### rappel de métrologie, incertitudes

1. **bi-prisme de Fresnel (révisions S1)** dans un dispositif optique destiné à créer des interférences lumineuses, la distance $i$ entre 2 franges de même nature (interfrange) est donnée par $i = \dfrac{\lambda D}{2(n-1)Ad}$, avec $\lambda$ la longueur d'onde de la lumière utilisée ; $n$ l'indice et $A$ l'angle du bi-prisme ; $d$ la distance entre source et bi-prisme et $D$ celle entre le bi-prisme et le plan d'observation des interférences.

   $n$ et $\lambda$ étant parfaitement connus ($n = 1,5$ ; $\lambda = 546$nm) on détermine la valeur de l'angle $A$ à partir des mesures de $i$, $d$ et $D$. On fixe $d = 20,0$cm et $D = 40,0$cm, mesurés avec une règle graduée de millimètre en millimètre.

   On mesure ensuite $i$ à l'aide d'un dispositif capable d'apprécier le micromètre et on relève les valeurs suivantes : $77$ ; $78$ ; $74$ ; $76$ ; $79$ ; $73$ ; $80$ ; $75$ ; $81$, exprimées en micromètres.

   Calculer la valeur supposée vraie de l'angle $A$ puis les incertitudes types absolue et relative sur cette valeur. Donner un intervalle de confiance à 95% pour cette mesure de $A$.

   corrigé succint :

   On calcule pour $i$ une moyenne $\boxed{\bar{i} = 77\,\mu\text{m}}$ avec un écart-type de $\boxed{\sigma(i) = 2,74\,\mu\text{m}.}$

   Comme $A = \dfrac{\lambda D}{2(n-1)id}$ on en déduit la valeur supposée vraie $A = 1,41818 \times 10^{-2}$ rad, on peut garder tous ces chiffres pour les calculs ultérieurs mais afficher un résultat à 3 chiffres $\boxed{A = 1,42 \times 10^{-2}\text{ rad.}}$ puisque tous les termes du calcul sont donnés avec 2 ou 3 chiffres dans l'énoncé. En mettre 2 serait plus rigoureux mais trop strict au vu de l'incertitude calculée plus loin...en donner 3 suppose que les valeurs données à 2 chiffres dans l'énoncé en ont en fait 3 - il n'y a pas de solution parfaitement satisfaisante.

   $A$ est produit et quotient de 3 variables $D$, $i$ et $d$ (supposées indépendantes), par conséquent l'incertitude relative sur $A$ est la somme quadratique des incertitudes relatives sur $D$, $i$ et $d$ :

   $$\frac{u(A)}{A} = \sqrt{\left(\frac{u(i)}{i}\right)^2 + \left(\frac{u(D)}{D}\right)^2 + \left(\frac{u(d)}{d}\right)^2}.$$

   $\boxed{u(D) \text{ et } u(d) \text{ valent } 1/(2\sqrt{3})\text{mm soit 0,28mm.}}$

   $u_A(i)$ vaut $2,74/\sqrt{9} = 0,91\mu$m

   $u_B(i)$ (incertitude de lecture) vaut $1/(2\sqrt{3})\mu$m, soit $u_B(i) = 0,28\mu$m, et donc

   $$\boxed{u_{tot}(i) = \sqrt{0,91^2 + 0,28^2} = 0,95\,\mu\text{m}.}$$

   Par ailleurs $\dfrac{u(A)}{A} = 1,25\%$, donc $\boxed{u(A) = 1,8 \times 10^{-4}\text{ rad}}$ (incertitude : on donne 2 chiffres en majorant le dernier, ici 1,77.. : en prend 1,8) et par conséquent

   $$\boxed{A = (14,2 \pm 0,18)10^{-3}\text{ rad.}}$$

   Enfin un intervalle de confiance 95% est obtenu en prenant un facteur d'élargissement de 1,96 : $1,96 \times 0,18 = 0,35$ donc on obtient l'invervalle $(14,2 \pm 0,35)10^{-3}$ rad soit

   $$\boxed{[13,8.10^{-3}\text{ rad}\ ;\ 14,5.10^{-3}\text{ rad}]}$$

2. **gaz parfait (révisions S1)**

   Une enceinte de volume $V$ contient une mole d'un gaz parfait.

   Pour calculer $V$ on mesure la pression interne $p$ et la température $T$ du gaz, relié par la relation $pV = nRT$ avec $R = 8,31451$J.mol$^{-1}$.K$^{-1}$.

   a) Donner l'expression de l'incertitude-type sur $V$ en fonction des incertitudes-type sur $p$ et $T$.

   b) Calculer la valeur de $V$ et celle de son incertitude-type sachant que :

   i. Les valeurs mesurées de $p$ et $T$ sont respectivement : $1,00$bar($10^5$Pa) et $324,6K$.

   ii. La lecture de la pression sur le manomètre consiste à repérer la position de l'aiguille devant une échelle graduée comportant 50 graduations entre la valeur 0 et la valeur maximum de 1 bar. La classe du manomètre est 1.

   iii. Le thermomètre utilisé pour mesurer $T$ est considéré comme parfaitement étalonné. C'est un thermomètre à mercure dont l'échelle est graduée de 0,2 en 0,2 degrés.

   corrigé succint :

   a) Loi des gaz parfaits : $pV = nRT$.

   On peut donc calculer le volume par $V = nRT/p$.

   L'énoncé indique que $n = 1$ et donne la valeur de $R$. On suppose ces deux quantités connues avec une incertitude nulle.

Pour calculer l'incertitude sur $V$, on va donc utiliser les incertitudes sur les quantités mesurées $p$ et $T$ (qui seront évaluées à la question suivante) par la formule de propagation des incertitudes :

$$\boxed{u(V) = \sqrt{\left(\frac{\partial V}{\partial p}u(p)\right)^2 + \left(\frac{\partial V}{\partial T}u(T)\right)^2}.}$$

Mais cette formule générale (à connaître bien sûr !) nécessite de calculer les dérivées partielles.

Ici, $V$ étant un quotient des deux variables mesurées, on peut aussi calculer plus simplement l'incertitude relative par $\boxed{\frac{u(V)}{V} = \sqrt{\left(\frac{u(p)}{p}\right)^2 + \left(\frac{u(T)}{T}\right)^2}.}$

b)

— Pour l'application numérique si on exprime les valeurs dans les unités de base du S.I, en particulier la pression en Pascals, on obtient : $\boxed{V = 2,70.10^{-2}\ \mathrm{m}^3.}$

— **incertitude sur la pression** : chaque graduation valant $10^5$ Pa (calibre) sur 50, l'incertitude de lecture est donc $\frac{10^5/50}{2\sqrt{3}}$ soit $577,35$ Pa.

Par ailleurs on a un calibre de $10^5$ Pa et un manomètre de classe 1, l'EMT vaut donc $a = 10^5/100 = 10^3$. L'incertitude de l'instrument est donc $10^3/\sqrt{3}$ soit $577,35$ Pa.

La somme quadratique de ces deux valeurs est l'incertitude de type B sur la pression

$$\boxed{u_B(p) = \sqrt{577,35^2 + 577,35^2} = 816,5\mathrm{Pa}.}$$

— **incertitude sur la température** : le thermomètre est supposé parfaitement étalonné, ce qui signifie qu'il n'y a pas d'incertitude sur l'instrument.

L'incertitude de lecture est $\frac{0,2}{2\sqrt{3}}$K soit 0.058K.

Ainsi, $\boxed{u_B(T) = 0.058\mathrm{K}.}$

— L'énoncé donne les mesures de $p$ et $T$, on vient de donner $u(p)$ et $u(T)$, on peut donc grâce à la formule de la question a) calculer $u(V)/V = 0.00816$ soit $0.816\%$.

En multipliant par la valeur de $V$ on obtient donc $\boxed{u(V) = 2,2.10^{-4}\mathrm{m}^3.}$

c) On peut présenter ainsi le résultat de la question précédente sous la forme :

$$\boxed{V = (2,70.10^{-2} \pm 2,2.10^{-4})\mathrm{m}^3}$$

ou encore $V = (2,70.10^{-2} \pm 0,022.10^{-2})\mathrm{m}^3$ ou encore $V = (2,70.10^{-2} \pm 0,02.10^{-2})\mathrm{m}^3$.

Ce n'est qu'une présentation pratique et synthétique de la valeur et de l'incertitude-type associée.

Pour donner un sens statistique à l'incertitude, on peut élargir l'écart-type en utilisant les propriétés de la loi normale. Ainsi en prenant un facteur 2, on sait que l'intervalle $\boxed{V = (2,70.10^{-2} \pm 2 \times 2,2.10^{-4})\mathrm{m}^3}$ aura une probabilité $95,45\%$ de contenir la valeur "vraie".

3. **éclairement : incertitudes et erreur systématique**

On désire étalonner un luxmètre, appareil qui mesure un éclairement $E$ (en lux).

Pour cela on utiliser une lampe qui délivre une intensité de $I = 390$ candelas (cd) avec une incertitude relative de $0.3\%$. La lampe est placée à la distance $d = 2$m du luxmètre, distance mesurée avec une incertitude-type de 1mm.

On admet la relation $E = \frac{I}{d^2}$.

(a) Calculer la valeur $E_{\text{calc}}$ correspondant aux valeurs de $I$ et $d$ mesurée, et l'incertitude $u(E_{\text{calc}})$ associée.

(b) On effectue 5 mesures successives d'éclairement, dans des conditions identiques, qui sont respectivement 96.502, 96.243, 96.361, 96.123 et 96.331 lux.

Calculer la valeur moyenne $\bar{E}$ et l'écart-type expérimental $s_E$.

(c) Quelle serait alors l'incertitude associée à une valeur $E_{\text{lue}}$ lue unique d'éclairement ?

(d) On constate que les valeurs $E_{\text{calc}}$ et $\bar{E}$ ne sont pas identiques, ce qui incite à introduire un terme correctif (offset) systématique $O = E_{\text{calc}} - \bar{E}$ pour toutes les mesures futures, de manière à avoir $E_{\text{supposée vraie}} = E_{\text{lue}} + O$.

Calculer la valeur numérique de $O$.

(e) Déterminer les incertitudes absolues et relatives sur $E_{\text{supposée vraie}}$ en fonction des incertitudes sur $E_{\text{lue}}$, $\bar{E}$ et $E_{\text{calc}}$. Effectuer les applications numériques.

corrigé succint :

(a) On trouve $E_{\text{calc}} = 97.5$ lum.

Par ailleurs par propagation des incertitudes (situation avec des puissances de variables indépendantes) : $u_{E_{\text{calc}}}/E_{\text{calc}} = \sqrt{(u(I)/I)^2 + 4(u(d)/d)^2} = 3.17.10^{-3} = 0.3\%$, donc $u_{E_{\text{calc}}} = 0.309$ lux.

(b) On calcule $\bar{E} = 96.3$ lux et $s_E = 0.1408$ lux, donc $u(\bar{E}) = s_E/\sqrt{5} = 0.063$ lux.

(c) Avec une seule mesure, $u(E_{\text{lue}}) = s_E = 0.1408$ lux.

(d) On trouve $O = 1.2$lux.

(e) $E_{\text{supposée vraie}} = E_{\text{lue}} + O = E_{\text{lue}} + E_{\text{calc}} - \bar{E}$, donc par propagation des incertitudes, l'incertitude sur $E_{\text{supposée vraie}}$ est la racine de la somme quadratique des trois incertitudes soit $u(E_{\text{lue}})^2 + u(\bar{E})^2 + u(E_{\text{calc}})^2 = s_E^2(1 + 1/5) + u_{E_c}^2$,

soit $u(E_{\text{supposée vraie}}) = 0.35$lux, l'incertitude relative est de $0.35/97.5$ soit $0.36\%$.

## dénombrement, probabilités élémentaires, variables aléatoires

4. Dans un groupe de 24 étudiants, est-ce courant d'en avoir deux qui fêtent leur anniversaire le même jour ?

corrigé succint : On simplifie le problème en considérant que tous les étudiants ont une chance sur 365 d'avoir un jour d'anniversaire donné (on ignore les 29 février et les variations saisonnières). Le résultat trouvé sera peu différent de la valeur exacte bien plus compliquée à calculer et qui nécessiterait de connaître la répartition des naissances en fonction du jour de l'année.

Le nombre de manière de choisir 24 jours d'anniversaires est $365^{24}$.

Le nombre de manière de choisir 24 jours d'anniversaires tous différents $365 \times 364 \times \ldots \times (365-24+1)$.

Ainsi la probabilité de ne pas avoir d'élèves qui fêtent leur anniversaire le même jour est $\dfrac{365 \times 364 \times \ldots \times (365-24+1)}{365^{24}}$, soit approximativement $46.17\%$.

5. Durant un jeu avec paquet de 32 cartes, chacun des deux joueurs reçoit 5 cartes : 2 connues de tous et 3 connues de lui seul.

   a) Quelle est la probabilité pour qu'un joueur ait un carré d'as ?

   b) Vous avez vu que votre adversaire a reçu deux as ; de votre côté vous n'en avez aucun. Quelle est la probabilité que votre adversaire ait un carré d'as en main ?

corrigé succint :

a) Il y a $\binom{32}{5}$ mains possibles. Parmi celles-ci, 28 contiennent un carré d'as (28 choix possibles pour l'unique carte qui n'est pas un as). Donc la probabilité est $\dfrac{28}{\binom{32}{5}} = 1/7192$ soit environ $1,39 \times 10^{-4}$ soit $0.013\%$.

b) On ne s'intéresse qu'aux 3 cartes inconnues, à choisir parmi $25 = 32 - 5 - 2$ (vos 5 cartes et les 2 as connus). Il y a donc $\binom{25}{3}$ choix possibles, et parmi ceux-ci, 23 contiennent les deux as manquants (choix des 2 as et d'une seule autre carte parmi les 23 qui ne sont pas des as). La probabilité est donc de $\dfrac{23}{\binom{25}{3}}$ soit exactement $1\%$.

6. On choisit 2 lampes dans une boîte de 24, dont 2 sont défectueuses.

   On note $X$ le nombre de lampes fonctionnelles parmi les 2 choisies.

   Donner la loi de $X$, son espérance, sa variance.

corrigé succint :

$X$ peut prendre les valeurs 0, 1, 2.

Il s'agit donc de calculer $p(X=0)$, $p(X=1)$ et $p(X=2)$.

On peut calculer plus facilement la probabilité que les deux lampes choisies fonctionnent : le nombre de choix est $\binom{22}{2} = \dfrac{22 \times 21}{2} = 231$, alors que le nombre de choix total est $\binom{24}{2} = \dfrac{24 \times 23}{2} = 276$.

Ainsi, $p(X=2)$ est $\dfrac{231}{276}$, soit approximativement 0.837.

$p(X=0)$ est la probabilité qu'aucune ne fonctionne soit $1/276$ (une seule manière de choisir les 2 lampes défectueuses) soit environ $0.36\%$.

Et enfin $p(X=1) = 1 - p(X=2) - p(X=0) = \dfrac{276-231-1}{276} = \dfrac{44}{276}$ soit environ $15.9\%$.

On peut alors calculer l'espérance

$E(X) = 0p(X=0) + 1p(X=1) + 2p(X=2) = \dfrac{44}{276} + 2\dfrac{231}{276} = \dfrac{506}{276} = \dfrac{11}{6} \simeq 1.83$,

$E(X^2) = 0^2p(X=0) + 1^2p(X=1) + 2^2p(X=2) = \dfrac{968}{276}$ et donc la variance vaut

$\text{Var}(X) = E(X^2) - E(X)^2 = \dfrac{121}{828} \simeq 0.146$.

7. $5\%$ des interrupteurs sortant d'une chaîne de production sont défectueux. On en prend deux au hasard. Soit $X$ la variable aléatoire « nombre d'interrupteurs défectueux dans l'échantillon prélevé ».

   Donner la loi de probabilité de X, calculer $E(X)$ et $\text{Var}(X)$.

corrigé succint : Les trois résultats possibles sont :

« deux interrupteurs défectueux », de probabilité $0.05^2 = 0.0025$,

« aucun interrupteur défectueux », de probabilité $0.95^2 = 0.9025$,

« un seul interrupteur défectueux », de probabilité $1 - 0.95^2 - 0.05^2 = 0.095$.

Alors $\boxed{E(X) = 0.0025 \times 2 + 0.095 = 0.1,}$ $E(X^2) = 0.0025 \times 4 + 0.095$, $E(X) = 0.105$, et donc $\boxed{\text{Var}(X) = E(X^2) - E(X)^2 = 0.095.}$

Remarque : on peut aussi voir que $X$ suit une loi binomiale de paramètres $n = 2$ et $p = 0.05$, donc $E(X) = np = 0.1$ et $\text{Var}(X) = np(1-p) = 0.1 \times 0.95 = 0.095$.

8. Deux machines $M_1$ et $M_2$ produisent quotidiennement respectivement 1000 et 3000 pièces. Le taux de pièces défectueuses est de $2\%$ pour $M_1$ et de $5\%$ pour $M_2$.

   Quelle est la probabilité qu'une pièce défectueuse provienne de $M_1$ ?

corrigé succint :

On tire une pièce au hasard : on note $D$ l'événement « la pièce est défectueuse », et $M_i$ l'événement « la pièce provient de la machine $M_i$ », $i = 1, 2$.

$p(D) = \dfrac{0.02 \times 1000 + 0.05 \times 3000}{1000 + 3000} = \dfrac{17}{400}$, $p(M_1) = \dfrac{1}{4}$, $p(M_2) = \dfrac{3}{4}$.

On sait que $p(D|M_1) = 0.02$, donc $p(D \cap M_1) = 0.02 \times p(M_1) = 0.005$. Et donc

$p(M_1|D) = \dfrac{p(D \cap M_1)}{p(D)} = \boxed{\dfrac{2}{17} \simeq 11.76\%.}$

9. On donne la distribution (loi) conjointe de deux variables discrètes :

| $X \vert Y$ | 0 | 1 | 2 |
|---|---|---|---|
| 10 | 0 | 0.25 | 0 |
| 12 | 0.25 | 0 | 0.25 |
| 14 | 0 | 0.25 | 0 |

   (a) déterminer les distributions (lois) marginales, puis $E(X)$ et $E(Y)$.

(b) Déterminer $E(X.Y)$ et la covariance de $X$ et $Y$.

(c) les deux variables sont-elles indépendantes ?

corrigé succint : Pour trouver les $p(X = i)$ on additionne les coefficients $p(X = i, Y = j)$ pour $j = 0, 1, 2$ : $\boxed{p(X = 10) = 0.25, p(X = 12) = 0.5, p(X = 14) = 0.25.}$

De même, $\boxed{p(Y = 0) = 0.25, p(Y = 1) = 0.5, p(Y = 2) = 0.25.}$

$E(X.Y) = 10p(X = 10, Y = 1) + 12p(X = 12, Y = 1) + 14p(X = 14, Y = 1) + 10p(X = 10, Y = 2) + 12p(X = 12, Y = 2) + 14p(X = 14, Y = 2) = 0.25 \times 10 + 0.25 \times 14 + 0.25 \times 24 = 12.$

$E(X) = 12$ et $E(Y) = 1$, et donc $E(X.Y) = E(X).E(Y)$, mais pourtant $\boxed{X \text{ et } Y \text{ ne sont pas indépendantes :}}$ on voit par exemple que $p(X = 10, Y = 0)$ n'est pas égal à $p(X = 10)p(Y = 0)$.

## lois usuelles

10. $X$ suit une loi normale de paramètres 2 et 4.

Calculer $p(X = 2)$, $p(X > 3)$, $p(-1 \leq X \leq 4)$, $p(-2 \leq X \leq 6)$.

corrigé succint :

Si $Z = \dfrac{X-2}{2}$, $Z$ suit une loi normale centrée réduite.

$p(X = 2) = 0$ car $X$ est continue.

$p(X > 3) = p(Z > 1/2) = 1 - p(Z < 1/2) = 1 - 0.6915 = 0.3085$,

$p(-1 \leq X \leq 4) = p(-3/2 \leq Z \leq 1) = p(Z \leq 1) - p(Z \leq -3/2) = p(Z \leq 1) - (1 - p(Z \leq 3/2)) = p(Z \leq 1) - 1 + p(Z \leq 3/2) = 0.8413 - 1 + 0.9332 = 0.7745$

Pour $p(-2 \leq X \leq 6)$ on peut refaire un calcul analogue au précédent, ou simplement remarquer que c'est la probabilité d'être dans un intervalle centré sur l'espérance et de "rayon" 2 écarts-types : la probabilité est donc $95.4\%$.

11. Dix boites sont posées dans une pièce. Un objet est caché dans l'une d'entre elles, et un joueur ouvre une boîte au hasard pour trouver l'objet.

Il effectue une série de 32 essais pour le retrouver (à chaque essai, réussi ou non, l'objet est caché à nouveau hors de la présence du joueur).

(a) Quelle est la loi de la variable $X$ « nombre de fois où l'objet est trouvé » ?

(b) Déterminer l'espérance et la variance de $X$.

(c) Calculer la probabilité d'avoir 0 succès, 3 succès, moins de 3 succès, plus de 10 succès.

(d) À partir de combien de succès peut-on considérer le résultat comme extraordinaire (c'est-à-dire qu'un tel nombre de succès n'est obtenu que pour $1\%$ des joueurs) ?

corrigé succint :

(a) Il s'agit d'une $\boxed{\text{loi binomiale de paramètres 32 et 0.1}}$ (on répète 32 fois l'opération qui consiste à choisir au hasard une boite, qui a une chance sur 10 de contenir l'objet, et on compte le nombre total de succès).

(b) Son espérance est $\boxed{E(X) = np = 3.2}$ et sa variance $\boxed{\text{Var}(X) = np(1 - p) = 2.88.}$

(c) $p(X = 0) = 0.034$, $p(X = 3) = 0.234$, $p(X \leq 3) = 0.600$, $p(X \geq 10) = 0.0008$.

Comme $n \geq 30$, $p \leq 0.1$ et $np = 3.2 \leq 5$, les conditions énoncées dans le cours sont vérifiées, et on utilisera effectivement l'approximation par une $\boxed{\text{loi de Poisson de paramètre } \lambda = np = 3 \text{ car on n'a pas la table pour la loi de paramètre 3.2.}}$

En utilisant l'approximation par une loi de Poisson de paramètre 3, on trouve donc :

— $\boxed{p(X = 0) = 0.050}$

— $\boxed{p(X = 3) = 0.022}$

— $\boxed{p(X \leq 3) = 0.647}$

— $p(X \geq 10) = 1 - p(X < 10) = 1 - p(X \leq 9) = 1 - 0.9989$, $\boxed{p(X \geq 10) = 0.001}$

— En utilisant la table de la loi de Poisson on constate que la probabilité d'obtenir moins de 7 essais est 0.988, d'obtenir moins de 8 essais est 0.996. Donc la probabilité d'obtenir 9 essais ou plus est inférieure à $1\%$.

12. Un ascenseur peut supporter une charge de 800 kilos.

Quel nombre $n$ le technicien chargé de rédiger l'écriteau "$n$ personnes autorisées au maximum" devrait-il choisir ?

13. Pour se prémunir contre les $10\%$ défections tardives habituellement constatées, une compagnie aérienne pratique la surréservation : elle vend 270 billets pour 250 sièges dans un avion.

Soit $X$ la variable aléatoire « nombre de personnes ayant réservé qui se présentent pour embarquer ».

(a) Montrer que $X$ suit une loi binomiale, et que l'approximation par une loi normale est justifiée.

(b) Quelle est la probabilité qu'exactement 250 personnes se présentent à l'embarquement ?

Quelle est la probabilité que toute personne ayant réservé et se présentant soit assurée d'un siège ?

corrigé succint :

(a) La probabilité qu'une personne ayant acheté un billet se présente à l'embarquement vaut $\boxed{p = 0.9,}$ puisque la compagnie constate 10% de défections.

Par conséquent, si on répète 270 fois un "tirage au sort" pour déterminer si un passager se présente (avec une probabilité de 0.9), et l'on compte le nombre de passagers qui se présentent : $X$ suit bien une loi binomiale de paramètres $n = 270$ et $p = 0.9$.

Comme $n > 50$ et et que $np(1-p) = 24.3$, on peut effectivement approcher la loi de $X$ par une $\boxed{\text{loi normale de paramètres } np = 243 \text{ et } np(1-p) = 24.3.}$

(b) On souhaite déterminer $p(X = 250)$ et $p(X \leq 250)$ ; si on utilise l'approximation par la loi normale, on calcule en fait $p(249.5 \leq X \leq 250.5)$ et $p(X \leq 250.5)$.

Pour cela, on se ramène à une loi normale centrée réduite : $p(X \leq 250.5) = p(\dfrac{X-243}{\sqrt{24.3}} \leq \dfrac{7.5}{\sqrt{24.3}}) = p(\dfrac{X-243}{\sqrt{24.3}} \leq 1.52) \simeq 0.9357$, de même $p(X \leq 249.5) \simeq 0.905$ Ainsi, $\boxed{\text{la probabilité que toute personne ayant réservé ait un siège est de } 93.57\%}$, et la probabilité qu'exactement 250 personnes se présentent est 0.03.

(remarque : un calcul exact donne, avec la loi binomiale, $p(X \leq 250) = 94.11\%$ ; par ailleurs un calcul de $p(X \leq 250)$ (et non 250.5) avec la loi normale donne une moins bonne approximation, 92.22%).

## estimation de paramètres

14. Dans un grand lot de pièces circulaires, on a prélevé au hasard 40 pièces dont on vérifie le diamètre. Les mesures (en cm) sont :

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 4.9 | 5.0 | 5.2 | 4.7 | 4.8 | 5.1 | 4.5 | 5.2 | 4.9 | 4.8 |
| 4.9 | 4.9 | 4.9 | 5.3 | 5.0 | 4.8 | 4.8 | 4.9 | 5.1 | 5.3 |
| 5.4 | 4.9 | 4.9 | 5.0 | 4.8 | 4.8 | 5.3 | 4.8 | 5.1 | 5.0 |
| 5.1 | 4.8 | 4.7 | 5.0 | 4.9 | 4.8 | 4.6 | 4.7 | 4.9 | 4.7 |

Estimer par un intervalle de confiance 95% le diamètre moyen des pièces.

corrigé succint :

On trouve les estimations ponctuelles $\boxed{\bar{x} = 4.93}$, $\boxed{s = 0.202}$ et $\boxed{s^2 = 0.0406.}$

L'effectif de l'échantillon étant « grand » (supérieur à 30), on peut considérer que $\sqrt{n}\dfrac{\bar{X}-\mu}{s}$ suit une loi normale centrée réduite.

Par conséquent, l'espérance de $\bar{X}$ sera dans 95% des cas dans l'intervalle $[\bar{x} - 1.96s/\sqrt{n}, \bar{x} + 1.96s/\sqrt{n}]$, soit ici dans l'intervalle $\boxed{[4.87, 4.99].}$

Estimation par intervalle de confiance de la variance :

**Si** $X$ **suit une loi normale**, on sait alors que $(n-1)s^2/\sigma^2$ suit une loi du $\chi^2$ à $n-1$ degrés de liberté, donc $\sigma^2$ est avec une probabilité de 95% dans l'intervalle $[(n-1)s^2/c_1^2, (n-1)s^2/c_2^2]$ pour $c_1^2 \simeq 59.342$ et $c_2^2 \simeq 24.433$ (lus dans la table du $\chi^2$ à 40 degrés de liberté, car le formulaire ne donne pas la table à 39 degrés de liberté), soit $\boxed{\sigma^2 \in [0.0267, 0.0648].}$

Si on ne sait pas que $X$ suit une loi normale, on ne peut rien dire...)

15. Des essais en laboratoire sur 20 lampes miniatures donnent les durées de vie suivantes, en heures : 451, 412, 412, 375, 407, 454, 375, 393, 355, 364, 414, 413, 345, 432, 392, 329, 439, 381, 451, 413.

On suppose la durée de vie distribuée normalement.

Estimer par un intervalle de confiance 95% la durée de vie moyenne.

corrigé succint : Les estimations ponctuelles de l'espéance, de l'écart-type et de la variance sont respectivement $\boxed{\bar{x} = 400.35}$, $\boxed{s = 36.01}$ et $\boxed{s^2 = 1297.}$

Alors $\sqrt{n}\dfrac{\bar{X}-\mu}{s}$ suit une loi de Student à 19 degrés de liberté, et donc l'espérance de $\bar{X}$ sera dans 95% des cas dans l'intervalle $[\bar{x} - t(0.95)s/\sqrt{n}, \bar{x} + t(0.95)s/\sqrt{n}]$, et on lit $t(0.95) = 2.093$ dans la table de la loi de Student à 19 degrés de liberté, donc l'intervalle cherché est $\boxed{[383.5, 417.2].}$

(complément hors-programme : estimation de l'écart-type par intervalle de confiance : $X$ suit une loi normale, donc $(n-1)s^2/\sigma^2$ suit une loi du $\chi^2$ à $n-1$ degré de liberté, donc $\sigma^2$ est avec une probabilité de 95% dans l'intervalle $[(n-1)s^2/c_1^2, (n-1)s^2/c_2^2]$ pour $c_1^2 \simeq 32.8523$ et $c_2^2 \simeq 8.9065$ (lus dans la table du $\chi^2$ à 19 degrés de liberté), soit $\sigma^2 \in [750.11, 2766.86]$, et donc l'écart-type a 95% de chances de vérifier $\boxed{\sigma \in [27.39, 52.6].}$)

16. (a) Sur un échantillon de 120 pièces fabriquées par une machine, 24 sont défectueuses. Trouver un intervalle de confiance à 98% de la proportion réelle de pièces défectueuses fabriquées par la machine.

(b) Même question avec 600 pièces défectueuses sur 3000.

corrigé succint :

(a) On admet que si la proportion de pièces défectueuses dans la population est $p$, la proportion $F$ de pièces défectueuses dans un échantillon d'effectif $n$ suit approximativement une loi normale $N(p, \dfrac{p(1-p)}{n})$. Par conséquent, $\dfrac{F-p}{\sqrt{p(1-p)/n}} \sim N(0,1)$, et donc dans 98% des cas, $-2.33 \leq \dfrac{F-p}{\sqrt{p(1-p)/n}} \leq 2.33$.

Si on approche l'écart-type $\sqrt{p(1-p)/n}$ par son estimation ponctuelle $\sqrt{0.2 * 0.8/120} = 0.0365$, on obtient $p \in [0.2 - 2.33 \times 0.0365, 0.2 + 2.33 \times 0.0365]$, soit l'intervalle $\boxed{[0.115, 0.285].}$

(b) De même, avec $n = 3000$, on trouve l'intervalle de confiance 98% $\boxed{[0.183, 0.217]}$.

Logiquement, l'augmentation de $n$ aboutit à un encadrement plus précis pour un taux de confiance fixé...

17. Le jour d'une élection, à la sortie des bureaux de vote, on interroge 1220 électeurs choisis au hasard : 615 d'entre-eux ont voté pour le candidat $A$.

a) Déterminer un intervalle de confiance 95% pour le résultat du candidat $A$ sur l'ensemble des électeurs.

b) Quelle est la probabilité pour que le candidat soit élu (c'est-à-dire que son score dépasse 50%) ?

corrigé succint :

a) Application directe du cours sur les estimations de proportion :
$[615/1220 \pm 1,96 \times \sqrt{(615/1220)(1-615/1220)/1220}$ soit $[0,476; 0,532]$.

b) On cherche $t$ tel que $0.5 = 615/1220 - t \times \sqrt{(615/1220)(1-615/1220)/1220}$ soit $t = 0.28$ soit $F(t) = 61\%$

**tests d'hypothèses**

18. On veut savoir si la résistance moyenne de composants produits dans une usine est $400\Omega$. On considère que la distribution des résistances est normale, et on mesure pour 16 composants les valeurs 392, 396, 386, 389, 388, 387, 403, 397, 401, 391, 400, 402, 394, 406, 406, 400.

Peut-on considérer, au seuil de signification $\alpha = 5\%$, que le lot respecte la norme de $400\Omega$ ? Même question avec un seuil de $\alpha = 1\%$.

corrigé succint : On détermine, à partir de l'échantillon, les estimations ponctuelles des espérance, variance et écart-type de la loi : on trouve $\boxed{\bar{x} = 396.125,\ s = 6.742 \text{ et } s^2 = 45.45.}$

Si l'on fait l'hypothèse $H_0$ : "le lot respecte la norme de $400\Omega$", alors dans 95% des cas la moyenne sur un échantillon d'effectif 16 se trouve dans l'intervalle $[400 - t * 6.742/4, 400 + t * 6.742/4]$, $t$ étant lu dans la table de la loi de Student à 15 degrés de liberté : $t = 2.1314$. Ainsi l'intervalle de confiance 95% pour la résistance est

$\boxed{[396.40, 403.59] \text{, et on peut donc, au risque } 5\%\text{, rejeter l'hypothèse.}}$

Au seuil $\alpha = 1\%$, on a dans l'hypothèse $H_0$ un intervalle de confiance pour la moyenne $[400 - t * 6.742/4, 400 + t * 6.742/4]$, avec $t = 2.9467$. Ainsi,

$\boxed{\text{l'intervalle est } [395.03, 404.97]\text{. Au risque } 1\%\text{, on ne rejette pas } H_0.}$

19. Un fabricant se vante de proposer des tubes à essai d'une durée de vie supérieure à 2000h de chauffage. A l'aide d'un échantillon de 100 tubes testés, on estime la durée de vie moyenne à 1975h et l'écart-type à 130h. Peut-on affirmer, au risque 5%, que le fabriquant ment ?

corrigé succint : Il s'agit ici d'un test unilatéral...$H_0$ est l'hypothèse : "la durée de vie moyenne vérifie $\mu \geq 2000$". On peut supposer, l'effectif de l'échantillon étudié étant grand, que $\sqrt{n}\dfrac{\bar{X}-\mu}{s}$ suit une loi normale centrée réduite.

Si $H_0$ est vérifiée, on cherche $t$ tel que $p(\mu - ts/\sqrt{n} \leq \bar{X}) = 0.95$, soit $p(-t \leq \sqrt{n}\dfrac{\bar{X}-\mu}{s}) = 0.95$, et donc $1 - F(-t) = F(t) = 0.95$ : $t = 1.64$.

Ainsi, dans l'hypothèse $H_0$, la durée de vie moyenne d'un échantillon d'effectif 100 se trouve, dans 95% des cas, dans l'intervalle $[2000 - 1.64 * 130/10, +\infty[= [1978.68, +\infty[$. La mesure de 1975h sur l'échantillon n'étant pas dans cet intervalle,

$\boxed{H_0 \text{ doit être rejetée : il est probable que le fabriquant mente.}}$

20. Un fabricant de médicaments affirme que les masses d'un composant dans ses comprimés sont réparties selon une loi normale d'espérance 75 mg. Les mesures pour le vérifier étant coûteuses, trois seulement sont réalisées, dont les résultats sont 70, 72 et 74 mg. Peut-on, au risque de 5% de se tromper, contester l'affirmation ?

corrigé succint : Notons $X$ la variable aléatoire correspondante et $\mu = E(X)$ : il s'agit donc ici d'effectuer un test bilatéral de l'hypothèse $H_0 : \mu = 75$.

On obtient sur un échantillon de 3 mesures : $n = 3$, $\bar{x} = 72$, $\sigma'^2 = 8/3$ et $s^2 = 8/2 = 4$, donc l'estimation ponctuelle de l'écart-type est $s = 2$.

On sait que $\sqrt{n}\dfrac{\overline{X}-\mu}{s}$ suit une loi de Student à 2 degrés de liberté, donc si $\alpha = 0.05$, $t(\alpha) = 4.3027$.

Ainsi, la moyenne des masses du composant, mesurées sur un échantillon d'effectif 3 sera, dans 95% des cas, dans l'intervalle $[75 - 4.3027 \times s/\sqrt{3}, 75 + 4.3027 \times s/\sqrt{3}] = [70.03, 79.97]$.

La valeur moyenne 72 mesurée sur l'échantillon étant bien dans cet intervalle,

$\boxed{\text{on n'a pas de raisons, au vu de ces mesures, de rejeter } H_0.}$

21. Un laboratoire pharmaceutique désire étudier les effets secondaires potentiels d'un médicament sur le taux de cholestérol des patients. Cent volontaires sont donc choisis pour tester le médicament.

(a) Avant l'expérience, leur taux de cholestérol moyen est de $2.02\text{g.L}^{-1}$, avec un écart-type de $0.2\text{g.L}^{-1}$

Le taux de cholestérol moyen dans la population étant de $2\text{g.L}^{-1}$, vérifier que cet échantillon est représentatif au risque 5%.

(b) Après un mois de traitement, seuls 97 volontaires reviennent faire un test. Leur taux moyen de cholestérol est passé à $2.09\text{g.L}^{-1}$ avec un écart-type d'échantillon de $0.25\text{g.L}^{-1}$.

La différence est-elle significative au risque 5% ? Au risque 1% ?

corrigé succint :

(a) Soit $X_1$ la variable aléatoire qui mesure le taux de cholestérol d'un individu ; $E(X_1) = \mu_1 = 2$.
$\overline{X_1}$ est le taux moyen mesuré sur un échantillon de taille $n_1 = 100$.
Alors $n_1$ étant plus grand que 30, on peut considérer que $\sqrt{n_1}\dfrac{\overline{X_1}-2}{s_1}$ suit une loi normale, avec $s_1 = 0.2$ estimation ponctuelle de l'écart-type de $X_1$.
Ainsi, dans 95% des cas le taux moyen observé sur un échantillon sera compris dans $[2 - 1.96 \times 0.2/10, 2 + 1.96 \times 0.2/10] = [1.961, 2.039]$.
Le taux de cholestérol moyen des volontaires étant bien dans cet intervalle, on peut considérer que cet échantillon est représentatif.

(b) Soit $X_2$ la variable aléatoire mesurant le taux de cholestérol d'un individu après un mois de traitement ; son espérance $\mu_2$ est inconnue. $\overline{X_2}$ est le taux moyen d'un échantillon de taille $n_2 = 97$.

On fait l'hypothèse $H_0$ : « les taux de cholestérol moyens sont les mêmes avant et après traitement ». Alors $\mu_1 = \mu_2$, et on peut considérer que $\dfrac{\overline{X_1}-\overline{X_2}}{\sqrt{s_1^2/n_1 + s_2^2/n_2}} \sim N(0,1)$ (avec $s_1 = 0.2$, $s_2 = 0.25$), et par conséquent on détermine l'intervalle de confiance au risque 5% de $\overline{X_1} - \overline{X_2}$ : $[-1.96\sqrt{s_1^2/n_1 + s_2^2/n_2}, 1.96\sqrt{s_1^2/n_1 + s_2^2/n_2}] = [-0.063, 0.063]$.
Comme la différence entre les taux moyens mesurés $2.02 - 2.09 = 0.07$ n'est pas dans cet intervalle, elle est significative, et on rejette $H_0$ donc

| on considère, au risque 5% de se tromper, que le médicament a un effet. |
|---|

En revanche, l'intervalle de confiance au risque 1% est
$[-2.57\sqrt{s_1^2/n_1 + s_2^2/n_2}, 2.57\sqrt{s_1^2/n_1 + s_2^2/n_2}] = [-0.083, 0.083]$, intervalle qui contient la valeur $2.02 - 2.09 = 0.07$, donc

| la différence n'est pas significative au risque 1%. |
|---|

22. Lors d'un DS commun de probabilités-statistiques :

les 30 alternants ont obtenu une moyenne de 9.43 avec un écart-type 5.06

les 78 S3 non-alternants, une moyenne de 8.84 avec un écart-type de 4.34

On veut tester l'hypothèse selon laquelle les populations d'alternants et de non-alternants ont le même niveau.

Au vu de ces échantillons peut-on, au risque 1%, refuser cette hypothèse ?

corrigé succint :

Si l'hypothèse est vérifiée, on doit avoir dans 99% des cas une différence entre les moyennes des deux échantillons comprises dans l'intervalle $[-t\sqrt{s_1^2/n_1 + s_2^2/n_2}, +t\sqrt{s_1^2/n_1 + s_2^2/n_2}]$, $t$ vérifiant $2F(t) - 1 = 0.99$.

On lit $t = 2.57$ dans la table de la loi normale centrée réduite, l'intervalle est donc $[-2.69, +2.69]$ : la différence mesurée sur nos échantillons, $9.43 - 8.84$, étant dans cet intervalle, n'est pas significative : on ne rejette pas l'hypothèse.

23. Dans cet exercice, on considère uniquement des familles de 5 enfants.

1) On suppose, pour chaque naissance, que la probabilité d'avoir une fille est de 49%, et 51% pour un garçon.

Donner la loi de la variable $X$ correspondant au nombre de filles dans la famille.

2) Tester, par un test du $\chi^2$ au risque 5%, que l'échantillon ci-dessous suit la répartition théorique définie à la question précédente :

| $X$ | 0 | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|
| Effectif | 18 | 56 | 110 | 88 | 40 | 8 |

corrigé succint :

1) $X$ est une loi binomiale de paramètres $n$ et 0.49.

2) Pour une loi binomiale de paramètre 5 et 0.49 les effectifs théoriques pour un échantillon de taille 320 sont :

| $X$ | 0 | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|
| Effectif | 18 | 56 | 110 | 88 | 40 | 8 |
| Effectif théorique | 11.04 | 53.04 | 101.92 | 97.92 | 47.07 | 9.04 |

Le $\chi_o^2$ vaut ainsi 7.38.

Le $c^2$ pour une loi à $6 - 2 = 4$ degrés de liberté et un risque de 5% vaut 9.488 : on accepte donc l'hypothèse.

24. On relève chaque jour pendant 200 jours le nombre d'atterrissages entre 14h et 15h dans un aéroport :

| Nb d'atterrissages | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Nb de jours | 11 | 28 | 43 | 47 | 32 | 28 | 7 | 0 | 2 | 1 | 1 |

(a) Soit X la variable « nombre d'atterrissages par jour entre 14h et 15h ». Donner les estimations ponctuelles de $E(X)$ et $\mathrm{Var}(X)$ et estimer $E(X)$ par un intervalle de confiance 95%. Ces résultats sont-ils compatibles avec une loi de Poisson ? Quel serait son paramètre ?

(b) Tester la validité de ce modèle (test du $\chi^2$ au risque 5%).

(c) Calculez la probabilité d'avoir dans cet aéroport, toujours entre 14h et 15h : 0 atterrissage un jour donné, 1 ou 2 atterrissages un jour donné, 2 atterrissages en tout sur 3 jours quelconques.

corrigé succint :

(a) L'effectif de l'échantillon est $n = 200$.
On détermine l'estimation ponctuelle de la moyenne $\bar{x} = 600/200 = 3$, et l'estimation ponctuelle de la variance $s^2 = 596/199 = 2.995$, soit $s = 1.73$.
L'effectif est suffisant pour assimiler la loi de $\sqrt{n}\dfrac{\bar{X}-\mu}{s}$ a une loi normale centrée réduite : l'intervalle de confiance 95% pour la moyenne est donc $[\bar{x} - 1.96\frac{s}{\sqrt{200}}, \bar{x} + 1.96\frac{s}{\sqrt{200}}]$, soit $[2.76, 3.24]$.

(**hors-programme :** On sait que l'intervalle de confiance pour la variance est $[\frac{199}{c_1^2}s^2, \frac{199}{c_2^2}s^2]$, avec $c_1^2$ et $c_2^2$ lus dans la table de la loi du $\chi^2$ à 199 degrés de liberté. En pratique on utilise la table de la loi du $\chi^2$ à 200 degrés de liberté et on lit dans les colonnes $0.025 = 0.05/2$ et $0.975 = 1 - 0.05/2$ : $c_1^2 = 241.1$ et $c_2^2 = 162.7$. Ainsi l'intervalle est $[0.825s^2, 1.223s^2] = [2.47, 3.66]$.)

L'espérance et la variance ont (quasiment) la même estimation ponctuelle, égale à trois : les résultats sont compatibles avec le fait que $X$ suive une loi de Poisson de paramètre 3.

(b) On définit l'hypothèse $H_0$ : « $X$ suit une loi de Poisson de paramètre 3 ».

Pour utiliser un test du $\chi^2$ pour accepter ou refuser $H_0$, on doit avoir des valeurs (ou classes) d'effectif "théorique" au moins égal à 5. Or ici ce n'est pas le cas : les valeurs 7, 8, 9, 10 ont des effectifs $np(X = 7)$, $np(X = 8)$, $np(X = 9)$, $np(X = 10)$ inférieurs : on doit regrouper en une seule classe les valeurs 7, 8, 9, 10.

Les probabilités des évènements $X = 0$, $X = 1$, ..., $X = 6$, $X \in [7, 10]$ sont déterminées par lecture de la table de la loi de Poisson : $p(X = 0) = 0.0498$, $p(X = 1) = 0.1494$, $p(X = 2) = 0.224$, $p(X = 3) = 0.224$, $p(X = 4) = 0.168$, $p(X = 5) = 0.1008$, $p(X = 6) = 0.0504$, $p(X \geq 7) = 1 - p(X \leq 6) = 1 - 0.9665 = 0.0335$. On obtient en multipliant ces valeurs par l'effectif total 200 les effectifs « théoriques » de chaque classe :

| Nb d'atterrissages | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7+ |
|---|---|---|---|---|---|---|---|---|
| Effectif mesuré | 11 | 28 | 43 | 47 | 32 | 28 | 7 | 4 |
| Effectif théorique | 9.96 | 29.88 | 44.8 | 44.8 | 33.6 | 20.16 | 10.08 | 6.7 |

et en en déduit la valeur de $\chi_o^2 = \frac{(11-9.96)^2}{9.96} + \frac{(28-29.88)^2}{29.88} + \ldots + \frac{(4-6.7)^2}{6.7} = 5.56$.

Le critère de décision pour accepter $H_0$ sera $\chi_o^2 \leq c^2$, avec $c^2$ lu dans la table du $\chi^2$ à 6 degrés de liberté (on a 8 classes, et le test porte sur une loi de Poisson : le nombre de d.d.l.à considérer est donc 8-2). On lit dans la table, dans la colonne $\alpha = 0.05$ et la ligne 5 : $c^2 = 12.59$.

Ainsi, on accepte $H_0$.

(c) Si on admet, suite au test du $\chi^2$, que $X$ suit une loi de Poisson de paramètre 3, on lit directement dans la table les valeurs $p(X = 0) = 4.98\%$, $p(1 \leq X \leq 2) = p(X = 1) + p(X = 2) = 37.34\%$.

Si on s'intéresse aux nombre d'atterissages entre 14h et 15h sur trois jours distincts, si on note $X_1$, $X_2$ et $X_3$ les variables donnant le nombre d'atterissage chacun des trois jours, on sait que, $X_1$, $X_2$ et $X_3$ étant indépendantes, $X_1 + X_2 + X_3$ suit une loi de Poisson de paramètre 9. Par conséquent, $p(X_1 + X_2 + X_3 = 2) = 0.5\%$.

## régression linéaire, droites de tendance

25. **moindres carrés verticaux, horizontaux, distance à la droite**

On se donne une série de trois points $A(0; 1)$, $B(1; 2.5)$, $C(2; 2.8)$.

(a) déterminer la droite minimisant la somme des carrés des distances verticales entre les points $A$, $B$, $C$ et les points de mêmes abscisses de la droite

(b) déterminer la droite minimisant la somme des carrés des distances horizontales entre les points $A$, $B$, $C$ et les points de mêmes ordonnées de la droite

(c) (*) on veut maintenant déterminer la droite $\Delta$, d'équation $y = ax + b$, qui minimise la somme des carrés des distances des points $A$, $B$ et $C$ à la droite $\Delta$.

i. montrer que la somme des carrés des distances des points $A$, $B$ et $C$ à la droite $\Delta$ vaut $Q(a, b) = \frac{(1-b)^2 + (2.5-a-b)^2 + (2.8-2a-b)^2}{a^2+1}$.

ii. calculer la dérivée $\frac{\partial d}{\partial b}$ et en déduire une relation que doivent vérifier $a$ et $b$

iii. calculer la dérivée $\frac{\partial d}{\partial a}$, remplacer grâce à la question précédente $b$ en fonction de $a$ dans cette expression. En déduire que $3.6a^2 + 0.28a - 3.6 = 0$

iv. Conclure.

corrigé succint :

(a) on calcule successivement $\bar{x} = 1$, $\bar{y} = 2.1$, $\sigma_x^2 = 0.816^2 = 0.666..$, $\sigma_{x,y} = 0.6$, et donc $y = 0.9x + 1.2$

(b) en permutant les $x$ et les $y$ on trouve la droite d'équation $x = 0.968y - 1.03$. Donc en isolant $y$ : $y = 1.03x + 1.07$.

Les coefficients ne sont pas très éloignés, mais néanmoins différents, de ceux de la question précédente.

(c) i. La distance du point $M(x, y)$ à la droite d'équation $y = ax + b$, qui passe par le point $A(0, b)$ et de vecteur directeur $\vec{u}(1, a)$ vaut $\frac{||\vec{AM} \wedge \vec{u}||}{||\vec{u}||}$, donc la distance au carré vaut $\frac{(y-ax-b)^2}{a^2+1}$.

Si on additionne ces quantités pour les 3 points étudiés, on trouve donc $Q(a, b) = \frac{(1-b)^2 + (2.5-a-b)^2 + (2.8-2a-b)^2}{a^2+1}$

ii. La dérivée par rapport à $b$ de $Q$ est $2\frac{-(1-b) - (2.5-a-b) - (2.8-2a-b)}{a^2+1}$, elle s'annule donc si et seulement si $-(1+2.5+2.8) + (0+1+2)a + 3b = 0$, on en déduit donc que $\bar{y} = a\bar{x} + b$.

Ainsi, $b = 2.1 - a$

iii. La dérivée par rapport à $a$ de $Q$ est $\frac{(a^2+1)(-2(2.5-a-b) - 4(2.8-(2*a+b)))}{(a^2+1)^2} - \frac{2a((1-b)^2 + (2.5-(a+b))^2 + (2.8-(2*a+b))^2)}{(a^2+1)^2}$.

Il faut ensuite utiliser l'expression $b = 2.1 - a$ de la formule précédente, développer puis simplifier, pour arriver à la condition $3.6a^2 + 0.28a - 3.6 = 0$.

iv. on trouve comme solutions de cette équation du second degré $a = 0.962$ ou une valeur négative (qui n'a pas de pertinence ici).

Ainsi, $a = 0.962$ puis $b = 1.138$.

Cette équation est en quelque sorte "intermédiaire" entre les équations trouvées aux deux précédentes questions.

26. **altitude et température** [Salle informatique] le 22 août à 13h50, on a noté les températures des stations météo autour de Grenoble (source : www.infoclimat.fr )

| station | altitude (m) | température (°C) |
|---|---|---|
| Gresse | 1245 | 28,8 |
| Croix de Chamrousse | 2250 | 17,8 |
| Vizille | 280 | 35,6 |
| Sechilienne | 550 | 35,6 |
| Theys – Sept-Laux | 2130 | 20,2 |
| Theys | 760 | 32,3 |
| Theys | 850 | 32,7 |
| Pinsot | 750 | 32,7 |
| Pinsot | 850 | 30,9 |
| La Ruchère | 1150 | 30,0 |
| Saint-Laurent-du-Pont | 410 | 35,8 |
| Le Versoud | 220 | 33,4 |
| Saint-Martin-d'Heres | 220 | 36,9 |
| Engins | 905 | 32,6 |
| Lans les Allières | 1450 | 27,7 |
| Lans | 990 | 33,4 |
| Méaudre | 976 | 31,8 |
| Grenoble CEA | 210 | 37,2 |

Étudier la dépendance entre température et altitude : droite de tendance, coefficient de corrélation, en utilisant un tableur et en utilisant les formules.

corrigé succinct :

Excel comme un calcul direct donne une droite de tendance d'équation $y = -0.0084x + 39.008$, avec $a$ en degrés par mètre, et $b$ en degrés.

27. **tests de normalité**

On dispose de données sur les masses (en kilogramme) de 76 étudiantes et étudiants de BUT2 il y a quelques années :

| masse (kg) | effectif |
|---|---|
| 44 | 1 |
| 45 | 2 |
| 52 | 3 |
| 53 | 2 |
| 55 | 3 |
| 58 | 1 |
| 59 | 2 |
| 60 | 2 |
| 61 | 2 |
| 62 | 3 |

| masse (kg) | effectif |
|---|---|
| 63 | 5 |
| 64 | 3 |
| 65 | 4 |
| 66 | 4 |
| 67 | 1 |
| 68 | 2 |
| 70 | 5 |
| 71 | 1 |
| 72 | 5 |
| 73 | 1 |

| masse (kg) | effectif |
|---|---|
| 74 | 2 |
| 75 | 2 |
| 76 | 2 |
| 77 | 1 |
| 78 | 5 |
| 80 | 5 |
| 81 | 1 |
| 82 | 1 |
| 83 | 1 |
| 84 | 1 |
| 87 | 1 |
| 106 | 1 |

(a) Calculer la moyenne et l'écart-type de la série statistique

(b) Répartir les données en une dizaine de classes, et tracer l'histogramme correspondant. Que peut-on supposer de la loi de répartition des masses ?

(c) Pour chaque classe, de centre $x_i$ : calculer les fréquences **cumulées** $f_i$ correspondantes, puis déterminer la valeur $t_i$ telle que $F(t_i) = f_i$, où $F$ est la fonction de répartition de la loi normale centrée réduite.

Puis tracer le nuage de points $(x_i, t_i)$.

Si la variable aléatoire suit une loi normale de paramètres $\bar{x}$ et $\sigma^2$, on admet que les points sont alignés sur une droite $t = \dfrac{x - \bar{x}}{\sigma}$.

Déterminer la droite de tendance du nuages de points, et ses caractéristiques. Conclusion ?

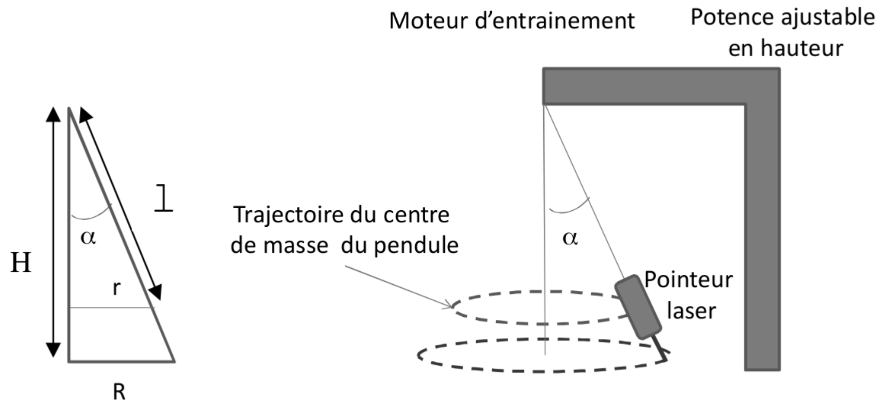
(d) Réaliser un test du $\chi^2$ de normalité de la série de données.

28. **estimation de $g$ et incertitudes** [Salle informatique]

Dans cette expérience (que vous avez réalisée l'an dernier en mécanique) on souhaite estimer la valeur de $g$ l'accélération de la pesanteur, et évaluer l'incertitude associée.

Pour cela on utilise un pendule composé d'une corde au bout de laquelle est attaché un pointeur laser.

Le pointeur tourne dans un plan horizontal. Un petit moteur permet d'éviter l'amortissement du mouvement du pendule et de garder une vitesse de rotation constante. La corde est attachée à l'axe du moteur, qui est fixé sur une potence de hauteur ajustable. Le pointeur laser permet de visualiser la trajectoire sur la table.




Représentation schématique du pendule

En régime permanent le pendule tourne autour d'un axe vertical et la corde fait un angle $\alpha$ constant avec cet axe. On note :

• $R$ le rayon de la trajectoire dessinée par le laser sur la table

• $H$ la distance entre la table et le point d'intersection entre l'axe de rotation du pendule et l'axe de la corde.

• $l$ la longueur du pendule, c'est-à-dire la distance entre le centre de masse du pointeur laser et le point d'intersection entre de l'axe du pendule et l'axe de la corde.

• $T$ la période de rotation du pendule.

• $y = \cos(\alpha) = \dfrac{H}{\sqrt{R^2 + H^2}}$.

• $\omega$ la vitesse angulaire,

• $x = \dfrac{1}{l\omega^2} = \dfrac{T^2}{4\pi^2 l}$.

On a réalisé une série de mesures, avec les valeurs expérimentales suivantes :

| $H$ (m) | $l$ (m) | $T$ (s) | $R$ (m) | $u(H)$ (m) | $u(l)$ (m) | $u(T)$ (s) | $u(R)$ (m) |
|---|---|---|---|---|---|---|---|
| 0.645 | 0.435 | 1.298 | 0.275 | 0.005 | 0.005 | 0.05 | 0.02 |
| 0.645 | 0.435 | 1.269 | 0.320 | 0.005 | 0.005 | 0.05 | 0.02 |
| 0.645 | 0.435 | 1.329 | 0.080 | 0.005 | 0.005 | 0.05 | 0.02 |
| 0.777 | 0.654 | 1.609 | 0.155 | 0.005 | 0.005 | 0.05 | 0.02 |
| 0.777 | 0.654 | 1.604 | 0.180 | 0.005 | 0.005 | 0.05 | 0.02 |
| 0.777 | 0.654 | 1.597 | 0.205 | 0.005 | 0.005 | 0.05 | 0.02 |

On admet (démontré en S2) que

$$y = gx$$

Par conséquent, on obtient une valeur expérimentale de $g$ en déterminant la valeur du coefficient de la droite de tendance (passant par l'origine) associée à ces valeurs $(x_i, y_i)$ correspondant aux mesures réalisées de $R$, $T$, $l$, $H$.

On demande alors à l'aide d'un tableur de :

• Tracer les $y_i$ en fonction des $x_i$. Comparer le résultat à la courbe théorique.

• En déduire une valeur expérimentale de $g$ l'accélération de la pesanteur. Quel est l'écart relatif avec la valeur exacte $9.81$m.s$^{-2}$ ?

• En utilisant la formule de propagation des incertitudes, exprimer $u(x)$ en fonction de $u(T)$ et $u(l)$.

• Montrer que l'on peut écrire assez simplement $\dfrac{u(x)}{x}$ en fonction de $\dfrac{u(T)}{T}$ et de $\dfrac{u(l)}{l}$.

• Quelle est la plus intéressante des deux formules ci-dessus pour calculer concrètement $u(x)$ avec un tableur ?

• En utilisant l'expression de $y(R, H)$ et la formule de propagation des incertitudes, donner une expression de $u(y)$.

• Calculer les dérivées partielles utiles puis les valeurs de $u(y)$ et enfin de $u(y)/y$ pour chaque série de mesures. Commenter.

• Calculer les dérivées partielles utiles puis les valeurs de $u(x)$ et enfin de $u(x)/x$ pour chaque série de mesures. Commenter.

• Le pente de la courbe de tendance linéaire est calculée par le tableur avec la formule $\dfrac{\sum_i x_i y_i}{\sum_i x_i^2}$. Vérifier que cette formule vous redonne la valeur déterminée précédemment.

• La formule de propagation des incertitudes appliquée à cette expression (fonction de $2n$ variables) montre que l'incertitude associée est alors

$$\frac{\sqrt{\sum_i y_i^2 u^2(x_i) + x_i^2 u^2(y_i)}}{\sum_i x_i^2}$$

.

Déterminer l'incertitude sur la valeur de $g$ associée à votre mesure.

Conclure.

corrigé succint :

On trouve $g = 9.65$m.s$^{-2}$, les u(x) sont de l'ordre de 0.007 usi et les u(y) de l'ordre de 0.01 usi. u(g) est autour de 0.279 m.s$^{-2}$.

29. **étalonnage d'un manomètre** [Salle informatique]

On considère un capteur de pression (manomètre) dans la gamme 0-110hPa, qui délivre une tension (que l'on lit) fonction de la pression.

On étalonne le manomètre avec les résultats ci-dessous (tension lues pour des pressions de référence). On suppose que l'incertitude sur ces pressions est négligeable.

| $x$ | $y$ |
|---|---|
| pression (hPa) | tension (V) |
| 10 | 0,261 |
| 40 | 1,008 |
| 70 | 1,758 |
| 90 | 2,257 |
| 110 | 2,759 |

On suppose que la tension varie linéairement avec la pression. Dans ce cas, on devrait avoir une relation de la forme $y = Ax$

(a) En utilisant un tableur, déterminer le coefficient de la droite de tendance (en forçant le passage par l'origine)

(b) Donner un estimateur de $A$ (noté $a$) au sens des moindres carrés.

Pour cela :

- écrire l'erreur quadratique en fonction de $A$
- déterminer la valeur $a$ pour laquelle la dérivée partielle par rapport à $A$ s'annule
- effectuer l'application numérique
- comparer au résultat donné par le tableur

(c) Le coefficient $a$ étant connu, donner l'erreur résiduelle pour chaque point et déduire ainsi une estimation de l'incertitude-type de mesure de la tension, notée $u_y$

(d) En déduire l'incertitude-type sur $a$ (expression puis valeur numérique)

(e) On met l'appareil en ligne dans une chaîne de mesure. On déduit donc la pression en fonction de la tension lue selon : $P_{\text{estimé}} = V_{\text{lue}}/a$.

La pression lue vaut 82,4294 hPa. Donner l'incertitude sur la pression

(f) On lit donc 82.4294 hPa. Critiquez cet affichage.

corrigé succint :

(a) On trouve $y = 0.0251x$ soit $V = 0.0251p$, l'unité du coefficient étant le V/hPa.

(b) L'erreur au niveau de chaque point est $e_i = y_i - ax_i$,
donc l'erreur quadratique totale est $Q(a) = \sum_i (y_i - ax_i)^2$.
$Q$ tend vers $+\infty$ quand $a$ tend vers $-\infty$ comme $+\infty$, donc elle admet un minimum.
En ce point la dérivée s'annule : $\sum_i -2x_i(y_i - ax_i) = 0$ donc $\sum_i x_i y_i - a\sum_i x_i^2 = 0$, et on retrouve (formule de cours) le fait que le meilleur estimateur de $a$ est $A = \dfrac{\sum_i x_i y_i}{\sum_i x_i^2}$.
L'application numérique donne $A = 0.025097...$ soit $A = 0.0251\text{V/hPa}$.
On retrouve la valeur calculée par Excel.

(c) Il n'y a pas d'incertitude sur les $x_i$ aussi (cf l'énoncé), et $a$ est constant dans tout le calcul.
Si on note l'erreur résiduelle $e = y - ax$, alors $u(e) = \sqrt{u(y)^2 + a^2u(x)^2} = u(y)$.
Donc : on calcule l'erreur résiduelle en prenant son écart-type, ici sans le diviser par $\sqrt{5}$ car on ne cherche pas l'incertitude sur une estimation par une moyenne, mais l'incertitude sur une éventuelle autre (sixième) mesure.
Ainsi, $u(y) = s(e) = 0.00492$ V.

(d) Comme $a = \dfrac{\sum_i x_i y_i}{\sum_i x_i^2}$, par propagation des incertitudes et en tenant compte du fait que les incertitudes sur $x_i$ sont nulles, $u(a)^2 = \sum_i \left(\dfrac{x_i}{\sum_i x_i^2}\right)^2 u(y_i)^2$.
En factorisant les $u(y_i) = u(y)$ constants : $u(a)^2 = u(y)^2 \sum_i \left(\dfrac{x_i}{\sum_i x_i^2}\right)^2 = u(y)^2 \dfrac{\sum_i x_i^2}{(\sum_i x_i^2)^2}$ donc au final, $u(a) = u(y)/\sqrt{\sum_i x_i^2}$.
Application numérique : $u(a) = 3.0 \times 10^{-5}$ et $u(a)/a = 0.0012 = 0.12\%$.

(e) Si $P = V/a$ alors $(u_p/p)^2 = (u(V)/V)^2 + (u(a)/a)^2$ donc $u(p)^2 = (1/a^2)u(V)^2 + (u(a)/a)^2p^2$,
application numérique $u(p) = 0.22$, donc on peut ne conserver que $p = 82.43 \pm 0.22\text{hPa}$.