

Corrigé et barème – Statistiques, probabilités et échantillonnage — niveau ★★★

Durée : 120 min – Calculatrice interdite – Sur 20 points · Mathématiques 2^{nde}

Exercice 1 – Calculs rapides

(3 pts)

- a) $P(A \cap B) = P(A) + P(B) - P(A \cup B) = 0,7 + 0,5 - 0,85 = 0,35$. A est la réunion disjointe de $A \cap B$ et $A \cap \overline{B}$, donc $P(A \cap \overline{B}) = 0,7 - 0,35 = 0,35$. (1 pt) (1)
- b) L'écart type devient $a \times 3 = 12$, donc $a = 4$; la moyenne devient $4 \times 8 + b = 50$, donc $b = 18$. (1 pt) (1)
- c) L'amplitude vaut $\frac{2}{\sqrt{n}} : \frac{2}{\sqrt{n}} \leq 0,05 \iff \sqrt{n} \geq 40 \iff n \geq 1\,600$. Le plus petit entier est $n = 1\,600$. (1 pt) (1)

Méthode — Pour imposer une moyenne et un écart type par une transformation affine, on commence par l'écart type, qui ne dépend que de a , puis on déduit b de la moyenne.

Exercice 2 – Démontrer

(4 pts)

- a) $\bar{y} = \frac{(ax_1 + b) + \dots + (ax_N + b)}{N} = \frac{a(x_1 + \dots + x_N) + Nb}{N} = a \times \frac{x_1 + \dots + x_N}{N} + b = a\bar{x} + b$. (0,5 pt) (0,8)
- b) Moyenne : $\frac{k \times 1 + (N - k) \times 0}{N} = \frac{k}{N} = f$. Variance (moyenne des carrés des écarts à la moyenne) : $V = \frac{k(1 - f)^2 + (N - k)(0 - f)^2}{N} = f(1 - f)^2 + (1 - f)f^2 = f(1 - f)[(1 - f) + f] = f(1 - f)$. (1 pt) (0,8)
- c) $\frac{1}{4} - \left(f - \frac{1}{2}\right)^2 = \frac{1}{4} - \left(f^2 - f + \frac{1}{4}\right) = f - f^2 = f(1 - f)$. Un carré étant positif, $V = f(1 - f) \leq \frac{1}{4}$, donc $\sigma = \sqrt{V} \leq \frac{1}{2}$, avec égalité si et seulement si $f = \frac{1}{2}$, c'est-à-dire autant de succès que d'échecs. (1 pt) (0,8)
- d) $g_i = 5x_i - 2(1 - x_i) = 7x_i - 2$ (vérification : $g_i = 5$ si $x_i = 1$ et $g_i = -2$ si $x_i = 0$). Ici $f = \frac{120}{400} = 0,3$. D'après a) et b), $\bar{g} = 7 \times 0,3 - 2 = 0,1$ euro par partie. D'après b) et la propriété admise, $\sigma_g = 7\sqrt{0,3 \times 0,7} = 7\sqrt{0,21} \approx 3,21$ euros. (1 pt) (0,8)
- e) D'après la loi des grands nombres, pour un très grand nombre de parties, la fréquence f des parties gagnées est proche de $p = \frac{2}{7}$; donc $\bar{g} = 7f - 2$ est proche de $7 \times \frac{2}{7} - 2 = 0$: le jeu est équitable à long terme, et le gain moyen de 0,1 euro observé en d) relève de la fluctuation. (0,5 pt) (0,8)

Astuce — Coder un succès par 1 et un échec par 0 transforme une fréquence en moyenne : tous les outils des statistiques descriptives s'appliquent alors à l'échantillonnage.

Exercice 3 – Filles, garçons et association sportive

(4 pts)

- a) Arbre : F (0,55) puis A (0,3) et $\bar{A}$ (0,7) ; G (0,45) puis A et $\bar{A}$ (probabilités à déterminer). $P(F \cap A) = P(F) \times P_F(A) = 0,55 \times 0,3 = 0,165$. (0,5 pt) (0,8)
- b) Probabilités totales : $P(A) = P(F \cap A) + P(G \cap A)$, donc $P(G \cap A) = 0,4 - 0,165 = 0,235$. Ainsi $P_G(A) = \frac{P(G \cap A)}{P(G)} = \frac{0,235}{0,45} = \frac{47}{90} \approx 0,522$. (1 pt) (0,8)
- c) $P_A(F) = \frac{P(F \cap A)}{P(A)} = \frac{0,165}{0,4} = 0,4125$. (0,5 pt) (0,8)
- d) Probabilités totales : $P(A) = P(F)P_F(A) + P(G)P_G(A)$ avec $P(F) + P(G) = 1$. On en déduit $P(A) - P_F(A) = P(G)(P_G(A) - P_F(A))$ et $P(A) - P_G(A) = P(F)(P_F(A) - P_G(A))$. Comme $P(F) > 0$ et $P(G) > 0$, ces deux différences sont de signes contraires (ou nulles) : $P(A)$ est compris entre $P_F(A)$ et $P_G(A)$. Autrement dit, $P(A)$ est une moyenne pondérée des deux probabilités conditionnelles. (1 pt) (0,8)
- e) $f = 0,4$; $\frac{1}{\sqrt{1600}} = \frac{1}{40} = 0,025$; intervalle $[0,36 - 0,025 ; 0,36 + 0,025] = [0,335 ; 0,385]$. Comme $0,4 > 0,385$, la fréquence observée est hors de l'intervalle : on peut considérer que ce lycée se distingue de la situation nationale, avec un risque d'erreur d'environ 5 %. (1 pt) (0,8)

Piège — $P_F(A)$ et $P_A(F)$ ne répondent pas à la même question : 0,3 est la part d'inscrites parmi les filles, 0,4125 la part de filles parmi les inscrits.

Exercice 4 – Problème — Contrôle de pots de yaourt

(4 pts)

- a) Centres : 120 ; 123 ; 125 ; 127 ; 130. $\bar{x} \approx \frac{16 \times 120 + 44 \times 123 + 70 \times 125 + 50 \times 127 + 20 \times 130}{1920 + 5412 + 8750 + 6350 + 2600} = \frac{25032}{200} = 125,16$ g. (1 pt) (0,8)
- b) Valeurs de d : -5 ; -2 ; 0 ; 2 ; 5. $\bar{d} = \frac{16 \times (-5) + 44 \times (-2) + 0 + 50 \times 2 + 20 \times 5}{200} = \frac{-80 - 88 + 100 + 100}{200} = \frac{32}{200} = 0,16$ (on retrouve $\bar{x} = 125 + 0,16$). Moyenne des carrés : $\frac{16 \times 25 + 44 \times 4 + 50 \times 4 + 20 \times 25}{200} = \frac{1276}{200} = 6,38$, donc $V_d = 6,38 - 0,16^2 = 6,3544$. Les d se déduisent des centres par une translation, qui ne modifie pas l'écart type : $\sigma_x = \sigma_d = \sqrt{6,3544} \approx 2,52$ g. L'intérêt : des nombres petits, donc des calculs plus simples et plus sûrs. (1 pt) (0,8)
- c) Fréquences cumulées croissantes : 0,08 ; 0,30 ; 0,65 ; 0,90 ; 1. La fréquence 0,5 est atteinte dans $[124 ; 126[$. Interpolation : $Me \approx 124 + 2 \times \frac{0,5 - 0,30}{0,65 - 0,30} = 124 + \frac{0,4}{0,35} \approx 125,14$ g. (0,5 pt) (0,8)
- d) Pots conformes : $44 + 70 + 50 = 164$, donc $P(C) = \frac{164}{200} = 0,82$. Parmi eux, $70 + 50 = 120$ pèsent au moins 124 g : probabilité $\frac{120}{164} = \frac{30}{41} \approx 0,732$. (1 pt) (0,8)
- e) $y = \frac{x - 5}{1000} = 0,001x - 0,005$. Moyenne : $\bar{y} = 0,001 \times 125,16 - 0,005 = 0,12016$ kg ; écart type : $\sigma_y = 0,001 \times \sigma_x \approx 0,00252$ kg. (0,5 pt) (0,8)

Méthode — Centrer les valeurs autour d'un nombre rond (ici 125) allège le calcul de la variance sans changer l'écart type ; il suffit ensuite d'ajouter 125 à la moyenne obtenue.

Exercice 5 – Pour aller plus loin — Au moins un six*(5 pts)*

- a) Dans le tableau à double entrée, les 36 couples sont équiprobables. L'événement contraire « aucun 6 » compte $5 \times 5 = 25$ couples, donc $p = 1 - \frac{25}{36} = \frac{11}{36} \approx 0,3056$. (0,5 pt) (0,833333)

b)

```
from random import randint

def essai():
    d1 = randint(1, 6)
    d2 = randint(1, 6)
    return d1 == 6 or d2 == 6
```

```
def frequence(n):
    s = 0
    for i in range(n):
        if essai():
            s = s + 1
    return s / n
```

(0,5 pt par fonction, soit 1 pt) (0,833333)

c)

```
from math import sqrt

def proportion(N, n):
    p = 11 / 36
    c = 0
    for j in range(N):
        f = frequence(n)
        if abs(f - p) <= 1 / sqrt(n):
            c = c + 1
    return c / N
```

(1 pt) (0,833333)

- d) Sur 1 000 échantillons simulés de 400 lancers, 95,8 % ont une fréquence à moins de $\frac{1}{\sqrt{400}} = 0,05$ de p , c'est-à-dire dans $[0,2556 ; 0,3556]$ environ. C'est conforme au résultat du cours : pour n assez grand, environ 95 % au moins des échantillons vérifient $|f - p| \leq \frac{1}{\sqrt{n}}$. (0,5 pt) (0,833333)
- e) $\frac{1}{3} = \frac{12}{36} > p$: il faut $\frac{1}{3} > p + \frac{1}{\sqrt{n}}$, soit $\frac{1}{\sqrt{n}} < \frac{12}{36} - \frac{11}{36} = \frac{1}{36}$, c'est-à-dire $\sqrt{n} > 36$, ou encore $n > 1\,296$. Pour $n = 1\,296$, $\frac{1}{3}$ est exactement la borne supérieure, donc encore dans l'intervalle fermé : la plus petite taille est $n = 1\,297$. (1 pt) (0,833333)
- f) q est hors de l'intervalle $\iff |q - p| > \frac{1}{\sqrt{n}} \iff \sqrt{n} > \frac{1}{|q - p|} \iff n > \frac{1}{(q - p)^2}$ (tous les nombres sont positifs et la fonction carré est croissante sur $[0 ; +\infty[$). Contrôle : avec $|q - p| = \frac{1}{36}$, on retrouve $n > 1\,296$. Si l'écart est divisé par 2, $\frac{1}{(q - p)^2}$ est multiplié par 4 : il faut environ quatre fois plus d'observations. (1 pt) (0,833333)

Erreur fréquente — Détecter un petit écart coûte cher : la taille d'échantillon nécessaire varie comme l'inverse du carré de l'écart, et non comme l'inverse de l'écart.

Ce document est la propriété intellectuelle de Ben, professeur particulier (benprofparticulier.fr), protégée par le Code de la propriété intellectuelle (art. L. 111-1 et L. 122-4). Il est mis à disposition pour un usage personnel de révision. Toute reproduction, diffusion, mise en ligne ou revente, totale ou partielle, sans autorisation écrite est interdite et constitue une contrefaçon (art. L. 335-2).