

Plaquette 5 - Fluctuation d'échantillonnage et simulation

Variables aléatoires — Première STI2D

Corrections détaillées

Boîte à outils

Fréquence observée. Sur un échantillon de taille n contenant k succès :

$$f = \frac{k}{n}$$

Nombre moyen de succès attendu. Si chaque pièce est un succès avec la probabilité p :

$$E(X) = n \times p$$

Fluctuation d'échantillonnage. Deux échantillons de même taille ne donnent pas la même fréquence : f **fluctue** autour de p . C'est normal, ce n'est pas une erreur de mesure.

Effet de la taille.

Taille de l'échantillon	Fréquence observée
petite ($n = 20$)	très variable, écart à p souvent grand
grande ($n = 2\,000$)	stable, écart à p petit

Plus n est grand, plus f se rapproche de p : c'est pourquoi on contrôle sur des lots importants.

Moyenne de plusieurs fréquences. La moyenne des fréquences de plusieurs échantillons est proche de p , même quand chacune s'en écarte.

Étendue. Pour mesurer la dispersion des simulations :
étendue = plus grande valeur – plus petite valeur.

Simuler. En Python, `random()` tire un nombre au hasard entre 0 et 1 : la condition `random() < p` est vraie avec la probabilité p , elle simule donc un succès.

⚠ Un écart entre f et p ne prouve rien à lui seul : on le compare à l'**intervalle des valeurs obtenues en simulation** avant de conclure.

Correction de l'exercice 1 - Une première fréquence

La fréquence de défauts est le nombre de pièces défectueuses divisé par l'effectif du lot :

$$f = \frac{k}{n} = \frac{15}{250} = 0,06$$

soit 6 % de pièces défectueuses dans ce lot.

Remarque. C'est une fréquence **observée** sur un lot : un autre lot de 250 pièces de la même production donnerait une valeur un peu différente.

Réponse. $f = 0,06$, soit 6 %.

Correction de l'exercice 2 - Quatre échantillons de même taille

a) On divise chaque effectif par 200 :

Échantillon	1	2	3	4
Fréquence	0,44	0,48	0,52	0,56

b) Étendue : plus grande moins plus petite fréquence,

$$0,56 - 0,44 = 0,12$$

c) En réunissant les quatre échantillons : $88 + 96 + 104 + 112 = 400$ pièces sur 800, d'où

$$f = \frac{400}{800} = 0,5$$

Les quatre fréquences fluctuent autour de cette valeur globale.

Réponse. a) 0,44 ; 0,48 ; 0,52 ; 0,56 — b) 0,12 — c) 0,5.

Correction de l'exercice 3 - L'effet de la taille

On calcule chaque fréquence, puis son écart à $p = 0,3$.

Taille n	20	200	2 000
Fréquence	$\frac{8}{20} = 0,4$	$\frac{66}{200} = 0,33$	$\frac{617}{2000} = 0,3085$
Écart à 0,3	0,1	0,03	0,0085

Constat. Plus l'échantillon est grand, plus la fréquence observée est proche de la proportion réelle 0,3 : la fluctuation d'échantillonnage diminue quand n augmente.

Réponse. 0,4 ; 0,33 ; 0,3085 — l'écart diminue quand n augmente.

Correction de l'exercice 4 - Comparer à la valeur attendue

a) Le nombre moyen attendu est l'effectif attendu :

$$n \times p = 400 \times 0,025 = 10$$

b) Fréquence observée :

$$f = \frac{14}{400} = 0,035$$

soit 3,5 % au lieu de 2,5 % attendus. Pour savoir si cet écart est inquiétant, il faudrait le comparer aux valeurs obtenues en simulant de nombreux échantillons de 400 pièces.

Réponse. a) 10 pièces — b) $f = 0,035$.

Correction de l'exercice 5 - Lire un tableau de simulation

a) Le plus grand nombre de succès est 29, le plus petit 21, sur 50 pièces :

$$f_{\max} = \frac{29}{50} = 0,58 \quad f_{\min} = \frac{21}{50} = 0,42$$

Étendue : $0,58 - 0,42 = 0,16$.

b) Total des succès : $22 + 26 + 24 + 29 + 21 + 27 + 25 + 23 = 197$. Nombre moyen par simulation :

$$\frac{197}{8} = 24,625$$

Fréquence correspondante : $\frac{24,625}{50} = 0,4925$, soit $\frac{197}{400}$ sur l'ensemble des 400 pièces simulées.

Réponse. a) 0,58 ; 0,42 ; étendue 0,16 — b) 24,625 succès, fréquence 0,4925.

Correction de l'exercice 6 – Remonter à l'effectif

L'effectif est la fréquence multipliée par la taille du lot : $k = f \times n$.

a) Calcul :

$$800 \times 0,035 = 28 \text{ pièces défectueuses}$$

b) Calcul :

$$1500 \times 0,012 = 18 \text{ pièces défectueuses}$$

Contrôle. $\frac{28}{800} = 0,035$ et $\frac{18}{1500} = 0,012$.

Réponse. a) 28 — b) 18.

Correction de l'exercice 7 – Un écart est-il surprenant ?

Au taux annoncé de 2 %, on attend en moyenne $1000 \times 0,02 = 20$ défauts.

Les 50 lots simulés à ce taux ont donné entre 9 et 32 défauts. Le lot reçu en contient 34 : c'est **plus que tous les lots simulés**.

Conclusion. Un tel résultat serait rare si le taux était vraiment de 2 % : l'annonce du fournisseur paraît **peu crédible**, et l'on peut demander des explications ou contrôler d'autres lots. Ce n'est pas une preuve : un lot exceptionnel reste possible.

Réponse. Peu crédible : 34 dépasse toutes les valeurs simulées (9 à 32).

Correction de l'exercice 8 – Simuler un échantillon en Python

a) La boucle répète 500 fois l'épreuve : `random() < 0.04` est vraie avec la probabilité 0,04 et ajoute alors 1 au compteur `s`. La fonction renvoie `s / 500`, c'est-à-dire la **fréquence de succès** dans un échantillon simulé de 500 pièces à 4 % de défauts. On obtient une valeur proche de 0,04, par exemple 0,036 ou 0,046.

b) Chaque appel tire de **nouveaux** nombres au hasard : on simule un autre échantillon, dont la fréquence est différente. C'est la fluctuation d'échantillonnage.

Réponse. a) Une fréquence simulée proche de 0,04 — b) chaque appel simule un nouvel échantillon.

Correction de l'exercice 9 – Compter les échantillons au-dessus d'un seuil

Les trois lignes ont chacune un rôle précis.

— **Ligne 1** : un succès augmente le compteur de l'échantillon, $s = s + 1$.

— **Ligne 2** : la fréquence de l'échantillon est le nombre de succès divisé par la taille, $f = s / n$:

$$f = \frac{s}{n}$$

— **Ligne 3** : si cette fréquence dépasse le seuil, on compte l'échantillon, $c = c + 1$.

La fonction renvoie `c`, le nombre d'échantillons (sur `nb`) dont la fréquence dépasse le seuil ; c / nb estime la probabilité de dépasser ce seuil.

Piège. Placer la ligne 2 à l'intérieur de la boucle `for i` : on calculerait la fréquence avant d'avoir fini l'échantillon.

Réponse. $s = s + 1$; $f = s / n$; $c = c + 1$.

Correction de l'exercice 10 – Une règle d'alerte en atelier

a) Nombre moyen de rebuts par heure :

$$500 \times 0,03 = 15$$

b) **Première heure** : 21 rebuts, ce qui ne dépasse pas 25 : pas d'alerte, malgré un écart à la moyenne.

Deuxième heure : 27 rebuts, soit plus de 25 : l'opérateur **déclenche l'alerte** et fait vérifier le réglage.

Nuance. Une alerte peut être due à la seule fluctuation ; mais deux heures consécutives au-dessus de 15 rendent un dérèglement plausible.

Réponse. a) 15 — b) pas d'alerte à 21 ; alerte à 27.

Correction de l'exercice 11 – Deux enquêtes de tailles différentes

a) Fréquences :

$$f_A = \frac{31}{50} = 0,62 \quad f_B = \frac{436}{800} = 0,545$$

Enquête	Taille	Réponses favorables	Fréquence
A	50	31	0,62
B	800	436	0,545

b) L'enquête **B** porte sur un échantillon bien plus grand : sa fréquence fluctue beaucoup moins d'un échantillon à l'autre. C'est donc elle qui estime le mieux l'opinion du personnel, environ 54,5 % d'avis favorables.

Réponse. a) 0,62 et 0,545 — b) l'enquête B, plus grande.

Correction de l'exercice 12 – La moyenne des fréquences

a) On divise chaque effectif par 400 :

Lot	1	2	3	4	5
Fréquence	0,1175	0,13	0,1525	0,1375	0,1125

Étendue : $0,1525 - 0,1125 = 0,04$.

b) Total des défauts : $47 + 52 + 61 + 55 + 45 = 260$ sur 2000 pièces :

$$f = \frac{260}{2000} = 0,13$$

Les lots ayant la même taille, c'est aussi la moyenne des cinq fréquences.

Réponse. a) Étendue 0,04 — b) 0,13.

Correction de l'exercice 13 – Espérance et moyenne observée

a) Espérance du gain :

$$E(G) = 0 \times 0,8 + 5 \times 0,15 + 20 \times 0,05 = 0,75 + 1 = 1,75$$

b) Sur 200 parties, on attend en moyenne

$$200 \times 1,75 = 350 \text{ €}$$

L'organisateur a versé 380 €, soit 30 € de plus, ou 1,90 € par partie au lieu de 1,75 €. Cet écart s'explique par la fluctuation : quelques gains de 20 € en plus suffisent.

Réponse. a) 1,75 € — b) 350 € attendus, 380 € versés : écart de fluctuation.

Correction de l'exercice 14 - Estimer une probabilité inconnue

a) Sur un très grand nombre d'essais, la fréquence observée est une bonne estimation de la probabilité :

$$p \approx \frac{2340}{10000} = 0,234$$

b) Sur 450 démarrages, on attend en moyenne

$$450 \times 0,234 = 105,3$$

soit environ 105 blocages. Le nombre réel fluctuera autour de cette valeur d'un jour à l'autre.

Réponse. a) $p \approx 0,234$ — b) environ 105 blocages.

Correction de l'exercice 15 - Faut-il régler la machine ?

Machine réglée : on attend en moyenne $600 \times 0,04 = 24$ défauts, et les simulations donnent entre 12 et 37 défauts.

a) 36 défauts est élevé, mais **dans l'intervalle** [12 ; 37] des simulations : cet écart peut s'expliquer par la seule fluctuation. On ne règle pas encore, mais on surveille.

b) 41 défauts est **en dehors** des valeurs simulées : c'est rare avec une machine réglée. On fait **vérifier le réglage**.

Remarque. Deux contrôles successifs élevés renforcent le soupçon de dérèglement.

Réponse. a) Dans l'intervalle simulé : pas de conclusion — b) hors intervalle : régler la machine.

Correction de l'exercice 16 - Synthèse : une campagne de contrôle

a) Nombre moyen de pièces défectueuses par lot :

$$1200 \times 0,05 = 60$$

Coût moyen de retouche : $60 \times 4,5 = 270$ €.

b) Fréquence de mars : $\frac{78}{1200} = 0,065$, soit 6,5 %. Coût réel : $78 \times 4,5 = 351$ €.

c) Sur douze lots : $12 \times 270 = 3240$ € en moyenne.

d) Au taux de 3,5 % : $1200 \times 0,035 = 42$ défauts, soit $42 \times 4,5 = 189$ € par mois. Économie moyenne : $270 - 189 = 81$ € par mois.

$$\frac{2500}{81} \approx 30,9$$

L'investissement est amorti en environ 31 mois, un peu plus de deux ans et demi.

Réponse. a) 60 pièces, 270 € — b) 0,065 ; 351 € — c) 3240 € — d) environ 31 mois.