Exercice 1 :

Dans la littérature, il est admis grâce à de nombreuses expériences que les résultats d'une épreuve psychologique suivent une loi normale d'espérance 100 et d'écart type 5.

Un doute existe ; un praticien reprend alors l'expérience sur un échantillon de 500 patients. Il a conservé les résultats de l'épreuve sous la forme suivante :

Intervalles	< 85	[85 - 90[	[90 - 95[	[95 - 100[	[100 - 105[	[105 - 110[	≥ 110	
Effectifs	9	12	71	168	158	67	15	500

Dans un premier temps on calcule la moyenne des observations et trouvant 99.65, l'expérimentateur n'est pas perturbé ! Il poursuit tout de même ses vérifications.

1. Donner la proportion de patients de l'échantillon dont le résultat est "≥ 110" et calculer la probabilité de ce même résultat pour la loi de probabilité supposée.

   La proportion sur cet intervalle est : 15/500 = 0.03
   La probabilité $P(X > 110)$ (on utilise la $\mathcal{N}(100, \sigma = 5)$) = 0.02275013

2. Même question pour "< 85".

   La proportion sur cet intervalle est : 9/500 = 0.018
   La probabilité $P(X < 85)$ (on utilise la $\mathcal{N}(100, \sigma = 5)$)= (ici avec la table = 0.001

3. Afin d'obtenir le résultat précédent avec plus de précision, on procède de la façon suivante : on calcule la probabilité d'être "< 84.55" et on admet que la probabilité d'être dans l'intervalle [84.55, 85] est de 0.00035. Quel est le nouveau résultat alors obtenu ? (C'est ce résultat que l'on utilisera dans la suite de l'exercice)

   $P(X < 85) = P(X < 84.55) + P(84.55 < X < 85) =$ **0.001 (question précédente) + 0.00035**
   On obtient donc un résultat plus précis : 0.00135

4. Dans l'un des 2 cas des questions 1 et 2, l'expérimentateur observe un écart entre l'observation et la théorie qui lui semble important. Lequel ?

   évidemment c'est la cas "< 85"

5. Il se propose alors, à l'aide d'un test statistique de confirmer ses doutes. Quel type de test peut-il mettre en œuvre compte tenu du tableau de données ? Quelle sera alors l'hypothèse nulle de ce test ? et l'hypothèse alternative ?

   Il peut mettre en œuvre un test du χ^2 d'adéquation à la loi normale $\mathcal{N}(100, \sigma = 5)$.
   l'hypothèse nulle de ce test H_0 : "Les observations sont issues d'une loi normale $\mathcal{N}(100, \sigma = 5)$"
   l'hypothèse alternative : "Les observations NE sont PAS issues d'une loi normale $\mathcal{N}(100, \sigma = 5)$".
   (par exemple une loi normale avec d'autres paramètres ou une loi d'autre nature)

6. Sans développer les calculs, décrire en détail la démarche et donner la zone de rejet du test.

   Les données sont fournies sous la forme d'un tableau de distribution avec un découpage en 7 intervalles.
   Pour chaque intervalle on calcule la probabilité sous H_0 (idem question 1 et 2).
   L'"effectif théorique" sous H_0 pour chaque intervalle sera donc cette proba *

500.
Et dans chacun des intervalles on calcule la contribution au χ^2 (voir ex question 7) que l'on additionne.
Le ddl = 6 et pour $\alpha = 0.001$ on a une zone de rejet de H_0 :
"rejet de H_0 si χ^2 calculé est > 22.46.

7. Pour finir, développer les calculs sur le cas déterminé à la question 4). Proposer une conclusion.

   Cas 1 (<85) : L'"effectif théorique" sous H_0 = 0.00135 * 500 = 0.674949.
   la contribution au $\chi^2 = (9 - 0.674949)^2/0.674949 = 102.6840$.
   Déjà très largement > 22.46.
   "On peut donc REJETER l'Hypothèse H_0 (sans état d'âme) et donc dire que Les observations NE SONT PAS issues d'une loi normale $\mathcal{N}(100, \sigma = 5)$"

8. Étant donné qu'il n'y a pas de raison de remettre en cause la nature de la loi (loi normale) ainsi que la valeur de l'espérance (=100) de l'épreuve étudiée, quels enseignements fournit alors ce test ?

   Si la nature de la loi (loi normale) ainsi que la valeur de l'espérance (=100) de l'épreuve étudiée ne sont pas à remettre en cause, c'est donc la valeur de la variance (supposée = 25). Comme les questions 1 et 2 le suggèrent (l'effectif observé est > à l'effectif théorique et se sont les 2 intervalles des "bords" de distribution), il semble donc que la variance des observations soit > à celle supposée (25).

Exercice 2 :

Partie A : Le service contentieux d'une banque étudie le montant des chèques sans provision émis durant le mois de décembre 2004 : elle en a comptabilisé 520, avec une moyenne égale à 75 euros et un écart-type de 30 euros.

1. Donner un intervalle de confiance à 92% sur le paramètre μ égal au montant moyen théorique d'un chèque sans provision.

   Soit X la v.a de la valeur d'un chèque sans provision, X n'est pas une v.a. gaussienne.
   Par contre la moyenne $\overline{X}$ l'est ("TCL") : c'est une $\mathcal{N}(\mu, \sigma^2/n)$: μ et σ^2 sont 2 paramètres inconnus.
   Ici $n = 520$, une observation de $\overline{X}$ est 75, une observation de σ est 30.
   Un "IC" du paramètre μ (représentant le montant moyen théorique d'un chèque sans provision) est : $[\overline{X} - l \times \sigma/\sqrt{n}, \overline{X} + l \times \sigma/\sqrt{n}]$
   Ici $l = 1.7506$ (IC à 92%) donc $IC = [72.697, 77.303]$

2. Depuis le passage à l'euro, le montant moyen des chèques sans provision s'était stabilisé à 72 euros. Le service de contentieux en déduit alors une augmentation du montant de ces chèques qu'il explique par une situation économique difficile. A-t-il raison dans son raisonnement avec une probabilité d'erreur de 5, 3 ou 1 % ? Pour répondre à cette question, on proposera un test statistique et on développera l'ensemble de la démarche et des calculs.

   Le test statistique à construire est le test d'une l'hypothèse nulle $\mu = 72$ (le montant moyen des chèques sans provision est resté stable en décembre 2004) contre l'alternative $\mu > 72$ (le montant moyen des chèques sans provision de décembre 2004 est en augmentation)
   Pour une erreur de 1re espèce α, la zone de rejet du test sera :
   Rejet de H_0 si $\overline{X} > 72 + l_\alpha \times 30/\sqrt{520}$, où

$\overline{X}$ est l'observation du montant moyen des chèques sans provision de décembre 2004, ici = 75.
$l_\alpha = 2.326348(\alpha = 0.01)$; $= 1.880794(\alpha = 0.03)$; $= 1.644854(\alpha = 0.05)$
Pour ces 3 erreurs de 1re espèce, les 3 bornes de rejet sont donc 75.06051 ; 74.47435 ; 74.16395 .
En conclusion, on a "rejet de H_0" pour $\alpha = 0.03$ *ou* $\alpha = 0.05$, et NON rejet pour $\alpha = 0.01$.
Si on souhaite une erreur "faible' (0.01), le service de contentieux N'A DONC PAS raison dans son raisonnement : l'augmentation du montant de ces chèques n'est PAS SIGNIFICATIF.

Partie B : Cette même banque a classiquement un taux de chèques sans provision de 2 % et tous les moyens sont mis en œuvre pour faire diminuer ce taux. Durant ce même mois de décembre 2004, la banque a traité 30000 chèques au total.

Peut-on considérer que le taux de chèques sans provision est en diminution (avec une probabilité d'erreur volontairement très faible) ?

Ici le test à mettre en place est le test du paramètre d'une loi de Bernouilli.
Soit T la v.a. "être un chèque sans provision" : T suit une loi de $\mathcal{B}(p)$ et on construit le test de $p = 0.02$ (le taux de chèques sans provision est resté stable) contre $p < 0.02$ (le taux de chèques sans provision est en diminution)
On a REJET de H_0 si $\overline{T} < 0.02 - l_\alpha\sqrt{(0.02)(1-0.02)/30000}$
On utilise ici le "TCL" pour la moyenne $\overline{T}$ d'une loi de Bernouilli, avec n=30000.
L'observation de $\overline{T}$ est dans ce cas de 520/30000 = 0.01733333.
Pour $\alpha = .01$ la borne de rejet est 0.01811964
Pour $\alpha = .001$ la borne de rejet est 0.01750219
Dans ces 2 cas il y a REJET de H_0 (0.017333 est plus petit que les bornes de rejet).
Le taux de chèques sans provision est SIGNIFICATIVEMENT en diminution

Partie C : Le conseil d'administration avait donné un objectif 0.017 pour le taux de chèques sans provision. Est-il en droit de dire que cet objectif n'est pas atteint ?

Même schéma que précédemment mais ici on construit le test de $p = 0.017$ (objectif atteint) contre $p > 0.017$ (objectif NON atteint)
On a REJET de H_0 si $\overline{T} > 0.017 + l_\alpha\sqrt{(0.017)(1-0.017)/30000}$
Pour $\alpha = .01$ la borne de rejet est 0.01873626
Pour $\alpha = .05$ la borne de rejet est 0.01822763
Dans ces 2 cas il y a NON REJET de H_0 (0.017333 n'est pas plus grand que les bornes de rejet).
Le conseil d'administration ne peut donc pas affirmer que l'objectif n'est pas atteint.