TD 1 : Estimation Ponctuelle

Pour estimer ponctuellement les paramètres d'une population, on utilise un échantillon de taille n : $\{x_1, x_2, \cdots, x_n\}$. Voici les formules nécessaires :

— La moyenne d'échantillon $\overline{x}$ est un estimateur de la moyenne de la population μ. Elle est donnée par :

$$\boxed{\overline{x} = \frac{x_1 + x_2 + \cdots + x_n}{n}}$$

— La variance d'échantillon s^2 est un estimateur de la variance de la population σ^2. Elle est donnée par :

$$\boxed{s^2 = \frac{(x_1 - \overline{x})^2 + (x_2 - \overline{x})^2 + \cdots + (x_n - \overline{x})^2}{n-1}}$$

— L'écart-type d'échantillon s est un estimateur de l'écart-type de la population σ. Il est donné par :

$$\boxed{s = \sqrt{s^2}}$$

— La proportion $\overline{p}$ des x éléments de l'échantillon caractérisés par un caractère donné est un estimateur de la proportion p des éléments de la population caractérisés par le caractère en question. Elle est donnée par :

$$\boxed{\overline{p} = \frac{x}{n}}$$

Exercice 1

On relève le nombre d'heures de travail hebdomadaire déclarées par un échantillon de 6 employés d'une start-up : 38, 42, 45, 36, 40 et 43.

1) Calculez l'estimation ponctuelle de la moyenne hebdomadaire de la population.

2) Que représente concrètement cette valeur ?

Solution 1

1) Moyenne d'échantillon :

$$\overline{x} = \frac{38 + 42 + 45 + 36 + 40 + 43}{6} = \frac{244}{6} \approx 40.67 \text{ heures}$$

2) $\overline{x} = 40.67$ h est l'estimation ponctuelle de μ, la moyenne réelle des heures travaillées par l'ensemble des employés de la start-up. On s'attend à ce que la population travaille en moyenne environ 40 heures et 40 minutes par semaine.

Exercice 2

Lors d'un sondage auprès de 150 ménages, 98 déclarent avoir réduit leurs dépenses non essentielles face à l'inflation.

1) Estimez ponctuellement la proportion de ménages ayant réduit leurs dépenses.

2) Exprimez ce résultat en pourcentage et commentez sa fiabilité immédiate.

Solution 2

1) Proportion d'échantillon :

$$\overline{p} = \frac{x}{n} = \frac{98}{150} \approx 0.6533$$

2) En pourcentage, cela représente 65,33 % des ménages. Cette estimation est ponctuelle : elle ne donne pas une marge d'erreur. Elle suppose que l'échantillon est aléatoire et représentatif. En pratique, un intervalle de confiance serait nécessaire pour quantifier la précision de cette estimation.

Exercice 3

Considérez l'échantillon suivant des rendements annuels (en %) de 5 fonds d'investissement : 8.2, 6.5, 9.1, 7.4 et 8.8.

1) Calculez l'écart-type d'échantillon s.

2) Pourquoi utilise-t-on le dénominateur $n-1$ plutôt que n ?

Solution 3

1) Avant de calculer l'écart-type d'échantillon s, il faut d'abord calculer la moyenne d'échantillon $\overline{x}$ et la variance d'échantillon s^2 :

$$\begin{aligned}
\overline{x} &= \frac{8.2 + 6.5 + 9.1 + 7.4 + 8.8}{5} = \frac{40}{5} = 8\,\% \\
s^2 &= \frac{(8.2-8)^2 + (6.5-8)^2 + (9.1-8)^2 + (7.4-8)^2 + (8.8-8)^2}{5-1} \\
&= \frac{0.04 + 2.25 + 1.21 + 0.36 + 0.64}{5} = \frac{4.5}{4} = 1.125
\end{aligned}$$

L'écart-type d'échantillon s est égal à :

$$s = \sqrt{s^2} = \sqrt{1.125} \approx 1.06\,\%$$

2) Le dénominateur $n-1$ (degrés de liberté) corrige le biais systématique qui sous-estimerait la variance de la population si on utilisait n. Cela rend s^2 un estimateur **sans biais** de σ^2.

Exercice 4

Un échantillon de 8 entreprises du secteur textile fournit les données suivantes sur leur marge brute (en %) : 12.4, 15.1, 11.8, 14.2, 16.0, 13.5, 14.8 et 12.9.

1) Estimez ponctuellement la marge moyenne et l'écart-type de la population.

2) Interprétez s dans ce contexte économique.

Solution 4

1) Estimation ponctuelle de la marge moyenne (calcul de la moyenne d'échantillon $\overline{x}$) :

$$\overline{x} = \frac{12.4 + 15.1 + 11.8 + 14.2 + 16.0 + 13.5 + 14.8 + 12.9}{8} = \frac{110.7}{8} = 13.8375 \approx 13.84\,\%$$

Estimation de la variance de la population (calcul de la variance d'échantillon s^2) :

$$\begin{aligned} s^2 &= \frac{(12.4 - 13.84)^2 + (15.1 - 13.84)^2 + (11.8 - 13.84)^2 + (14.2 - 13.84)^2 + \cdots + (12.9 - 13.84)^2}{8 - 1} \\ &= \frac{2.0736 + 1.5876 + 4.1616 + 0.1296 + 4.6656 + 0.1156 + 0.9216 + 0.8836}{7} = \frac{14.5388}{7} \approx 2.08 \end{aligned}$$

Estimation de l'écart-type de la population (calcul de l'écart-type d'échantillon s) :

$$s = \sqrt{s^2} = \sqrt{2.08} \approx 1.44\,\%$$

2) L'estimation de l'écart-type indique une dispersion moyenne d'environ ± 1.44 point de pourcentage autour de la marge moyenne de 13.84 %. Cela reflète l'hétérogénéité des performances ou des modèles économiques au sein du secteur textile.

Exercice 5

Lors d'une enquête auprès de 300 diplômés en économie, 210 travaillent dans le secteur privé, 60 dans le public, et 30 sont au chômage ou en reprise d'études.

1) Estimez la proportion de diplômés travaillant dans le secteur privé.

2) Estimez la proportion de diplômés travaillant dans le secteur *non public* (privé + autres).

Solution 5

1) Estimation de la proportion de diplômés travaillant dans le secteur privé :

$$\overline{p}_{\text{privé}} = \frac{210}{300} = 0.70 = 70\,\%$$

2) Estimation de la proportion de diplômés travaillant dans le secteur *non public* :

$$\overline{p}_{\text{non public}} = \frac{210 + 30}{300} = \frac{240}{300} = 0.80 = 80\,\%$$

Exercice 6

Un responsable qualité relève le poids (en g) de 6 paquets de café : 248, 252, 249, 251, 250, 254.

1) Calculez la moyenne d'échantillon $\overline{x}$ et l'écart-type d'échantillon s.

2) Si le 6$^{\text{e}}$ paquet avait pesé 270 g au lieu de 254 g, comment $\overline{x}$ et s évolueraient-ils ? Que cela indique-t-il sur la robustesse de ces estimateurs ?

Solution 6

1) Calcul de la moyenne d'échantillon :

$$\overline{x} = \frac{248 + 252 + 249 + 251 + 250 + 254}{6} = \frac{1504}{6} \approx 250.67\,g$$

Calcul de l'écart-type d'échantillon :

$$\begin{aligned} s^2 &= \frac{(248 - 250.67)^2 + (252 - 250.67)^2 + (249 - 250.67)^2 + \cdots + (254 - 250.67)^2}{6-1} \\ &= \frac{7.1289 + 1.7689 + 2.7889 + 0.1089 + 0.4489 + 11.0889}{5} = \frac{23.3334}{5} \approx 4.67 \\ s &= \sqrt{s^2} = \sqrt{4.67} \approx 2.16\,g \end{aligned}$$

2) Calcul de la nouvelle moyenne d'échantillon :

$$\overline{x}' = \frac{248 + 252 + 249 + 251 + 250 + 270}{6} = \frac{1520}{6} \approx 253.33\,g$$

Calcul du nouvel écart-type d'échantillon :

$$\begin{aligned} s'^2 &= \frac{(248 - 253.33)^2 + (252 - 253.33)^2 + (249 - 253.33)^2 + \cdots + (270 - 253.33)^2}{6-1} \\ &= \frac{28.4089 + 1.7689 + 18.7489 + 5.4289 + 11.0889 + 277.8889}{5} = \frac{343.3334}{5} \approx 68.67 \\ s' &= \sqrt{s'^2} = \sqrt{68.67} \approx 8.29\,g \end{aligned}$$

La moyenne $\overline{x}$ est passée de 250.67 g à 253.33 g. L'écart-type s, basé sur les carrés des écarts à la moyenne, est passé de 2.16 g à 8.29 g. La moyenne $\overline{x}$ et l'écart-type s sont sensibles (non robustes) aux valeurs extrêmes. En effet, la valeur 270 g est bien éloignée des autres valeurs, elle est de ce fait considérée comme une valeur extrême.

Exercice 7

Une banque centrale souhaite estimer le taux d'inflation mensuel moyen μ. Un échantillon aléatoire de 12 régions donne les taux suivants (%) : 1.8, 2.1, 1.9, 2.3, 2.0, 1.7, 2.2, 2.1, 1.9, 2.0, 2.4 et 1.8.

1) Calculez l'estimation ponctuelle de μ.

2) La banque annonce officiellement 2,0 %. L'estimation ponctuelle confirme-t-elle cette annonce ?

Solution 7

1) Estimation ponctuelle de μ :

$$\overline{x} = \frac{1.8 + 2.1 + 1.9 + 2.3 + 2.0 + 1.7 + 2.2 + 2.1 + 1.9 + 2.0 + 2.4 + 1.8}{12} = \frac{24.2}{12} \approx 2.02\,\%$$

2) Oui, l'estimation ponctuelle 2,02 % est extrêmement proche de 2,00 %. Cependant, une estimation ponctuelle ne permet pas de conclure statistiquement à une égalité exacte. Une différence aussi faible est très probablement due au hasard d'échantillonnage. Un test d'hypothèse ou un intervalle de confiance serait requis pour une validation formelle.

Exercice 8

Sur un lot de 500 pièces électroniques inspectées, 18 présentent un défaut de soudure.

1) Estimez la proportion de pièces défectueuses dans la population.

2) Si le seuil d'acceptation du fournisseur est de 5 %, que pouvez-vous dire à ce stade ?

Solution 8

1) Estimation de la proportion de pièces défectueuses dans la population :

$$\overline{p} = \frac{18}{500} = 0.036 = 3.6\,\%$$

2) Ponctuellement, 3,6 % < 5 %, donc le lot semble conforme. Mais l'estimation ponctuelle ignore la variabilité d'échantillonnage. Il est possible que la vraie proportion p soit légèrement supérieure à 3,6 %. Sans intervalle de confiance, on ne peut pas garantir que $p < 0{,}05$ avec un niveau de risque contrôlé.

Exercice 9

On vous donne la série suivante de dépenses mensuelles en carburant (€) pour 7 ménages : 120, 135, 110, 145, 125, 130 et140.

1) Calculez la moyenne d'échantillon $\overline{x}$ et l'écart-type d'échantillon s.

2) Vérifiez numériquement que $\sum(x_i - \overline{x}) = 0$.

Solution 9

1) Calcul de la moyenne d'échantillon $\overline{x}$ et de l'écart-type d'échantillon s :

$$\begin{aligned}
\overline{x} &= \frac{120 + 135 + 110 + 145 + 125 + 130 + 140}{7} = \frac{905}{7} \approx 129.29\,\text{€} \\
s^2 &= \frac{(120 - 129.29)^2 + (135 - 129.29)^2 + (110 - 129.29)^2 + \cdots + (140 - 129.29)^2}{7 - 1} \\
&= \frac{86.3041 + 32.6041 + 372.1041 + 246.8041 + 18.4041 + 0.5041 + 117.7041}{6} \\
&= \frac{871.4287}{6} \approx 145.24 \\
s &= \sqrt{s^2} = \sqrt{145.24} \approx 12.05\,\text{€}
\end{aligned}$$

2) Somme des écarts à la moyenne $\overline{x}$:

$$\begin{aligned}
\sum(x_i - \overline{x}) &= (120 - 129.29) + (135 - 129.29) + (110 - 129.29) + \cdots + (140 - 129.29) \\
&= -9.29 + 5.71 - 19.29 + 15.71 - 4.29 + 0.71 + 10.71 = -0.03 \approx 0
\end{aligned}$$

La somme des écarts à la moyenne arithmétique est toujours (aux erreurs d'arrondis près) nulle. C'est la raison pour laquelle on utilise les carrés des écarts à la moyenne pour mesurer la dispersion.

Exercice 10

Une entreprise souhaite estimer le temps moyen de traitement d'une commande. Elle prend un échantillon de 10 commandes et obtient $\overline{x} = 4.2$ jours et $s = 0.8$ jour.

1) Que signifie $s = 0.8$ jour ?

2) Si l'entreprise double la taille de l'échantillon à 20 (toujours aléatoire), l'estimation ponctuelle $\overline{x}$ sera-t-elle nécessairement plus précise ? Et l'erreur type ?

Solution 10

1) L'écart-type s estime la variabilité des délais entre commandes dans la population. Environ 68% des commandes (si distribution normale) ont un délai dans $[3{,}4; 5{,}0]$ jours.

2) Non, $\overline{x}$ n'est pas *nécessairement* plus proche de μ sur une réalisation unique. Cependant, *l'erreur type* $\sigma_{\overline{x}} = \sigma/\sqrt{n}$ diminue avec n. La distribution d'échantillonnage se resserre, augmentant la *probabilité* que $\overline{x}$ soit proche de μ. La précision statistique augmente, mais pas la garantie sur un seul échantillon.

Exercice 11

On étudie les salaires annuels (k€) de 5 cadres : 52, 58, 49, 61 et 55.

1) Calculez le salaire moyen $\overline{x}$ et l'écart-type s.

2) Si tous les salaires augmentent de 5k€ (négociation collective), que deviennent le salaire moyen $\overline{x}$ et l'écart-type s ?

3) Si tous les salaires sont multipliés par 1.02 (indexation 2%), que deviennent le salaire moyen $\overline{x}$ et l'écart-type s ?

Solution 11

1) Calcul du salaire moyen :

$$\overline{x} = \frac{52 + 58 + 49 + 61 + 55}{5} = \frac{275}{5} = 55\,\text{k€}$$

Calcul de l'écart-type :

$$\begin{aligned} s^2 &= \frac{(52-55)^2 + (58-55)^2 + (49-55)^2 + (61-22)^2 + (55-55)^2}{5-1} \\ &= \frac{9 + 9 + 36 + 36 + 0}{4} = \frac{90}{4} = 22.5 \\ s &= \sqrt{s^2} = \sqrt{22.5} \approx 4.74\,\text{k€} \end{aligned}$$

2) Si tous les salaires augmentent de 5k€, le salaire moyen $\overline{x}'$ devient :

$$\overline{x}' = \frac{57 + 63 + 54 + 66 + 60}{5} = \frac{300}{5} = 60\,\text{k€}$$

L'écart-type s' devient :

$$\begin{aligned} s'^2 &= \frac{(57-60)^2 + (63-60)^2 + (54-60)^2 + (66-60)^2 + (60-60)^2}{5-1} \\ &= \frac{9 + 9 + 36 + 36 + 0}{4} = \frac{90}{4} = 22.5 \\ s' &= \sqrt{s'^2} = \sqrt{22.5} \approx 4.74\,\text{k€} \end{aligned}$$

La dispersion des salaires ne change pas lorsqu'ils sont augmentés du même montant.

3) Si tous les salaires sont multipliés par 1.02, le salaire moyen $\overline{x}$" devient :

$$\overline{x}" = \frac{62.4 + 69.6 + 58.8 + 73.2 + 66}{5} = \frac{330}{5} = 66\,\text{k€}$$

L'écart-type s" devient :

$$\begin{aligned}
s"^2 &= \frac{(62.4-66)^2 + (69.6-66)^2 + (58.8-66)^2 + (73.2-66)^2 + (66-66)^2}{5-1} \\
&= \frac{12.96 + 12.96 + 51.84 + 51.84}{4} = \frac{129.6}{4} = 32.4 \\
s" &= \sqrt{s"^2} = \sqrt{32.4} \approx 5.69\,\text{k€}
\end{aligned}$$

La dispersion des salaires augmente lorsqu'ils sont multipliés par un nombre supérieur à 1.

Exercice 12

Un économiste collecte des données sur le ratio (en %) dette/PIB de 8 pays émergents : 45, 52, 48, 55, 50, 47, 53 et 49.

1) Calculez l'estimation ponctuelle de la moyenne et de la variance.

2) Un journal titre : « La dette moyenne dépasse 50 % du PIB. » Votre calcul contredit-il ce titre ?

Solution 12

1) Estimation ponctuelle de la moyenne :

$$\overline{x} = \frac{45 + 52 + 48 + 55 + 50 + 47 + 53 + 49}{8} = \frac{399}{8} = 49.88\,\%$$

Estimation ponctuelle de la variance :

$$\begin{aligned}
s^2 &= \frac{(45-49.88)^2 + (52-49.88)^2 + (48-49.88)^2 + (55-49.88)^2 + \cdots + (49-49.88)^2}{8-1} \\
&= \frac{23.8144 + 4.4944 + 3.5344 + 26.2144 + 0.0144 + 8.2944 + 9.7344 + 0.7744}{7} \\
&= \frac{76.8752}{7} \approx 10.98
\end{aligned}$$

2) Non, 49.88 % est très proche de 50 %. La différence est de 0.12 point, probablement due au hasard d'échantillonnage. Le titre est peut-être excessif ou arrondi, mais les données ne le contredisent pas formellement sans test statistique.

Exercice 13

Une étude porte sur le temps de livraison (jours) de 12 commandes e-commerce : 3, 4, 2, 5, 4, 3, 6, 4, 3, 5, 4 et 3.

1) Calculez $\overline{x}$, s et l'estimation ponctuelle de la proportion de livraisons ≤ 3 jours.

2) Rédigez un paragraphe d'interprétation économique adressé au directeur logistique.

3) Citez deux limites inhérentes à l'utilisation d'estimations ponctuelles seules.

Solution 13

1) Calcul de $\overline{x}$:

$$\overline{x} = \frac{3+4+2+5+4+3+6+4+3+5+4+3}{12} = \frac{46}{12} \approx 3.83\,\text{jours}$$

Calcul de s :

$$\begin{aligned}
s^2 &= \frac{(3-3.83)^2+(4-3.83)^2+(2-3.83)^2+(5-3.83)^2+(4-3.83)^2+\cdots+(3-3.83)^2}{12-1} \\
&= \frac{0.6889+0.0289+3.3489+1.3689+0.0289+0.6889+4.7089+0.0289+\cdots+06889}{11} \\
&= \frac{13.67}{11} \approx 1.24 \\
s &= \sqrt{s^2} = \sqrt{1.24} = 1.11 \text{ jour}
\end{aligned}$$

Estimation ponctuelle de la proportion de livraisons ≤ 3 jours $\{3, 2, 3, 3, 3\}$:

$$\overline{p} = \frac{5}{12} \approx 0.4167 = 41.67\%$$

2) Interprétation économique adressée au directeur logistique :
Sur la base de cet échantillon, le délai moyen de livraison est estimé à 3.8 jours avec une variabilité d'environ 1.11 jour. Près de 42 % des commandes sont livrées en 3 jours ou moins. Ces indicateurs suggèrent une relative efficacité, mais la dispersion indique que certains clients attendent significativement plus longtemps (jusqu'à 6 jours observés). Une analyse complémentaire des causes de retard est recommandée.

3) Limites inhérentes à l'utilisation d'estimations ponctuelles seules :

— Erreur d'échantillonnage inconnue : On ne connaît pas la distance exacte entre l'estimation et le paramètre réel.

— Sensibilité au plan d'échantillonnage : Si l'échantillon n'est pas aléatoire ou si la population cible/échantillonnée diverge, l'estimation peut être systématiquement biaisée.

— Absence de niveau de confiance : Contrairement à un intervalle, une estimation ponctuelle ne quantifie pas le risque de se tromper.