

CORRIGÉS

Chapitre 1

1 Variables discrètes et variables continues

- Faux, une variable discrète peut aussi prendre des valeurs négatives.
- Vrai.
- Faux, le chiffre d'affaires est une variable continue.
- Vrai.
- Vrai.

2 Caractères et modalités

- Faux. Par exemple, le caractère « sexe » n'a que deux modalités.
- Faux, les modalités doivent être incompatibles.
- Faux, car si la taille est bien un caractère quantitatif, l'état matrimonial est un caractère qualitatif.
- Vrai.
- Vrai.

3 Graphiques et centre de classe

- Vrai.
- Vrai.
- Faux, il s'agit de la courbe des fréquences cumulées.
- Faux, il est nécessaire de corriger les amplitudes afin que l'aire de chaque rectangle composant l'histogramme soit bien proportionnelle à l'effectif (ou la fréquence).
- Vrai.

4 Le mode

- Faux, il s'agit de la médiane.
- Faux.
- Faux, il s'agit de la moyenne.
- Faux, le mode est une caractéristique de tendance centrale.
- Vrai.

5 Étude de la répartition des notes : 18, 15, 8, 12, 8, 15, 4

- Faux, il convient de pondérer par les effectifs : $\frac{18 + 15 \times 2 + 8 \times 2 + 12 + 4}{7} = 11,42$.
- Vrai.
- Faux, la distribution est bimodale, les deux valeurs du mode étant 8 et 15.
- Faux, l'étendue est égale à $18 - 4 = 14$.
- Vrai, le moment simple d'ordre 1 étant égal à la moyenne.

6 Caractéristiques d'une distribution

Le tableau 1.1 reporte les différents calculs nécessaires pour répondre aux diverses questions posées.

▼ Tableau 1.1 Location de voitures

Classes	n_i	x_i	f_i	F_i	$n_i x_i$	$n_i x_i^2$
[0,10[2	5	0,01	0,01	10	50
[10,20[19	15	0,11	0,13	285	4 275
[20,30[28	25	0,17	0,30	700	17 500
[30,40[54	35	0,33	0,62	1 890	66 150
[40,50[31	45	0,19	0,81	1 395	62 775
[50,60[21	55	0,13	0,93	1 155	63 525
[60,70[11	65	0,07	1,00	715	46 475
Total	166		1		6 150	260 750

- La variable étant groupée en classes, il s'agit d'une variable continue.
- Les résultats sont reportés dans le tableau 1.1, x_i désignant les centres de classes, f_i les fréquences et F_i les fréquences cumulées.
- Il est possible de représenter la série au moyen d'un histogramme.
- La classe modale est celle pour laquelle la fréquence est la plus élevée, il s'agit donc de la classe [30,40[. On peut calculer la valeur du mode au moyen de la relation (1.14) :

$$Mode = 30 + 10 \times \frac{0,33 - 0,17}{(0,33 - 0,17) + (0,33 - 0,19)} = 35,42$$

La valeur du mode est donc égale à 35 voitures.

- La moyenne peut être calculée à l'aide de l'équation (1.24) :

$$\bar{x} = \frac{1}{166} \times 6\,150 = 37,05$$

En moyenne, 37 voitures sont louées par jour.

- On utilise la colonne des fréquences cumulées afin de repérer la classe médiane, il s'agit ainsi de la classe [30,40[. La valeur précise de la médiane est obtenue à l'aide de l'équation (1.17) :

$$M = 30 + \frac{10}{0,33} \times [0,5 - 0,30]$$

La valeur médiane est ainsi égale à 36,30 : il y a autant de sociétés louant moins de 36 voitures que de sociétés en louant plus.

- On calcule l'écart-type corrigé :

$$S_x = \sqrt{\frac{1}{N-1} \sum_{i=1}^N n_i x_i^2 - \frac{N}{N-1} \bar{x}^2}$$

soit :

$$S_x = \sqrt{\frac{1}{166-1} \times 260\,750 - \frac{166}{166-1} \times 37,05^2}$$

L'écart-type est ainsi égal à 14,12 voitures.

7 Sujet d'examen

▼ **Tableau 1.2** Répartition des salariés par tranche de salaire

Classes de salaires	n_i	x_i	a_i	f_i	F_i	$n_i x_i$	$\frac{n_i x_i}{\sum_{i=1}^5 n_i x_i}$	$\sum_{i=1}^5 \frac{n_i x_i}{\sum_{i=1}^5 n_i x_i}$
[1200,1400[3	1 300	200	0,015	0,015	3 900	0,011	0,011
[1400,1600[6	1 500	200	0,03	0,045	9 000	0,027	0,038
[1600,1800[182	1 700	200	0,91	0,955	309 400	0,909	0,947
[1800,2000[5	1 900	200	0,025	0,98	9 500	0,028	0,975
[2000,2200[4	2 100	200	0,02	1	8 400	0,025	1
Total	200					340 200		1

1. L'effectif total de l'entreprise est donné par :

$$N = \sum_{i=1}^5 n_i = 3 + 6 + 182 + 5 + 4 = 200$$

2. Les valeurs demandées sont reportées dans le tableau 1.2.
3. La moyenne est donnée par :

$$\bar{x} = \frac{1}{200} \times 340\,200 = 1701$$

Le salaire moyen au sein de l'entreprise considérée est de 1701 euros.

4. La classe modale est la classe pour laquelle l'effectif (ou la fréquence) est le plus élevé, il s'agit en conséquence de la classe [1600,1800[. La valeur du mode est obtenue par la formule :

$$Mode = e_{i-1} + a_m \times \frac{d_1}{d_1 + d_2}$$

où e_{i-1} désigne la valeur de l'extrémité inférieure de la classe modale (1600), a_m l'amplitude de cette même classe (200), d_1 la différence entre l'effectif de la classe modale (182) et l'effectif de la classe précédente (6) et d_2 la différence entre l'effectif de la classe modale (182) et l'effectif de la classe suivante (5). Dans le cas de notre entreprise, le mode est ainsi égal à 1 699,72 euros :

$$Mode = 1\,600 + 200 \times \frac{182 - 6}{(182 - 6) + (182 - 5)} = 1\,699,72$$

5. L'étendue est donnée par la différence entre la valeur maximale et la valeur minimale prise par le salaire. Dans notre cas, les observations étant groupées par classes, l'étendue est égale à la différence entre l'extrémité supérieure de la dernière classe et l'extrémité inférieure de la première classe, soit :

$$Etendue = 2\,200 - 1\,200 = 1\,000$$

L'étendue des salaires au sein de l'entreprise considérée est ainsi égale à 1 000 euros.

6. Au regard des valeurs prises par les fréquences cumulées, on constate que la proportion de salariés gagnant moins de 1 600 euros est de 4,5 % et que celle touchant moins de 1 800 euros est de 95,5 %. On en déduit que la valeur de la médiane est comprise entre 1 600 et 1 800 euros et que la classe médiane est la classe [1600,1800[. La valeur de la médiane est donnée par :

$$M = e_{i-1} + \frac{a_i}{f_i} \times [0,5 - F_{i-1}]$$

où e_{i-1} est l'extrémité inférieure de la classe médiane (1600), a_i l'amplitude de la classe médiane (200), f_i la fréquence de la classe médiane (0,91) et F_{i-1} désigne la fréquence cumulée de la classe au dessus de la classe médiane dans le tableau (0,045). D'où :

$$M = 1\,600 + \frac{200}{0,91} \times [0,5 - 0,045] = 1\,700$$

La médiane est ainsi égale à 1 700 euros : dans l'entreprise considérée, il y a autant de salariés gagnant moins de 1 700 euros par mois que de salariés percevant un salaire mensuel supérieur à 1 700 euros.

7. Trois caractéristiques de tendance centrale ont été calculées : la moyenne égale à 1701 euros, le mode égal à 1 699,72 euros et la médiane valant 1 700 euros. Ces trois valeurs étant très proches, on en déduit que la distribution des salaires au sein de l'entreprise est symétrique.
8. La détermination de la classe médiale est similaire à celle de la classe médiane, le calcul étant basé non plus sur les seuls effectifs n_i mais sur le produit $n_i x_i$ représentant la masse salariale. Par un raisonnement identique et au regard des valeurs obtenues dans la dernière colonne du tableau 1.2, on en déduit que la classe médiale est la classe [1600,1800[. De même, la valeur de la médiale, notée MI et exprimée en euros, est donnée par :

$$MI = 1\,600 + \frac{200}{0,909} \times [0,5 - 0,038] = 1\,701,65$$

9. On constate que les valeurs de la médiane et de la médiale sont très proches. Plus précisément, on a :

$$\frac{MI - M}{Etendue} = \frac{1\,701,65 - 1\,700}{1\,000} = 0,00165$$

Le rapport calculé est ainsi très proche de zéro, l'écart entre la médiale et la médiane étant très faible par rapport à l'étendue. Cela correspond à une concentration nulle des salaires, c'est-à-dire à une parfaite équipartition. Ce résultat était prévisible au regard des données figurant dans le tableau 1.13 : la distribution des observations fait en effet ressortir que la moyenne est très proche de 1 700 euros, témoignant d'une très faible dispersion des salaires autour du salaire moyen.

Chapitre 2

1 Distributions marginales et conditionnelles

- a. Faux, la distribution marginale d'une variable ne tient compte que de l'information contenue dans cette variable.
b. Vrai.
c. Faux, la somme des fréquences, quel que soit le type de fréquence considéré, est toujours égale à 1.
d. Faux, il existe toujours une relation entre fréquences marginales et conditionnelles (► équation (2.15)).
e. Vrai.

2 Liaison entre deux variables

- a. Faux, il s'agit d'un cas de corrélation négative.

- b. Faux, deux variables corrélées négativement évoluent en sens contraire.
- c. Vrai.
- d. Vrai.
- e. Vrai.

3 Droite de régression et ajustement

- a. Faux.
- b. Vrai.
- c. Faux, il s'agit de minimiser la somme des carrés des écarts et non la somme des écarts.
- d. Faux, b est l'ordonnée à l'origine et a est la pente de la droite de régression.
- e. Vrai.

4 Coefficient de corrélation linéaire

- a. Faux, il s'agit d'une corrélation négative.
- b. Faux, un coefficient de corrélation est compris entre -1 et 1 .
- c. Faux, il ne faut pas confondre corrélation et causalité.
- d. Faux, les variables peuvent être corrélées de façon non-linéaire.
- e. Vrai.

5 Analyse de la variance et coefficient de détermination

- a. Vrai.
- b. Vrai.
- c. Faux, c'est le rapport entre la variance expliquée et la variance totale.
- d. Faux, un coefficient de détermination est compris entre 0 et 1 .
- e. Vrai.

6 Étude de la liaison taux de chômage/taux d'inflation

1. La figure 2.10 montre que le nuage de points est relativement allongé et semble se répartir autour d'une droite d'allure décroissante. Cela laisse présager l'existence d'une liaison négative entre les deux variables. Un tel résultat est cohérent d'un point de vue économique en vertu de la relation de Phillips selon laquelle les deux variables tendent à évoluer en sens contraire.

En toute rigueur, notons que nous considérons ici une relation linéaire entre le taux de chômage et le taux d'inflation alors que la courbe de Phillips renvoie à une relation non-linéaire entre les deux variables.

2. Rappelons que la covariance entre deux variables x et y est donnée par :

$$\text{Cov}(x, y) = \frac{1}{N} \sum_{i=1}^N x_i y_i - \bar{x} \bar{y}$$

D'où :

$$\text{Cov}(x, y) = \frac{1}{30} \times 881,01 - 6,71 \times 4,74 = -2,43$$

La covariance est négative, comme attendu.

3. La valeur du coefficient de corrélation linéaire entre les deux variables est donnée par :

$$r(x, y) = \frac{\text{Cov}(x, y)}{\sigma_x \sigma_y}$$

D'où :

$$r(x, y) = \frac{-2,43}{\sqrt{4,64} \times \sqrt{1,51}} = -0,92$$

Le coefficient de corrélation linéaire entre le taux de chômage et le taux d'inflation est négatif et proche de -1 . Il existe donc bien une relation décroissante entre les deux variables, conformément à la relation de Phillips.

4. Les valeurs des coefficients de la droite de régression sont donnés par :

$$a = \frac{\text{Cov}(x, y)}{V(x)} = \frac{-2,43}{4,64} = -0,52$$

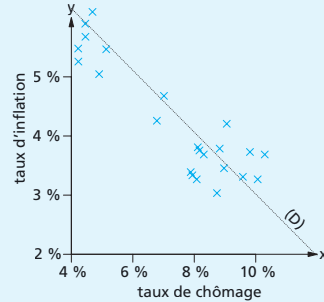
et, sachant que la droite de régression passe par le point moyen :

$$b = \bar{y} - a\bar{x} = 4,74 + 0,52 \times 6,71 = 8,23$$

On en déduit donc que la droite de régression a pour expression :

$$\hat{y} = -0,52x + 8,23$$

Le coefficient de pente de la droite de régression est bien négatif, confirmant à nouveau la relation inverse entre le taux d'inflation et le chômage. La figure 2.1 reproduit l'ajustement du nuage de points par cette droite de régression.



▲ Figure 2.1 Ajustement du nuage de points par la droite de régression.

5. On peut calculer le coefficient de détermination comme suit :

$$R^2 = \frac{\text{Cov}(x, y)^2}{V(x)V(y)} = \frac{(-2,43)^2}{4,64 \times 1,51} = 0,84$$

On en déduit que le pourcentage de variance expliquée par la régression est de 84 %.

7 Sujet d'examen

1. Les coefficients de variation CV , exprimés en pourcentage, sont donnés par le rapport entre l'écart-type et la moyenne pour chacune des trois séries. On a donc : $CV_{EUR} = 104,1565$, $CV_{DKK} = 6,5569$ et $CV_{ATS} = 5,2164$. La série la plus volatile est donc la série EUR , suivie par DKK puis ATS .

2. On rappelle que les coefficients de corrélation sont donnés par :

$$r(EUR, DKK) = \frac{Cov(EUR, DKK)}{\sigma_{EUR} \sigma_{DKK}}$$

$$r(ATS, DKK) = \frac{Cov(ATS, DKK)}{\sigma_{ATS} \sigma_{DKK}}$$

On obtient ainsi :

$$r(EUR, DKK) = \frac{\frac{1}{172}(-37,3106) - (-0,1227) \times 1,8942}{0,1278 \times 0,1242} = 0,9763$$

$$r(ATS, DKK) = \frac{\frac{1}{172}(817,2116) - 2,4998 \times 1,8942}{0,1304 \times 0,1242} = 0,9947$$

On constate que les deux coefficients de corrélation sont positifs : les deux variables considérées évoluent dans le même sens. Ils sont en outre très proches de 1, ce qui témoigne d'une forte corrélation entre les deux variables ; la corrélation la plus forte étant obtenue entre les variables DKK et ATS .

3. a. L'application de la méthode des moindres carrés ordinaires conduit aux résultats suivants :

$$a = \frac{Cov(EUR, DKK)}{V(EUR)} = \frac{\frac{1}{172}(-37,3106) - (-0,1227) \times 1,8942}{(0,1278)^2} = 0,9488$$

$$b = 1,8942 - 0,9488 \times (-0,1227) = 2,0106$$

La droite de régression s'écrit donc :

$$\widehat{DKK}_t = 0,9488 \times EUR_t + 2,0106$$

b. Toutes choses égales par ailleurs, une augmentation de 1 % de EUR s'accompagne d'une hausse de 0,9488 % du taux de change du dollar vis-à-vis de la couronne danoise.

c. L'équation d'analyse de la variance est donnée par :

$$\text{Variance totale} = \text{Variance expliquée} + \text{Variance résiduelle} \quad (2.1)$$

La variance totale est égale à $0,1242^2$ et la variance résiduelle vaut $0,1169/172 = 0,0007$. On en déduit donc que la variance expliquée est égale à : $0,1242^2 - 0,0007 = 0,0147$. La part de la variance résiduelle dans la variance totale est en conséquence égale à $0,0007/0,1242^2 = 0,0454$, soit 4,54 %.

d. Le coefficient de détermination est donné par le rapport entre la variance expliquée et la variance totale, soit :

$$R^2 = \frac{0,0147}{(0,1242)^2} = 0,9530$$

La variable EUR explique donc 95,30 % des variations du taux de change du dollar vis-à-vis de la couronne danoise.

e. Si l'on multiplie par 10 toutes les valeurs observées des variables EUR et DKK , cela n'a aucun impact sur la pente de la droite de régression. En effet, soient : $DKK'_t = 10 \times DKK_t$ et $EUR'_t = 10 \times EUR_t$. On peut écrire :

$$\frac{DKK'_t}{10} = a \times \frac{EUR'_t}{10} + b$$

soit encore :

$$DKK'_t = a \times EUR'_t + 10b$$

Si l'on ajoute 10 à chacune des valeurs observées des deux variables, le coefficient de la pente de la droite de régression n'est pas modifié. Posons : $DKK''_t = 10 + DKK'_t$ et $EUR''_t = 10 + EUR'_t$. On a :

$$DKK''_t - 10 = a(EUR''_t - 10) + b$$

soit encore :

$$DKK''_t = a \times EUR''_t + 10 - 10a + b$$

4. On sait que la variance totale est égale à $0,1242^2$ et que la variance expliquée vaut 0,0153. On en déduit, d'après l'équation d'analyse de la variance, la valeur de la variance résiduelle : $0,1242^2 - 0,0153 = 1,2564 \cdot 10^{-4}$. La somme des carrés des résidus est en conséquence égale à : $172 \times 1,2564 \cdot 10^{-4} = 0,0216$. Le coefficient de détermination est quant à lui donné par :

$$R^2 = \frac{0,0153}{(0,1242)^2} = 0,9919$$

On en déduit que la variable ATS explique 99,19 % de la variation de DKK .

5. Les deux modèles ayant la même variable expliquée, DKK , il est possible de les comparer. Le coefficient de détermination du deuxième modèle est supérieur à celui du premier modèle. La variance résiduelle est plus faible dans le deuxième modèle. Au regard de ces résultats, on retiendra donc le deuxième modèle pour expliquer les fluctuations de DKK .

Chapitre 3

1 Propriétés sur les indices

- a. Vrai.
b. Faux. On a en effet : $I_{t/t'}^g = \frac{I_{t/0}^g}{I_{t'/0}^g}$.
c. Faux, aucun des deux indices ne vérifie ces propriétés.
d. Vrai.
e. Vrai.

2 Indices synthétiques

- a. Faux, il s'agit de l'indice des prix de Laspeyres.
b. Faux, l'indice de Paasche est une moyenne harmonique.
c. Faux, les indices élémentaires doivent être basés à la même date.
d. Vrai, la structure du panier est celle de la date courante.
e. Vrai, la structure du panier est celle de la date de référence.

3 Évolution du prix d'un quotidien de la presse écrite

- Faux.
- Faux.
- Vrai. En effet, en utilisant les propriétés de circularité et de réversibilité, on a :

$$I_{2013/2005} = I_{2013/2000} \times I_{2000/2005} = \frac{I_{2013/2000}}{I_{2005/2000}} = \frac{1,30}{1,50} = 0,867$$

- Faux.
- Faux.

4 Évolution d'un indice de prix

- Faux.
- Faux.
- Vrai.
- Faux.
- Faux.

5 Formule d'un indice synthétique

- Faux.
- Faux.
- Vrai.
- Faux.
- Faux.

6 Indices élémentaires, synthétiques et effet qualité

▼ **Tableau 3.1** Consommation de pommes d'un ménage : calcul des indices

Variété i	p_0^i	q_0^i	p_t^i	q_t^i	$p_0^i q_0^i$	$p_t^i q_t^i$	$p_0^i q_t^i$	$p_t^i q_0^i$
Golden	2,8	10	3,1	15	28	46,5	42	31
Pink Lady	3,9	5	4,4	15	19,5	66	58,5	22
Royal Gala	2,1	15	2,2	10	31,5	22	21	33
Somme		30		40	79	134,5	121,5	86

▼ **Tableau 3.2** Consommation de pommes d'un ménage : coefficients de pondération

Variété i	α_0^i	α_t^i	$\alpha_0^i \frac{p_t^i}{p_0^i}$	$\alpha_0^i \frac{p_t^i}{p_0^i}$	$\alpha_0^i \frac{q_t^i}{q_0^i}$	$\alpha_t^i \frac{q_t^i}{q_0^i}$
Golden	0,35	0,35	0,31	0,39	0,53	0,23
Pink Lady	0,25	0,49	0,43	0,28	0,74	0,16
Royal Gala	0,40	0,16	0,16	0,42	0,27	0,25
Somme	1	1	0,90	1,09	1,54	0,64

- Le prix moyen du kilogramme de pommes en 2010 et en 2013 est donné par :

$$\bar{p}_0 = \frac{\sum_i p_0^i q_0^i}{\sum_i q_0^i} = \frac{79}{30} = 2,63$$

$$\bar{p}_t = \frac{\sum_i p_t^i q_t^i}{\sum_i q_t^i} = \frac{134,5}{40} = 3,36$$

On en déduit l'indice élémentaire du prix moyen :

$$I_{t/0}^g(\bar{p}) = \frac{\bar{p}_t}{\bar{p}_0} = \frac{3,36}{2,63} = 1,278$$

Le prix moyen du kilogramme de pommes a augmenté de 27,8 % entre 2010 et 2013.

- L'indice élémentaire des quantités s'obtient comme suit :

$$I_{t/0}^q(q) = \frac{q_t}{q_0} = \frac{40}{30} = 1,333$$

et montre que la consommation de pommes de notre ménage s'est accrue de 33,3 % entre 2010 et 2013.

- Le tableau 3.1 fournit les calculs nécessaires à l'obtention des indices des prix de Laspeyres et de Paasche donnés par les relations suivantes :

$$L_{t/0}^g(\bar{p}) = \frac{\sum_i p_t^i q_0^i}{\sum_i p_0^i q_0^i} = \frac{86}{79} = 1,089$$

$$P_{t/0}^g(\bar{p}) = \frac{\sum_i p_0^i q_t^i}{\sum_i p_0^i q_0^i} = \frac{134,5}{121,5} = 1,107$$

Il est également possible d'effectuer les calculs à l'aide des coefficients de pondération (► tableau 3.2) :

- Pour l'indice de Laspeyres, rappelons que les poids sont donnés par :

$$\alpha_0^i = \frac{p_0^i q_0^i}{\sum_i p_0^i q_0^i}$$

D'où :

$$L_{t/0}^g(\bar{p}) = \sum_i \alpha_0^i \frac{p_t^i}{p_0^i} = 1,089$$

- Pour l'indice de Paasche, rappelons que les poids sont donnés par :

$$\alpha_t^i = \frac{p_t^i q_t^i}{\sum_i p_t^i q_t^i}$$

et l'on a donc :

$$P_{t/0}^g(\bar{p}) = \frac{1}{\sum_i \alpha_t^i \frac{p_0^i}{p_t^i}} = \frac{1}{0,903} = 1,107$$

On retrouve bien entendu les mêmes valeurs. La hausse du prix moyen du kilogramme de pommes entre 2010 et 2013 est comprise entre 8,9 % et 10,7 % selon que l'on retient l'indice de Laspeyres ou de Paasche.

- Les indices des quantités de Laspeyres et de Paasche s'obtiennent à l'aide des calculs figurant dans le tableau 3.1 :

$$L_{t/0}^g(q) = \frac{\sum_i p_0^i q_t^i}{\sum_i p_0^i q_0^i} = \frac{121,5}{79} = 1,538$$

$$P_{t/0}^g(q) = \frac{\sum_i p_t^i q_0^i}{\sum_i p_t^i q_0^i} = \frac{134,5}{86} = 1,564$$

- Les indices de Fisher sont donnés par :

$$F_{t/0}^g(\bar{p}) = \sqrt{L_{t/0}^g(\bar{p}) P_{t/0}^g(\bar{p})} = \sqrt{1,089 \times 1,107} = 1,098$$

et

$$F_{t/0}^g(q) = \sqrt{L_{t/0}^g(q) P_{t/0}^g(q)} = \sqrt{1,538 \times 1,564} = 1,551$$

On retrouve bien le fait que la valeur de l'indice de Fisher est comprise entre les valeurs prises par les indices de Laspeyres et de Paasche. En revanche, on constate que la valeur prise par l'indice de Paasche est supérieure à celle de l'indice de Laspeyres, contrairement au cas fréquemment rencontré. Ce résultat n'est toutefois pas surprenant au regard des valeurs associées aux coefficients de pondération α_0^i et α_t^i . En effet, rappelons que l'inégalité selon laquelle l'indice de Paasche est généralement inférieur à l'indice de Laspeyres repose sur l'hypothèse de coefficients de pondération égaux entre les deux dates, notamment pour les variétés Pink Lady et Royal Gala. On notera que les indices élémentaires des prix associés aux variétés Golden et surtout Pink Lady sont ceux qui ont le plus augmenté, ces deux variétés étant celles pour lesquelles les coefficients de pondération se sont accrus. Dans ces conditions, l'inégalité n'a plus de raison d'être vérifiée.

6. Calculons les indices de qualité :

$$S_t = \frac{I_{t/0}^g(\bar{p})}{L_{t/0}^g(\bar{p})} = \frac{P_{t/0}^g(q)}{I_{t/0}^g(q)} = \frac{1,278}{1,089} = \frac{1,564}{1,333} = 1,173$$

et :

$$S_t' = \frac{I_{t/0}^g(\bar{p})}{P_{t/0}^g(\bar{p})} = \frac{L_{t/0}^g(q)}{I_{t/0}^g(q)} = \frac{1,278}{1,107} = \frac{1,538}{1,333} = 1,153$$

L'indice de valeur est donné par :

$$I_{t/0}^V = I_{t/0}^g(\bar{p}) \times S_t \times I_{t/0}^g(q) = P_{t/0}^g(\bar{p}) \times S_t' \times I_{t/0}^g(q)$$

soit, avec nos données :

$$I_{t/0}^V = 1,089 \times 1,173 \times 1,333 = 1,107 \times 1,153 \times 1,333 = 1,703$$

La hausse de 70,3 % de la valeur du panier de fruits du ménage se décompose comme suit :

- une hausse du prix moyen comprise entre 8,9 % et 10,7 % ;
- une hausse de la quantité consommée de 33,3 % ;
- un effet de structure, pour 15,3 % à 17,3 %.

Cet effet de structure provient d'une baisse de la consommation de la variété la moins chère (Royal Gala) au profit de pommes plus onéreuses, la Golden et surtout la Pink Lady. Ce résultat était attendu au vu des valeurs prises par les coefficients de pondération : alors qu'en 2010 le coefficient de pondération le plus élevé était associé à la variété Royal Gala, il est associé à la variété Pink Lady en 2013.

7 Sujet d'examen

▼ Tableau 3.3 Prix et quantités des biens A, B, C et D

Bien h	p_0^h	p_t^h	q_0^h	q_t^h	$i_{t/0}^h$	$p_t^h q_0^h$	$p_0^h q_t^h$	$p_t^h q_t^h$	$p_0^h q_0^h$
A	169	610	210	160	3,61	128 100	35 490	97 600	27 040
B	81	265	220	230	3,27	58 300	17 820	60 950	18 630
C	1 023	2 470	30	18	2,41	74 100	30 690	44 460	18 414
D	32	64	470	470	2,00	30 080	15 040	30 080	15 040
Total						290 580	99 040	233 090	79 124

1. Les quantités consommées pour chacun des groupes de biens en 2014 (notées q_t^h) sont reportées dans la cinquième colonne du tableau 3.3. Elles s'obtiennent comme suit :

- Pour le bien A : $210 \times (1 - 0,238) = 160$
- Pour le bien B : $220 \times (1 + 0,0454) = 230$
- Pour le bien C : $30 \times (1 - 0,40) = 18$
- Pour le bien D : $470 \times (1 + 0) = 470$

2. Les indices élémentaires des prix pour chaque groupe de biens, base 1 en 2000, notés $I_{t/0}^h$, sont donnés par :

$$I_{t/0}^h = \frac{p_t^h}{p_0^h}$$

et les valeurs correspondantes sont reportées dans la sixième colonne du tableau 3.3. Les prix des 4 biens ont augmenté entre 2000 et 2014. Le bien dont le prix a le plus augmenté est le bien A, celui dont le prix a le moins augmenté est le bien D.

3. Les calculs nécessaires à la détermination des indices des prix de Laspeyres et de Paasche figurent dans les quatre dernières colonnes du tableau 3.3. On a ainsi :

$$L_{t/0}^h(p) = \frac{290 580}{99 040} = 2,934$$

$$P_{t/0}^h(p) = \frac{233 090}{79 124} = 2,946$$

On en déduit la valeur de l'indice des prix de Fisher :

$$F_{t/0}^h(p) = \sqrt{293,4 \times 294,6} = 2,940$$

On relève ainsi une augmentation des prix de l'ordre de 194 % entre 2000 et 2014. On constate que la valeur de l'indice de Fisher est bien comprise entre les valeurs des deux autres indices.

8 Sujet d'examen

▼ Tableau 3.4 Prix, quantités et indices élémentaires – Biens a, b, c

h	p_0^h	q_0^h	p_t^h	q_t^h	$p_0^h q_0^h$	$p_t^h q_t^h$	α_0^h (en %)	α_t^h (en %)
a	4	5	8	2,25	20	18	20	17
b	8	5	6	5	40	30	40	28
c	10	4	12	5	40	60	40	55
Total	14	12,25	100	108	100	100	100	100

1. Au regard des valeurs des indices élémentaires des prix figurant dans le tableau 3.8, on constate que sur la période 2010-2014, le prix du bien a a augmenté de 100 %, celui du bien b a diminué de 25 % et celui du bien c a augmenté de 20 %.
2. Les prix et les quantités des trois types de produits vendus en 2014 sont reportés dans le tableau 3.4 et sont obtenus comme suit. On sait que :

$$i_{t/0}^h(p) = \frac{p_t^h}{p_0^h} \times 100$$

et :

$$i_{t/0}^h(q) = \frac{q_t^h}{q_0^h} \times 100$$

On déduit donc de ces relations les expressions suivantes pour les prix et quantités en 2014 :

$$p_t^h = \frac{i_{t/0}^h(p)p_0^h}{100}$$

et :

$$q_t^h = \frac{i_{t/0}^h(q)q_0^h}{100}$$

3. Afin de comparer les structures des chiffres d'affaires en 2010 et en 2014 selon les trois types de produits, il convient de calculer les coefficients budgétaires α_0^h et α_t^h :

$$\alpha_0^h = \frac{p_0^h q_0^h}{\sum_h p_0^h q_0^h}$$

$$\alpha_t^h = \frac{p_t^h q_t^h}{\sum_h p_t^h q_t^h}$$

Au regard des valeurs figurant dans le tableau 3.4, on constate que les parts que représentent les produits *a* et *b* dans le chiffre d'affaires diminuent, alors que celle du produit *c* augmente.

4. Le nombre total de produits vendus est égal à 14 en 2010 et 12,25 en 2014. On en déduit la valeur de l'indice élémentaire de la quantité totale de produits vendus (base 1 en 2010) :

$$I_{t/0}(q) = \frac{12,25}{14} = 0,875$$

La quantité totale vendue a diminué de 12,5 % entre 2010 et 2014.

5. On sait que l'indice de valeur est donné par :

$$I_{t/0}^V = I_{t/0}(\bar{p}) \times I_{t/0}(q)$$

On en déduit la valeur de l'indice élémentaire du prix moyen $I_{t/0}(\bar{p})$ des produits considérés :

$$I_{t/0}(\bar{p}) = \frac{1,080}{0,875} = 1,234$$

Le prix moyen des produits vendus a augmenté de 23,4 % entre 2010 et 2014. Par ailleurs, le rapport $\frac{L_{t/0}(q)}{I_{t/0}(q)}$ correspond à l'indice de qualité S'_t . Cet indice étant égal à 1,133, on en conclut que la variation de la structure des ventes (effet qualité) vers les produits les plus chers a pour conséquence une augmentation du chiffre d'affaires de 13,3 %.

Chapitre 4

1 Schéma de décomposition additif et coefficients saisonniers

- Faux.
- Faux.
- Faux.
- Vrai, encore appelée série désaisonnalisée.
- Faux.

2 Schéma de décomposition additif

- Faux, c'est le cas dans un schéma multiplicatif.
- Faux.
- Vrai.
- Faux.
- Vrai.

3 Principe de conservation des aires et moyennes mobiles

- Faux.
- Vrai.
- Vrai.
- Faux.
- Vrai.

4 Lissage exponentiel

- Faux, le LES ne s'applique qu'au cas de séries avec tendance constante.
- Faux, le LED ne s'applique pas au cas de séries présentant une composante saisonnière.
- Faux, plus le paramètre de lissage est proche de 1, plus le poids des observations récentes est important.
- Faux, le LES et le LED ne peuvent être appliqués si la série comporte une composante saisonnière.
- Vrai.

5 Indice Euro Stoxx

Dans la mesure où les deux mois ne comportent pas le même nombre de jours, il faut calculer les valeurs corrigées des jours ouvrés. En considérant une moyenne de 21 jours par mois, les valeurs corrigées sont données par : $(3093,124/23) \times 21 = 2824,157$ en janvier et $(3085,865/20) \times 21 = 3240,158$ en février. On en déduit donc : $(3240,158 - 2824,157)/2824,157 = 0,1473$.

- Faux.
- Faux.
- Faux.
- Vrai.
- Faux.

6 Étude de l'évolution de l'indice Euro Stoxx 50

Notons préalablement que les techniques de lissage doivent être appliquées à des séries dites stationnaires, c'est-à-dire à des séries dont la moyenne est stable au cours du temps. L'indice boursier étant ici considéré sur une période très courte (un mois), cette propriété est vérifiée.

- L'application de la formule du LES :

$$L_t = \alpha Y_t + (1 - \alpha)L_{t-1}$$

nous permet d'obtenir les valeurs lissées L_t de la série Y_t ainsi qu'elles sont reportées dans le tableau 4.1 pour chacune des trois valeurs considérées du paramètre de lissage α . Nous avons retenu comme valeur de départ pour L_0 la première valeur de la série Y_1 , c'est-à-dire la valeur observée au 01/04/2014.

▼ **Tableau 4.1** Indice Euro Stoxx 50. Application du LES

t	Date	Y_t	L_t	$\alpha = 0,2$		L_t	$\alpha = 0,5$		L_t	$\alpha = 0,74$	
				u_t	u_t^2		u_t	u_t^2		u_t	u_t^2
1	01/04/2014	3186,336	3186,336	0,000	0,000	3186,336	0,000	0,000	3186,336	0,000	0,000
2	02/04/2014	3187,450	3186,559	1,114	1,241	3186,893	1,114	1,241	3187,160	1,114	1,241
3	03/04/2014	3206,759	3190,599	20,200	408,048	3196,826	19,866	394,658	3201,663	19,599	384,107
4	04/04/2014	3230,332	3198,545	39,733	1578,724	3213,579	33,506	1122,652	3222,878	28,669	821,891
5	07/04/2014	3185,967	3196,030	-12,578	158,218	3199,773	-27,612	762,423	3195,564	-36,911	1362,433
6	08/04/2014	3177,658	3192,355	-18,372	337,522	3188,716	-22,115	489,073	3182,314	-17,906	320,621
7	09/04/2014	3182,793	3190,443	-9,562	91,440	3185,754	-5,923	35,076	3182,668	0,479	0,230
8	10/04/2014	3152,864	3182,927	-37,579	1412,177	3169,309	-32,890	1081,769	3160,613	-29,804	888,299
9	11/04/2014	3116,540	3169,650	-66,387	4407,254	3142,925	-52,769	2784,581	3127,999	-44,073	1942,441
10	14/04/2014	3131,566	3162,033	-38,084	1450,370	3137,245	-11,359	129,017	3130,639	3,567	12,723
11	15/04/2014	3091,524	3147,931	-70,509	4971,516	3114,385	-45,721	2090,436	3101,694	-39,115	1529,951
12	16/04/2014	3139,264	3146,198	-8,667	75,120	3126,824	24,879	618,983	3129,496	37,570	1411,521
13	17/04/2014	3155,806	3148,119	9,608	92,319	3141,315	28,982	839,938	3148,965	26,310	692,229
14	18/04/2014	3155,806	3149,657	7,687	59,084	3148,561	14,491	209,984	3154,027	6,841	46,795
15	21/04/2014	3155,806	3150,887	6,149	37,814	3152,183	7,245	52,496	3155,344	1,779	3,163
16	22/04/2014	3199,686	3160,646	48,799	2381,384	3175,935	47,503	2256,507	3188,157	44,342	1966,251
17	23/04/2014	3175,973	3163,712	15,327	234,903	3175,954	0,038	0,001	3179,141	-12,184	148,449
18	24/04/2014	3189,809	3168,931	26,097	681,066	3182,881	13,855	191,966	3187,035	10,668	113,810
19	25/04/2014	3147,397	3164,624	-21,534	463,722	3165,139	-35,484	1259,143	3157,703	-39,638	1571,193
20	28/04/2014	3165,837	3164,867	1,213	1,470	3165,488	0,698	0,487	3163,722	8,134	66,163
21	29/04/2014	3208,685	3173,631	43,818	1920,026	3187,087	43,197	1865,972	3196,995	44,963	2021,658
22	30/04/2014	3198,387	3178,582	24,756	612,883	3192,737	11,300	127,700	3198,025	1,392	1,939
Somme				21376,30		16314,10		15307,11			

2. La valeur prévue \hat{Y}_t de Y_t est donnée par :

$$\hat{Y}_t = L_t$$

et l'erreur de prévision u_t est donnée par :

$$u_t = Y_t - \hat{d}_{t-1} = Y_t - L_{t-1}$$

On déduit, d'après les résultats reportés dans le tableau 4.1, les valeurs de la somme des carrés des erreurs de prévision :

- Pour $\alpha = 0,2$, on a : $\sum u_t^2 = 21376,30$
 - Pour $\alpha = 0,5$, on a : $\sum u_t^2 = 16314,10$
 - Pour $\alpha = 0,74$, on a : $\sum u_t^2 = 15307,11$
3. Si l'on retient pour α la valeur qui minimise la somme des carrés des erreurs de prévision, il convient de choisir la valeur $\alpha = 0,74$. Ce résultat, qui consiste à retenir une valeur relativement élevée de α , n'est pas surprenant dans la mesure où nous étudions une série boursière, qui constitue un exemple typique de série relativement heurtée.
4. La valeur prévue de la série pour le 01/05/2014, c'est-à-dire pour $t = 23$ est donnée par la dernière valeur obtenue pour la série lissée :

$$\hat{Y}_{23} = L_{22}$$

soit 3198,025 pour une valeur du paramètre de lissage égale à 0,74. Dans la mesure où la prévision est indépendante de l'horizon dans le cas du LES, on a :

$$\hat{Y}_{24} = L_{22}$$

et la valeur prévue au 30/04/2014 de la série pour le 02/05/2014 est donc égale à 3198,025 pour $\alpha = 0,74$.

7 Étude des ventes de voitures neuves en France

- Le tableau 4.2 est un tableau à double entrée dans lequel on a reporté en ligne les années et en colonne les trimestres. Ce tableau est appelé tableau de Buys-Ballot.
- Il est possible de classer, pour chaque année, les trimestres en fonction des valeurs décroissantes des ventes de voitures. Le tableau 4.3 reporte les trimestres ainsi classés, le chiffre figurant dans chaque case correspondant au numéro du trimestre. On constate que le deuxième trimestre est en général classé en première position alors que le troisième trimestre est systématiquement classé en dernière position. Cela témoigne d'un pic dans les ventes de voitures durant le deuxième trimestre de chaque année et d'un creux au cours du troisième trimestre, indiquant l'existence d'une saisonnalité dans la série des ventes de voitures neuves en France. Ce résultat est cohérent avec ce que l'on observe sur le marché automobile français puisque les ventes de voitures neuves sont en effet caractérisées par une forte hausse au mois de juin et un creux durant le mois d'août.
- Les valeurs demandées de la moyenne et de l'écart-type sont reportées dans les deux dernières colonnes du tableau de Buys-Ballot (tableau 4.2).
- Notons \bar{Y}_i la moyenne des ventes de voitures pour chaque année i et σ_i l'écart-type de cette même série. Le coefficient a de la droite de régression de σ_i sur \bar{Y}_i est donné par :

$$a = \frac{\text{Cov}(\bar{Y}_i, \sigma_i)}{V(\bar{Y}_i)} = \frac{340\,638\,745}{1\,651\,883\,555} = 0,2062$$

Le coefficient de la pente de la droite de régression étant significativement différent de zéro, cela témoigne de l'existence d'un lien entre la moyenne et l'écart-type des ventes

de voitures. Cela peut être confirmé par le calcul du coefficient de corrélation entre les deux séries :

$$\begin{aligned} r(\bar{Y}_i, \sigma_i) &= \frac{\text{Cov}(\bar{Y}_i, \sigma_i)}{\sigma_{\bar{Y}_i} \times \sigma_{\sigma_i}} \\ &= \frac{340\,638\,745}{40\,643,3704 \times 11\,497,1203} = 0,7290 \end{aligned}$$

L'existence d'un lien entre les valeurs annuelles de la moyenne et celles de l'écart-type conduit à retenir un schéma de décomposition multiplicatif.

▼ **Tableau 4.2** Tableau de Buys-Ballot

	Trimestre 1	Trimestre 2	Trimestre 3	Trimestre 4	Moyenne	Écart-type
2006	526 502	582 727	409 138	482 182	500 137,25	63 481,14
2007	519 191	561 417	442 753	541 182	516 135,75	44 922,35
2008	526 121	602 775	446 869	474 518	512 570,75	59 340,43
2009	505 456	625 859	482 186	655 170	567 167,75	74 530,69
2010	590 869	597 383	432 029	589 905	552 546,5	69 640,24
2011	642 627	557 981	422 728	537 592	540 232	78 444,12
2012	501 879	520 836	370 006	464 292	464 253,25	58 094,43
2013	431 388	480 730	367 112	477 721	439 237,75	46 006,28
Moyenne	530 504,13	566 213,50	421 602,63	527 820,25		
Écart-type	58 789,33	44 152,99	36 491,19	63 063,71		

▼ **Tableau 4.3** Classement des trimestres en fonction des valeurs décroissantes des ventes

Année	T1	T2	T3	T4
2006	2	1	4	3
2007	2	4	1	3
2008	2	1	4	3
2009	4	2	1	3
2010	2	1	4	3
2011	1	2	4	3
2012	2	1	4	3
2013	2	4	1	3

Chapitre 5

1 Expérience aléatoire

- Vrai.
- Vrai.
- Vrai.
- Faux. L'univers des possibles se compose d'événements élémentaires, mais aussi d'événements composés (union ou intersection de singletons).
- Faux. L'univers des possibles peut être fini ou infini (dénombrable ou non dénombrable).

2 Événement

- Vrai.
- Vrai.
- Faux. L'intersection de deux événements disjoints est l'ensemble vide.
- Faux. L'union d'un événement et son complémentaire ne correspond pas nécessairement à l'univers des possibles. Elle peut correspondre à un autre événement.
- Vrai. On peut toujours considérer que l'ensemble vide fait partie d'un ensemble d'événements quelconque.

3 Ensemble des parties et tribu

- Faux. Ceci n'est vrai que dans le cas particulier où l'univers des possibles ne comprend qu'un seul résultat.
- Vrai.
- Vrai.
- Vrai. C'est la propriété de stabilité par passage au complémentaire.
- Faux. D'après la propriété de stabilité par réunion dénombrable, c'est l'union d'événements disjoints appartenant à une tribu, qui appartient elle aussi à cette tribu.

4 Mesure de probabilité

- Faux. Une mesure de probabilité peut être définie sur un univers probabilisable.
- Vrai.
- Vrai. Dans le cas d'un univers fini ou dénombrable, l'ensemble des parties est une tribu. L'univers et l'ensemble des parties forment alors un univers probabilisable.
- Vrai.
- Faux. Si $B \subset A$, alors $\Pr(B) \leq \Pr(A)$.

5 Probabilité et dénombrement

- Une grille-réponses est une suite de 4 réponses. Pour chaque question il y a 5 réponses possibles. Il y a donc $5^4 = 625$ grilles-réponses possibles.
- L'événement A « répondre au hasard correctement à au moins 2 questions » est réalisé si l'étudiant répond correctement à 2, 3 ou 4 questions. Notons A_i pour $i \in \{0, 1, 2, 3, 4\}$ l'événement « répondre au hasard correctement à i questions ». On a donc $A = A_2 \cup A_3 \cup A_4$. Si l'étudiant répond au hasard, la probabilité associée à l'événement A_i est égale à :

$$\Pr(A_i) = \frac{C_4^i \times 4^{4-i}}{5^4}$$

En effet, si l'étudiant répond correctement à i questions parmi 4, il a $4 - i$ questions fausses. Pour chaque question

fausse, il 4 réponses (fausses) possibles sur les 5 choix proposés. Ainsi, on obtient

$$\Pr(A_4) = \frac{C_4^4 \times 4^{4-4}}{5^4} = \frac{1}{625}$$

$$\Pr(A_3) = \frac{C_4^3 \times 4^{4-3}}{5^4} = \frac{16}{625}$$

$$\Pr(A_2) = \frac{C_4^2 \times 4^{4-2}}{5^4} = \frac{96}{625}$$

Ainsi, la probabilité que l'étudiant obtienne au moins la moyenne sur cet exercice est égale à environ 18 %.

$$\Pr(A) = \Pr(A_2) + \Pr(A_3) + \Pr(A_4) = \frac{113}{625} = 0,1808$$

6 Probabilité et dénombrement

L'univers des possibles Ω correspond à l'ensemble des combinaisons de 4 billets parmi les 50. Il y en a :

$$\text{card}(\Omega) = C_{50}^4 = \binom{50}{4} = \frac{50!}{(50-4)! \times 4!} = 230\,300$$

On ne gagne rien si les 4 billets achetés figurent parmi les 48 billets perdants. La probabilité de ne rien gagner, événement noté A , est donc égale à :

$$\Pr(A) = \frac{C_{48}^4}{C_{50}^4} = \frac{1}{230\,300} \times \frac{48!}{(48-4)! \times 4!} = \frac{194\,580}{230\,300} = 0,8449$$

L'événement « gagner au moins un lot » correspond à l'événement complémentaire de l'événement « ne rien gagner ». Par conséquent, la probabilité associée est égale à :

$$\Pr(\overline{A}) = 1 - 0,8449 = 0,1551$$

Il y a 15,51 % de chances de gagner au moins un lot dans cette loterie.

7 Suite d'événements et probabilité totale

- On note F_n l'événement « fumer le jour n » et \overline{F}_n l'événement complémentaire « ne pas fumer le jour n ». L'ensemble d'événements $\{F_n, \overline{F}_n\}$ est un système complet, on peut donc appliquer la formule des probabilités totales pour déterminer la probabilité le jour $n+1$:

$$\begin{aligned} \Pr(F_{n+1}) &= \Pr(F_{n+1} | F_n) \times \Pr(F_n) + \Pr(F_{n+1} | \overline{F}_n) \\ &\quad \times \Pr(\overline{F}_n) \\ &= 0,8 \times \Pr(F_n) + 0,1 \times (1 - \Pr(F_n)) \end{aligned}$$

On obtient donc une formule de récurrence du type :

$$\Pr(F_{n+1}) = 0,7 \times \Pr(F_n) + 0,1$$

- On considère la suite $P_{n+1} = 0,7P_n + 0,1$. Une solution particulière est donnée par la solution de l'équation :

$$x = 0,7x + 0,1 \iff x = \frac{1}{3}$$

La solution générale est la solution de $P_{n+1} = 0,7P_n$, c'est-à-dire $P_n = 0,7^{n-1}P_1$ où P_1 désigne la probabilité fumer le premier jour. La probabilité de fumer le jour n est donc donnée par :

$$\Pr(F_n) = 0,7^{n-1} \times \Pr(F_1) + \frac{1}{3}$$

Lorsque $n \rightarrow \infty$, on obtient $\Pr(F_n) = 1/3$. Il y a donc 66,66 % de chances que cette personne finisse par s'arrêter de fumer.

8 Probabilité et indépendance

1. Pour un individu de la population, la probabilité d'achat des deux produits est égale à :

$$\Pr(A \cap B) = p_A \times p_B$$

car les ventes des deux produits sont indépendantes.

2. Pour un individu de la population, la probabilité d'achat de l'un ou de l'autre produit est égale à :

$$\begin{aligned}\Pr(A \cup B) &= \Pr(A) + \Pr(B) - \Pr(A \cap B) \\ &= p_A + p_B - p_A \times p_B\end{aligned}$$

9 Probabilité et dénombrement

1. Soit $p = N_R/N$ la probabilité de tirer une boule rouge. À chaque tirage il y a deux éventualités : soit tirer une boule rouge (événement R), soit tirer une boule blanche (événement B) avec $\Pr(R) = p$ et $\Pr(B) = 1 - p$ ce qui définit un système complet sur l'univers des possibilités $\Omega = \{B, R\}$. Faire n tirages avec remise revient à considérer l'espace produit probabilisable $\Omega^n = \{(x_1, \dots, x_n) ; x_i = B \text{ ou } R \forall i\}$ muni de la probabilité p^n . Soit l'événement E_k « on tire k boules rouges » composé des éléments $\omega = (x_1, \dots, x_n)$ tels que parmi les singletons x_i il y a k fois la lettre R et $n - k$ fois la lettre B . Il y a donc C_n^k événements correspondants, c'est-à-dire C_n^k façons de disposer k boules rouges à n places numérotées. Chacune de ces éventualités a pour probabilité $p^k (1 - p)^{n-k}$ car les tirages sont indépendants.

$$P(E_k) = C_n^k \left(\frac{N_R}{N}\right)^k \left(1 - \frac{N_R}{N}\right)^{n-k}$$

2. On vérifie que la probabilité de tirer $k = 10$ boules rouges est égale à 9,85 %, puisque :

$$P(E_{10}) = C_{50}^{10} \left(\frac{25}{100}\right)^{10} \left(1 - \frac{25}{100}\right)^{50-10} = 9,85 \%$$

10 Probabilité conditionnelle

On note D l'événement « lot défectueux » et \bar{D} l'événement « lot valide ». On note T l'événement « le test conduit au rejet du lot » et \bar{T} l'événement « le test ne conduit pas au rejet du lot ».

1. La probabilité qu'un lot soit effectivement défectueux si le test conduit au rejet du lot est égale à :

$$\Pr(D|T) = \frac{\Pr(D \cap T)}{\Pr(T)} = \frac{\Pr(T|D) \times \Pr(D)}{\Pr(T)}$$

Or la probabilité qu'un test conduise au rejet d'un lot est égale à :

$$\begin{aligned}\Pr(T) &= \Pr(T|D) \times \Pr(D) + \Pr(T|\bar{D}) \times \Pr(\bar{D}) \\ &= 0,98 \times 0,05 + 0,04 \times 0,95 = 0,087\end{aligned}$$

On en déduit que :

$$\Pr(D|T) = \frac{\Pr(T|D) \times \Pr(D)}{\Pr(T)} = \frac{0,98 \times 0,05}{0,087} = 0,5632$$

2. La probabilité qu'un lot soit valide si le test conduit au rejet du lot est égale à la probabilité complémentaire :

$$\Pr(\bar{D}|T) = 1 - \Pr(D|T) = 1 - 0,5632 = 0,4368$$

3. La probabilité qu'un lot soit défectueux si le test ne conduit pas au rejet du lot est égale à :

$$\begin{aligned}\Pr(D|\bar{T}) &= \frac{\Pr(D \cap \bar{T})}{\Pr(\bar{T})} = \frac{\Pr(\bar{T}|D) \times \Pr(D)}{\Pr(\bar{T})} \\ &= \frac{0,02 \times 0,05}{1 - 0,087} = 0,0011\end{aligned}$$

4. La probabilité qu'un lot soit valide si le test ne conduit pas au rejet du lot est égale à la probabilité complémentaire :

$$\Pr(\bar{D}|\bar{T}) = 1 - \Pr(D|\bar{T}) = 1 - 0,0011 = 0,9989$$

11 Probabilité conditionnelle

Soit A l'événement « l'étudiant est issu de la formation A » et B l'événement « l'étudiant est issu de la formation B ». Soit M l'événement « l'étudiant a obtenu une mention bien ». La probabilité qu'un étudiant ayant obtenu une mention bien soit issu de la formation A est égale à :

$$\Pr(A|M) = \frac{\Pr(A \cap M)}{\Pr(M)} = \frac{\Pr(M|A) \times \Pr(A)}{\Pr(M)}$$

La probabilité d'obtenir une mention est égale à :

$$\begin{aligned}\Pr(M) &= \Pr(M|A) \times \Pr(A) + \Pr(M|B) \times \Pr(B) \\ &= 0,30 \times 0,70 + 0,20 \times 0,30 = 0,27\end{aligned}$$

Par conséquent :

$$\Pr(A|M) = \frac{0,30 \times 0,70}{0,27} = 0,7778$$

12 Probabilité conditionnelle

1. Nous avons :

$$\begin{aligned}\Pr(A_{n+1}) &= \Pr(A_{n+1}|A_n) \times \Pr(A_n) + \Pr(A_{n+1}|\bar{A}_n) \\ &\quad \times \Pr(\bar{A}_n) \\ &= \frac{1}{8} \Pr(A_n) + \frac{3}{8} (1 - \Pr(A_n)) \\ &= \frac{3}{8} - \frac{2}{8} \Pr(A_n)\end{aligned}$$

2. La relation précédente définit une suite géométrique. Une solution particulière est donnée par

$$x = \frac{3}{8} - \frac{2}{8}x \iff x = \frac{3}{10}$$

On en déduit que la probabilité de l'événement « l'individu se rend au cinéma le jour i » est égale à :

$$\Pr(A_n) = \frac{3}{10} + p_1 \left(-\frac{2}{8}\right)^{n-1}$$

avec $\Pr(A_1) = p_1$.

Chapitre 6

1 Variable aléatoire

- a. Faux. Une variable aléatoire est une application mesurable.
- b. Faux. Seules les variables quantitatives peuvent être continues.
- c. Vrai.
- d. Vrai. Une variable aléatoire continue est définie sur un support infini non dénombrable.
- e. Faux. Pas nécessairement, une variable aléatoire discrète peut être définie sur un support de dimension finie.

2 Fonction de densité, de masse et de répartition

- a. Vrai.
- b. Faux. La densité est toujours positive, mais elle peut être supérieure à 1 en un point.
- c. Faux. La fonction de densité est la dérivée de la fonction de répartition.
- d. Vrai. C'est la définition de la fonction de répartition.
- e. Faux. La fonction de masse ne correspond pas à la dérivée de la fonction de répartition.

3 Fonction de répartition et quantile

- a. Faux. Un quantile est une réalisation.
- b. Vrai.
- c. Faux. Les quantiles sont définis sur le support de la loi de probabilité.
- d. Faux. Pour de nombreuses lois usuelles, la fonction de répartition n'a pas d'expression analytique.
- e. Vrai.

4 Indépendance

- a. Faux. La covariance n'est qu'une mesure de dépendance linéaire. Il peut exister d'autres formes de dépendance non-linéaire entre ces variables.
- b. Faux. Si la densité jointe est égale au produit des densités marginales, les variables sont indépendantes.
- c. Vrai.
- d. Vrai.
- e. Vrai.

5 Fonction de répartition

Soit X une variable aléatoire réelle admettant pour fonction de répartition :

$$F_X(x) = \begin{cases} 0 & \text{si } x < 0 \\ 1/4 & \text{si } 0 \leq x < 1 \\ 3/4 & \text{si } 1 \leq x < 2 \\ 1 & \text{si } x \geq 2 \end{cases}$$

1.

$$\begin{aligned} \Pr\left(-\frac{1}{2} < X < \frac{1}{2}\right) &= \Pr\left(X < \frac{1}{2}\right) - \Pr\left(X \leq -\frac{1}{2}\right) \\ &= F_X\left(\frac{1}{2}\right) - F_X\left(-\frac{1}{2}\right) \\ &= \frac{1}{4} - 0 = \frac{1}{4} \end{aligned}$$

2.

$$\begin{aligned} \Pr\left(-\frac{1}{2} < X < \frac{3}{2}\right) &= \Pr\left(X < \frac{3}{2}\right) - \Pr\left(X \leq -\frac{1}{2}\right) \\ &= F_X\left(\frac{3}{2}\right) - F_X\left(-\frac{1}{2}\right) \\ &= \frac{3}{4} - 0 = \frac{3}{4} \end{aligned}$$

3.

$$\begin{aligned} \Pr(X > 3) &= 1 - \Pr(X \leq 3) \\ &= 1 - F_X(3) \\ &= 1 - 1 \\ &= 0 \end{aligned}$$

6 Espérance

1. Par définition :

$$\begin{aligned} \mathbb{E}((X - a)^2) &= \mathbb{V}(X - a) + (\mathbb{E}(X - a))^2 \\ &= \mathbb{V}(X) + (\mathbb{E}(X - a))^2 \end{aligned}$$

2. On pose $f(a) = \mathbb{E}((X - a)^2)$. Cette fonction atteint son minimum lorsque $\partial f(a)/\partial a = 0$ et $\partial^2 f(a)/\partial a^2 > 0$.

$$\frac{\partial f(a)}{\partial a} = \mathbb{E}(-2(X - a)) = \mathbb{E}(-2X + 2a) = -2\mathbb{E}(X) + 2a = 0$$

$$\frac{\partial^2 f(a)}{\partial a^2} = 2 > 0$$

On en déduit que $a = \mathbb{E}(X)$ est un minimum. Donc la valeur de a pour laquelle l'espérance $\mathbb{E}((X - a)^2)$ est minimum est égale à $\mathbb{E}(X)$.

7 Fonction de densité

1. La fonction $f_X(x)$ définie par :

$$f_X(x) = \begin{cases} 0 & \text{si } x \leq \theta \\ \exp(-(x - \theta)) & \text{si } x > \theta \end{cases}$$

est positive $\forall x \in \mathbb{R}$. Montrons qu'elle intègre à l'unité sur le support $X(\Omega) = \mathbb{R}$.

$$\begin{aligned} \int_{-\infty}^{+\infty} f_X(x) dx &= \int_{\theta}^{+\infty} \exp(-(x - \theta)) dx \\ &= \int_0^{+\infty} \exp(-x) dx \\ &= \Gamma(1) = 0! = 1 \end{aligned}$$

où $\Gamma(\cdot)$ désigne la fonction gamma. Donc $f_X(x)$ satisfait les conditions d'une fonction de densité.

2. Par définition de la fonction de répartition, $\forall x \in \mathbb{R}$:

$$\begin{aligned} F_X(x) &= \int_{-\infty}^x f_X(u) du = \int_{\theta}^x \exp(-(u - \theta)) du \\ &= \exp(\theta) \int_{\theta}^x \exp(-u) du \\ &= \exp(\theta) \times [-\exp(-u)]_{\theta}^x \\ &= 1 - \exp(-(x - \theta)) \end{aligned}$$

3. Par définition, l'espérance $\mathbb{E}(X)$ est égale à :

$$\mathbb{E}(X) = \int_{-\infty}^{+\infty} x f_X(x) dx = \int_{\theta}^{+\infty} x \exp(-(x-\theta)) dx$$

En posant $u = x - \theta \iff x = u + \theta$, $du = dx$ il vient :

$$\begin{aligned} \mathbb{E}(X) &= \int_{\theta}^{+\infty} (u + \theta) \exp(-u) du \\ &= \int_{\theta}^{+\infty} u \exp(-u) du + \int_{\theta}^{+\infty} \theta \exp(-u) du \\ &= \Gamma(2) + \theta \Gamma(1) = 1! + \theta \times 0! = 1 + \theta \end{aligned}$$

4. De la même façon :

$$\mathbb{E}(X^2) = \int_{-\infty}^{+\infty} x^2 f_X(x) dx = \int_{\theta}^{+\infty} x^2 \exp(-(x-\theta)) dx$$

En posant $u = x - \theta \iff x = u + \theta$, $du = dx$ il vient :

$$\begin{aligned} \mathbb{E}(X^2) &= \int_{\theta}^{+\infty} (u + \theta)^2 \exp(-u) du \\ &= \int_{\theta}^{+\infty} (u^2 + 2u\theta + \theta^2) \exp(-u) du \\ &= \int_{\theta}^{+\infty} u^2 \exp(-u) du + 2\theta \int_{\theta}^{+\infty} u \exp(-u) du + \theta^2 \int_{\theta}^{+\infty} \exp(-u) du \\ &= \Gamma(3) + 2\theta \Gamma(2) + \theta^2 \Gamma(1) \\ &= \theta^2 + 2\theta + 2 \end{aligned}$$

On en déduit la variance :

$$\begin{aligned} \mathbb{V}(X) &= \mathbb{E}(X^2) - \mathbb{E}(X)^2 \\ &= \theta^2 + 2\theta + 2 - (1 + \theta)^2 \\ &= \theta^2 + 2\theta + 2 - 1 - 2\theta - \theta^2 = 1 \end{aligned}$$

8 Transformée de variable aléatoire

1. Puisque les variables X et Y sont indépendantes :

$$\Pr(Z > k) = \Pr((X > k) \cap (Y > k)) = \Pr(X > k) \times \Pr(Y > k)$$

Par ailleurs, nous savons que :

$$\begin{aligned} \Pr(X > k) &= \Pr(X \geq k + 1) \\ &= \Pr(X = k + 1) + \Pr(X = k + 2) + \dots \\ &= \sum_{i=k+1}^{+\infty} \Pr(X = i) \\ &= \sum_{i=k+1}^{+\infty} q^{i-1} p \end{aligned}$$

En posant $j = i - k - 1$, on obtient :

$$\Pr(X > k) = \sum_{j=0}^{+\infty} q^{j+k} p = q^k p \sum_{j=0}^{+\infty} q^j = \frac{q^k p}{1 - q} = q^k$$

De la même façon $\Pr(Y > k) = q^k$. On en conclut que :

$$\Pr(Z > k) = q^{2k} \quad \forall k \in \mathbb{N}^*$$

2. Par définition :

$$\Pr(Z \geq k) = \Pr((Z > k) \cup (Z = k))$$

Puisque les deux événements $(Z > k)$ et $(Z = k)$ sont incompatibles, on a :

$$\Pr(Z \geq k) = \Pr(Z > k) + \Pr(Z = k)$$

ou encore :

$$\Pr(Z = k) = \Pr(Z \geq k) - \Pr(Z > k)$$

Or puisque Z est définie sur \mathbb{N}^* , on a $\Pr(Z \geq k) = \Pr(Z > k - 1)$. Par conséquent :

$$\Pr(Z = k) = \Pr(Z > k - 1) - \Pr(Z > k)$$

3. D'après les résultats précédents, on montre que :

$$\begin{aligned} \Pr(Z = k) &= \Pr(Z > k - 1) - \Pr(Z > k) \\ &= q^{2(k-1)} - q^{2k} \\ &= q^{2k-2} (1 - q^2) \\ &= (q^2)^{k-1} (1 - q^2) \end{aligned}$$

Donc Z suit une loi géométrique de paramètre $(1 - q^2)$.

9 Variable aléatoire discrète

1. Au premier tour, il y a 2 rois rouges dans le paquet de $2n$ cartes. La probabilité de l'événement E_1 « le premier roi rouge obtenu est la 1^{re} carte retournée » est donc égale à :

$$\Pr(E_1) = \frac{2}{2n}$$

Au second tour, la probabilité conditionnelle à l'échec du premier tour, c'est-à-dire à \overline{E}_1 , de tirer un roi rouge parmi les $2n - 1$ cartes restantes est égale à :

$$\Pr(E_2 | \overline{E}_1) = \frac{2}{2n - 1}$$

La probabilité d'avoir échoué au premier tirage est égale à :

$$\Pr(\overline{E}_1) = 1 - \Pr(E_1) = \frac{2n - 2}{2n}$$

Dès lors, la probabilité de l'événement E_2 est égale à :

$$\begin{aligned} \Pr(E_2) &= \Pr(E_2 | \overline{E}_1) \times \Pr(\overline{E}_1) \\ &= \frac{2}{2n - 1} \times \frac{2n - 2}{2n} = \frac{2n - 2}{n(2n - 1)} \end{aligned}$$

Au troisième tour, la probabilité de l'événement E_3 peut s'écrire sous la forme :

$$\begin{aligned} \Pr(E_3) &= \Pr(E_3 | \overline{E}_2, \overline{E}_1) \times \Pr(\overline{E}_2) \times \Pr(\overline{E}_1) \\ &= \frac{2}{2n - 2} \times \frac{2n - 3}{2n - 1} \times \frac{2n - 2}{2n} = \frac{2n - 3}{n(2n - 1)} \end{aligned}$$

De façon générale, on montre que :

$$\begin{aligned} \Pr(E_k) &= \Pr(E_k | \overline{E}_{k-1}, \dots, \overline{E}_1) \times \Pr(\overline{E}_{k-1}) \times \dots \times \Pr(\overline{E}_1) \\ &= \frac{2n - k}{n(2n - 1)} \end{aligned}$$

On vérifie que la probabilité de tirer le premier roi rouge au $k^{\text{ème}}$ tour est nulle, puisque cet événement est impossible sachant qu'il y a deux rois rouges dans le jeu de cartes, et que nécessairement le premier roi rouge a été tiré entre le premier tour et le $k - 1^{\text{ème}}$ tour.

$$\Pr(E_{2n}) = \frac{2n - 2n}{n(2n - 1)} = 0$$

2. Soit $E(\Omega) = \{E_1, \dots, E_{2n}\}$ l'univers des résultats possibles. On définit une variable aléatoire X correspondant au gain : à chaque tour le joueur perd 1 euro et il gagne a euros s'il tire un roi rouge. S'il tire le roi au 1^{er} tour, il gagne $a - 1$ euros, s'il tire le roi au deuxième tour il gagne $a - 2$ euros. Ainsi, s'il tire un roi rouge au $k^{\text{ème}}$ tour, il gagne $a - k$ euros. La variable X est donc définie sur le support $X(\Omega) = \{a - 1, a - 2, \dots, a - 2n\}$.
3. Par définition :

$$\Pr(X = a - k) = \Pr(E_k) = \frac{2n - k}{n(2n - 1)}$$

4. On vérifie que :

$$\begin{aligned} \sum_{k=1}^{2n} \Pr(X = a - k) &= \frac{1}{n(2n - 1)} \sum_{k=1}^{2n} (2n - k) \\ &= \frac{1}{n(2n - 1)} \left(4n^2 - \frac{2n(2n + 1)}{2} \right) \\ &= \frac{1}{n(2n - 1)} \left(\frac{4n^2 - 2n}{2} \right) \\ &= 1 \end{aligned}$$

Le système défini par les réalisations $X(\Omega)$ est un système complet.

10 Variable aléatoire continue

1. Pour que la fonction $f_X(x)$ soit une fonction de densité, elle doit être positive $\forall x \in \mathbb{R}$ et son intégrale sur \mathbb{R} doit être égale à 1. La positivité est assurée dès lors que $a > 0$. Par ailleurs :

$$\begin{aligned} \int_{-\infty}^{+\infty} f_X(x) dx &= a \int_{-\infty}^{+\infty} \exp(-|x|) dx \\ &= 2a \int_0^{+\infty} \exp(-x) dx = 1 \end{aligned}$$

On doit donc avoir $a = 1/2$. La fonction $f_X(x)$ est une densité si :

$$f_X(x) = \frac{1}{2} \exp(-|x|)$$

2. La fonction de répartition est définie par :

$$F_X(x) = \Pr(X \leq x) = \frac{1}{2} \int_{-\infty}^x \exp(-|u|) du \quad \forall x \in \mathbb{R}$$

Lorsque $x \leq 0$, on a :

$$\begin{aligned} F_X(x) &= \frac{1}{2} \int_{-\infty}^x \exp(-|u|) du = \frac{1}{2} \int_{-\infty}^x \exp(u) du \\ &= \frac{1}{2} [\exp(u)]_{-\infty}^x = \frac{\exp(x)}{2} \end{aligned}$$

Lorsque $x > 0$, on a :

$$\begin{aligned} F_X(x) &= \frac{1}{2} \int_{-\infty}^x \exp(-|u|) du \\ &= \frac{1}{2} \int_{-\infty}^0 \exp(u) du + \frac{1}{2} \int_0^x \exp(-u) du \\ &= \frac{1}{2} [\exp(u)]_{-\infty}^0 + \frac{1}{2} [-\exp(-u)]_0^x \\ &= \frac{1}{2} - \frac{\exp(-x)}{2} + \frac{1}{2} = 1 - \frac{\exp(-x)}{2} \end{aligned}$$

On obtient au final :

$$F_X(x) = \begin{cases} \frac{\exp(x)}{2} & \text{si } x \leq 0 \\ 1 - \frac{\exp(-x)}{2} & \text{si } x > 0 \end{cases}$$

3. Par définition :

$$\begin{aligned} \mathbb{E}(X) &= \frac{1}{2} \int_{-\infty}^{+\infty} x \exp(-|x|) dx \\ &= \frac{1}{2} \int_{-\infty}^0 x \exp(x) dx + \frac{1}{2} \int_0^{+\infty} x \exp(-x) dx \end{aligned}$$

En posant un changement de variable $z = -x$ sur la première intégrale, il vient :

$$\mathbb{E}(X) = -\frac{1}{2} \int_0^{+\infty} z \exp(-z) dz + \frac{1}{2} \int_0^{+\infty} x \exp(-x) dx = 0$$

Chapitre 7

1 Loi usuelles discrètes

- a. Faux. La loi binomiale correspond au nombre de succès obtenus dans n tirages successifs et indépendants d'une expérience de Bernoulli.
- b. Vrai.
- c. Vrai.
- d. Vrai.
- e. Vrai.

2 Loi usuelles continues

- a. Vrai.
- b. Faux. Cette propriété n'est vraie que pour la loi exponentielle.
- c. Faux. Sa skewness est positive.
- d. Faux. Ceci n'est vrai que si $\mu = 0$ et $\sigma^2 = 1$.
- e. Vrai.

3 Loi normale

- a. Vrai.
- b. Faux. $\Phi^{-1}(0,025) = -1,96$. De façon générale $\Phi^{-1}(\alpha) = -\Phi^{-1}(1 - \alpha)$.
- c. Vrai. La fonction de répartition inverse de la loi normale centrée réduite $\Phi^{-1}(\alpha)$ est croissante sur $[0, 1]$.
- d. Vrai. Puisque la fonction de densité de la loi normale centrée réduite est symétrique par rapport à 0, on a $\Phi(x) = -\Phi(-x)$.
- e. Vrai.

4 Loi de Student

- Faux. La kurtosis peut être égale à 3 si le nombre de degrés de liberté de la loi de Student tend vers l'infini.
- Vrai.
- Faux. Puisque la fonction de densité de la loi de Student est symétrique par rapport à 0, on a $F_X(x) = -F_X(-x)$.
- Vrai.
- Faux. Car si $\alpha < 0,5$ le quantile d'ordre α est nécessairement négatif. Si $X \sim t(4)$, le vrai quantile est $F_X^{-1}(0,05) = -2,1318$.

5 Loi usuelles discrètes

- Par définition $\Pr(X = 4) = 0$ car $4 \notin X(\Omega)$.

$$\Pr(X = 1) = \frac{1}{3 - (-2) + 1} = \frac{1}{6}$$

$$\Pr(X \leq 2) = F_X(2) = \frac{2 - (-2) + 1}{3 - (-2) + 1} = \frac{5}{6}$$

Ou de façon équivalente :

$$\Pr(X \leq 2) = \sum_{i=-2}^2 \Pr(X = i) = \sum_{i=-2}^2 \frac{1}{6} = 5 \times \frac{1}{6} = \frac{5}{6}$$

- Si $X \sim \mathcal{B}(10; 0,4)$ alors son support est $X(\Omega) = \{0, 1, \dots, 10\}$. La réalisation $x = 2,57$ n'appartient pas à $X(\Omega)$ donc $\Pr(X = 2,57) = 0$.

$$\Pr(X = 2) = \binom{10}{2} \times 0,4^2 \times 0,6^{10-2} = 0,1209$$

Soit $F_X(x)$ la fonction de répartition de la loi $\mathcal{B}(10; 0,4)$:

$$\begin{aligned} \Pr(2,57 < X \leq 4) &= \Pr(X \leq 4) - \Pr(X \leq 2,57) \\ &= F_X(4) - F_X(2,57) \end{aligned}$$

Nous savons que :

$$\begin{aligned} F_X(2,57) &= \sum_{k=0}^{[2,57]} \binom{10}{k} \times 0,4^k \times (1-0,4)^{10-k} \\ &= \sum_{k=0}^2 \binom{10}{k} \times 0,4^k \times (1-0,4)^{10-k} \\ &= F_X(2) \end{aligned}$$

Pour une loi $\mathcal{B}(10; 0,4)$, on lit dans la table fournie en annexe que :

$$\begin{aligned} F_X(4) &= \Pr(X \leq 4) = 0,6331 \\ F_X(2) &= \Pr(X \leq 2) = 0,1673 \end{aligned}$$

On en déduit que :

$$\Pr(2,57 < X \leq 4) = 0,6331 - 0,1673 = 0,4658$$

- Si $X \sim \mathcal{P}(4)$ alors son support est $X(\Omega) = \mathbb{N}$. La réalisation $x = 2,57$ n'appartient pas à $X(\Omega)$ donc $\Pr(X = 2,57) = 0$.

$$\Pr(X = 5) = \frac{4^5 \times \exp(-4)}{5!} = 0,1563$$

Soit $F_X(x)$ la fonction de répartition de la loi $\mathcal{P}(4)$:

$$\begin{aligned} \Pr(2,57 < X \leq 4) &= \Pr(X \leq 4) - \Pr(X \leq 2,57) = F_X(4) \\ &\quad - F_X(2,57) \end{aligned}$$

Par définition :

$$F_X(2,57) = \exp(-4) \sum_{i=0}^{[2,57]} \frac{4^i}{i!} = \exp(-4) \sum_{i=0}^2 \frac{4^i}{i!} = F_X(2)$$

Pour une loi $\mathcal{P}(4)$, on lit dans la table fournie en annexe que :

$$F_X(2) = 0,2381 \quad F_X(4) = 0,6288$$

On en déduit que :

$$\Pr(2,57 < X \leq 4) = 0,6288 - 0,2381 = 0,3907$$

6 Loi normale

- La loi normale est une loi continue et donc la probabilité associée à une réalisation particulière est nulle $\Pr(X = 2,57) = 0$.

$$\begin{aligned} \Pr(|X| \geq 1) &= 1 - \Pr(|X| \leq 1) \\ &= 1 - \Pr(-1 \leq X \leq 1) \\ &= 1 - \Pr(X \leq 1) + \Pr(X \leq -1) \end{aligned}$$

Si $X \sim \mathcal{N}(2, 2)$ alors $(X - 2)/\sqrt{2}$ suit une loi normale centrée réduite $\mathcal{N}(0, 1)$. Par conséquent :

$$\begin{aligned} \Pr(|X| \geq 1) &= 1 - \Pr(X \leq 1) + \Pr(X \leq -1) \\ &= 1 - \Pr\left(\frac{X-2}{\sqrt{2}} \leq \frac{1-2}{\sqrt{2}}\right) \\ &\quad + \Pr\left(\frac{X-2}{\sqrt{2}} \leq \frac{-1-2}{\sqrt{2}}\right) \\ &= 1 - \Phi(-0,7071) + \Phi(-2,1213) \end{aligned}$$

où $\Phi(\cdot)$ désigne la fonction de répartition de la loi normale centrée réduite. Par lecture de la table de la loi normale centrée réduite, il vient :

$$\begin{aligned} \Phi(-0,7071) &= 1 - \Phi(0,7071) \simeq 1 - \Phi(0,71) \\ &\simeq 1 - 0,761148 = 0,238852 \end{aligned}$$

La valeur précise, obtenue avec un logiciel (Matlab), est égale à $\Phi(-0,7071) = 0,2398$. De la même façon, par lecture de la table de la loi normale centrée réduite, il vient :

$$\begin{aligned} \Phi(-2,1213) &= 1 - \Phi(2,1213) \simeq 1 - \Phi(2,12) \\ &\simeq 1 - 0,982997 = 0,017003 \end{aligned}$$

La valeur précise obtenue avec un logiciel est égale à $\Phi(-2,1213) = 0,0169$. Par conséquent :

$$\Pr(|X| \geq 1) \simeq 1 - 0,238852 + 0,017003 = 0,778151$$

Avec un logiciel, on obtient $\Pr(|X| \geq 1) = 0,7772$.

- Soit $F_X^{-1}(\alpha)$ le quantile d'ordre α de la loi normale $\mathcal{N}(2, 2)$ tel que :

$$\Pr(X \leq F_X^{-1}(\alpha)) = F_X(F_X^{-1}(\alpha)) = \alpha$$

Étant données les propriétés de la fonction de répartition de la loi normale, on sait que $F_X(-\infty) = 0$. Donc le quantile d'ordre $\alpha = 0$ de la loi normale $\mathcal{N}(2, 2)$ est égal à $-\infty$, i.e. $F_X^{-1}(0) = -\infty$. La définition du quantile à $\alpha = 95\%$ est la suivante :

$$\Pr(X \leq F_X^{-1}(0,95)) = 0,95$$

Si $X \sim \mathcal{N}(2,2)$ alors $(X-2)/\sqrt{2}$ suit une loi normale $\mathcal{N}(0,1)$. Dès lors :

$$\begin{aligned}\Pr(X \leq F_X^{-1}(0,95)) &= \Pr\left(\frac{X-2}{\sqrt{2}} \leq \frac{F_X^{-1}(0,95)-2}{\sqrt{2}}\right) \\ &= \Phi(F_X^{-1}(0,95)) = 0,95\end{aligned}$$

On a donc une relation entre le quantile à 95 % de la loi normale $\mathcal{N}(2,2)$ et le quantile à 95 % de la loi normale centrée réduite :

$$F_X^{-1}(0,95) = 2 + \sqrt{2} \times \Phi^{-1}(0,95)$$

Par lecture de la table de la loi normale centrée réduite, il vient $\Phi^{-1}(0,95) \approx 1,65$. Avec un logiciel on obtient $\Phi^{-1}(0,95) = 1,6449$. Dès lors :

$$F_X^{-1}(0,95) \approx 2 + \sqrt{2} \times 1,65 = 4,333$$

ou de façon plus précise avec un logiciel $F_X^{-1}(0,95) = 4,3262$.

3. La définition du quantile à $\alpha = 1\%$ de la loi normale $\mathcal{N}(2,2)$ est la suivante :

$$\Pr(X \leq F_X^{-1}(0,01)) = 0,01$$

Si $X \sim \mathcal{N}(2,2)$ alors $(X-2)/\sqrt{2}$ suit une loi normale centrée réduite $\mathcal{N}(0,1)$. On a donc une relation entre le quantile à 1 % de la loi normale $\mathcal{N}(2,2)$ et le quantile à 1 % de la loi normale centrée réduite :

$$F_X^{-1}(0,01) = 2 + \sqrt{2} \times \Phi^{-1}(0,01) = 2 - \sqrt{2} \times \Phi^{-1}(0,99)$$

Par lecture de la table de la loi normale centrée réduite, il vient $\Phi^{-1}(0,99) \approx 2,33$. Avec un logiciel on obtient $\Phi^{-1}(0,99) = 2,3263$. Dès lors :

$$F_X^{-1}(0,01) \approx 2 - \sqrt{2} \times 2,33 = -1,2951$$

ou de façon plus précise avec un logiciel $F_X^{-1}(0,01) = -1,2900$.

7 Loi Binomiale

1. La variable X_n suit une loi Binomiale $\mathcal{B}(n,p)$ avec $p = 2/8 = 1/4$, donc

$$\mathbb{E}(X_n) = np = \frac{n}{4}$$

$$\mathbb{V}(X_n) = np(1-p) = \frac{n}{4} \left(1 - \frac{1}{4}\right) = \frac{3n}{16}$$

On en déduit que :

$$\mathbb{E}(F_n) = \mathbb{E}\left(\frac{X_n}{n}\right) = \frac{1}{n} \mathbb{E}(X_n) = \frac{1}{4}$$

$$\mathbb{V}(F_n) = \mathbb{V}\left(\frac{X_n}{n}\right) = \frac{1}{n^2} \mathbb{V}(X_n) = \frac{3}{16n}$$

2. On suppose que $n = 10\,000$. Dès lors, il vient :

$$\mathbb{E}(F_n) = \frac{1}{4} \quad \mathbb{V}(F_n) = \frac{3}{160\,000}$$

On en déduit que :

$$\begin{aligned}\Pr(F_n \in]0,22; 0,26[) &= \Pr(F_n - \mathbb{E}(F_n) \in]-0,03; 0,01[) \\ &> \Pr(|F_n - \mathbb{E}(F_n)| < 0,01)\end{aligned}$$

ou encore :

$$\Pr(F_n \in]0,22; 0,26[) > 1 - \Pr(|F_n - \mathbb{E}(F_n)| \geq 0,01)$$

D'après l'inégalité de Tchebychev (► chapitre 8), nous avons :

$$\Pr(|F_n - \mathbb{E}(F_n)| \geq 0,01) \leq \frac{\mathbb{V}(F_n)}{0,01^2}$$

On obtient :

$$\Pr(F_n \in]0,22; 0,26[) \geq 1 - \frac{\mathbb{V}(F_n)}{0,01^2} = 1 - \frac{3}{160\,000 \times 0,01^2}$$

ou encore :

$$\Pr(F_n \in]0,22; 0,26[) \geq \frac{13}{16}$$

3. Le nombre minimal n de tirages nécessaires pour que la probabilité de l'événement $F_n \in]0,22; 0,26[$ soit au moins égale à 0,99 est tel que :

$$\Pr(F_n \in]0,22; 0,26[) \geq 1 - \frac{\mathbb{V}(F_n)}{0,01^2} = 0,99$$

ou encore :

$$1 - \frac{3}{0,01^2 \times 16n} = 0,99$$

On en déduit la valeur minimale de n :

$$n = \frac{3}{16 \times 0,01^3} = 187\,500$$

8 Loi normale

Soit X une variable aléatoire dont la loi de probabilité est la loi normale de moyenne 12 et d'écart-type 4.

$$X \sim \mathcal{N}(12,16) \Leftrightarrow \frac{X-12}{4} \sim \mathcal{N}(0,1)$$

On note $\Phi(\cdot)$ la fonction de répartition de la loi normale centrée réduite.

1. Les probabilités de ces événements sont respectivement égales à :

$$\Pr(X = 2) = 0$$

puisque, pour une variable aléatoire continue, la probabilité d'être en un point est nulle.

$$\Pr(X < 16) = \Pr\left(\frac{X-12}{4} < \frac{16-12}{4}\right) = \Phi(1) = 0,8413$$

$$\begin{aligned}\Pr(X > 20) &= 1 - \Pr(X \leq 20) \\ &= 1 - \Pr\left(\frac{X-12}{4} < \frac{20-12}{4}\right) \\ &= 1 - \Phi(2) \\ &= 0,0228\end{aligned}$$

$$\Pr(X < 0) = \Pr\left(\frac{X-12}{4} < \frac{0-12}{4}\right) = \Phi(-3) = 0,0013$$

2. On cherche le nombre e tel que la probabilité de réalisation de l'événement $(|X - 12| > e)$ soit égale à 0,01.

$$\begin{aligned}\Pr(|X - 12| > e) &= 1 - \Pr(|X - 12| \leq e) \\ &= 1 - \Pr(-e \leq X - 12 \leq e) \\ &= 1 - \Pr\left(-\frac{e}{4} \leq \frac{X - 12}{4} \leq \frac{e}{4}\right) \\ &= 1 - \Pr\left(\frac{X - 12}{4} \leq \frac{e}{4}\right) \\ &\quad + \Pr\left(\frac{X - 12}{4} \leq -\frac{e}{4}\right) \\ &= 1 - \Phi\left(\frac{e}{4}\right) + \Phi\left(-\frac{e}{4}\right)\end{aligned}$$

Puisque $\Phi(-x) = 1 - \Phi(x)$, il vient

$$\Pr(|X - 12| > e) = 2 - 2\Phi\left(\frac{e}{4}\right) = 0,01$$

On en déduit que :

$$\Phi\left(\frac{e}{4}\right) = \frac{2 - 0,01}{2} \iff e = 4 \times \Phi^{-1}\left(\frac{2 - 0,01}{2}\right)$$

La constante e est égale à :

$$e = 4 \times \Phi^{-1}(0,9950) = 4 \times 2,5758 = 10,3033$$

3. La variable Y suit une loi normale avec :

$$\mathbb{E}(Y) = \mathbb{E}(aX + b) = a\mathbb{E}(X) + b = 12a + b$$

$$\mathbb{V}(Y) = \mathbb{V}(aX + b) = a^2\mathbb{V}(X) = 16a^2$$

Ainsi, on obtient :

$$Y \sim \mathcal{N}(12a + b, 16a^2) \iff \frac{Y - 12a - b}{4a} \sim \mathcal{N}(0, 1)$$

Nous savons que :

$$\Pr(Y < 24) = \Phi\left(\frac{24 - 12a - b}{4a}\right) = 0,2266$$

$$\Pr(Y > 42) = 1 - \Pr(Y \leq 42)$$

$$= 1 - \Phi\left(\frac{42 - 12a - b}{4a}\right) = 0,0068$$

On obtient un système à deux équations et deux inconnues :

$$\frac{24 - 12a - b}{4a} = \Phi^{-1}(0,2266) \simeq -0,75$$

$$\frac{42 - 12a - b}{4a} = \Phi^{-1}(0,9320) \simeq 1,5$$

On en déduit que :

$$a = 2, \quad b = 6$$

9 Loi binomiale et loi de Poisson

On considère les nombres x_0 et x_1 définis comme suit :

- x_0 est la plus grande valeur entière de x telle que :

$$\Pr(X \leq x) \leq 0,05$$

- x_1 est la plus petite valeur entière de x telle que :

$$\Pr(X \geq x) \leq 0,05$$

1. On suppose que la variable X suit une loi binomiale $\mathcal{B}(40; 0,08)$. Déterminons x_0 , la plus grande valeur entière de x telle que :

$$\Pr(X \leq x) \leq 0,05$$

Nous savons que :

$$\Pr(X \leq 0) = 0,0356$$

$$\Pr(X \leq 1) = 0,1594$$

Donc $x_0 = 0$. De la même façon, déterminons x_1 , la plus petite valeur entière de x telle que :

$$\Pr(X \geq x) \leq 0,05 \iff 1 - \Pr(X < x) \leq 0,05$$

Ce qui peut se réécrire sous la forme :

$$\Pr(X \leq x - 1) \geq 0,95$$

puisque X est définie sur le support $X(\Omega) = \mathbb{N}$. Nous savons que :

$$\Pr(X \leq 5) = 0,9033$$

$$\Pr(X \leq 6) = 0,9624$$

Donc $x - 1 \geq 6$ ou $x \geq 7$. La plus petite valeur entière de x telle que $\Pr(X \geq x) \leq 0,05$ est égale à $x_1 = 7$.

2. On suppose que la variable X suit une loi de Poisson $\mathcal{P}(12)$. Déterminons x_0 , la plus grande valeur entière de x telle que :

$$\Pr(X \leq x) \leq 0,05$$

Nous savons que :

$$\Pr(X \leq 6) = 0,0458$$

$$\Pr(X \leq 7) = 0,0895$$

Donc $x_0 = 6$. De la même façon, déterminons x_1 , la plus petite valeur entière de x telle que :

$$\Pr(X \geq x) \leq 0,05 \iff 1 - \Pr(X < x) \leq 0,05$$

Ce qui peut se réécrire sous la forme :

$$\Pr(X \leq x - 1) \geq 0,95$$

puisque X est définie sur le support $X(\Omega) = \mathbb{N}$. Nous savons que :

$$\Pr(X \leq 17) = 0,9370$$

$$\Pr(X \leq 18) = 0,9626$$

Donc $x - 1 \geq 18$ ou $x \geq 19$. La plus petite valeur entière de x telle que $\Pr(X \geq x) \leq 0,05$ est égale à $x_1 = 19$.

10 Loi de Fisher et loi du khi-deux

On considère les nombres x_0 et x_1 définis comme suit :

$$\Pr(X < x_0) = 0,05$$

$$\Pr(X > x_1) = 0,05$$

1. On suppose que la variable X suit une de Fisher $\mathcal{F}(4, 20)$. Soit $F_X(x; n, m)$ la fonction de répartition de la loi $\mathcal{F}(n, m)$. En utilisant un logiciel de statistique, on obtient directement :

$$\Pr(X < x_0) = 0,05 \iff x_0 = F_X^{-1}(0,05; 4, 20) = 0,1723$$

$$\Pr(X > x_1) = 0,05 \iff x_1 = F_X^{-1}(0,95; 4, 20) = 2,8661$$

2. On suppose que la variable X suit une loi du khi-deux $\chi^2(8)$. Soit $F_X(x; \nu)$ la fonction de répartition de la loi $\chi^2(\nu)$. En utilisant un logiciel de statistique, on obtient directement :

$$\Pr(X < x_0) = 0,05 \iff x_0 = F_X^{-1}(0,05; 8) = 2,7326$$

$$\Pr(X > x_1) = 0,05 \iff x_1 = F_X^{-1}(0,95; 8) = 15,5073$$

Chapitre 8

1 Convergence

- Faux. La convergence en probabilité implique la convergence en loi, mais la réciproque n'est pas toujours vraie.
- Vrai. La convergence en probabilité implique la convergence en loi au sens où si $Y_n \xrightarrow{p} Y \iff Y_n - Y \xrightarrow{p} 0$, alors $Y_n \xrightarrow{d} Y$.
- Vrai. La convergence presque sûre implique la convergence en probabilité et donc la convergence en loi.
- Vrai. La convergence presque sûre implique la convergence en probabilité, mais la réciproque n'est pas toujours vraie.
- Faux.

2 Si Y_1, \dots, Y_n sont des variables indépendantes et identiquement distribuées, alors :

- Faux. Sans hypothèses supplémentaires sur les variables Y_i , on ne peut pas invoquer la loi forte des grands nombres.
- Vrai. D'après la loi faible des grands nombres, la moyenne empirique converge en probabilité vers l'espérance.
- Faux. Cette phrase ne veut rien dire.
- Faux. La moyenne empirique converge en probabilité vers l'espérance, et donc sa distribution asymptotique est dégénérée (au sens strict).
- Vrai. D'après le théorème central limite, $\sqrt{n}(\bar{Y}_n - \mathbb{E}(Y_i))$ converge en loi vers une loi normale.

3 Si une variable Z_n est asymptotiquement normalement distribuée, cela signifie que :

- Faux. Rien ne permet de dire que la loi asymptotique est $\mathcal{N}(0,1)$.
- Vrai. Au sens strict, cela signifie que $\sqrt{n}(Z_n - \mathbb{E}(Z_n)) \xrightarrow{d} \mathcal{N}(m, \sigma^2)$.
- Vrai. Lorsque $\sqrt{n}(Z_n - \mathbb{E}(Z_n)) \xrightarrow{d} \mathcal{N}(0, \sigma^2)$, on dit que Z_n est asymptotiquement normalement distribuée.
- Faux. Rien ne permet de dire que la loi asymptotique est $\mathcal{N}(0,1)$.
- Vrai. Si la variable Z_n est asymptotiquement normalement distribuée :

$$\sqrt{n}(Z_n - m) \xrightarrow{d} \mathcal{N}(0, \sigma^2)$$

Pour une dimension n grande et finie, on peut écrire que :

$$Z_n \stackrel{asy}{\approx} \mathcal{N}\left(m, \frac{\sigma^2}{n}\right)$$

4 Soit Y_1, \dots, Y_n une suite de n variables aléatoires et soit \bar{Y}_n la moyenne empirique :

- Vrai. C'est le théorème de Lindeberg-Levy qui s'applique dans ce cas.
- Vrai. C'est le théorème de Lyapounov qui s'applique dans ce cas.
- Vrai. Si les espérances sont identiques, c'est le théorème de Lindeberg-Feller qui s'applique. Si les espérances sont différentes, on applique le théorème de Lyapounov.
- Vrai. C'est le théorème de Lyapounov qui s'applique dans ce cas.
- Faux. Sans hypothèse particulière sur la forme de la dépendance (dépendance faible), le théorème central limite ne s'applique pas.

5 Convergences

- Les variables X_1/σ et X_2/σ sont indépendantes et suivent toutes deux une loi normale centrée réduite. Dès lors, les variables X_1^2/σ^2 et X_2^2/σ^2 sont elles aussi indépendantes et suivent une loi du khi-deux à un degré de liberté (► chapitre 6). La variable Y^2/σ^2 est une somme de deux khi-deux indépendants, elle suit une distribution du khi-deux à 2 degrés de liberté :

$$\frac{Y^2}{\sigma^2} = \frac{X_1^2}{\sigma^2} + \frac{X_2^2}{\sigma^2} \sim \chi^2(2)$$

D'après les propriétés de la loi du khi-deux, on sait que si $Z \sim \chi^2(\nu)$ alors $\mathbb{E}(Z) = \nu$ et $\mathbb{V}(Z) = 2\nu$. On en déduit que :

$$\mathbb{E}\left(\frac{Y^2}{\sigma^2}\right) = \frac{1}{\sigma^2} \mathbb{E}(Y^2) = 2 \iff \mathbb{E}(Y^2) = 2\sigma^2$$

$$\mathbb{V}\left(\frac{Y^2}{\sigma^2}\right) = \frac{1}{\sigma^4} \mathbb{V}(Y^2) = 4 \iff \mathbb{V}(Y^2) = 4\sigma^4$$

- Calculons $\mathbb{E}(\hat{\sigma}^2)$ et $\mathbb{V}(\hat{\sigma}^2)$. Sous l'hypothèse d'un échantillon i.i.d., il vient :

$$\mathbb{E}(\hat{\sigma}^2) = \frac{1}{2n} \sum_{i=1}^n \mathbb{E}(Y_i^2) = \frac{n \times 2\sigma^2}{2n} = \sigma^2$$

$$\mathbb{V}(\hat{\sigma}^2) = \mathbb{V}\left(\frac{1}{2n} \sum_{i=1}^n Y_i^2\right) = \frac{1}{4n^2} \sum_{i=1}^n \mathbb{V}(Y_i^2) = \frac{n \times 4\sigma^4}{4n^2} = \frac{\sigma^4}{n}$$

Dès lors :

$$\mathbb{E}(\hat{\sigma}^2) = \sigma^2$$

$$\lim_{n \rightarrow \infty} \mathbb{V}(\hat{\sigma}^2) = \lim_{n \rightarrow \infty} \frac{\sigma^4}{n} = 0$$

On en déduit que $\hat{\sigma}^2$ converge vers σ^2 :

$$\hat{\sigma}^2 \xrightarrow{p} \sigma^2$$

- Les variables Y_1^2, \dots, Y_n^2 sont i.i.d. avec $\mathbb{E}(Y_i^2) = 2\sigma^2$ et $\mathbb{V}(Y_i^2) = 4\sigma^4$. Par application du théorème central limite, il vient :

$$\sqrt{n}\left(\frac{1}{n} \sum_{i=1}^n Y_i^2 - 2\sigma^2\right) \xrightarrow{d} \mathcal{N}(0, 4\sigma^4)$$

Or, l'estimateur $\hat{\sigma}^2$ est défini par :

$$\hat{\sigma}^2 = \frac{1}{2n} \sum_{i=1}^n Y_i^2$$

En utilisant la méthode delta pour une fonction $g(x) = x/2$, il vient :

$$\sqrt{n}(\hat{\sigma}^2 - \sigma^2) \xrightarrow{d} \mathcal{N}(0, \sigma^4)$$

6 Théorème central limite

1. Soit X_i la variable binaire valant 1 si l'abonné i est connecté et 0 sinon. La variable X_i suit une distribution de Bernouilli de probabilité $p = \Pr(X_i = 1) = 0,20$. Les variables X_i sont indépendantes, donc la variable X définie par :

$$X = \sum_{i=1}^n X_i$$

avec $n = 5\,000$ suit une loi binomiale $\mathcal{B}(5\,000; 0,20)$. Par définition :

$$\mathbb{E}(X) = np = 5\,000 \times 0,2 = 1\,000$$

$$\mathbb{V}(X) = np(1-p) = 5\,000 \times 0,2 \times 0,8 = 800$$

2. Les variables X_i sont i.i.d. avec $\mathbb{E}(X_i) = p = 0,2$ et $\mathbb{V}(X_i) = p(1-p) = 0,16$, par application du théorème central limite de Lindeberg-Levy il vient :

$$\sqrt{n} \left(\frac{1}{n} \sum_{i=1}^n X_i - p \right) \xrightarrow{d} \mathcal{N}(0, p(1-p))$$

Pour n grand mais fini, on a donc :

$$\sqrt{n} \left(\frac{X}{n} - p \right) \approx \mathcal{N}(0, p(1-p))$$

ou encore :

$$\frac{X}{n} \approx \mathcal{N}\left(p, \frac{p(1-p)}{n}\right)$$

Finalement, il vient :

$$X \approx \mathcal{N}(np, np(1-p)) \Leftrightarrow \frac{X - np}{\sqrt{np(1-p)}} \approx \mathcal{N}(0, 1)$$

Si l'on admet que $n = 5\,000$ est une taille suffisamment importante pour que cette approximation soit valide, la variable Y suit asymptotiquement une loi normale centrée réduite :

$$Y = \frac{X - 1\,000}{\sqrt{800}} \approx \mathcal{N}(0, 1)$$

3. On cherche un seuil N tel que :

$$\Pr(X \geq N) \leq 0,975$$

On sait que :

$$\Pr(X \geq N) = \Pr\left(Y \geq \frac{N - 1\,000}{\sqrt{800}}\right) \leq 0,025$$

Si l'on admet que Y suit approximativement une loi normale centrée réduite, il vient :

$$1 - \Pr\left(Y < \frac{N - 1\,000}{\sqrt{800}}\right) \leq 0,025$$

$$\Leftrightarrow \Phi\left(\frac{N - 1\,000}{\sqrt{800}}\right) \geq 0,975$$

où $\Phi(\cdot)$ désigne la fonction de répartition de la loi normale centrée réduite. Nous savons que :

$$0,975 = \Phi(\Phi^{-1}(0,975)) = \Phi(1,96)$$

Ainsi, on obtient une relation du type :

$$\Phi\left(\frac{N - 1\,000}{\sqrt{800}}\right) \geq \Phi(1,96)$$

Puisque la fonction de répartition $\Phi(\cdot)$ est croissante :

$$\frac{N - 1\,000}{\sqrt{800}} \geq 1,96$$

On en déduit :

$$N \geq 1,96 \times \sqrt{800} + 1\,000 = 1\,054$$

7 Convergences

Soit (X_n) une suite de variables aléatoires discrètes telles que $\forall n \geq 2$:

$$\Pr(X_n = -n) = \Pr(X_n = n) = \frac{1}{2n^2}$$

$$\Pr(X_n = 0) = 1 - \frac{1}{n^2}$$

1. On fait l'hypothèse que la limite en probabilité de la suite (X_n) est égale à 0.

$$X_n \xrightarrow{p} 0$$

Cette hypothèse est vraie si et seulement si $\forall \varepsilon$:

$$\lim_{n \rightarrow \infty} \Pr(|X_n - 0| > \varepsilon) = 0$$

ou de façon équivalente lorsque :

$$\lim_{n \rightarrow \infty} \Pr(|X_n - 0| < \varepsilon) = 1$$

Or, par définition, si l'on fait tendre ε vers 0, il vient :

$$\Pr(|X_n - 0| < \varepsilon) = \Pr(|X_n| < \varepsilon) = \Pr(X_n = 0) = 1 - \frac{1}{n^2}$$

Ainsi, on obtient :

$$\lim_{n \rightarrow \infty} \Pr(|X_n - 0| < \varepsilon) = \lim_{n \rightarrow \infty} 1 - \frac{1}{n^2} = 1$$

On en déduit que :

$$X_n \xrightarrow{p} 0$$

2. La suite (X_n) converge en moyenne quadratique vers 0 si et seulement si :

$$\lim_{n \rightarrow \infty} \mathbb{E}(|X_n - 0|^2) = \lim_{n \rightarrow \infty} \mathbb{E}(|X_n|^2) = \lim_{n \rightarrow \infty} \mathbb{E}(X_n^2) = 0$$

Or, par définition :

$$\begin{aligned} \mathbb{E}(X_n^2) &= \sum i^2 \Pr(X = i) \\ &= 0^2 \Pr(X = 0) + \sum_{i>0} i^2 \Pr(X = i) + \sum_{i<0} i^2 \Pr(X = i) \\ &= \sum_{i>0} i^2 \Pr(X = i) + \sum_{i>0} i^2 \Pr(X = -i) \\ &= \sum_{i>0} i^2 \frac{2}{2i^2} > 0 \end{aligned}$$

Donc il n'y a pas de convergence en moyenne quadratique vers 0.

8 Théorème central limite

1. Soit X_i la variable binaire valant 1 si le passager i se présente à l'embarquement et 0 sinon. La variable X_i suit une distribution de Bernoulli de probabilité $p = \Pr(X_i = 1) = 0,90$. Les variables X_i sont indépendantes, donc la variable S_n définie par :

$$S_n = \sum_{i=1}^n X_i$$

suit une loi binomiale $\mathcal{B}(n; 0,90)$. Par définition :

$$\mathbb{E}(S_n) = np = 0,9n$$

$$\mathbb{V}(S_n) = np(1-p) = 0,09n$$

2. Les variables X_i sont i.i.d. avec $\mathbb{E}(X_i) = p = 0,9$ et $\mathbb{V}(X_i) = p(1-p) = 0,09$. Par application du théorème central limite de Lindeberg-Levy il vient :

$$\sqrt{n} \left(\frac{S_n}{n} - p \right) \xrightarrow{d} \mathcal{N}(0, p(1-p))$$

Pour n grand mais fini, on a donc :

$$\sqrt{n} \left(\frac{S_n}{n} - p \right) \overset{asy}{\approx} \mathcal{N}(0, p(1-p))$$

ou encore :

$$\frac{S_n}{n} \overset{asy}{\approx} \mathcal{N} \left(p, \frac{p(1-p)}{n} \right)$$

Finalement, il vient :

$$S_n \overset{asy}{\approx} \mathcal{N}(np, np(1-p)) \Leftrightarrow \frac{S_n - np}{\sqrt{np(1-p)}} \overset{asy}{\approx} \mathcal{N}(0,1)$$

ou encore

$$\frac{S_n - 0,9n}{0,3\sqrt{n}} \overset{asy}{\approx} \mathcal{N}(0,1)$$

Le directeur commercial de la compagnie aimerait connaître la valeur maximale de n telle que :

$$\Pr(S_n \leq 300) \geq 0,99$$

Par définition :

$$\begin{aligned} \Pr(S_n \leq 300) &= \Pr \left(\frac{S_n - 0,9n}{0,3\sqrt{n}} \leq \frac{300 - 0,9n}{0,3\sqrt{n}} \right) \\ &= \Phi \left(\frac{300 - 0,9n}{0,3\sqrt{n}} \right) \end{aligned}$$

où $\Phi(\cdot)$ désigne la fonction de répartition de la loi normale centrée réduite. On a donc :

$$\Phi \left(\frac{300 - 0,9n}{0,3\sqrt{n}} \right) \geq \Phi(\Phi^{-1}(0,99)) = \Phi(2,3263)$$

Puisque la fonction $\Phi(\cdot)$ est croissante :

$$\frac{300 - 0,9n}{0,3\sqrt{n}} \geq 2,3263$$

On obtient l'équation :

$$-0,9n - 0,6979\sqrt{n} + 300 \geq 0$$

En posant $x = \sqrt{n}$, on obtient une fonction du second degré à étudier et l'on cherche le plus grand x pour lequel elle est positive. Le plus grand n pour lequel cette quantité est positive est 316.

Chapitre 9

1 Propriétés d'un estimateur

- Faux. Un estimateur sans biais n'est pas nécessairement efficace.
- Faux. Un estimateur sans biais n'est pas nécessairement convergent. Il n'est convergent que si variance tend vers zéro lorsque la taille d'échantillon tend vers l'infini.
- Vrai. Puisque la convergence presque sûre implique la convergence en probabilité.
- Faux. Un estimateur biaisé peut être convergent s'il est asymptotiquement non biaisé.
- Vrai. La comparaison de la variance de l'estimateur à la borne FDCR ne s'applique que pour des estimateurs sans biais. Un estimateur efficace est donc sans biais.

2 Variance empirique

- Vrai. La variance empirique (corrigée ou non corrigée) est un estimateur de la variance.
- Vrai.
- Faux. Seule une transformée de la variance empirique corrigée a une distribution exacte du khi-deux.
- Faux. Le nombre de degrés de liberté est $n-1$ et non n .
- Faux. En raison de la correction de petit échantillon, on a

$$S_n^2 = (n-1)^{-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2.$$

3 Comparaison d'estimateurs

- Vrai. L'estimateur $\hat{\theta}_1$ est préféré à $\hat{\theta}_2$ si sa variance est plus faible.
- Faux. Ce n'est pas parce que $\mathbb{V}(\hat{\theta}_1) < \mathbb{V}(\hat{\theta}_2)$, que la variance de $\hat{\theta}_1$ atteint la borne FDCR.
- Vrai. Si la variance de $\hat{\theta}_1$ atteint la borne FDCR, alors la variance de tout estimateur sans biais est nécessairement égale ou supérieure à celle de l'estimateur $\hat{\theta}_1$.
- Vrai. L'optimalité implique généralement l'efficacité, mais la réciproque n'est pas vraie puisque l'efficacité requiert d'effectuer un certain nombre d'hypothèses sur la distribution des variables de l'échantillon.
- Faux. La convergence au sens faible n'implique pas nécessairement l'efficacité.

4 Intervalle de confiance

- Vrai. C'est la définition d'un intervalle de confiance.
- Vrai.
- Vrai. En effet, plus l'estimateur est précis, plus la réalisation de l'intervalle de confiance sera concentrée autour de la vraie valeur du paramètre.
- Vrai. En général, le niveau de risque α est inférieur à 50 %.
- Faux. C'est la réalisation d'un intervalle de confiance pour un échantillon donné qui se ramène à un segment de deux valeurs réelles.

5 Estimation et loi de Rayleigh

1. La variable Y^2/σ^2 peut s'écrire sous la forme :

$$\left(\frac{Y}{\sigma}\right)^2 = \left(\frac{X_1}{\sigma}\right)^2 + \left(\frac{X_2}{\sigma}\right)^2$$

Les variables X_1/σ et X_2/σ suivent des lois normales centrées réduites $\mathcal{N}(0,1)$. Par conséquent, les variables X_1^2/σ^2 et X_2^2/σ^2 ont une distribution du khi-deux à un degré de liberté. Puisque ces variables sont indépendantes, la somme $X_1^2/\sigma^2 + X_2^2/\sigma^2$ admet une distribution du khi-deux à deux degrés de liberté.

$$\left(\frac{Y}{\sigma}\right)^2 \sim \chi^2(2)$$

Etant données les propriétés de la loi du khi-deux, on sait que si $X \sim \chi^2(v)$, alors $\mathbb{E}(X) = v$ et $\mathbb{V}(X) = 2v$. On en déduit que :

$$\mathbb{E}\left(\frac{Y^2}{\sigma^2}\right) = \frac{1}{\sigma^2} \mathbb{E}(Y^2) = 2 \iff \mathbb{E}(Y^2) = 2\sigma^2$$

$$\mathbb{V}\left(\frac{Y^2}{\sigma^2}\right) = \frac{1}{\sigma^4} \mathbb{V}(Y^2) = 4 \iff \mathbb{V}(Y^2) = 4\sigma^4$$

2. On étudie la quantité $\mathbb{E}(\hat{\sigma}^2)$:

$$\mathbb{E}(\hat{\sigma}^2) = \mathbb{E}\left(\frac{1}{2n} \sum_{i=1}^n Y_i^2\right) = \frac{1}{2n} \sum_{i=1}^n \mathbb{E}(Y_i^2) = \frac{n \times 2\sigma^2}{2n} = \sigma^2$$

puisque les variables Y_1, \dots, Y_n sont i.i.d., de même loi que Y , avec $\mathbb{E}(Y_i^2) = 2\sigma^2$. L'estimateur $\hat{\sigma}^2$ est sans biais.

3. Nous savons que l'estimateur est sans biais :

$$\mathbb{E}(\hat{\sigma}^2) = \sigma^2$$

Par ailleurs :

$$\mathbb{V}(\hat{\sigma}^2) = \mathbb{V}\left(\frac{1}{2n} \sum_{i=1}^n Y_i^2\right) = \frac{1}{4n^2} \sum_{i=1}^n \mathbb{V}(Y_i^2) = \frac{n \times 4\sigma^4}{4n^2} = \frac{\sigma^4}{n}$$

puisque les variables Y_1, \dots, Y_n sont i.i.d. avec $\mathbb{V}(Y_i^2) = 4\sigma^4$. Dès lors :

$$\lim_{n \rightarrow \infty} \mathbb{V}(\hat{\sigma}^2) = \lim_{n \rightarrow \infty} \frac{\sigma^4}{n} = 0$$

On peut en déduire que l'estimateur est convergent (au sens faible) :

$$\hat{\sigma}^2 \xrightarrow{p} \sigma^2$$

4. Les variables Y_1^2, \dots, Y_n^2 sont i.i.d. avec $\mathbb{E}(Y_i^2) = 2\sigma^2$ et $\mathbb{V}(Y_i^2) = 4\sigma^4$. Par application du théorème central limite de Lindeberg-Levy, il vient :

$$\sqrt{n} \left(\frac{1}{n} \sum_{i=1}^n Y_i^2 - 2\sigma^2 \right) \xrightarrow{d} \mathcal{N}(0, 4\sigma^4)$$

Or, l'estimateur $\hat{\sigma}^2$ vérifie :

$$\hat{\sigma}^2 = \frac{1}{2n} \sum_{i=1}^n Y_i^2 = g\left(\frac{1}{n} \sum_{i=1}^n Y_i^2\right)$$

où la fonction $g(\cdot)$ est définie par $g(x) = x/2$. Cette fonction est continue et continument différentiable, avec $\partial g(x)/\partial x = 1/2$. Par application de la méthode delta (► chapitre 6), il vient :

$$\sqrt{n} \left(g\left(\frac{1}{n} \sum_{i=1}^n Y_i^2\right) - g(2\sigma^2) \right) \xrightarrow{d} \mathcal{N}\left(0, \left(\frac{\partial g(z)}{\partial z}\right)_{2\sigma^2} \times 4\sigma^4\right)$$

ou de façon équivalente :

$$\sqrt{n}(\hat{\sigma}^2 - \sigma^2) \xrightarrow{d} \mathcal{N}\left(0, \left(\frac{1}{2}\right)^2 \times 4\sigma^4\right)$$

Après simplification, on obtient :

$$\sqrt{n}(\hat{\sigma}^2 - \sigma^2) \xrightarrow{d} \mathcal{N}(0, \sigma^4)$$

5. Le résultat précédent signifie que pour une taille d'échantillon n suffisamment grande, mais finie :

$$\hat{\sigma}^2 \approx^{asy} \mathcal{N}\left(\sigma^2, \frac{\sigma^4}{n}\right)$$

La variance asymptotique de l'estimateur $\hat{\sigma}^2$ est donc égale à :

$$\mathbb{V}_{asy}(\hat{\sigma}^2) = \frac{\sigma^4}{n}$$

6. Ici nous avons :

$$\mathbb{V}(\hat{\sigma}^2) = \frac{\sigma^4}{n} = I_n^{-1}(\theta)$$

La variance de l'estimateur $\hat{\sigma}^2$ atteint la borne FDCR : l'estimateur est efficace.

6 Loi exacte et loi asymptotique

1. Puisque $Z_i \sim \mathcal{N}(0,1)$, la variable Z_i^2 suit une loi du khi-deux à un degré de liberté. Les variables Z_i^2 sont indépendantes, donc la variable $D_n = \sum_{i=1}^n Z_i^2$ suit une loi du khi-deux à n degrés de liberté : $D_n \sim \chi^2(n)$.
2. Puisque $Z_i^2 \sim \chi^2(1)$, on a $\mathbb{E}(Z_i^2) = 1$ et $\mathbb{V}(Z_i^2) = 2$. Les variables Z_i^2 sont i.i.d., donc nous pouvons appliquer le théorème central limite de Lindeberg-Levy :

$$\sqrt{n} \left(\frac{1}{n} \sum_{i=1}^n Z_i^2 - 1 \right) \xrightarrow{d} \sqrt{n} \left(\frac{D_n}{n} - 1 \right) \xrightarrow{d} \mathcal{N}(0, 2)$$

3. On note $F_n(x)$ la fonction de répartition d'une loi $\chi^2(n)$. En utilisant la loi exacte $\chi^2(100)$, on obtient :

$$\begin{aligned} \Pr(D_n > 118,49) &= 1 - \Pr(D_n \leq 118,49) \\ &= 1 - F_{100}(118,49) \\ &= 1 - 0,90 = 0,1 \end{aligned}$$

où $F_{100}(\cdot)$ désigne la fonction de répartition d'une loi du khi-deux à 100 degrés de liberté. La loi asymptotique de D_n peut s'écrire sous la forme suivante :

$$\frac{D_n}{n} \approx^{asy} \mathcal{N}\left(1, \frac{2}{n}\right) \quad \text{ou} \quad D_n \approx^{asy} \mathcal{N}(n, 2n)$$

Si l'on admet que pour $n = 100$ cette approximation est valide, il vient :

$$\begin{aligned}\Pr(D_n > 118,49) &= 1 - \Pr(D_n \leq 118,49) \\ &= 1 - \Pr\left(\frac{D_n - n}{\sqrt{2n}} \leq \frac{118,49 - n}{\sqrt{2n}}\right) \\ &= 1 - \Phi\left(\frac{118,49 - 100}{\sqrt{200}}\right) \\ &= 1 - \Phi(1,308) \\ &\approx 0,0954\end{aligned}$$

où $\Phi(\cdot)$ est la fonction de répartition de la loi normale centrée réduite.

7 Comparaison d'estimateurs

1. Calculons $\mathbb{E}(\widehat{\theta}_1)$ et $\mathbb{E}(\widehat{\theta}_2)$:

$$\begin{aligned}\mathbb{E}(\widehat{\theta}_1) &= \frac{1}{n} \sum_{i=1}^n \mathbb{E}(X_i) = \frac{1}{n} \sum_{i=1}^n \mathbb{E}(X) = \theta \\ \mathbb{E}(\widehat{\theta}_2) &= \frac{1}{n} \sum_{i=1}^n \mathbb{E}(X_i^2) = \frac{1}{n} \sum_{i=1}^n \mathbb{E}(X^2)\end{aligned}$$

Or on sait que

$$\mathbb{E}(X^2) = \mathbb{V}(X) + \mathbb{E}(X)^2 = \theta - \theta^2 + \theta^2 = \theta$$

Il vient donc :

$$\mathbb{E}(\widehat{\theta}_1) = \theta \quad \mathbb{E}(\widehat{\theta}_2) = \theta$$

Les deux estimateurs sont sans biais.

2. Calculons $\mathbb{V}(\widehat{\theta}_1)$ et $\mathbb{V}(\widehat{\theta}_2)$. Puisque les variables X_i sont i.i.d., on a :

$$\begin{aligned}\mathbb{V}(\widehat{\theta}_1) &= \frac{1}{n^2} \sum_{i=1}^n \mathbb{V}(X_i) = \frac{1}{n^2} \sum_{i=1}^n \mathbb{V}(X) = \frac{\theta - \theta^2}{n} \\ \mathbb{V}(\widehat{\theta}_2) &= \frac{1}{n^2} \sum_{i=1}^n \mathbb{V}(X_i^2) = \frac{1}{n^2} \sum_{i=1}^n \mathbb{V}(X^2) = \frac{2\theta^2 - 2\theta^4}{n}\end{aligned}$$

Dans les deux cas, on a :

$$\begin{aligned}\mathbb{E}(\widehat{\theta}_1) &= \mathbb{E}(\widehat{\theta}_2) = \theta \\ \lim_{n \rightarrow \infty} \mathbb{V}(\widehat{\theta}_1) &= \lim_{n \rightarrow \infty} \mathbb{V}(\widehat{\theta}_2) = 0\end{aligned}$$

Les deux estimateurs sont donc convergents au sens de la convergence en probabilité.

3. On sait que

$$\begin{aligned}\mathbb{V}(\widehat{\theta}_1) &= \frac{\theta - \theta^2}{n} \\ \mathbb{V}(\widehat{\theta}_2) &= \frac{2\theta^2 - 2\theta^4}{n}\end{aligned}$$

Dès lors :

$$\frac{\mathbb{V}(\widehat{\theta}_1)}{\mathbb{V}(\widehat{\theta}_2)} = \frac{\theta(1 - \theta)}{2\theta^2(1 - \theta^2)} = \frac{1}{2\theta(1 + \theta)}$$

avec $\theta \in [0, 1]$. Donc on ne peut pas déterminer si $\mathbb{V}(\widehat{\theta}_1) > \mathbb{V}(\widehat{\theta}_2)$ ou si $\mathbb{V}(\widehat{\theta}_1) < \mathbb{V}(\widehat{\theta}_2)$ puisque la position du ratio $\mathbb{V}(\widehat{\theta}_1)/\mathbb{V}(\widehat{\theta}_2)$ par rapport à 1 dépend de la vraie valeur de θ qui est inconnue.

4. D'après le théorème central limite, on a :

$$\sqrt{n} \left(\frac{1}{n} \sum_{i=1}^n X_i - \mathbb{E}(X) \right) \xrightarrow{d} \mathcal{N}(0, \mathbb{V}(X))$$

Donc

$$\sqrt{n}(\widehat{\theta}_1 - \alpha) \xrightarrow{d} \mathcal{N}(0, \theta - \theta^2)$$

De la même façon :

$$\sqrt{n} \left(\frac{1}{n} \sum_{i=1}^n X_i^2 - \mathbb{E}(X^2) \right) \xrightarrow{d} \mathcal{N}(0, \mathbb{V}(X^2))$$

Donc

$$\sqrt{n}(\widehat{\theta}_2 - \alpha) \xrightarrow{d} \mathcal{N}(0, 2\theta(1 + \theta))$$

8 Estimation

1. Calculons l'espérance de l'estimateur $\widehat{\theta}$:

$$\begin{aligned}\mathbb{E}(\widehat{\theta}) &= \mathbb{E} \left(\ln(c) - \frac{1}{n} \sum_{i=1}^n \ln(X_i) \right) \\ &= \ln(c) - \frac{1}{n} \sum_{i=1}^n \mathbb{E}(\ln(X_i)) \\ &= \ln(c) - \frac{1}{n} \sum_{i=1}^n (\ln(c) - \theta) \\ &= \theta\end{aligned}$$

L'estimateur $\widehat{\theta}$ est sans biais.

2. La séquence de variables aléatoires i.i.d. $\ln(X_1), \dots, \ln(X_n)$ vérifie $\mathbb{E}(\ln(X_i)) = \ln(c) - \theta$. D'après la loi faible des grands nombres :

$$\frac{1}{n} \sum_{i=1}^n \ln(X_i) \xrightarrow{p} \mathbb{E}(\ln(X_i)) = \ln(c) - \theta$$

En utilisant le *continuous mapping theorem* pour une fonction $g(z) = \ln(c) - z$, on obtient :

$$g \left(\frac{1}{n} \sum_{i=1}^n \ln(X_i) \right) \xrightarrow{p} g(\ln(c) - \theta)$$

ou de façon équivalente :

$$\ln(c) - \frac{1}{n} \sum_{i=1}^n \ln(X_i) \xrightarrow{p} \ln(c) - \ln(c) + \theta$$

Par conséquent :

$$\widehat{\theta} \xrightarrow{p} \theta$$

L'estimateur $\widehat{\theta}$ est convergent (au sens faible).

Chapitre 10

1 Vraisemblance et log-vraisemblance

- a. Vrai. La log-vraisemblance d'un échantillon correspond au logarithme de la vraisemblance de l'échantillon.

- b. Faux. Pour une variable continue, la vraisemblance d'un échantillon correspond à la densité jointe des variables de l'échantillon.
- c. Faux. La log-vraisemblance d'un échantillon est égale à la somme des vraisemblances individuelles associées à chaque observation de cet échantillon.
- d. Faux. La log-vraisemblance d'un échantillon est une constante puisqu'elle dépend des paramètres (quantités déterministes) et de la réalisation de l'échantillon.
- e. Vrai. La vraisemblance d'un échantillon dépend de deux arguments : le vecteur de paramètres et les données de l'échantillon.

2 Estimateur du maximum de vraisemblance

- a. Faux. L'estimateur du maximum de vraisemblance, comme tout estimateur, est une variable aléatoire.
- b. Vrai. L'estimateur du maximum de vraisemblance est une fonction des variables aléatoires de l'échantillon.
- c. Faux. Au sens strict, la solution du programme de maximisation de la log-vraisemblance correspond à une réalisation de l'estimateur du maximum de vraisemblance (estimation). On en déduit ensuite un estimateur (variable aléatoire).
- d. Vrai. Le gradient de l'échantillon, évalué au point de la réalisation de l'estimateur du maximum de vraisemblance, est égal à zéro.
- e. Faux. La hessienne associée à l'échantillon est alors une matrice de dimension 3×3 .

3 Score, hessienne et information de Fisher

- a. Vrai. On peut définir le gradient comme une réalisation du score.
- b. Vrai. Par définition, l'espérance du score est nulle quelle que soit valeur du paramètre.
- c. Vrai. C'est l'une des deux définitions de l'information de Fisher associée à l'échantillon.
- d. Vrai. Par construction, l'information de Fisher de l'échantillon est égale à l'information de Fisher moyenne multipliée par la taille de l'échantillon.
- e. Faux. Ceci n'est vrai que dans le cas de lois marginales. Dans le cas de modèles économétriques (lois conditionnelles), l'information de Fisher moyenne correspond à l'espérance (par rapport aux variables explicatives X) de l'information de Fisher associée à une observation de l'échantillon.

4 Propriétés de l'estimateur du maximum de vraisemblance

- a. Faux. L'estimateur du maximum de vraisemblance est sans biais, uniquement lorsque les hypothèses de régularité sont satisfaites.
- b. Faux. Sous les hypothèses de régularité usuelles, l'estimateur du maximum de vraisemblance est convergent au sens faible.
- c. Vrai. Sous les hypothèses de régularité, l'estimateur du maximum de vraisemblance est asymptotiquement normalement distribué.
- d. Faux. Sa variance asymptotique est égale à la borne FDCR.
- e. Faux. La variance asymptotique de l'estimateur du maximum de vraisemblance est égale à l'inverse de la matrice d'information de Fisher associée à l'échantillon.

5 Maximum de vraisemblance

1. Puisque le paramètre μ est connu, nous exprimons le logarithme de la densité de X en fonction uniquement du paramètre inconnu σ^2 :

$$\ln f_X(x; \sigma^2) = -\ln x - \frac{1}{2} \ln(\sigma^2) - \frac{1}{2} \ln(2\pi) - \frac{(\ln x - \mu)^2}{2\sigma^2}$$

Puisque les variables X_1, \dots, X_n sont indépendantes, la log-vraisemblance de l'échantillon (x_1, \dots, x_n) est :

$$\ell_n(\sigma^2; x) = \ln L_n(\sigma^2; x) = \sum_{i=1}^n \ln f_X(x_i; \sigma^2)$$

On obtient :

$$\begin{aligned} \ell_n(\sigma^2; x) &= - \sum_{i=1}^n \ln x_i - \frac{n}{2} \ln(\sigma^2) \\ &\quad - \frac{n}{2} \ln(2\pi) - \frac{1}{2\sigma^2} \sum_{i=1}^n (\ln x_i - \mu)^2 \end{aligned}$$

2. L'estimateur du maximum de vraisemblance est défini par :

$$\hat{\sigma}^2 = \arg \max_{\sigma^2 \in \mathbb{R}^+} \ell_n(\sigma^2; x)$$

Pour simplifier la dérivation de la log-vraisemblance, on pose $\theta = \sigma^2$.

$$\begin{aligned} \ell_n(\sigma^2; x) &= - \sum_{i=1}^n \ln x_i - \frac{n}{2} \ln(\theta) - \frac{n}{2} \ln(2\pi) \\ &\quad - \frac{1}{2\theta} \sum_{i=1}^n (\ln x_i - \mu)^2 \end{aligned}$$

La condition du premier ordre (équation de log-vraisemblance) est :

$$\begin{aligned} \text{CN} : g_n(\theta, x) &= \frac{\partial \ell_n(\sigma^2; x)}{\partial \theta} \bigg|_{\theta} = \\ &= -\frac{n}{2\theta} + \frac{1}{2\theta^2} \sum_{i=1}^n (\ln x_i - \mu)^2 = 0 \end{aligned}$$

D'où l'on tire que :

$$\hat{\theta} = \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (\ln x_i - \mu)^2$$

On vérifie qu'il s'agit d'un maximum :

$$CS : H_n(\widehat{\theta}, x) = \frac{\partial^2 \ell_n(\sigma^2; x)}{\partial \theta^2} \Big|_{\widehat{\theta}} = \frac{n}{2\widehat{\theta}^2} - \frac{1}{\widehat{\theta}^3} \sum_{i=1}^n (\ln x_i - \mu)^2$$

Sachant que $\sum_{i=1}^n (\ln x_i) - \mu^2 = n \times \widehat{\theta}$, on obtient :

$$\frac{\partial^2 \ell_n(\sigma^2; x)}{\partial \theta^2} \Big|_{\widehat{\theta}} = \frac{n}{2\widehat{\theta}^2} - \frac{n \times \widehat{\theta}}{\widehat{\theta}^3} = -\frac{n}{2\widehat{\theta}^2} = -\frac{n}{2\sigma^4} < 0$$

On a bien un maximum. Par conséquent, l'estimateur du maximum de vraisemblance $\widehat{\sigma}^2$ du paramètre σ^2 est défini par :

$$\widehat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (\ln X_i - \mu)^2$$

3. Calculons l'espérance de l'estimateur $\widehat{\sigma}^2$:

$$\begin{aligned} \mathbb{E}(\widehat{\sigma}^2) &= \mathbb{E}\left(\frac{1}{n} \sum_{i=1}^n (\ln X_i - \mu)^2\right) = \frac{1}{n} \sum_{i=1}^n \mathbb{E}((\ln X_i - \mu)^2) \\ &= \frac{1}{n} \sum_{i=1}^n \mathbb{E}((\ln X_i - \mathbb{E}(\ln X_i))^2) = \frac{1}{n} \sum_{i=1}^n \mathbb{V}(\ln X_i) \end{aligned}$$

Puisque $\mathbb{V}(\ln X_i) = \sigma^2$, nous obtenons :

$$\mathbb{E}(\widehat{\sigma}^2) = \frac{n \times \sigma^2}{n} = \sigma^2$$

L'estimateur $\widehat{\sigma}^2$ est sans biais.

4. Il y a plusieurs façons de montrer ce résultat. Une façon consiste à utiliser la loi faible des grands nombres (► chapitre 8). La séquence de variables $(\ln X_i - \mu)^2$ est i.i.d. avec une espérance égale à :

$$\mathbb{E}((\ln X_i - \mu)^2) = \mathbb{E}((\ln X_i - \mathbb{E}(\ln X_i))^2) = \mathbb{V}(\ln X_i) = \sigma^2$$

D'après la loi faible des grands nombres, la moyenne empirique de variables i.i.d. converge en probabilité vers l'espérance :

$$\widehat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (\ln X_i - \mu)^2 \xrightarrow{p} \mathbb{E}((\ln X_i - \mu)^2) = \sigma^2$$

L'estimateur $\widehat{\sigma}^2$ est convergent au sens faible.

5. Le score associé à l'échantillon est défini par :

$$S_n(\sigma^2, X) = \frac{\partial \ell_n(\sigma^2; x)}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4} \sum_{i=1}^n (\ln X_i - \mu)^2$$

Par conséquent :

$$\begin{aligned} \mathbb{E}(S_n(\sigma^2, X)) &= -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4} \sum_{i=1}^n \mathbb{E}((\ln X_i - \mu)^2) \\ &= -\frac{n}{2\sigma^2} + \frac{n \times \sigma^2}{2\sigma^4} = 0 \quad \forall \sigma^2 \in \mathbb{R}^+ \end{aligned}$$

6. La quantité d'information de Fisher associée à l'échantillon est égale à :

$$\begin{aligned} I_n(\sigma^2) &= \mathbb{E}(-H_n(\sigma^2; X)) \\ &= \mathbb{E}\left(-\frac{n}{2\sigma^4} + \frac{1}{\sigma^6} \sum_{i=1}^n (\ln X_i - \mu)^2\right) \\ &= -\frac{n}{2\sigma^4} + \frac{1}{\sigma^6} \sum_{i=1}^n \mathbb{E}((\ln X_i - \mu)^2) \\ &= -\frac{n}{2\sigma^4} + \frac{n \times \sigma^2}{\sigma^6} = \frac{n}{2\sigma^4} \end{aligned}$$

La quantité moyenne d'information de Fisher est égale à :

$$I(\sigma^2) = \frac{1}{n} \times I_n(\sigma^2) = \frac{1}{2\sigma^4}$$

7. La fonction $\ln f_X(x; \sigma^2)$ satisfait les hypothèses de régularité. Par conséquent, l'estimateur du maximum de vraisemblance est asymptotiquement normalement distribué :

$$\sqrt{n}(\widehat{\sigma}^2 - \sigma^2) \xrightarrow{d} \mathcal{N}(0, I^{-1}(\sigma^2))$$

où $I(\sigma^2)$ correspond à la quantité moyenne d'information de Fisher. Ainsi, nous avons :

$$\sqrt{n}(\widehat{\sigma}^2 - \sigma^2) \xrightarrow{d} \mathcal{N}(0, 2\sigma^4)$$

Pour une taille d'échantillon suffisamment grande mais finie, ce résultat implique que :

$$\widehat{\sigma}^2 \approx \mathcal{N}\left(\sigma^2, \frac{2\sigma^4}{n}\right)$$

8. La variance asymptotique de l'estimateur $\widehat{\sigma}^2$ est égale à :

$$\mathbb{V}_{asy}(\widehat{\sigma}^2) = \frac{2\sigma^4}{n}$$

La borne FDCR est définie par l'inverse de la quantité d'information de Fisher associée à l'échantillon :

$$I_n^{-1}(\sigma^2) = \frac{2\sigma^4}{n}$$

L'estimateur $\widehat{\sigma}^2$ est asymptotiquement efficace puisque :

$$\mathbb{V}_{asy}(\widehat{\sigma}^2) = I_n^{-1}(\sigma^2) = \frac{2\sigma^4}{n}$$

9. Un estimateur convergent de cette variance asymptotique est donné par :

$$\widehat{\mathbb{V}}_{asy}(\widehat{\sigma}^2) = \frac{2\widehat{\sigma}^4}{n}$$

Un autre estimateur basé sur la hessienne (dérivée seconde de la log-vraisemblance) est donné par :

$$\widehat{\mathbb{V}}_{asy}(\widehat{\sigma}^2) = \left(-\sum_{i=1}^n H_i(\widehat{\theta}, x_i)\right) = -\frac{n}{2\widehat{\theta}^2} + \frac{1}{\widehat{\theta}^3} \sum_{i=1}^n (\ln x_i - \mu)^2$$

Asymptotiquement, les deux estimateurs donnent la même valeur, mais à distance finie ils peuvent fournir des estimations différentes.

6 Maximum de vraisemblance

1. On sait que les variables (X_1, \dots, X_n) sont i.i.d. de même loi que X . Dès lors, la log-vraisemblance associée au n -échantillon (X_1, \dots, X_n) s'écrit :

$$\begin{aligned}\ell_n(\beta; x) &= \ln L_n(\beta; x) \\ &= \sum_{i=1}^n \ln f_X(x_i; \alpha, \beta) \\ &= (\alpha - 1) \sum_{i=1}^n \ln(x_i) - \frac{1}{\beta} \sum_{i=1}^n x_i \\ &\quad - n \ln(\Gamma(\alpha)) - n\alpha \ln(\beta)\end{aligned}$$

2. Le gradient est défini par :

$$g_n(\beta; x) = \frac{\partial \ell_n(\beta; x)}{\partial \beta} = \frac{1}{\beta^2} \sum_{i=1}^n x_i - \frac{n\alpha}{\beta}$$

La hessienne est définie par :

$$H_n(\beta; x) = \frac{\partial^2 \ell_n(\beta; x)}{\partial \beta^2} = -\frac{2}{\beta^3} \sum_{i=1}^n x_i + \frac{n\alpha}{\beta^2}$$

3. Le score associé à l'échantillon est égal à :

$$S_n(\beta; X) = \frac{\partial \ell_n(\beta; X)}{\partial \beta} = \frac{1}{\beta^2} \sum_{i=1}^n X_i - \frac{n\alpha}{\beta}$$

Son espérance est égale à :

$$\begin{aligned}\mathbb{E}(S_n(\beta; X)) &= \mathbb{E}\left(\frac{1}{\beta^2} \sum_{i=1}^n X_i - \frac{n\alpha}{\beta}\right) \\ &= \frac{1}{\beta^2} \sum_{i=1}^n \mathbb{E}(X_i) - \frac{n\alpha}{\beta} \\ &= \frac{n\alpha\beta}{\beta^2} - \frac{n\alpha}{\beta} = 0\end{aligned}$$

4. Soit $\hat{\beta}$ l'estimateur du maximum de vraisemblance du paramètre β . Ce dernier vérifie :

$$\hat{\beta} = \arg \max_{\beta \in \mathbb{R}^+} \ell_n(\beta; x)$$

La condition nécessaire du programme d'optimisation de la log-vraisemblance (équation de log-vraisemblance) s'écrit :

$$g_n(\hat{\beta}; x) = \frac{1}{\hat{\beta}^2} \sum_{i=1}^n x_i - \frac{n\alpha}{\hat{\beta}} = 0$$

D'où l'on tire que :

$$\hat{\beta}(x) = \frac{1}{n\alpha} \sum_{i=1}^n x_i$$

La condition suffisante se déduit de $H_n(\hat{\beta}; x)$:

$$\begin{aligned}H_n(\hat{\beta}; x) &= -\frac{2}{\hat{\beta}^3} \sum_{i=1}^n x_i + \frac{n\alpha}{\hat{\beta}^2} \\ &= -\frac{2n\alpha\hat{\beta}}{\hat{\beta}^3} + \frac{n\alpha}{\hat{\beta}^2} = -\frac{n\alpha}{\hat{\beta}^2} < 0\end{aligned}$$

puisque le paramètre α est positif et $\hat{\beta}$ est positif (moyenne de réalisations x_i positives). L'estimateur du maximum de vraisemblance (variable aléatoire) est défini par :

$$\hat{\beta} = \frac{1}{n\alpha} \sum_{i=1}^n X_i$$

5. Calculons $\mathbb{E}(\hat{\beta})$:

$$\mathbb{E}(\hat{\beta}) = \mathbb{E}\left(\frac{1}{n\alpha} \sum_{i=1}^n X_i\right) = \frac{1}{n\alpha} \sum_{i=1}^n \mathbb{E}(X_i) = \frac{n\alpha\beta}{n\alpha} = \beta$$

L'estimateur $\hat{\beta}$ est sans biais.

6. Puisque les X_i sont i.i.d. de même loi que X , il vient :

$$\mathbb{V}(\hat{\beta}) = \mathbb{V}\left(\frac{1}{n\alpha} \sum_{i=1}^n X_i\right) = \frac{1}{\alpha^2 n^2} \sum_{i=1}^n \mathbb{V}(X_i) = \frac{n\alpha\beta^2}{\alpha^2 n^2} = \frac{\beta^2}{\alpha n}$$

Par conséquent :

$$\lim_{n \rightarrow \infty} \mathbb{V}(\hat{\beta}) = \lim_{n \rightarrow \infty} \frac{\beta^2}{\alpha n} = 0$$

Sachant que par ailleurs $\mathbb{E}(\hat{\beta}) = \beta$, on en déduit que $\hat{\beta}$ est un estimateur convergent (au sens faible) de β :

$$\hat{\beta} \xrightarrow{p} \beta$$

7. La quantité d'information de Fisher associée à l'échantillon est définie par :

$$I_n(\beta) = \mathbb{V}(S_n(\beta; X)) = \mathbb{E}(-H_n(\beta; X))$$

D'après les résultats de la question 2, il vient :

$$\begin{aligned}I_n(\beta) &= \mathbb{E}\left(\frac{2}{\beta^3} \sum_{i=1}^n X_i - \frac{n\alpha}{\beta^2}\right) = \frac{2}{\beta^3} \sum_{i=1}^n \mathbb{E}(X_i) - \frac{n\alpha}{\beta^2} \\ &= \frac{2n\alpha\beta}{\beta^3} - \frac{n\alpha}{\beta^2} = \frac{n\alpha}{\beta^2}\end{aligned}$$

Dès lors :

$$\mathbb{V}(\hat{\beta}) = I_n^{-1}(\beta) = \frac{\beta^2}{\alpha n}$$

L'estimateur du maximum de vraisemblance $\hat{\beta}$ est efficace au sens de la borne FDCR.

8. Les variables du n -échantillon (X_1, \dots, X_n) sont i.i.d. avec $\mathbb{E}(X_i) = \alpha\beta$ et $\mathbb{V}(X_i) = \alpha\beta^2$. D'après le théorème central limite :

$$\sqrt{n}\left(\frac{1}{n} \sum_{i=1}^n X_i - \mathbb{E}(X_i)\right) \xrightarrow{d} \mathcal{N}(0, \mathbb{V}(X))$$

ou encore :

$$\sqrt{n}\left(\frac{1}{n} \sum_{i=1}^n X_i - \alpha\beta\right) \xrightarrow{d} \mathcal{N}(0, \alpha\beta^2)$$

On en déduit la loi asymptotique de l'estimateur $\hat{\beta} = (\alpha n)^{-1} \sum_{i=1}^n X_i$ par la méthode delta en posant $\hat{\beta} = g(\bar{X}_n)$ avec $g(x) = x/\alpha$ et $\partial g(x)/\partial x = 1/\alpha$.

$$\sqrt{n}(g(\bar{X}_n) - g(\alpha\beta)) \xrightarrow{d} \mathcal{N}\left(0, \alpha\beta^2 \left(\frac{\partial g(x)}{\partial x}\right)_{\alpha\beta}^2\right)$$

On obtient finalement :

$$\sqrt{n}(\hat{\beta} - \beta) \xrightarrow{d} \mathcal{N}\left(0, \frac{\beta^2}{\alpha}\right)$$

9. Le problème est régulier, dès lors :

$$\sqrt{n}(\widehat{\beta} - \beta) \xrightarrow{d} \mathcal{N}(0, I(\beta)^{-1})$$

où $I(\beta)$ désigne la quantité d'information moyenne de Fisher, définie par :

$$I(\beta) = \frac{1}{n} I_n(\beta) = \frac{\alpha}{\beta^2}$$

On en conclut que :

$$\sqrt{n}(\widehat{\beta} - \beta) \xrightarrow{d} \mathcal{N}\left(0, \frac{\beta^2}{\alpha}\right)$$

On retrouve la même loi asymptotique qu'à la question précédente.

10. On sait que pour $\alpha = 2$:

$$\widehat{\beta} = \frac{1}{\alpha n} \sum_{i=1}^n X_i = \frac{\bar{X}_n}{2}$$

Sachant que dans cet échantillon, on obtient $\bar{X}_N = 4$, la réalisation de l'estimateur du maximum de vraisemblance est égale à :

$$\widehat{\beta}(x) = 2$$

Chapitre 11

1 Règle de décision d'un test statistique

- Faux. Un test statistique est une règle de décision qui porte sur le rejet ou le non rejet de l'hypothèse nulle.
- Faux. Le rejet de l'hypothèse nulle n'implique pas nécessairement l'acceptation de l'hypothèse alternative.
- Faux. Même si l'on ne rejette pas l'hypothèse nulle, cela ne signifie pas nécessairement que l'on accepte l'hypothèse nulle ou que cette hypothèse est vraie.
- Vrai. C'est précisément ce que l'on peut conclure d'un test.
- Faux. Le rejet ou le non rejet de l'hypothèse nulle dépend du niveau du test.

2 Région critique d'un test

- Vrai. C'est la définition de la région critique.
- Vrai. Comme toute statistique, une statistique de test est une fonction des variables aléatoires de l'échantillon, c'est donc une variable aléatoire.
- Faux. Ce n'est pas toujours le cas. On peut définir un test à partir d'une transformée d'un estimateur du paramètre dont on connaît la loi sous l'hypothèse nulle.
- Vrai.
- Vrai. C'est précisément la règle de décision d'un test.

3 Niveau et puissance d'un test

- Faux. Le niveau d'un test correspond à la probabilité de rejeter à tort l'hypothèse nulle.
- Faux. Cela dépend de la forme de la région critique.
- Vrai. C'est ainsi que l'on résout l'arbitrage entre le risque de première espèce et le risque de deuxième espèce en fonction de la valeur critique du test.
- Vrai. C'est la définition de la puissance qui correspond à un moins la probabilité de risque de deuxième espèce.

- Faux. C'est la puissance d'un test convergent qui tend vers l'unité lorsque la taille d'échantillon tend vers l'infini.

4 Test paramétrique et lemme de Neyman-Pearson

- Vrai. C'est la définition de l'objet du lemme de Neyman-Pearson.
- Faux. La région critique sera la forme $W = \{x : \widehat{\theta}(x) < c\}$. C'est lorsque la réalisation $\widehat{\theta}(x)$ est relativement petite, par rapport à la valeur critique, que l'on rejette l'hypothèse nulle.
- Faux. La région critique est de la forme $W = \{x : |\widehat{\theta}(x) - \theta_0| < c\}$. Lorsque la distance $|\widehat{\theta}(x) - \theta_0|$ (en valeur absolue) excède la valeur critique, on rejette l'hypothèse nulle.
- Faux. Il n'existe pas de test UPP dans le cas d'un test bilatéral.
- Faux. C'est la région de non rejet d'un test bilatéral de niveau α qui correspond à l'intersection des régions de non rejet des tests unilatéraux associés de niveau $\alpha/2$.

5 Tests paramétriques

- La log-vraisemblance associée à l'échantillon (x_1, \dots, x_n) s'écrit :

$$\ell_n(\sigma^2; x) = \ln L_n(\sigma^2; x) = \sum_{i=1}^n \ln(x_i) - n \ln(\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^n x_i^2$$

D'après le lemme de Neyman-Pearson, la région critique du test UPP de niveau α est déterminée par :

$$W = \left\{ x \mid \frac{L_n(\sigma_0^2; x)}{L_n(\sigma_1^2; x)} < k \right\}$$

où k est une constante déterminée par le niveau de risque de première espèce α . Cette inégalité peut se réécrire sous la forme :

$$\ell_n(\sigma_0^2; x) - \ell_n(\sigma_1^2; x) < \ln(k)$$

$$\iff n \left(\log(\sigma_1^2) - \log(\sigma_0^2) \right) + \left\{ \frac{1}{\sigma_1^2} - \frac{1}{\sigma_0^2} \right\} \frac{1}{2} \sum_{i=1}^n x_i^2 < \ln(k)$$

$$\iff \left\{ \frac{\sigma_0^2 - \sigma_1^2}{\sigma_0^2 \sigma_1^2} \right\} \frac{1}{2} \sum_{i=1}^n x_i^2 < k_1$$

avec $k_1 = \ln(k) - n \left(\log(\sigma_1^2) - \log(\sigma_0^2) \right)$. Puisque dans notre cas $\sigma_1^2 > \sigma_0^2$, on a :

$$\frac{1}{2n} \sum_{i=1}^n x_i^2 > c$$

où $c = k_1 \sigma_0^2 \sigma_1^2 / ((\sigma_0^2 - \sigma_1^2)n)$. La région critique du test UPP de niveau α est de la forme :

$$W = \{x \mid T_n(x) > c\}$$

où c est une constante déterminée par le risque de première espèce.

2. La distribution asymptotique de la statistique de test T_n sous l'hypothèse nulle $H_0 : \sigma^2 = \sigma_0^2$ est :

$$T_n \underset{H_0}{\overset{asy}{\approx}} \mathcal{N}\left(\sigma_0^2, \frac{\sigma_0^4}{n}\right) \iff \frac{T_n - \sigma_0^2}{\sigma_0^2/\sqrt{n}} \underset{H_0}{\overset{asy}{\approx}} \mathcal{N}(0,1)$$

Par définition du risque de première espèce :

$$\begin{aligned} \alpha &= \Pr(W|H_0) = \Pr(T_n > c|H_0) \\ &= 1 - \Pr\left(\frac{T_n - \sigma_0^2}{\sigma_0^2/\sqrt{n}} \leq \frac{c - \sigma_0^2}{\sigma_0^2/\sqrt{n}} \middle| H_0\right) \\ &= 1 - \Phi\left(\frac{c - \sigma_0^2}{\sigma_0^2/\sqrt{n}}\right) \end{aligned}$$

où $\Phi(\cdot)$ désigne la fonction de répartition de la loi normale centrée réduite. D'où l'on tire que :

$$1 - \alpha = \Phi\left(\frac{c - \sigma_0^2}{\sigma_0^2/\sqrt{n}}\right)$$

La valeur critique associée à un niveau $\alpha = 5\%$ est égale à :

$$c = \sigma_0^2 + \Phi^{-1}(1 - \alpha) \frac{\sigma_0^2}{\sqrt{n}} \quad (*)$$

Il vient :

$$c = 2 + \Phi^{-1}(0,95) \times \frac{2}{\sqrt{500}} = 2 + 1,64 \times \frac{2}{\sqrt{500}} = 2,14$$

La région critique du test UPP de niveau $\alpha = 5\%$ est définie par :

$$W = \{x | T_n(x) > 2,14\}$$

3. Sous l'hypothèse alternative $H_1 : \sigma^2 = \sigma_1^2 = 2,1$, la distribution asymptotique de la statistique de test T_n devient :

$$T_n \underset{H_1}{\overset{asy}{\approx}} \mathcal{N}\left(\sigma_1^2, \frac{\sigma_1^4}{n}\right) \iff \frac{T_n - \sigma_1^2}{\sigma_1^2/\sqrt{n}} \underset{H_1}{\overset{asy}{\approx}} \mathcal{N}(0,1)$$

Par définition de la puissance, $\Pr(W|H_1)$, on a :

$$\begin{aligned} \text{Puissance} &= \Pr(T_n > c|H_1) \\ &= 1 - \Pr\left(\frac{T_n - \sigma_1^2}{\sigma_1^2/\sqrt{n}} \leq \frac{c - \sigma_1^2}{\sigma_1^2/\sqrt{n}} \middle| H_1\right) \\ &= 1 - \Phi\left(\frac{c - \sigma_1^2}{\sigma_1^2/\sqrt{n}}\right) \end{aligned} \quad (**)$$

On a donc une puissance égale à :

$$\begin{aligned} \text{Puissance} &= 1 - \Phi\left(\frac{2,14 - 2,1}{2,1/\sqrt{500}}\right) = 1 - \Phi(0,42) \\ &= 1 - 0,66 = 0,34 \end{aligned}$$

Avec cette règle de décision, il y a 34 % de chances de rejeter l'hypothèse $H_0 : \sigma^2 = \sigma_0^2 = 2$ lorsque l'hypothèse $H_1 : \sigma^2 = \sigma_1^2 = 2,1$ est vraie dans la population.

4. En remplaçant c par son expression (équation (*)) dans la puissance (équation (**)), il vient :

$$\text{Puissance} = 1 - \Phi\left(\frac{\sigma_0^2 - \sigma_1^2}{\sigma_1^2/\sqrt{n}} + \frac{\sigma_0^2}{\sigma_1^2} \Phi^{-1}(1 - \alpha)\right)$$

On vérifie que si σ_1^2 tend vers σ_0^2 , alors la puissance tend vers le risque de première espèce α (le test est non biaisé).

De plus, puisque $\sigma_0^2 < \sigma_1^2$, si n tend vers l'infini, la puissance tend vers l'unité. En effet :

$$\lim_{n \rightarrow \infty} \frac{\sigma_0^2 - \sigma_1^2}{\sigma_1^2/\sqrt{n}} = -\infty$$

Par conséquent, $\forall \sigma_1^2 > \sigma_0^2$, on a :

$$\begin{aligned} 1 - \lim_{n \rightarrow \infty} \Phi\left(\frac{\sigma_0^2 - \sigma_1^2}{\sigma_1^2/\sqrt{n}} + \frac{\sigma_0^2}{\sigma_1^2} \Phi^{-1}(1 - \alpha)\right) &= 1 - \Phi(-\infty) \\ &= 1 - 0 = 1 \end{aligned}$$

Le test est convergent.

5. La réalisation de la statistique de test $T_n(x)$ est égale à :

$$T_n(x) = \frac{1}{2n} \sum_{i=1}^n x_i^2 = \frac{2\,070}{2 \times 500} = 2,07$$

La région critique du test pour un niveau $\alpha = 5\%$ est $W = \{x | T_n(x) > 2,14\}$. La réalisation de la statistique de test n'appartient pas à la région critique. Par conséquent, on ne peut pas rejeter l'hypothèse nulle $H_0 : \sigma^2 = 2$ pour un seuil de risque de 5 %.

6. La région critique du test $H_0 : \sigma^2 = \sigma_0^2$ contre $H_1 : \sigma^2 = \sigma_1^2$ ne dépend pas de la valeur de σ_1^2 . La région critique du test UPP unilatéral $H_0 : \sigma^2 = \sigma_0^2$ contre $H_1 : \sigma^2 > \sigma_0^2$ est identique :

$$\begin{aligned} W &= \left\{x | T_n(x) > \sigma_0^2 + \Phi^{-1}(1 - \alpha) \frac{\sigma_0^2}{\sqrt{n}}\right\} \\ &= \{x | T_n(x) > 2,14\} \end{aligned}$$

7. La région de non rejet du test bilatéral de niveau α est définie par l'intersection des régions de non rejet des tests unilatéraux UPP de niveau $\alpha/2$:

$$\text{Test A : } H_0 : \sigma^2 = \sigma_0^2 \text{ contre } H_1 : \sigma^2 > \sigma_0^2$$

$$\text{Test B : } H_0 : \sigma^2 = \sigma_0^2 \text{ contre } H_1 : \sigma^2 < \sigma_0^2$$

Soient \overline{W}_A et \overline{W}_B les régions de non rejet associées :

$$\begin{aligned} \overline{W}_A &= \left\{x | T_n(x) < \sigma_0^2 + \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \frac{\sigma_0^2}{\sqrt{n}}\right\} \\ &= \{x | T_n(x) < 2,17\} \end{aligned}$$

$$\begin{aligned} \overline{W}_B &= \left\{x | T_n(x) > \sigma_0^2 + \Phi^{-1}\left(\frac{\alpha}{2}\right) \frac{\sigma_0^2}{\sqrt{n}}\right\} \\ &= \{x | T_n(x) > 1,83\} \end{aligned}$$

La région de non rejet $\overline{W} = \overline{W}_A \cap \overline{W}_B$ et la région critique W du test bilatéral de niveau $\alpha = 5\%$ sont donc respectivement définies par :

$$\overline{W} = \{x | 1,83 < T_n(x) < 2,17\}$$

$$W = \{x | T_n(x) \notin [1,83; 2,17]\}$$

De façon générale, la région de non rejet peut aussi s'écrire sous la forme :

$$\overline{W} = \left\{x \middle| \left| \frac{T_n(x) - \sigma_0^2}{\sigma_0^2/\sqrt{n}} \right| < \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)\right\}$$

La région critique du test bilatéral de niveau α est alors définie par :

$$W = \left\{ x \mid \left| \frac{T_n(x) - \sigma_0^2}{\sigma_0^2 / \sqrt{n}} \right| \geq \Phi^{-1} \left(1 - \frac{\alpha}{2} \right) \right\}$$

6 Tests paramétriques

1. Les variables X_1, \dots, X_n sont i.i.d., par conséquent la log-vraisemblance associée à l'échantillon (x_1, \dots, x_n) est égale à :

$$\ell_n(\theta; x) = \sum_{i=1}^n \ln f_X(x_i; \theta)$$

avec :

$$\ln f_X(x_i; \theta) = -\ln(\theta) - \frac{1}{\theta} \ln(c) + \left(\frac{1-\theta}{\theta} \right) \ln(x_i)$$

Ainsi, on obtient :

$$\ell_n(\theta; x) = -n \ln(\theta) - \frac{n}{\theta} \ln(c) + \left(\frac{1-\theta}{\theta} \right) \sum_{i=1}^n \ln(x_i)$$

2. L'estimateur du maximum de vraisemblance $\hat{\theta}$ du paramètre θ vérifie :

$$\hat{\theta} = \arg \max_{\theta \in \mathbb{R}^+} \ell_n(\theta; x)$$

La condition nécessaire (équation de log-vraisemblance) est :

$$\begin{aligned} g_n(\hat{\theta}; x) &= \left. \frac{\partial \ell_n(\theta; x)}{\partial \theta} \right|_{\hat{\theta}} \\ &= -\frac{n}{\hat{\theta}} + \frac{n}{\hat{\theta}^2} \ln(c) - \frac{1}{\hat{\theta}^2} \sum_{i=1}^n \ln(x_i) = 0 \end{aligned}$$

En résolvant cette équation, il vient :

$$\hat{\theta} = \ln(c) - \frac{1}{n} \sum_{i=1}^n \ln(x_i)$$

La condition suffisante de ce programme est :

$$H_n(\hat{\theta}; x) = \left. \frac{\partial^2 \ell_n(\theta; x)}{\partial \theta^2} \right|_{\hat{\theta}} = \frac{n}{\hat{\theta}^2} - \frac{2n}{\hat{\theta}^3} \left(\ln(c) - \frac{1}{n} \sum_{i=1}^n \ln(x_i) \right)$$

En utilisant les résultats de la condition nécessaire, il vient :

$$\begin{aligned} H_n(\hat{\theta}; x) &= \frac{n}{\hat{\theta}^2} - \frac{2n}{\hat{\theta}^3} \left(\ln(c) - \frac{1}{n} \sum_{i=1}^n \ln(x_i) \right) \\ &= \frac{n}{\hat{\theta}^2} - \frac{2n\hat{\theta}}{\hat{\theta}^3} \\ &= -\frac{n}{\hat{\theta}^2} < 0 \end{aligned}$$

Nous avons donc un maximum. L'estimateur (variable aléatoire) du maximum de vraisemblance $\hat{\theta}$ du paramètre θ est égal à :

$$\hat{\theta} = \ln(c) - \frac{1}{n} \sum_{i=1}^n \ln(x_i)$$

3. La séquence de variables aléatoires i.i.d. $\ln(X_1), \dots, \ln(X_n)$ vérifie $\mathbb{E}(\ln(X_i)) = \ln(c) - \theta$. D'après la loi faible des grands nombres :

$$\frac{1}{n} \sum_{i=1}^n \ln(X_i) \xrightarrow{p} \mathbb{E}(\ln(X_i)) = \ln(c) - \theta$$

En utilisant le *continuous mapping theorem* pour une fonction $g(z) = \ln(c) - z$, on obtient :

$$g\left(\frac{1}{n} \sum_{i=1}^n \ln(X_i)\right) \xrightarrow{p} g(\ln(c) - \theta)$$

ou de façon équivalente :

$$\ln(c) - \frac{1}{n} \sum_{i=1}^n \ln(X_i) \xrightarrow{p} \ln(c) - \ln(c) + \theta$$

Par conséquent :

$$\hat{\theta} \xrightarrow{p} \theta$$

L'estimateur $\hat{\theta}$ est convergent (au sens faible).

4. Puisque le problème est régulier, nous avons :

$$\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{d} \mathcal{N}(0, \mathbf{I}^{-1}(\theta))$$

où θ désigne la vraie valeur du paramètre et $\mathbf{I}(\theta_0)$ correspond à la quantité d'information de Fisher moyenne.

$$\mathbf{I}(\theta) = \frac{1}{n} \times \mathbf{I}_n(\theta)$$

avec :

$$\begin{aligned} \mathbf{I}_n(\theta) &= \mathbb{E}(-H_n(\theta; X)) \\ &= \mathbb{E}\left(-\frac{n}{\theta^2} + \frac{2n}{\theta^3} \left(\ln(c) - \frac{1}{n} \sum_{i=1}^n \ln(X_i) \right)\right) \\ &= -\frac{n}{\theta^2} + \frac{2n}{\theta^3} \left(\ln(c) - \frac{1}{n} \sum_{i=1}^n \mathbb{E}(\ln(X_i)) \right) \\ &= -\frac{n}{\theta^2} + \frac{2n}{\theta^3} (\ln(c) - \ln(c) + \theta) \\ &= -\frac{n}{\theta^2} + \frac{2n\theta}{\theta^3} \\ &= \frac{n}{\theta^2} \end{aligned}$$

Par conséquent :

$$\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{d} \mathcal{N}(0, \theta^2)$$

ou de façon équivalente :

$$\hat{\theta}^{asy} \approx \mathcal{N}\left(\theta, \frac{\theta^2}{n}\right)$$

5. D'après le lemme de Neyman Pearson, la région critique du test UPP de niveau α est donné par :

$$\frac{L_n(\theta_0; x)}{L_n(\theta_1; x)} < K$$

où K est une constante déterminée par le niveau α ou de façon équivalente :

$$\ell_n(\theta_0; x) - \ell_n(\theta_1; x) < \ln(K)$$

On obtient :

$$\begin{aligned} -n \ln(\theta_0) - \frac{n}{\theta_0} \ln(c) + \left(\frac{1}{\theta_0} - 1\right) \sum_{i=1}^n \ln(x_i) + n \ln(\theta_1) \\ + \frac{n}{\theta_1} \ln(c) - \left(\frac{1}{\theta_1} - 1\right) \sum_{i=1}^n \ln(x_i) < \ln(K) \\ \iff \left(\frac{1}{\theta_0} - \frac{1}{\theta_1}\right) \sum_{i=1}^n \ln(x_i) < K_2 \end{aligned}$$

où $K_2 = \ln(K) + n(\ln(\theta_0) - \ln(\theta_1)) + n \ln(c) (\theta_0^{-1} - \theta_1^{-1})$ est un terme constant.

$$\left(\frac{\theta_1 - \theta_0}{\theta_0 \theta_1}\right) \sum_{i=1}^n \ln(x_i) < K_2$$

Puisque $\theta_1 < \theta_0$, on a :

$$\sum_{i=1}^n \ln(x_i) > K_3$$

avec $K_3 = K_2 \theta_0 \theta_1 / (\theta_1 - \theta_0)$. Cette inégalité peut se réécrire :

$$\ln(c) - \frac{1}{n} \sum_{i=1}^n \ln(x_i) < K_4$$

avec $K_4 = \ln(c) - K_3/n$. La région critique du test UPP de niveau α est de la forme :

$$W = \left\{x : \widehat{\theta}(x) < A\right\}$$

où A est une constante déterminée par le niveau α .

6. D'après la définition du risque de première espèce :

$$\alpha = \Pr(W | H_0)$$

Donc :

$$\alpha = \Pr\left(\widehat{\theta} < A \mid \widehat{\theta} \overset{asy}{\approx} \mathcal{N}\left(\theta_0, \frac{\theta_0^2}{n}\right)\right)$$

ou de façon équivalente :

$$\begin{aligned} \alpha &= \Pr\left(\frac{\widehat{\theta} - \theta_0}{\theta_0/\sqrt{n}} < \frac{A - \theta_0}{\theta_0/\sqrt{n}} \mid \frac{\widehat{\theta} - \theta_0}{\theta_0/\sqrt{n}} \overset{asy}{\approx} \mathcal{N}(0, 1)\right) \\ &= \Phi\left(\frac{A - \theta_0}{\theta_0/\sqrt{n}}\right) \end{aligned}$$

où $\Phi(\cdot)$ désigne la fonction de répartition de la loi normale centrée réduite. A partir de cette expression, on peut déduire la valeur critique du test UPP de niveau α :

$$A = \theta_0 + \frac{\theta_0}{\sqrt{n}} \Phi^{-1}(\alpha)$$

7. On considère le test d'hypothèse simple contre hypothèse simple :

$$H_0 : \theta = \theta_0 \text{ contre } H_1 : \theta = \theta_1$$

avec $\theta_1 < \theta_0$. La région critique du test UPP de niveau α est :

$$W = \left\{x : \widehat{\theta}(x) < \theta_0 + \frac{\theta_0}{\sqrt{n}} \Phi^{-1}(\alpha)\right\}$$

Cette région critique ne dépend pas de la valeur de θ_1 . Elle correspond donc à celle du test unilatéral :

$$H_0 : \theta = \theta_0 \text{ contre } H_1 : \theta < \theta_0$$

8. On considère les tests unilatéraux :

$$\text{Test A : } H_0 : \theta = \theta_0 \text{ contre } H_1 : \theta < \theta_0$$

$$\text{Test B : } H_0 : \theta = \theta_0 \text{ contre } H_1 : \theta > \theta_0$$

Les régions de non rejet des tests unilatéraux UPP de niveau $\alpha/2$ sont égales à :

$$\overline{W}_A = \left\{x : \widehat{\theta}(x) > \theta_0 + \frac{\theta_0}{\sqrt{n}} \Phi^{-1}\left(\frac{\alpha}{2}\right)\right\}$$

$$\overline{W}_B = \left\{x : \widehat{\theta}(x) < \theta_0 + \frac{\theta_0}{\sqrt{n}} \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)\right\}$$

Donc la région de non rejet du test bilatéral de niveau α est définie par :

$$\overline{W} = \left\{x : \theta_0 + \frac{\theta_0}{\sqrt{n}} \Phi^{-1}\left(\frac{\alpha}{2}\right) < \widehat{\theta}(x) < \theta_0 + \frac{\theta_0}{\sqrt{n}} \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)\right\}$$

Puisque $\Phi^{-1}(\alpha/2) = -\Phi^{-1}(1 - \alpha/2)$, cette région peut se réécrire sous la forme :

$$\overline{W} = \left\{x : |\widehat{\theta}(x) - \theta_0| < \frac{\theta_0}{\sqrt{n}} \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)\right\}$$

La région critique du test bilatéral de niveau α est donnée par :

$$W = \left\{x : |\widehat{\theta}(x) - \theta_0| > \frac{\theta_0}{\sqrt{n}} \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)\right\}$$

9. Par définition de la fonction puissance :

$$P(\theta) = \Pr(W | H_1) \quad \forall \theta \neq \theta_0$$

Sous l'hypothèse alternative :

$$\widehat{\theta} \overset{asy}{\underset{H_1}{\approx}} \mathcal{N}\left(\theta, \frac{\theta^2}{n}\right)$$

avec $\theta \neq \theta_0$. La fonction puissance est égale à :

$$\begin{aligned} P(\theta) &= 1 - \Pr(\overline{W} | H_1) \\ &= 1 - \Pr\left(\theta_0 + \frac{\theta_0}{\sqrt{n}} \Phi^{-1}\left(\frac{\alpha}{2}\right) < \widehat{\theta} < \theta_0 + \frac{\theta_0}{\sqrt{n}} \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)\right) \\ &= 1 - \Pr\left(\widehat{\theta} < \theta_0 + \frac{\theta_0}{\sqrt{n}} \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)\right) \\ &\quad + \Pr\left(\widehat{\theta} < \theta_0 + \frac{\theta_0}{\sqrt{n}} \Phi^{-1}\left(\frac{\alpha}{2}\right)\right) \\ &= 1 - \Phi\left(\frac{\theta_0 + \frac{\theta_0}{\sqrt{n}} \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) - \theta}{\theta/\sqrt{n}}\right) \\ &\quad + \Phi\left(\frac{\theta_0 + \frac{\theta_0}{\sqrt{n}} \Phi^{-1}\left(\frac{\alpha}{2}\right) - \theta}{\theta/\sqrt{n}}\right) \end{aligned}$$

On obtient :

$$P(\theta) = 1 - \Phi\left(\frac{\theta_0 - \theta}{\theta/\sqrt{n}} + \frac{\theta_0}{\theta}\Phi^{-1}\left(1 - \frac{\alpha}{2}\right)\right) \\ + \Phi\left(\frac{\theta_0 - \theta}{\theta/\sqrt{n}} + \frac{\theta_0}{\theta}\Phi^{-1}\left(\frac{\alpha}{2}\right)\right)$$

Lorsque n tend vers l'infini, deux cas doivent être considérés. Si $\theta > \theta_0$, alors :

$$\lim_{n \rightarrow \infty} P(\theta) = 1 - \Phi(-\infty) + \Phi(-\infty) = 1$$

Si $\theta < \theta_0$, on a :

$$\lim_{n \rightarrow \infty} P(\theta) = 1 - \Phi(+\infty) + \Phi(+\infty) = 1 - 1 + 1 = 1$$

Quelle que soit la valeur de θ , la fonction puissance tend vers 1 :

$$\lim_{n \rightarrow \infty} P(\theta) = 1$$

Le test est convergent.