paramètres d'intérêt de chacun des deux sondages.
- c) Combien de personnes faut-il interroger au cours de la seconde enquête, si le degré de confiance retenu est de 95% et la précision désirée, 6% ?
- d) À l'issue du second sondage, il a été constaté 43,8% d'auditeurs. Donner une estimation et un intervalle de confiance du paramètre faisant l'objet de l'étude avec un niveau de confiance de 95%.
- e) Pouvez-vous affirmer que l'audience du segment étudiant a augmenté d'une enquête à l'autre ? (Les deux enquêtes sont séparées de six mois et la station a revu entretemps sa grille de programmes.)

**Exercice 2. Prix d'un appartement.**

Nous avons relevé quelques annonces d'appartements à vendre dans la presse parisienne. Les données sont reproduites dans le tableau suivant :

1. CENSIER, bas de R. Mouffetard, pied-à-terre, 28m <sup>2</sup> , tt confort. Visite vendredi, samedi, dim. 650.000 F à discuter. Facilités
2. CONTRESCARPE, imm. Ancien, pierre de taille, beau duplex caractère, 50m <sup>2</sup> , poutres, refait neuf, 1.400.000 F
3. R. St-Simon, en pleine verdure, calme, plein soleil, superbe appt 4p., 106m <sup>2</sup> , cuis. aménagée, s. de bains moderne, chff. cent. Parfait état. Px 3.250.000 F à discuter. Agence s'abstenir. Direct propriétaire.
4. RAPP 7P., 196m <sup>2</sup> standing, 9 fenêtres plein soleil, 4.000.000 F
5. R. St André-des-Arts, beau liv + chbre, imm. XVIIIe siècle, 55m <sup>2</sup> , 1.350.000 F.
6. 5e PRES QUAIS, 7 pièces, 190m <sup>2</sup> caractère, standing, 3.950.000 F
7. GOBELINS, Beau 5p., 110m <sup>2</sup> , gd cft, soleil, 250.000 F
8. GOBELINS, et. élevé, calme, asc., 2 pièces, 60m <sup>2</sup> , 1.600.000 F
9. CENSIER, très grand studio + entrée 48m <sup>2</sup> , tt cft, ensoleillé, calme, bel imm., 1.250.000 F
10. PANTHEON, 7e étage, ascenseur, grand studio 35m <sup>2</sup> + terrasse. Vue. 1.250.000 F
11. RUE MADAME, 3P. + Serv., 86m <sup>2</sup> , 1.750.000 F
12. RUE DE SEINE, 3P., tt cft, 65m <sup>2</sup> , calme, soleil, 1.500.000 F
13. PANTHEON, bel imm., verdure, magnifique studio 32m <sup>2</sup> , caractère, 775.000 F.
14. SEVRES BAB, 1er ét., 2P., grande cuis., bns 52m <sup>2</sup> , état neuf, 1.225.000 F
15. MONTARNASSE, Part. vend atelier d'artiste 40m <sup>2</sup> , duplex, vue imprenable, tout confort, Prix 1.000.000 F
16. RUE D'ASSAS, imm. gd standing, bel appart 260m <sup>2</sup> , triple récept. + 5 ch., tt cft (travaux) 2 park., 2 ch. Serv., Prix 7.500.000 F à déb.
17. BD St-GERMAIN, 4P., 70m <sup>2</sup> , à amén., 4e ét., 1.625.000 F.
18. ILE St-LOUIS, Lux. appt., 117m <sup>2</sup> , en duplex, grande récept., grande chambre, 2 sdb, Terras., parf. et., décor tr. bon goût, 4.750.000 F
19. JUSSIEU, Charme, gd 3pcs, 90m <sup>2</sup> , 1.890.000 F
20. QUARTIER LATIN, 30m <sup>2</sup> à aménager, prix 390.000 F
21. MONTARNASSE, Imm. p.d.t., 4-5 P., 105m <sup>2</sup> , bon état, 1.875.000 F
22. RUE MAZARINE, 4e ét., sans ascens., 52m <sup>2</sup> à rénover. Prix total 1.000.000 F
23. CENSIER, Bel imm., 4P. 80m <sup>2</sup> , tt cft, petits travaux, 1.350.000 F
24. ASSAS LUXEMBOURG, 3P. 60m <sup>2</sup> s/arbres, imm. caractère, 1.475.000 F
25. SUR JARDINS OBSERVATOIRE, 140m <sup>2</sup> , grand charme, 4.950.000 F
26. RUE DE SAVOIE, 4e ét., Studio 20m <sup>2</sup> , dche, 425.000 F crédit possible.
27. PRES LUXEMBOURG, Bel imm., pierre de taille, Appartement 100m <sup>2</sup> , salon, salle à manger, 2 chambres, office, cuis., bains, chf. cent., asc., prix! : 2.475.000 F
28. Mo GOBELINS, studio, cuis., s. de bains, 28m <sup>2</sup> , calme. Prix 425.000 F

- a) Un économiste souhaite utiliser le modèle linéaire simple pour modéliser le lien entre le prix d'un appartement et sa surface. Spécifier ce modèle (donner les coefficients) et bien identifier chacune des composantes du modèle (la variable explicative et la variable à expliquer) dans le contexte de ce problème.
- b) Tester l'hypothèse nulle suivante avec un test approprié (vous donnerez le nom de ce test) :

$$\mathcal{H}_0 : \beta_1 = 0$$

contre

$$\mathcal{H}_1 : \beta_1 \neq 0.$$

- c) Donner la valeur de la variation qui est expliquée par la droite des MCO et la variation qui est inexpliquée par la même droite. En déduire le pourcentage de variation qui est expliqué par la droite des MCO.
- d) Calculer une estimation du prix d'un appartement pour les superficies suivantes :  $X_1 = 50, X_2 = 100, X_3 = 200$ . Pour quelle superficie l'estimation du prix d'un appartement serait-elle la plus précise parmi les trois valeurs précédentes ?
- e) Entre quelles valeurs peut se situer la vraie valeur du prix d'un appartement pour les superficies dont le prix a été déterminé à la question précédente ? Utiliser un niveau de confiance de 95%. Quelle est la marge d'erreur dans l'estimation effectuée à la question précédente ?

```
> surface = c(28, 50, 106, 196, 55, 190, 110, 60, 48, 35, 86, 65, 32,
52, 40, 260, 70, 117, 90, 30, 105, 52, 80, 60, 140, 20, 100, 28)
> prix = c(650, 1400, 3250, 4000, 1340, 3950, 2500, 1600, 1250, 1250,
1750, 1500, 775, 1225, 1000, 7500, 1625, 3750, 1890, 390, 1875, 1000,
1350, 1475, 4950, 425, 2475, 425)
> prix_au_m2 = c(23.21, 28, 30.66, 20.41, 24.36, 20.79, 22.73, 26.67,
26.04, 35.71, 20.35, 23.08, 24.22, 23.56, 25, 28.85, 23.21, 40.6,
21, 13, 17.86, 19.23, 16.88, 24.58, 35.36, 21.25, 24.75, 15.18)
> appart = data.frame(surface, prix, prix_au_m2)
> appart
  surface prix prix_au_m2
1      28  650     23.21
2      50 1400     28.00
3     106 3250     30.66
4     196 4000     20.41
5      55 1340     24.36
6     190 3950     20.79
7     110 2500     22.73
8      60 1600     26.67
9      48 1250     26.04
10     35 1250     35.71
11     86 1750     20.35
12     65 1500     23.08
13     32  775     24.22
```

14	52	1225	23.56
15	40	1000	25.00
16	260	7500	28.85
17	70	1625	23.21
18	117	3750	40.60
19	90	1890	21.00
20	30	390	13.00
21	105	1875	17.86
22	52	1000	19.23
23	80	1350	16.88
24	60	1475	24.58
25	140	4950	35.36
26	20	425	21.25
27	100	2475	24.75
28	28	425	15.18

```
> appart.lm = lm(appart$prix~appart$surface)
> appart.lm
```

Call:

```
lm(formula = appart$prix ~ appart$surface)
```

Coefficients:

```
(Intercept) appart$surface
      -150.67         26.37
```

```
> plot(appart$surface,appart$prix,xlab = "surface", ylab = "prix",
col.axis = "peru",col.lab = "royalblue",col.main = "royalblue",
main = " étude de 28 appartements")
> abline(appart.lm, col = "green3")
> residus<-residuals(appart.lm)
> shapiro.test(residus)
```

Shapiro-Wilk normality test

data: residus

W = 0.9519, p-value = 0.2216

```
> summary(appart.lm)
```

Call:

```
lm(formula = appart$prix ~ appart$surface)
```

Residuals:

Min	1Q	Median	3Q	Max
-1018.36	-250.36	22.24	143.17	1408.51

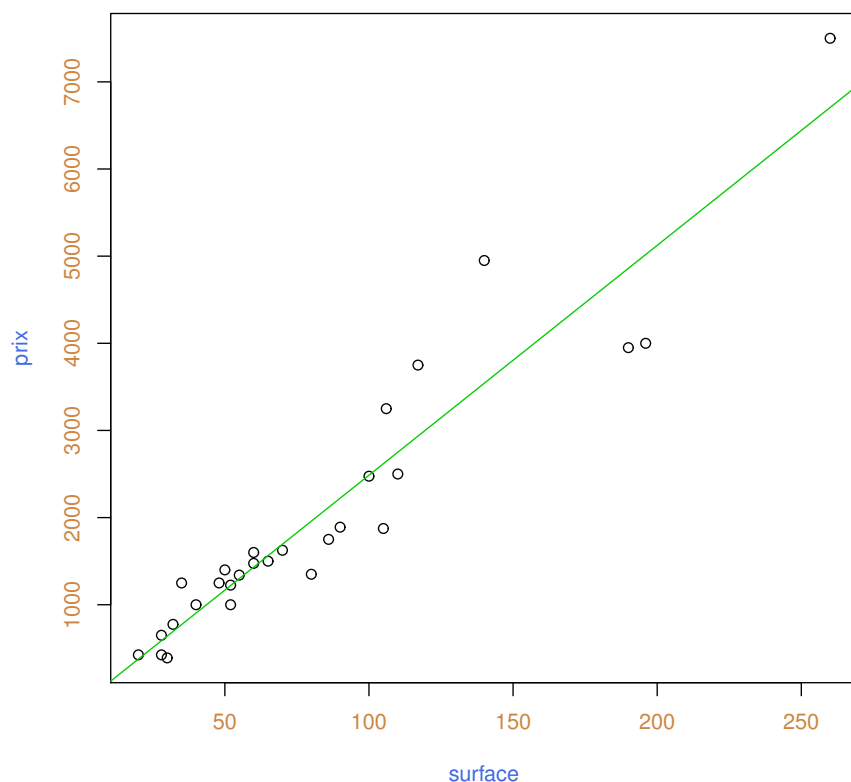
Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	-150.672	177.079	-0.851	0.403
appart\$surface	26.373	1.777	14.838	3.32e-14 ***

---

Residual standard error: 527.8 on 26 degrees of freedom  
 Multiple R-squared: 0.8944, Adjusted R-squared: 0.8903  
 F-statistic: 220.2 on 1 and 26 DF, p-value: 3.321e-14

### étude de 28 appartements



```
> anova(appart.lm)
Analysis of Variance Table
```

```
Response: appart$prix
      Df Sum Sq Mean Sq F value    Pr(>F)
appart$surface  1 61337423 61337423  220.17 3.321e-14 ***
Residuals      26  7243273   278587
---
```

```
> predict(appart.lm,interval="confidence")
```

```
      fit      lwr      upr
1  587.7602 302.41148 873.1090
```

2	1167.9573	931.35148	1404.5631
3	2644.8225	2422.28646	2867.3585
4	5018.3558	4555.19080	5481.5208
5	1299.8202	1071.78079	1527.8597
6	4860.1203	4416.50456	5303.7360
7	2750.3128	2521.69927	2978.9264
8	1431.6832	1211.02729	1652.3391
9	1115.2121	874.87630	1355.5479
10	772.3684	504.17554	1040.5612
11	2117.3706	1911.89685	2322.8444
12	1563.5462	1348.96883	1778.1235
13	693.2506	417.86512	968.6361
14	1220.7025	987.65720	1453.7477
15	904.2313	647.43374	1161.0289
16	6706.2018	6025.46354	7386.9400
17	1695.4091	1485.49202	1905.3262
18	2934.9210	2693.90190	3175.9401
19	2222.8610	2015.91700	2429.8050
20	640.5054	360.18926	920.8216
21	2618.4499	2397.30843	2839.5913
22	1220.7025	987.65720	1453.7477
23	1959.1351	1753.92593	2164.3442
24	1431.6832	1211.02729	1652.3391
25	3541.4906	3247.47967	3835.5016
26	376.7795	70.38314	683.1758
27	2486.5869	2271.62117	2701.5526
28	587.7602	302.41148	873.1090

**Exercice 3. Trouvez-vous votre évier trop bas ?**

Avez-vous déjà fait la vaisselle dans un évier situé dans une cité U construite dans les années 70 ? Si oui, vous avez remarqué qu'il était très bas. Imaginez que vous soyez un constructeur de cuisines et que vous commenciez à entendre des plaintes sur la hauteur de vos éviers. Vous les construisiez jusqu'à présent pour une taille moyenne de 165 centimètres (vos données vous disaient que la taille moyenne des femmes était justement de 165 centimètres, et vos analyses marketing montrent que ce sont elles les prescriptrices d'achat dans les couples en ce qui concerne la cuisine). Vous conduisez donc une étude et mesurez toutes les femmes qui se présentent à votre magasin (en échange d'un petit cadeau). 64 femmes se prêtent au jeu, cela vous donne les hauteurs  $x_1, \dots, x_{64}$  dont la distribution parente suit une loi normale. La hauteur moyenne relevée est égale à  $\bar{x}_{64} = 169,5$  centimètres et l'écart-type sur les données est égal à  $s_{64,c} = 16,0$  centimètres.

Faut-il faire remonter cette information au syndicat des constructeurs de cuisines afin qu'il diligente une étude plus approfondie ?

**Information dont il vous faudra tenir compte :** vous voulez vous présenter à la présidence du syndicat l'an prochain.

Pour répondre à la question, vous allez suivre la démarche suivante :

- a) allez-vous utiliser un test paramétrique ou non ? Justifier votre réponse.
- b) Donner le nom du test que vous allez utiliser.
- c) Donner les hypothèses du test. S'agit-il d'un test bilatéral ou d'un test unilatéral ? Justifier votre réponse.
- d) Donner la statistique du test et calculer-la.
- e) Conclure. Calculer le risque d'erreur associé à cette décision.



Nous rappelons également quelques valeurs de quantiles qui peuvent vous aider à répondre aux trois exercices :

```
> qnorm(0.975)
[1] 1.959964
> qnorm(0.95)
[1] 1.644854
> qt(0.975,64)
[1] 1.99773
> qt(0.95,64)
[1] 1.669013
> qt(0.975,63)
[1] 1.998341
> qt(0.95,63)
[1] 1.669402
```