Exercice 1 (Estimation, 4pts)

Une enquête a été menée sur les salaires pratiqués dans un secteur donné en DA :

30 000	50 000	60 000	70 000	75 000
90 000	100 000	150 000	155 000	200 000

On connaît l'écart type des salaires : elle vaut σ = 20 000 DA.

a) (2p) Calculer un intervalle de confiance à 95 % pour la moyenne.

b) (2p) Quel niveau correspond à un intervalle de longueur 10 000 DA ?

Exercice 2 (Convergence, 4pts)

Dans un réseau informatique, le nombre de bits reçus par erreur est modélisé par une variable aléatoire binomiale, et supposons que la probabilité qu'un bit soit reçu par erreur soit de 1×10^{-5} . Si 16 millions de bits sont transmis :

1) (2p) Exprimer la probabilité que 150 erreurs ou moins se produisent en utilisant la loi binomiale ? Expliquer les difficultés pour la calculer.
2) (2p) Evaluer cette probabilité avec une approximation appropriée.

N.B./ Loi binomiale $P(X=k) = C_n^k p^k (1-p)^{n-k}, C_n^k = \frac{n!}{k!(n-k)!}$

Exercice 3 (Test hypothèse, 7pts)

Dans un secteur de l'électroménager, on a constaté que les appareils les plus fiables sont dans une moyenne de puissance 3 000 Watts et un écart type de 1 000 Watts. On soupçonne que la puissance moyenne en Watts est inférieure à cette norme, mais il faudrait le démontrer avec une conclusion forte pour apporter des changements à la production. On veut effectuer un test d'hypothèse à partir d'un échantillon aléatoire de 16 appareils en utilisant un seuil de α = 0.05.

a) (1p) Quelle serait la bonne formulation pour l'hypothèse alternative ?

b) (2p) Déterminez les régions de rejet et d'acceptation en utilisant le seuil de α = 0.05 sous H0 : μ = 3 000 Watts. Tracez la courbe qui la représente et identifiez la zone de rejet de H0.

c) *(1p) Supposons que l'échantillon a une puissance de 2 600 Watts. Quelle est votre conclusion ?*

d) *(3p) Sur le graphique tracé en b),* ajoutez la courbe représentant la loi de la moyenne échantillonnée lorsque la vraie valeur μ est de 2 000 Watts. Evaluer l'erreur β et la puissance du test. Hachurez la zone correspondant à la probabilité d'erreur de deuxième espèce. Interpréter toutes les plages de valeurs dans ces courbes !

Exercice 4 (Test Khi2, 5pts)

On a interrogé des habitants de trois quartiers (A, B, C) d'une grande ville sur le nombre de smartphones achetés par catégories (S1 (bas de gamme), S2 (moyenne gamme), S3 (haut de gamme)) :

	A	B	C	Total
S1	3	7	10	20
S2	2	6	7	15
S3	2	14	29	45
Total	7	27	46	80

Déterminer si les catégories de smartphones dépendent ou non des quartiers avec un risque 10% ?

Stat2 — Jan. 2025 — EF

Exercice 1 a) m = moyenne = $98000 = \frac{70000+50000+\dots+200000}{10}$

$\sigma = 20000$

Intervalle de confiance à 95% IC :

$$IC = \left[m - t \cdot \frac{\sigma}{\sqrt{n}}, m + t \cdot \frac{\sigma}{\sqrt{n}}\right]$$

$$= \left[98000 - 1.96 \times \frac{20000}{\sqrt{10}}, 98000 + 1.96 \times \frac{20000}{\sqrt{10}}\right]$$

$$= [85\,603,87 , 110\,396.12]$$

b) Niveau de confiance à 10 000 DA :

$$t \cdot \frac{\sigma}{\sqrt{n}} = 5000 \Rightarrow t = \frac{\sqrt{10} \times 5000}{20000}$$

$$\Rightarrow t = 0.79$$

$$P(Z < 0.79) = 0.5 + 0.2852 = 0.7852 = 1 - \frac{\alpha}{2}$$

$$\Rightarrow \alpha = (1 - 0.7852) \times 2 = 0.4296 = 43\%$$

$\Rightarrow$ Niveau de confiance est donc 57%

Exercice 2 1) Probabilité 150 erreurs ou moins :

$$P(X \leq 150) = \sum_{x=0}^{150} C_{150}^{x} \times (10^{-5}) \cdot (1 - 10^{-5})^{16 \cdot 10^6 - x}$$

Difficultés : 1) Calcul couteux. 2) Instabilité numérique.

2) Evaluation par approximation :

Par le fait $n.p = (16.10^6) \times (1 - 10^{-5}) = 160 > 5$ et $n(1-p) > 5$,

on peut approximer la variable binomiale X (nombre de bits)

par une variable normale $Z = \frac{X - np}{\sqrt{n.p.(1-p)}}$

$$P_{Binomiale}(X \leq 150) = P_{Normale}(X < 150.5) = P\left(Z \leq \frac{150.5 - 160}{\sqrt{160 \times (1 - 10^{-5})}}\right)$$

$$= P(Z \leq -0.75) = 0.227 = 22.7\%$$

EF — Stats — Jan-2025

Exercice 3 a) Hypothèse alternative : $H_0 : \mu = 3000$

$H_a : \mu < 3000$

b) Région de rejet et d'acceptation en utilisant $\alpha = 0.05$:

$$P\left(Z < \frac{k - 3000}{1000/\sqrt{16}}\right) = 0.05 \Rightarrow \frac{k - 3000}{250} = 1.64$$

$\Rightarrow k = 2590$

Zone de rejet H_0 : $X \leq 2590$

— d'acceptation de H_0 : $X > 2590$

c) Conclusion si puissance 2600.

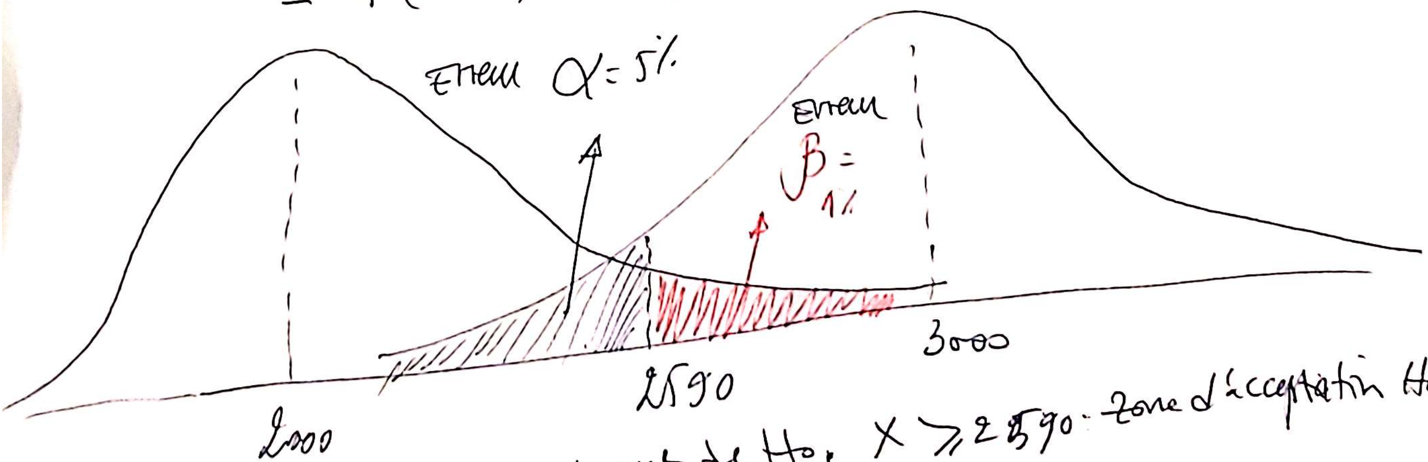
Puisque $2600 > 2590 \Rightarrow H_0$ est acceptée

$\Rightarrow$ Appareil fiable.

d) Erreur de type 2 : Il est question d'évaluer l'erreur de type 2, d'acceptation de H_0 sachant que H_0 est rejetée par la moyenne $\mu = 2000$ (même σ).

$$\beta = P(\mu \geq 2590 \mid \mu = 2000) = P\left(Z \geq \frac{2590 - 2000}{1000/4}\right)$$

$$= P(Z \geq 2.36) = 0.5 - 0.4909 = 0.01 = 1\%.$$



$X \geq 2590$: Zone d'acceptation H_0

$X \leq 2590$: Erreur 1, de rejet de H_0.

Puissance de test $= 1 - \beta = 99\%$.

L'erreur de type 2 est très faible.

Exercice 4 — EF - Statis - Jan - 2025

Test d'indépendance entre quartiers et catégories.

Soit le tableau P2 d'indépendance :

	A	B	C
S_1	$\frac{20\times 7}{80}$	$\frac{20\times 27}{80}$	$\frac{20\times 48}{80}$
S_2	$\frac{15\times 7}{80}$	$\frac{15\times 27}{80}$	$\frac{15\times 46}{80}$
S_3	$\frac{45\times 7}{80}$	$\frac{45\times 27}{80}$	$\frac{45\times 46}{80}$

$$P2(i,j) = \frac{(\text{Somme } i)\times(\text{Somme } j)}{\text{Total}}$$

	A	B	C
S_1	1.75	6.75	11.50
S_2	1.31	5.06	8.63
S_3	4.02	15.19	25.88

$$\chi^2 = \sum \frac{(\text{valeur observée} - \text{valeur calculée})^2}{\text{valeur calculée}}$$

$= 3.36$

α 10%, nb. lignes = 3, " colonnes = 3 — Degrés de liberté $= (3-1)\times(3-1) = 4$

χ^2 de la table $= 7.78$

$3.36 < 7.78 \Longrightarrow$ Indépendance entre les deux, entre quartiers et les catégories.