CORRIGE du DEVOIR

Exercice 1

$\mathcal{P}$ ={français âgés de 15 à 24 ans}
X = fumer, variable qualitative dichotomique : oui, non
p = proportion de fumeurs parmi les français de 15 à 24 ans (dans $\mathcal{P}$), p connue dans $\mathcal{P}$: p = 0,53 = 53%
Echantillon de taille n=40 de X issu de $\mathcal{P}$.

1) Puisque n =40 ≥ 30, n(1–p) = 40×0,47 =18,8 ≥ 5, d'où np = 40×0,53 ≥ 5, la fréquence empirique F_{40} a une distribution approximativement normale de moyenne p = 0,53 et d'écart-type
$\sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{0,53\times 0,47}{40}} = \sqrt{0,006228} \approx 0,0789 \approx 0,08$ car la proportion de fumeurs dans $\mathcal{P}$ est p = 0,53.

2) Observer 24 fumeurs ou plus sur un échantillon de taille n=40 correspond à observer une fréquence d'au moins f = 24/40 = 0,6. La probabilité correspondante s'écrit :
$$P(F_{40} > f) = P(F_{40} > 0,6) = 1 - P(F_{40} \le 0,6) \approx 1 - P\left(Z \le \frac{0,6-0,53}{0,08}\right) = 1 - F\left(\frac{0,07}{0,08}\right) = 1 - F(0,875)$$
d'où $P(F_{40} > f) \approx 1 - F(0,88) = 1 - 0,8106 = 0,1894$ où F est la fonction de répartition de la loi $\mathcal{N}(0,1)$
(calcul exact : 1–F(0,887) ≈ 1–F(0,89) = 1–0,8133 = 0,1867; écart-type arrondi à 0,079 : 1–F(0,886) ≈ 1–F(0,89)).
➔ environ 19% des échantillons de taille 40 de X issus de $\mathcal{P}$ ont une fréquence de fumeurs supérieure à 60%, soit plus de 26 fumeurs sur 40 français de 15 à 24 ans.

3) Observer 16 fumeurs ou moins sur un échantillon de taille n=40 correspond à observer une fréquence d'au plus f = 16/40 = 0,4. La probabilité correspondante s'écrit :
$$P(F_{40} \le f) = P(F_{40} \le 0,4) \approx P\left(Z \le \frac{0,4-0,53}{0,08}\right) \approx F\left(\frac{-0,13}{0,08}\right) = F(-1,625) \approx 1 - F(1,63) = 1 - 0,9484 = 0,0516$$
où F est la fonction de répartition de la loi $\mathcal{N}(0,1)$ (calcul exact : 1–F(1,647) ≈ 1–F(1,65) = 1–0,9505 = 0,0495; écart-type arrondi à 0,079 : 1–F(1,646) ≈ 1–F(1,65)).
➔ environ 5% des échantillons de taille 40 de X issus de $\mathcal{P}$ ont une fréquence de fumeurs inférieure à 40%, soit moins de 16 fumeurs sur 40 français de 15 à 24 ans.

4) Intervalle de variation au niveau 90% (au risque α=10%) de la fréquence empirique sur les échantillons de taille 40 de X issus de $\mathcal{P}$:
$$I_{90\%}(F_{40}) = [Q_{0,05}\,;Q_{0,95}] \approx \left[0,53 \pm z_{0,95} \times \sqrt{\frac{0,53\times0,47}{40}}\right] = [0,53 \pm 1,645 \times 0,08] \approx [0,53 \pm 0,132] = [0,398\,;0,662]$$
$I_{90\%}(F_{40})$ ≈ [53% ± 13,2%] ≈ [39,8% ; 66,2%] car $z_{1-(\alpha/2)}=z_{0,95}=1,645$ est le quantile d'ordre 0,95 de la loi $\mathcal{N}(0,1)$
(calcul exact : $I_{90\%}(F_{40})$ = [0,53 ± 0,13] =[0,40 ; 0,66] ; même résultat avec écart-type arrondi à 0,079).
➔ 90% des échantillons de taille 40 de X issus de $\mathcal{P}$ ont une fréquence observée de fumeurs comprise entre 39,8% et 66,2% environ (soit entre 16 et 26 fumeurs environ sur 40 français de 15 à 24 ans).

La précision (marge d'erreur) à 90% dans l'estimation de la proportion de fumeurs parmi les français de 15 à 24 ans est égale à la demi-longueur de l'intervalle de variation au niveau 90% (au risque α=10%) $I_{90\%}(F_{40})$ c'est à dire 0,132 ou 13,2%
➔ la précision (marge d'erreur) à 90% dans l'estimation de la proportion de fumeurs parmi les français de 15 à 24 ans sur les échantillons de taille 40 de X issus de $\mathcal{P}$ est d'environ 13,2%.

5) Observer 26 fumeurs sur 40 français de 15 à 24 ans c'est observer une fréquence de f=26/40=0,65=65%, plus élevée que la proportion p=53% mais qui appartient à l'intervalle de variation à 90% trouvé précédemment, c'est à dire que la probabilité d'observer cette fréquence de 65% ou plus est supérieure à 5%.
(Elle est égale à : $P(F_{40} > f) = 1 - P(F_{40} \le 0,65) \approx 1 - P\left(Z \le \frac{0,65-0,53}{0,08}\right) = 1 - F(1,5) = 1 - 0,9332 = 0,0668$ donc environ 7% des échantillons de taille 40 de X issus de $\mathcal{P}$ ont une fréquence de fumeurs supérieure à 65%)
➔ plus de 5% des échantillons de taille 40 de X issus de $\mathcal{P}$ ont une fréquence de fumeurs supérieure à 65%, soit plus de 26 fumeurs sur 40 français de 15 à 24 ans.

Bien que l'échantillon observé d'étudiants français de moins de 24 ans ne soit a priori pas représentatif de la population des jeunes français considérée (de 15 à 24 ans) car, d'une part les étudiants sont en moyenne plus âgés, et d'autre part ils ne représentent pas les autres catégories socio-professionnelles, la fréquence observée de 65% de fumeurs n'est pas incompatible avec la proportion de 53% de fumeurs dans la population étudiée.

Exercice 2

$\mathcal{P}$={français âgés de 18 ans et plus}
X= qualifier de "vulgaire" les émissions de "téléréalité", variable qualitative dichotomique : oui, non
p = proportion de personnes qualifiant de "vulgaire" les émissions de "téléréalité", dans $\mathcal{P}$, p inconnue dans $\mathcal{P}$
Echantillon de taille n=1 002 de X issu de $\mathcal{P}$

1) L'estimation ponctuelle de p est donnée par la fréquence (empirique) observée $f = \frac{565}{1\,002} = 0{,}564 = 56{,}4\%$

➔ la proportion de français de 18 ans et plus qui qualifient de "vulgaire" les émissions de "téléréalité est estimée à 56,4%.

2) L'estimation par intervalle de confiance au risque α=5% (au niveau 95%) de la proportion p dans $\mathcal{P}$ s'écrit :

$$IC_{95\%}(p) \approx \left[0{,}564 \pm z_{0,975}\sqrt{\frac{0{,}564\times(1-0{,}564)}{1\,002}}\right] = \left[0{,}564 \pm z_{0,975}\sqrt{\frac{0{,}564\times 0{,}436}{1\,002}}\right] \approx [0{,}564 \pm 1{,}96\times 0{,}01567]$$

$IC_{95\%}(p) \approx [0{,}564 \pm 0{,}031] = [0{,}533;\ 0{,}595] = [56{,}4\% \pm 3{,}1\%] = [53{,}3\%\ ;\ 59{,}5\%]$

où $z_{1-(\alpha/2)} = z_{0,975} = 1{,}96$ est le quantile d'ordre 0,975 de la loi $\mathcal{N}(0,1)$.

(calcul exact même résultat, ou arrondi à 2 décimales f = 0,56 : $IC_{95\%}(p) \approx [0{,}56 \pm 0{,}03] = [0{,}53;\ 0{,}59]$).

Cette approximation est justifiée puisque $n = 1\,002 \geq 30$ et $n\times(1-0{,}595) = 1\,002\times 0{,}405 = 406{,}2 \geq 5$ donc $n\times 0{,}533 \geq 5$, $n\times 0{,}595 \geq 5$ et $n\times(1-0{,}533) = n\times 0{,}467 \geq 5$, donc la fréquence empirique F_n suit approximativement une loi normale.

➔ l'estimation par intervalle de confiance au risque 5% (au niveau 95%) de la proportion de français de 18 ans ou plus qui qualifient de "vulgaire" les émissions de "téléréalité" est d'environ 53,3% à 59,5% ; la précision (marge d'erreur) de l'estimation à 95% est d'environ 3,1%.

3) La demi-longueur de l'intervalle précédent $IC_{95\%}(p)$ est d'environ 3,1% ; pour avoir une marge d'erreur (demi-longueur) de 2% donc plus faible, il faudrait plus de sujets.
Pour calculer la valeur minimum de n pour f=0,564 et α=5% connus, la demi-longueur de l'intervalle $IC_{95\%}(p)$ s'écrit : $z_{0,975}\sqrt{\frac{f(1-f)}{n}} = 1{,}96\sqrt{\frac{0{,}564(1-0{,}564)}{n}}$. D'où n est tel que : $1{,}96\sqrt{\frac{0{,}564\times 0{,}436}{n}} \leq 0{,}02$ c'est à dire

$$\frac{1{,}96}{0{,}02}\sqrt{0{,}564\times 0{,}436} \leq \sqrt{n} \text{ d'où } n \geq \left(\frac{1{,}96}{0{,}02}\right)^2 \times 0{,}564\times 0{,}436 = 2\,361{,}66$$

(calcul exact même résultat, ou arrondi à 2 décimales f = 0,56 : n ≥2 366,4).

➔ on choisira donc une taille d'échantillon au moins égale à 2 362 pour que la demi-longueur de l'intervalle de confiance à 95% de la proportion soit inférieure à 2%.

4) $\mathcal{P}_1$={français âgés de 34 ans ou moins} et $\mathcal{P}_2$={français âgés de 35 ans ou plus}
X= qualifier de "vulgaire" les émissions de "téléréalité", variable qualitative dichotomique : oui, non
p_1 = proportion de français de 34 ans ou moins qui qualifient de "vulgaire" les émissions de "téléréalité" dans $\mathcal{P}_1$, p_1 inconnue dans $\mathcal{P}_1$, échantillon de taille n_1 de X issu de $\mathcal{P}_1$
p_2 = proportion de français de 35 ans ou plus qui qualifient de "vulgaire" les émissions de "téléréalité" dans $\mathcal{P}_2$, p_2 inconnue dans $\mathcal{P}_2$, échantillon de taille n_2 = 778 de X issu de $\mathcal{P}_2$
d'où $n_1 = 1\,002 - 778 = 224$ et on observe sur cet échantillon $(565-421) = 144$ français de 34 ans ou moins qui qualifient de "vulgaire" les émissions de "téléréalité

L'estimation ponctuelle de p_1 est donnée par la fréquence (empirique) observée : $f_1 = \frac{144}{224} \approx 0{,}643 = 64{,}3\%$

➔ la proportion de français de 34 ans ou moins qui qualifient de "vulgaire" les émissions de "téléréalité est estimée à 64,3%.

Exercice 3

$\mathcal{P}$={candidats à l'admission dans une université américaine}

X = résultat aux tests, variable quantitative $X \sim \mathcal{N}(\mu = 500, \sigma = 50)$ dans $\mathcal{P}$. On note F la fonction de répartition de la loi $\mathcal{N}(0,1)$.

1) $P(X \leq 512) = P\left(Z \leq \frac{512-500}{50}\right) = P(Z \leq 0{,}24) = F(0{,}24) = 0{,}5948$

➔ environ 59,5% des candidats ont un résultat inférieur à 512.

2) $P(425 \leq X \leq 575) = P(X \leq 575) - P(X \leq 425) = 0{,}9332 - 0{,}0668 = 0{,}8664$ car

$$P(X \leq 575) = P\left(Z \leq \frac{575-500}{50}\right) = P(Z \leq 1{,}5) = F(1{,}5) = 0{,}9332$$

$$P(X \leq 425) = P\left(Z \leq \frac{425-500}{50}\right) = P(Z \leq -1{,}5) = 1 - F(1{,}5) = 1 - 0{,}9332 = 0{,}0668$$

➔ environ 86,6% des candidats ont un résultat compris entre 425 et 575.

3) $P(X \geq 400) = 1 - P(X \leq 400) = 1 - P\left(Z \leq \frac{400-500}{50}\right) = 1 - P(Z \leq -2) = 1 - F(-2) = 1 - (1 - F(2)) = F(2) = 0{,}9772$

➔ environ 97,7% des candidats ont un résultat supérieur à 400 et sont donc admis.

4) 5% des candidats ont un résultat supérieur au résultat minimum cherché, donc 95% des candidats ont un résultat inférieur au résultat minimum cherché c'est donc par définition le quantile d'ordre 0,95 de X :

$$x_{0,95} = 500 + (50 \times z_{0,95}) = 500 + (50 \times 1{,}645) = 500 + 82{,}25 = 582{,}25$$

car $z_{0,95}$=1,645 est le quantile d'ordre 0,95 de la loi $\mathcal{N}(0,1)$.

➔ 95% des candidats ont un résultat inférieur à 582,25 donc 5% des candidats ont un résultat supérieur à 582,25 : le candidat doit donc avoir un résultat supérieur à 582,25 pour être dans les 5% des meilleurs.

5) 80% des candidats ont un résultat supérieur au résultat minimum requis cherché, donc 20% des candidats ont un résultat inférieur au résultat minimum cherché , qui est donc par définition le quantile d'ordre 0,2 de X :

$$x_{0,2} = 500 + (50 \times z_{0,2}) = 500 - (50 \times z_{0,8}) = 500 - (50 \times 0{,}84) = 500 - 42 = 458$$

car $z_{0,8}$=0,84 est le quantile d'ordre 0,8 de la loi $\mathcal{N}(0,1)$.

➔ 20% des candidats ont un résultat inférieur à 458 donc 80% des candidats ont un résultat supérieur à 458, résultat minimum requis pour être admis.

6) L'intervalle cherché est l'intervalle de variation à 95% (au risque α=5%) du résultat X dans $\mathcal{P}$:

$$I_{95\%}(X) = [x_{0,025} ; x_{0,975}] = [500 \pm 50 \times z_{0,975}] = [500 \pm 50 \times 1{,}96] = [500 \pm 98] = [402 ; 598]$$

car $z_{1-(\alpha/2)} = z_{0,975}$=1,96 est le quantile d'ordre 0,975 de la loi $\mathcal{N}(0,1)$.

➔ 95% des candidats ont un résultat compris entre 402 et 598.

7) Chaque classe contient la même proportion de résultats donc 20%
- la première classe contient les 20% des résultats les plus faibles donc ceux des non admis, sa borne étant le quantile d'ordre 20% de X calculé à la question 5 : $x_{0,2} = 458$
- la seconde classe contient les 20% des résultats suivants donc sa borne est le quantile d'ordre 40% de X : $x_{0,4}$
- la troisième classe contient les 20% des résultats suivants donc sa borne est le quantile d'ordre 60% de X : $x_{0,6}$
- la quatrième classe contient les 20% des résultats suivants donc sa borne est le quantile d'ordre 80% de X : $x_{0,8}$
- et la cinquième (dernière) classe contiendra donc les 20% des résultats les plus élevés.

$x_{0,8}$ est symétrique de $x_{0,2}$ par rapport à μ=500 donc : $x_{0,8} = 500 + 42 = 542$ ou

$x_{0,8} = 500 + (50 \times z_{0,8}) = 500 + (50 \times 0{,}84) = 500 + 42 = 542$ car $z_{0,8}$=0,84 est le quantile d'ordre 0,8 de la loi $\mathcal{N}(0,1)$.

$x_{0,6} = 500 + (50 \times z_{0,6}) = 500 + (50 \times 0{,}25) = 500 + 12{,}5 = 512{,}5$ car $z_{0,6}$=0,25 est le quantile d'ordre 0,6 de la loi $\mathcal{N}(0,1)$

et $x_{0,4}$ symétrique de $x_{0,6}$ par rapport à μ=500 : $x_{0,4} = 500 - 12{,}5 = 487{,}5$

ou $x_{0,4} = 500 - (50 \times z_{0,6}) = 500 - (50 \times 0{,}25) = 500 - 12{,}5 = 487{,}5$

classe 1 (non admis) : résultats inférieurs à 458
classe 2 : résultats compris entre 458 et 487,5
classe 3 : résultats compris entre 487,5 et 512,5
classe 4 : résultats compris entre 512,5 et 542
classe 5 : résultats supérieurs à 542

Echantillon de taille n=65 de X issu de $\mathcal{P}$ pour lequel le résultat moyen observé $\bar{x} = 487{,}5$.

8) Puisque le résultat X suit une loi normale, la moyenne empirique du résultat $\overline{X}_{65}$ a une distribution normale de moyenne μ=500, de variance $\frac{\sigma^2}{n} = \frac{50^2}{65} \approx 38{,}46$ et d'écart-type $\frac{\sigma}{\sqrt{n}} = \frac{50}{\sqrt{65}} \approx 6{,}2$ car X a une distribution de moyenne μ=500 et d'écart-type σ=50.

9) $P(\overline{X}_{65} < \bar{x}) = P(\overline{X}_{65} < 487{,}5) \approx P\left(Z \leq \frac{487{,}5 - 500}{6{,}2}\right) = F(-2{,}016) \approx F(-2{,}02) = 1 - F(2{,}02) = 1 - 0{,}9783 = 0{,}0217$ où F est la fonction de répartition de la loi $\mathcal{N}(0,1)$.

➔ environ 2,2% (peu) des échantillons de taille 65 de X issus de $\mathcal{P}$ ont un résultat moyen inférieur à 487,5 résultat moyen observé.

10) 5% des échantillons de taille 65 ont un résultat moyen inférieur au résultat cherché, qui est donc par définition le quantile d'ordre 5% de $\overline{X}_{65}$:

$$Q_{0,05} \approx 500 + \left(z_{0,05} \times \frac{50}{\sqrt{65}}\right) = 500 - (z_{0,95} \times 6{,}2) = 500 - (1{,}645 \times 6{,}2) \approx 500 - 10{,}2 \approx 489{,}8$$

car $z_{0,95}$=1,645 est le quantile d'ordre 0,95 de la loi $\mathcal{N}(0,1)$.

➔ 5% des échantillons de taille 65 de X issus de $\mathcal{P}$ ont un résultat moyen inférieur à 489,8 environ.

11) Intervalle de variation à 95% (au risque 5%) du résultat moyen sur les échantillons de taille 65 de X issus de $\mathcal{P}$:

$$I_{95\%}(\overline{X}_n) = [Q_{0,025} ; Q_{0,975}] \approx \left[500 \pm z_{0,975} \times \frac{50}{\sqrt{65}}\right] = [500 \pm 1{,}96 \times 6{,}2] \approx [500 \pm 12{,}15] = [487{,}85 ; 512{,}15]$$

car $z_{1-(\alpha/2)} = z_{0,975}$=1,96 est le quantile d'ordre 0,975 de la loi $\mathcal{N}(0,1)$.

➔ 95% des échantillons de taille 65 de X issus de $\mathcal{P}$ ont un résultat moyen compris entre 487,85 et 512,15 environ.

12) Le résultat moyen observé $\bar{x} = 487{,}5$ sur l'échantillon de taille 65 est tel que :

2,2% des échantillons de taille 65 de X issus de $\mathcal{P}$ ont un résultat moyen inférieur à celui observé (cf question 9) ou $\bar{x} = 487{,}5$ n'appartient pas à l'intervalle de variation à 95% du résultat moyen sur les échantillons de taille 65 puisque $I_{95\%}(\overline{X}_n) = [487{,}85 ; 512{,}15]$ (cf question 11).

➔ l'échantillon considéré ayant été obtenu par tirage au sort dans $\mathcal{P}$, il est représentatif de la population des candidats : il est donc surprenant d'observer un tel résultat moyen (on a observé par hasard un des 5% des échantillons de taille 65 qui, par construction, ont un résultat observé en dehors de l'intervalle de variation à 95% de $\overline{X}_n$).

Exercice 4

A.

1) $\mathcal{P}$={garçons âgés de 6 à 12 ans, diagnostiqués hyperactifs (critères du DSM-II-R) et adressés en consultation psychiatrique}

2) X= score de résistance à la "distractibilité" sur l'échelle d'intelligence de Wechsler (WISC-R), variable quantitative de moyenne μ et d'écart-type σ inconnus dans $\mathcal{P}$

Echantillon de X issu de $\mathcal{P}$ de taille n=31

L'estimation ponctuelle du score moyen de résistance à la "distractibilité" μ est donnée par la moyenne (empirique) observée $\bar{x} = \frac{283}{31} \approx 9{,}129 \approx 9{,}1$

➔ le score moyen de résistance à la "distractibilité" sur l'échelle d'intelligence de Wechsler des garçons hyperactifs de 6 à 12 ans est estimé à 9,1 (points).

3) L'estimation ponctuelle sans biais de la variance du score de résistance à la "distractibilité" σ² est donnée par la variance (empirique) observée sans biais $s^2{*} = \frac{2\,682 - 31 \times 9{,}1^2}{30} = \frac{2\,682 - 2567{,}11}{30} = \frac{114{,}89}{30} \approx 3{,}83$

(ou $s^2 = \frac{2\,682}{31} - 9{,}1^2 = 86{,}52 - 82{,}81 \approx 3{,}71$ et $s^2{*} = \frac{31}{30} \times 3{,}71 \approx 1{,}033 \times 3{,}71 \approx 3{,}83$)

L'estimation ponctuelle sans biais de l'écart-type du score de résistance à la "distractibilité" σ est donnée par l'écart-type (empirique) observée sans biais $s^* = \sqrt{3{,}83} = 1{,}957 \approx 2$

(calcul exact : s^{2*}=3,28 et s*=1,81 ; arrondi à 2 décimales $\bar{x} \approx 9{,}13$ s^{2*}=3,26 et s*=1,81)

➔ la variance du score de résistance à la "distractibilité" des garçons de 6 à 12 ans hyperactifs est estimée à 3,83
l'écart-type du score de résistance à la "distractibilité" des garçons de 6 à 12 ans hyperactifs est estimé à 2 (points).

4) L'estimation par intervalle de confiance au niveau 90% (au risque α=10%) du score moyen μ dans $\mathcal{P}$ s'écrit :

$$IC_{90\%}(\mu) \approx \left[\bar{x} \pm z_{1-\frac{\alpha}{2}} \frac{s^*}{\sqrt{n}}\right] \approx \left[9{,}1 \pm z_{0{,}95} \frac{2}{\sqrt{31}}\right] \approx [9{,}1 \pm 1{,}645 \times 0{,}359] = [9{,}1 \pm 0{,}59] \approx [9{,}1 \pm 0{,}6] = [8{,}5;\, 9{,}7]$$

où $z_{1-(\alpha/2)}=z_{0{,}95}$=1,645 est le quantile d'ordre 0,95 de la loi $\mathcal{N}(0,1)$

(calcul exact : $IC_{90\%}(\mu) \approx$ [9,129±0,535] ≈ [8,594 ; 9,664] ; arrondi à 2 décimales pour $\bar{x} \approx 9{,}13$ et s*=1,81 : $IC_{90\%}(\mu) \approx$ [9,13±0,53] ≈ [8,60 ; 9,66]).

Cette approximation est justifiée car n=31≥30 donc la moyenne empirique $\overline{X}_n$ suit approximativement une loi normale.

➔ l'estimation par intervalle de confiance au niveau 90% du score moyen de résistance à la "distractibilité" des garçons de 6 à 12 ans hyperactifs est d'environ 8,5 à 9,7 (points); la précision (ou marge d'erreur) de l'estimation à 90% est d'environ 0,6 (point).

5) La marge d'erreur dans l'estimation du score moyen de résistance à la "distractibilité" au risque 10% (au niveau 90%), donnée par la demi-longueur de l'intervalle $IC_{90\%}(\mu)$, est d'environ 0,6 (point) pour un échantillon de taille n=31 ; pour avoir une marge d'erreur (demi-longueur) de 0,5 point donc plus faible, il faudrait plus de sujets.

Pour calculer la valeur minimum de n, pour s*=2 et α=10% connus, la demi-longueur de l'intervalle $IC_{90\%}(\mu)$ s'écrit :

$$z_{0{,}95} \frac{s^*}{\sqrt{n}} = 1{,}645 \frac{2}{\sqrt{n}} \text{ et on cherche n tel que : } 1{,}645 \frac{2}{\sqrt{n}} \le 0{,}5 \text{ c'est à dire } \frac{1{,}645 \times 2}{0{,}5} \le \sqrt{n} \text{ d'où } n \ge (1{,}645 \times 4)^2 = 43{,}3$$

(calcul exact : n ≥ 35,53 ; calcul arrondi à 2 décimales n ≥ 35,29).

➔ on choisirait donc une taille d'échantillon au moins égale à 44 pour que la marge d'erreur dans l'estimation du score moyen de résistance à la "distractibilité" à 90% soit inférieure à 1 point.

B.

1) $\mathcal{P}$={garçons âgés de 6 à 12 ans, diagnostiqués hyperactifs (critères du DSM-II-R) et adressés en consultation psychiatrique}

X= score d'aptitude à l'organisation perceptive sur l'échelle d'intelligence de Wechsler (WISC-R), variable quantitative de moyenne μ et d'écart-type σ inconnus dans $\mathcal{P}$

Echantillon de X issu de $\mathcal{P}$ de taille n=31

L'estimation ponctuelle du score moyen d'aptitude à l'organisation perceptive μ est donnée par la moyenne (empirique) observée $\bar{x} = 10{,}5$

➔ le score moyen d'aptitude à l'organisation perceptive sur l'échelle d'intelligence de Wechsler des garçons hyperactifs de 6 à 12 ans est estimé à 10,5 (points).

2) L'estimation ponctuelle de l'écart-type du score d'aptitude à l'organisation perceptive σ est donnée par l'écart-type observé sans biais $s^* = \sqrt{\frac{31}{30}} \times 1{,}48 = 1{,}01653 \times 1{,}48 = 1{,}504 \approx 1{,}5$

➔ l'écart-type du score d'aptitude à l'organisation perceptive des garçons de 6 à 12 ans hyperactifs est estimé à 1,5 (point).

3) L'estimation par intervalle de confiance au risque α=1% (au niveau 99%) du score moyen μ dans $\mathcal{P}$ s'écrit :

$$IC_{99\%}(\mu) \approx \left[\bar{x} \pm z_{1-\frac{\alpha}{2}} \frac{s^*}{\sqrt{n}}\right] \approx \left[10{,}5 \pm z_{0{,}995} \frac{1{,}5}{\sqrt{31}}\right] \approx [10{,}5 \pm 2{,}575 \times 0{,}269] \approx [10{,}5 \pm 0{,}7] = [9{,}8;\, 11{,}2]$$

où $z_{1-(\alpha/2)}=z_{0{,}995}$=2,575 est le quantile d'ordre 0,995 de la loi $\mathcal{N}(0,1)$.

Cette approximation est justifiée car n=31≥30 donc la moyenne empirique $\overline{X}_n$ suit approximativement une loi normale.

➔ l'estimation par intervalle de confiance au risque 1% (au niveau 99%) du score moyen d'aptitude à l'organisation perceptive des garçons de 6 à 12 ans hyperactifs est d'environ 9,8 à 11,2 (points); la précision (ou marge d'erreur) de l'estimation à 99% est d'environ 0,7 (point).