Document d'accompagnement thématique

MINISTÈRE DE L'AGRICULTURE ET DE L'ALIMENTATION
Liberté Égalité Fraternité

ENSEIGNEMENT AGRICOLE L'AVENTURE DU VIVANT LES MÉTIERS GRANDEUR NATURE

Inspection de l'Enseignement Agricole

Diplôme: BTSA Anabiotec

Thème : Exemples d'utilisation des mathématiques dans des situations favorisant l'acquisition de capacités.

Commentaires, recommandations pédagogiques.

L'enseignement des mathématiques doit contribuer, notamment en lien avec les disciplines professionnelles, à l'acquisition des capacités :

C4.1 – Réaliser des analyses ou des essais dans le secteur de la santé

C4.2 – Réaliser des analyses ou des essais dans le secteur agro-alimentaire

C4.3 – Réaliser des analyses ou des essais dans les secteurs agricoles et de l'environnement

C4.4 – Piloter un procédé biotechnologique

C7.3 – Choisir les analyses et contrôles adaptés aux objectifs fixés

C8.2 – Valider des résultats.

L'enseignant veillera à s'appuyer sur les acquis des apprenants pour développer de nouveaux outils mathématiques principalement dans le but de répondre à des problématiques professionnelles. La mobilisation de ces outils dans le cadre de la résolution de problèmes concourt à l'obtention des capacités professionnelles susvisées. Cela donne du sens, puis montre l'importance de mobiliser de nouveaux outils mathématiques au service de l'acquisition des capacités professionnelles.

L'enseignement des mathématiques est intégratif et en lien avec ce qui est fait dans les disciplines professionnelles est un appui qui permet d'ancrer durablement les apprentissages. Les contextes doivent varier en fonction des situations techniques et provenir de documents issus de sources multiples : l'ANSES, l'OMS, la HAS, le CNRS, l'INRA, l'IFREMER, l'INC, les normes ISO, l'INSEE, AGRESTE, données issues de l'exploitation ou de l'atelier technologique de l'établissement, documentations, résultats issus de projets (dossier expérimental du bloc 8 de compétences, le stage en milieu professionnel, les projets éventuels avec des partenaires privés ou publics…).

La progression construite par le professeur de mathématiques devra être en lien direct avec celle proposée par les collègues de disciplines professionnelles.

La résolution de problèmes demande de mobiliser des techniques calculatoires. Les calculs, pour une grande partie, peuvent être délégués à un outil de calcul numérique (calculatrice, tableur, logiciel de calcul, …). Il ne s'agit pas ici de développer une virtuosité technique mais plutôt de se positionner comme observateur et de se questionner sur les processus mis en œuvre dans le domaine professionnel. La recherche de réponse amènera naturellement à élaborer des démarches, mener des calculs à l'aide d'un outil adapté, s'assurer de la cohérence de résultats et prendre des décisions.

L'institutionnalisation des notions, phase indispensable dans le processus d'apprentissage, a pour but d'expliciter les savoirs et les savoir-faire, de donner des repères simples aux apprenants. Ce temps doit être court et synthétique. Les développements théoriques sont réduits à l'essentiel et toujours présentés dans un cadre simple.

Des mathématiques transversales à tous les blocs de compétences.

L'acquisition des capacités professionnelles demande d'aborder de nouvelles notions qui s'appuient de façon implicite sur des connaissances mathématiques vues dans les classes antérieures du collège et du lycée. Certaines difficultés d'apprentissage de ces nouveaux concepts proviennent d'un manque de maîtrise de ces prérequis. Il est indispensable d'y consacrer régulièrement du temps afin de réactiver et consolider ces savoirs sans entrer dans un schéma de révision. Le choix de réinvestir les notions transversales suivantes sera décidé en fonction de la progression choisie:

- Proportion, pourcentage, proportionnalité, rendement, concentration, densité, ...
- Sens des opérations, application de formule, résolution d'une équation à une inconnue ou deux équations à deux inconnues, changement d'échelle ou d'unité, représentation graphique de fonctions et exploitation graphique.
- Fonctions affines, puissances, logarithmes népérien et décimal, exponentielles. Calcul de pente, de nombre dérivé, de vitesses moyenne et instantanée, d'aires.
- Représentations de diagrammes statistiques pertinents, interprétation et utilisation d'indicateurs statistiques.
- Probabilités élémentaires, lien entre fréquences et probabilités, arbres de probabilités.

Afin que les apprenants soient aguerris aux pratiques calculatoires élémentaires favorisant l'acquisition des capacités, des automatismes mathématiques doivent être développés par un travail régulier, afin d'obtenir une aisance suffisante. La pratique de l'ensemble de ces items doit être très régulière, principalement sur des situations en lien avec les disciplines professionnelles.

Au-delà d'une pratique dans toutes les activités de la classe, il est aussi important d'entretenir ces automatismes par des rituels de début de séance, très régulièrement sur l'ensemble des deux années, sous forme de « questions flash » privilégiant l'activité mentale avec un recours à des connaissances, des procédures, des méthodes et des stratégies fondamentales dans la pratique professionnelle. Cela ne doit pas faire l'objet d'un chapitre d'enseignement spécifique car les notions qui les sous-tendent ont été travaillées dans les classes antérieures. Cette pratique, propre à chaque enseignant, doit s'adapter aux besoins du diplôme.

Les exemples ci-dessous ne sont pas exhaustifs mais donnent une orientation de ce qui peut être fait.

Parmi eux, certains doivent être propices au calcul mental.

- Sens des opérations qui permet d'effectuer des calculs courants.
- Calculer une moyenne, une moyenne pondérée.
- Passer d'une proportion (1/2, 3/4, 1/5, …) à un pourcentage (50%, 75%, 20%, …) et inversement.
- Calcul de pourcentages, calcul de prix TTC à partir d'un prix HT et inversement, avec des taux de TVA différents.
- Lier augmentation et diminution en pourcentage avec coefficient multiplicateur et les utiliser en situation.
- Comparer en situation des proportions et des pourcentages.
- Application de formules et détermination de la valeur numérique d'une grandeur connaissant les autres.
- Calculs géométriques élémentaires s'appuyant sur les objets géométriques élémentaires : rectangle, carré, triangle, cube, pavé, cylindre.
- Conversions de mesures et capacités usuelles (cm^3 en L, ha en km^2,…)
- Reconnaitre graphiquement des fonctions de référence, en décrire les variations et les extremums.

- Choisir une représentation graphique adaptée pour représenter des données, des proportions ou des pourcentages (graphique, diagramme circulaire, semi-circulaire, diagramme en bâton ou en barres, barres empilées,…).
- Inversement, interpréter des diagrammes et retrouver des données statistiques à partir de représentations.

Les outils numériques doivent être intégrés à l'enseignement des mathématiques. Ils apportent une plus-value permettant d'aborder de véritables problèmes issus des disciplines professionnelles. L'usage des outils numériques tels que le tableur, les logiciels de traitement de données statistiques, de sondage, de cartographie, … doit être pensé dans l'optique de résoudre des problèmes qui n'auraient pas été accessibles sans. La maîtrise de ces outils numériques n'est pas un but de l'enseignement des mathématiques. La calculatrice reste aussi un outil facilement mobilisable en classe. Cela n'est pas contradictoire avec une pratique du calcul mental régulière mais raisonnée, tant par la difficulté des questions posées que le contexte de sa pratique.

Intentions majeures du référentiel de mathématiques.

Le cours de mathématiques intervient essentiellement dans le cadre de trois blocs de compétences :
-le bloc 7 qui organise les contrôles et les analyses. L'apprenant doit être capable de faire des choix pertinents et adaptés dans un but de contrôle et d'analyse. Il mène, entre autres, une réflexion quant au choix des variables, d'un échantillonnage et d'un test d'hypothèses adéquats en amont de la mise en place des analyses ;
-le bloc 4 qui met en œuvre les analyses, essais et procédés. L'apprenant doit être capable d'obtenir de façon autonome des résultats valides qui découlent des choix opérés en amont. Ces résultats doivent permettre dans la majorité des cas une prise de décision et, dans tous les cas, servir de support ultérieur à une communication vers l'extérieur;
-le bloc 8 qui valorise les résultats d'activités. L'apprenant doit être capable de valider les résultats d'analyses précédemment effectuées et d'en faire une interprétation à visée professionnelle ou grand public suivant les interlocuteurs rencontrés.

Le cours de mathématiques forme un ensemble. Il n'est donc pas possible de le décomposer clairement capacité par capacité ou de proposer un découpage chronologique unique qui suivrait, par exemple, la numérotation des capacités. Ce référentiel est conçu pour que les connaissances mises en œuvre soient utilisées pendant l'ensemble de la formation, à l'intérieur de la discipline mais surtout au service d'une réflexion globale en équipe pédagogique et, le plus souvent possible, dans un contexte professionnel. On développe donc l'aptitude des apprenants à mobiliser des compétences de mathématiques en situation professionnelle. En revanche, ce document précise les domaines en mathématiques sur lesquels peuvent porter l'évaluation certificative. Cette évaluation ne couvre pas forcément de façon exhaustive la totalité des items cités et doit plutôt répondre globalement à la capacité visée. Enfin, les points abordés ont vocation à être mobilisés dans chaque sous-capacité du bloc de compétence où ils sont nécessaires.

C4.1, C4.2, C4.3 et C4.4 Mettre en œuvre des analyses, des essais et des procédés biotechnologiques -

On attend de l'apprenant qu'il soit capable de donner des résultats valides ainsi que de vérifier la conformité des valeurs obtenues dans le but de répondre de manière fiable au commanditaire. Les outils numériques sont mobilisés à cette fin.

Mobilisation de l'outil mathématique

L'enseignement est à répartir dans les différentes capacités C4.1, C4.2, C4.3 et C4.4 sur l'ensemble de la formation en fonction de la progression pédagogique choisie par l'équipe pédagogique et en soutien des enseignements techniques.

- Mise en place d'un rituel (ou automatismes).
- Utilisation des notions de statistique en vue d'une modélisation à priori.
- Utilisation de notions de statistique et de probabilité en vue d'une estimation.
- Mise en œuvre d'un test statistique en vue d'une prise de décision.

- **Utilisation des notions de statistique en vue d'une modélisation à priori.**

Réalisation d'une modélisation simple en utilisant un ajustement.

L'apprenant doit être capable de réaliser une modélisation simple à l'aide d'un ajustement affine nécessitant éventuellement un ou des changements de variables pour obtenir un ajustement utilisant des fonctions logarithmes, exponentielles ou puissances.

Les situations étudiées sont variées et essentiellement issues du domaine professionnel. Elles doivent être réfléchies au sein de l'équipe pédagogique (courbe d'étalonnage, dosage de solutions en chimie ou en biochimie, croissance de populations microbiennes, régression pour des résultats d'électrophorèse…). Le cours de mathématiques est l'occasion d'une première approche qui pourra être exploitée et complétée ultérieurement dans d'autres enseignements.

La phase de compréhension réalisée, qui doit s'appuyer sur une approche graphique, les calculs en situation sont exclusivement effectués à l'aide d'un outil numérique (calculatrice, courbe de tendance du tableur, logiciel R, applications dédiées).

Aucune technicité inutile, détaillant les calculs pas à pas, pour obtenir un ajustement ne seront à maîtriser par l'apprenant. Le résultat obtenu avec l'outil numérique et l'interprétation faite des résultats sont à privilégier.

La qualité de l'ajustement obtenu devra être questionnée en distinguant la variable explicative et la variable expliquée et en étudiant les résidus. Les coefficients de corrélation linéaire et de détermination apportent des informations pour l'interprétation des résultats, pour autant, on soulignera les limites éventuelles de ces indicateurs.

Corrélation et causalité.

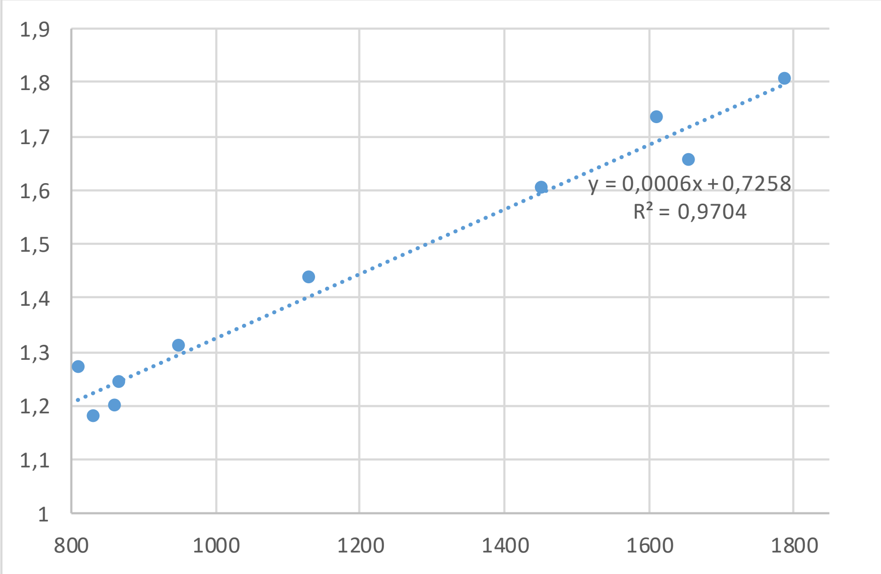
Au-delà de la mise en évidence d'une corrélation, le fait que, sur un relevé statistique, il puisse exister une relation entre deux grandeurs (corrélation), ne signifie pas pour autant qu'il existe entre elles un lien de causalité.

Exemple :

Il a été relevé sur plusieurs années, aux USA, le nombre de nouveaux doctorats en informatique et les revenus générés par les jeux d'arcades.

Année	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
Doctorats décernés en science informatique aux USA	861	830	809	867	948	1129	1453	1656	1787	1611
Revenus générés par les jeux d'arcade	1,196	1,176	1,269	1,24	1,307	1,435	1,601	1,654	1,803	1,734

La corrélation est évidente, la causalité beaucoup moins !



<u>Repérer des résultats aberrants.</u>
La détermination d'un ajustement affine ou la comparaison de méthodes fournissent des situations intéressantes où la recherche de résultats aberrants peut s'avérer utile. L'apprenant doit être capable de construire une gamme étalon cohérente, ce qui peut conduire à enlever une mesure dans le cadre d'un étalonnage. On met en place un test (Grubbs, Dixon, ...) selon le contexte pour décider si une valeur est aberrante.
Un document de mai 2021, « Mesure et incertitudes au lycée », disponible sur Eduscol, peut servir de ressource sur la question de la mesure et des incertitudes. Bien que destiné initialement à l'enseignement de la physique-chimie pour les classes de lycée, il apporte des axes de réflexion pertinents pour l'enseignant de mathématiques.
https://eduscol.education.fr/document/7067/download

<u>Lois de probabilité.</u>
Les notions de variable aléatoire, d'espérance mathématique, de variance, d'écart type et l'interprétation de ces paramètres pour une loi discrète ou une loi continue sont abordées en situation avant d'être formalisées.
La loi binomiale et la loi normale sont étudiées principalement. Un échange avec les collègues qui interviennent dans les enseignements professionnels est nécessaire afin de contextualiser l'étude des lois (pics de chromatographie, dénombrements microbiens exprimés en UFC, situation d'enzymologie…).
D'autres lois en liaison avec le domaine professionnel peuvent être introduites si elles répondent à un besoin spécifique, comme la loi de Poisson dans le cas des évènements rares, par exemple lors du dénombrement microbien en milieu liquide par la méthode du nombre le plus probable (NPP).
Les lois de probabilité peuvent être aussi l'occasion de procéder à une différenciation des apprentissages suivant le bagage et la poursuites d'études envisagée de chaque apprenant en liant l'étude de variables aléatoires continues au calcul de probabilité, calcul intégral et calcul d'aire en privilégiant le contexte professionnel en chromatographie ou en enzymologie par exemple.
Les changements de variables pour revenir à un calcul de probabilité sur la loi normale centrée réduite, l'approximation d'une loi binomiale par une loi normale et la correction de continuité qui en résulte ou la lecture directe ou inverse d'une table de la loi normale centrée réduite pourraient être abordées et expliquées mais les outils numériques actuels doivent permettre ensuite la détermination d'une valeur exacte rapidement obtenus dans le contexte de la situation choisie.
La phase de compréhension réalisée, les calculs en situation sont exclusivement effectués à l'aide d'un outil numérique (calculatrice, tableur, logiciel R, applications dédiées). On insistera alors sur l'interprétation des résultats. Le traitement de simulations avec des outils numériques pour illustrer les notions précédentes devra être recherché dès que cela est possible.

<u>Test d'indépendance du χ^2.</u>
L'apprenant doit être capable de mettre en place un test d'indépendance de deux caractères qualitatifs, le test d'indépendance du χ^2.

La mise en place de ce test est aussi l'occasion d'étudier quelques propriétés d'un couple de variables aléatoires discrètes afin de reconnaître une situation de dépendance ou d'indépendance (somme de deux variables, espérance, et variance dans le cas où les variables sont indépendantes). Il est essentiel de développer d'abord la compréhension de la situation à partir de simulations. L'étude porte, si possible, sur des données issues d'un travail collaboratif avec l'équipe pédagogique intervenant dans les enseignements professionnels. On s'assure que les hypothèses permettant l'application du test sont réalisées.

Exemple: Introduire le test d'indépendance du χ^2.

Vous êtes technicien(ne) dans un laboratoire d'analyse des sols et vous participez à l'optimisation d'un couvert végétal pour enrichir le sol en azote et limiter le lessivage des nitrates. Vous mesurez le dosage des nitrates pas chromatographie ionique dans les nappes phréatiques. Une étude a été menée sur 144 parcelles pour évaluer l'influence du couvert végétal (moutarde, phacélie, raygrass) sur les conséquences environnementales. L'utilisation du logiciel R ou d'un tableur en installant un utilitaire d'analyse facilite grandement les calculs. L'utilisation d'une calculatrice reste possible.

type de couvert végétal / conséquence environnementale	moutarde	phacélie	raygrass	Total
Faible	8	15	19	42
Moyenne	33	17	5	55
Forte	15	31	1	47
Total	56	63	25	144

Le tableau ci-dessus peut être saisi dans R via l'instruction

```
>tableau=data.frame(moutarde=c(8,33,15),phacélie=c(15,17,31),raygrass=c(19,5,1),
        row.names=c('conséquence environnementale faible','conséquence environnementale moyenne','conséquence environnementale forte'))
```

Dans le cas où les variables seraient indépendantes, on s'attend à obtenir le tableau des effectifs suivants.

type de couvert végétal / conséquence environnementale	moutarde	phacélie	raygrass	Total
Faible	16.33333	18.3750	7.291667	42
Moyenne	21.38889	24.0625	9.548611	55
Forte	18.27778	20.5625	8.159722	47
Total	56	63	25	144

Celui-ci s'obtient en récupérant le résultat du test dans une variable que l'on va nommer khi puis l'affichage des effectifs attendus.

```
>khi=chisq.test(tableau)
>khi$expected
```

Dans le cas de l'indépendance « parfaite », le tableau des effectifs attendus devrait être identique au tableau des effectifs observés. La mesure de l'indépendance amène donc à mesurer l'écart existant entre les deux tableaux via la formule :

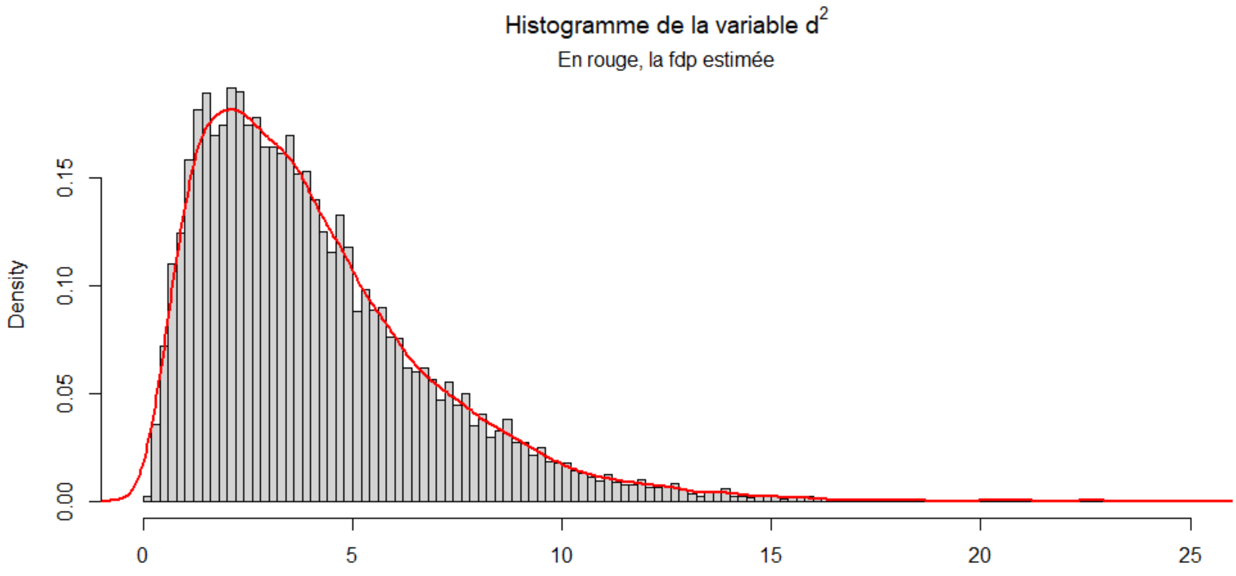
$$d^2 = \sum_{i,j} \frac{\left(O_{i,j} - E_{i,j}\right)^2}{E_{i,j}}$$

où les $O_{i,j}$ correspondent aux effectifs observés et les $E_{i,j}$ aux effectifs attendus sous l'hypothèse d'indépendance.

Pour obtenir une idée de la distribution de la variable d^2, on simule des tableaux dont les lignes et les colonnes sont indépendantes et ayant les mêmes marges que le tableau initial. Ce type de simulation étant difficile à mettre en œuvre, on pourra plutôt pour faire émerger la loi du χ^2 s'intéresser à l'adéquation à une loi. (cf annexe1). Pour les plus curieux, on pourra consulter l'algorithme RCont dû à Boyett[1] sur la génération des tableaux de contingence de marges en ligne et en colonne fixées.

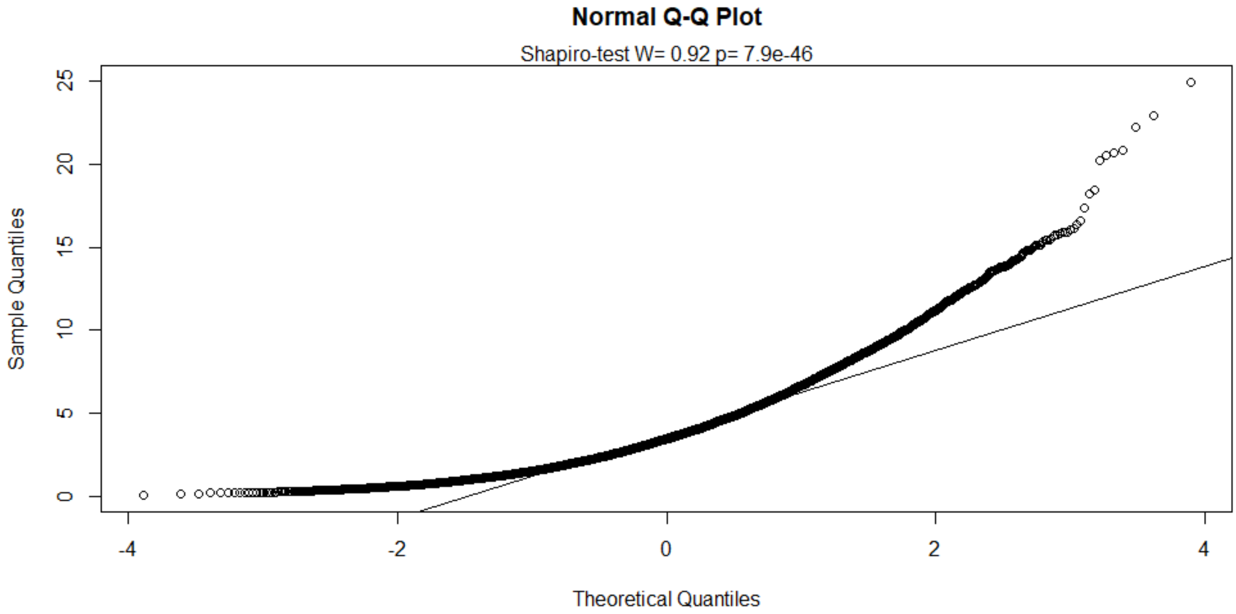
Les lignes de commande ci-dessous permettent d'obtenir l'histogramme où la variable informatique d contient la valeur observée de d^2 pour 5000 simulations de tableaux, ainsi que la fonction de densité de probabilité estimée.

```
>hist(d,breaks=100,prob=T)
>lines(density(d),col='red',lwd=2)
```

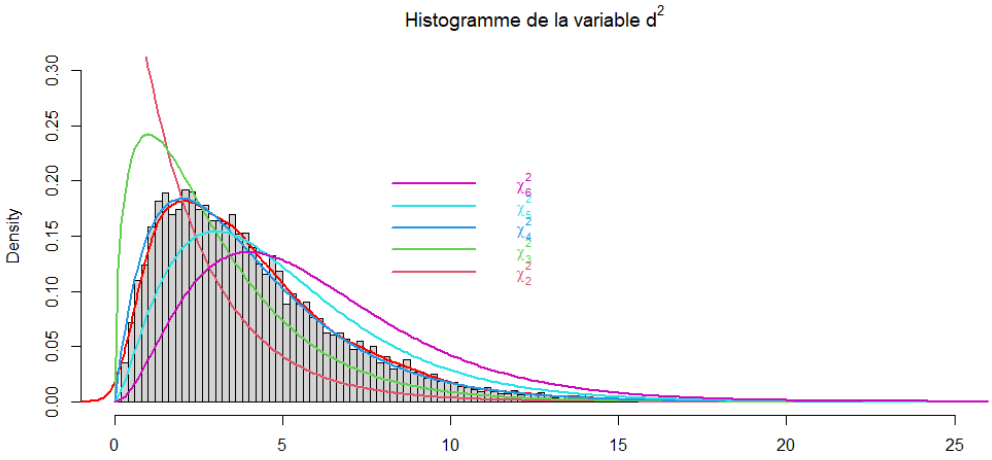


On peut alors tester la normalité de cette distribution. Avec les commande qqnorm(d), qqline(d) et shapiro.test(d)

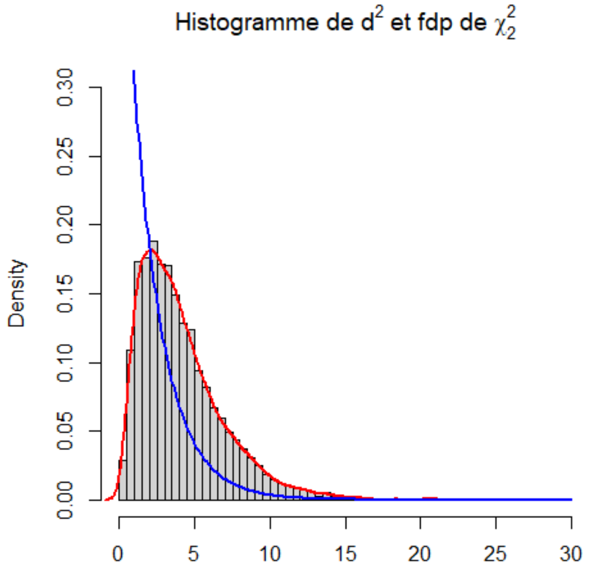
[1] James M. Boyett *Journal of the Royal Statistical Society. Series C (Applied Statistics)* Vol. 28, No. 3 (1979), pp. 329-332



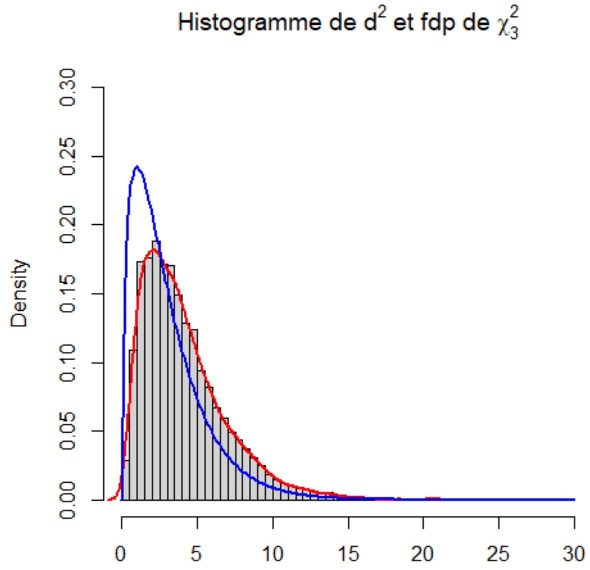
On rejette donc la normalité, ce qui amène à chercher une loi approchant d^2.
La forme de l'histogramme incite à essayer une loi du Khi2 avec un certain degré de liberté.



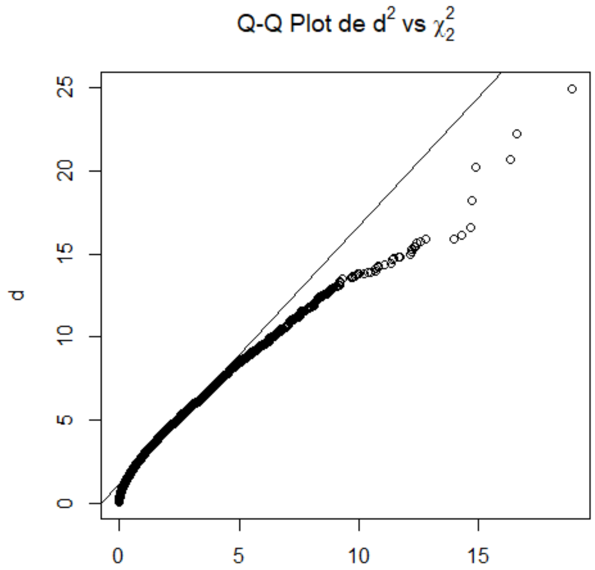
Etudions alors l'adéquation aux différentes lois du χ^2, cas par cas.



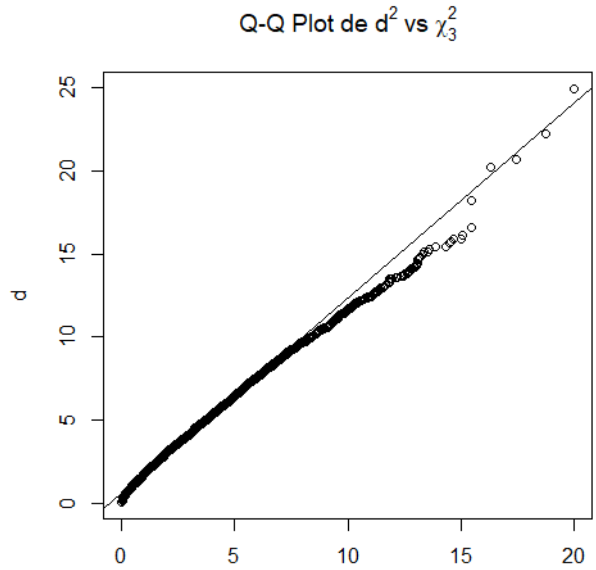
Histogramme de d^2 et fdp de χ_2^2
Density
0.00 0.05 0.10 0.15 0.20 0.25 0.30
0 5 10 15 20 25 30



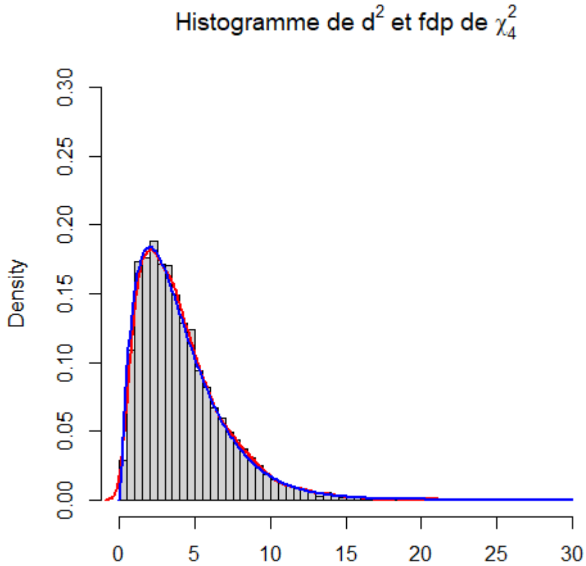
Histogramme de d^2 et fdp de χ_3^2
Density
0.00 0.05 0.10 0.15 0.20 0.25 0.30
0 5 10 15 20 25 30

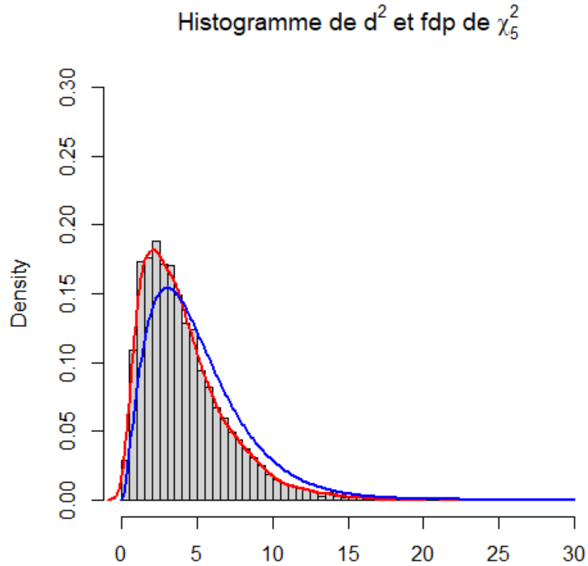


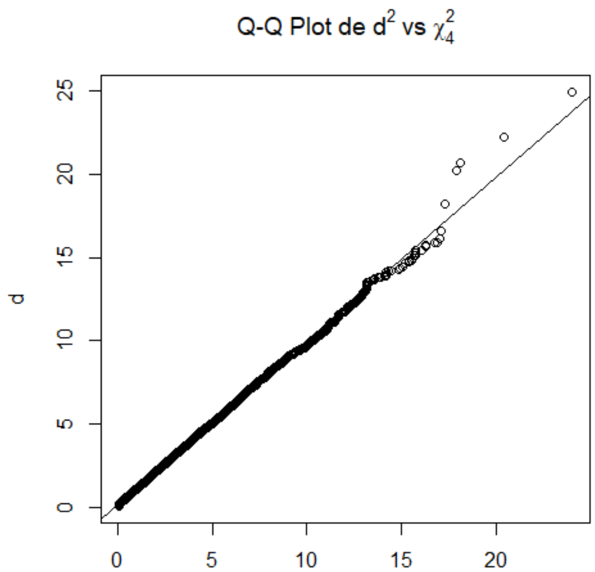
Q-Q Plot de d^2 vs χ_2^2
d
0 5 10 15 20 25
0 5 10 15

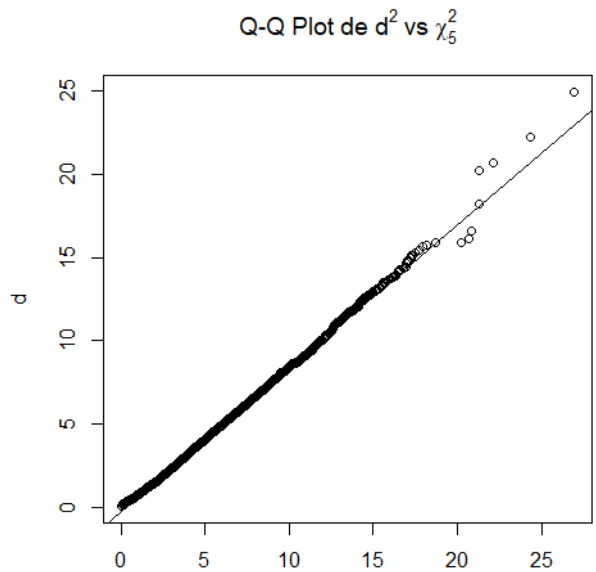


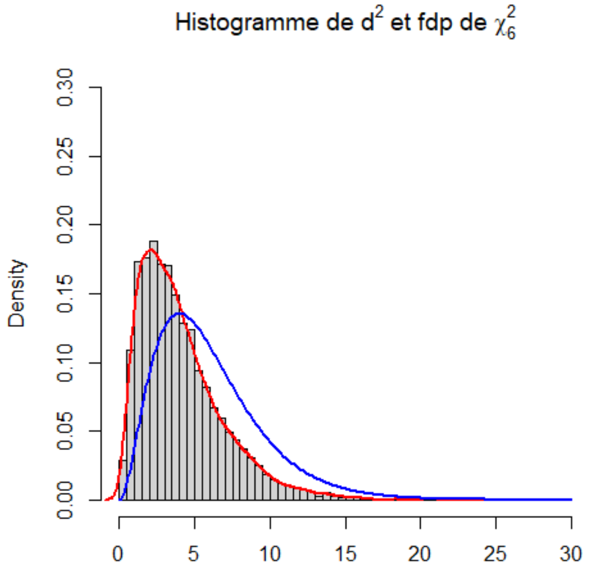
Q-Q Plot de d^2 vs χ_3^2
d
0 5 10 15 20 25
0 5 10 15 20

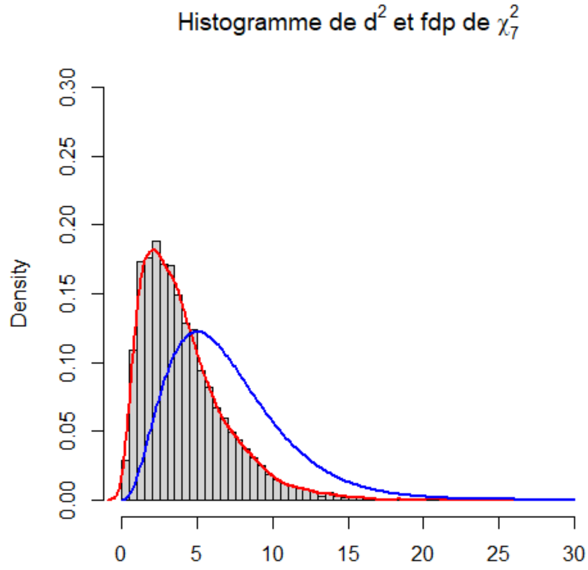

Histogramme de d^2 et fdp de χ^2_4
Density
0.00 0.05 0.10 0.15 0.20 0.25 0.30
0 5 10 15 20 25 30

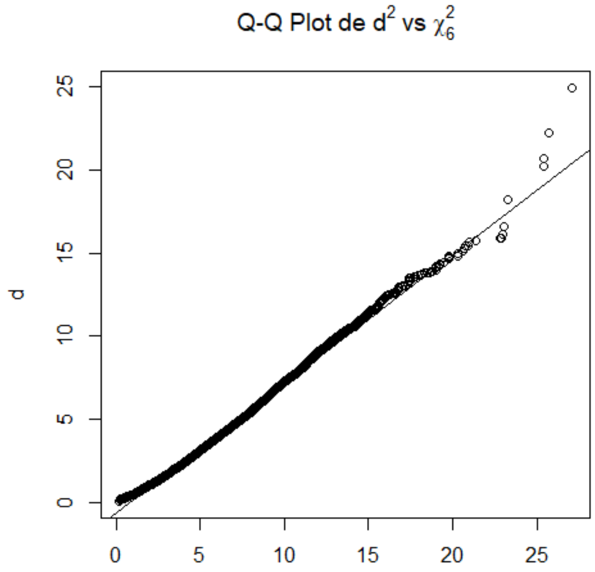

Histogramme de d^2 et fdp de χ^2_5
Density
0.00 0.05 0.10 0.15 0.20 0.25 0.30
0 5 10 15 20 25 30

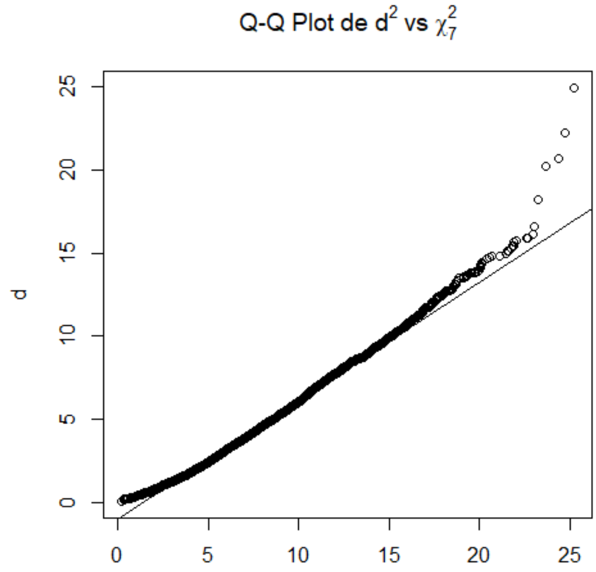

Q-Q Plot de d^2 vs χ^2_4
d
0 5 10 15 20 25
0 5 10 15 20


Q-Q Plot de d^2 vs χ^2_5
d
0 5 10 15 20 25
0 5 10 15 20 25



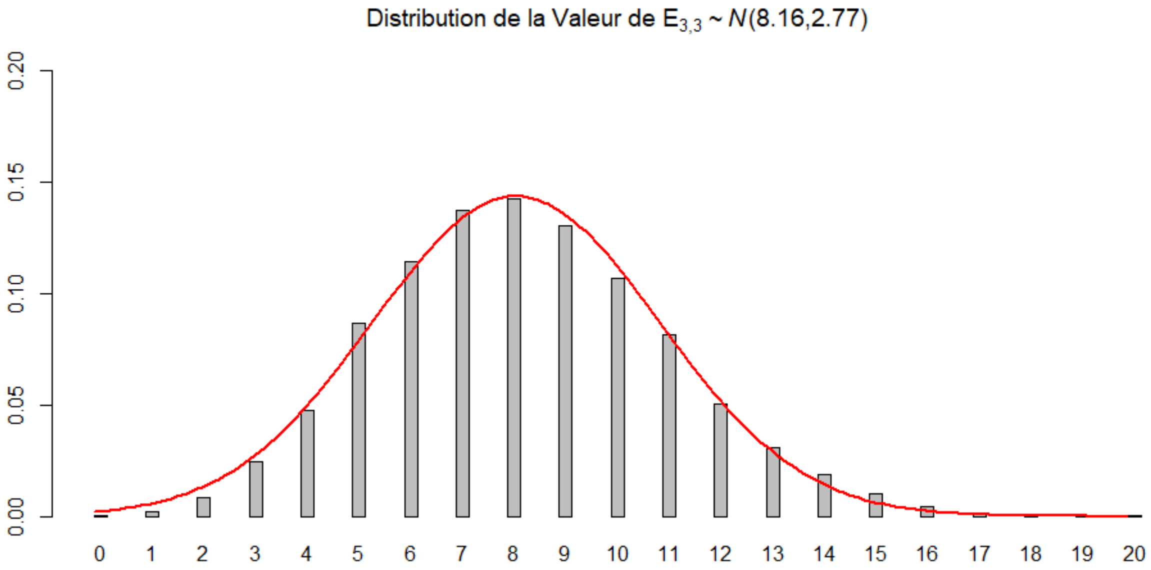






On peut alors émettre la conjecture que d^2 semble suivre une loi du χ^2 à 4 ou 5 degrés de liberté. Cette approche par simulation sera complétée par la mise en place du test du χ^2 puis l'institutionnalisation de la loi asymptotique de d^2.

Dans le cas du rejet de l'indépendance, l'étude devra se poursuivre par l'analyse des écarts à l'indépendance. On pourra à cet effet s'intéresser à une unique cellule du tableau. Par exemple, choisissons la cellule correspondant à une forte conséquence environnementale et au raygrass. Sous l'hypothèse d'indépendance, la valeur attendue est $E_{3,3} \approx 8{,}16$ et la probabilité attendue est environ égale à 0,057. On peut alors par simulation étudier la distribution des valeurs de cette cellule et en déduire « l'écart » à la valeur attendue.



La valeur observée est 1, qui sous les conditions de l'indépendance, a peu de chance d'être observée, de probabilité inférieure à 0,05.

D'après le théorème central limite, la distribution des valeurs de cette cellule suit approximativement une loi normale de moyenne $E_{3,3} \approx 144 \times 0{,}057 \approx 8{,}16$ et d'écart type $\sqrt{144 \times 0{,}057(1-0{,}057)} \approx \sqrt{E_{3,3} \times (1-0{,}057)}$.

On peut alors considérer les résidus de Pearson standardisées égaux à $\frac{O_{i,j}-E_{i,i}}{\sqrt{E_{i,j}}}$ que l'on retrouve dans l'expression de la variable statistique d^2. Les résidus de Pearson standardisées vont suivre approximativement une loi normale centrée et d'écart type inférieur à 1. Ils mesurent l'attraction (positif) ou la répulsion (négatif) par rapport à la valeur attendue. Les résidus, correspondant à des écarts statistiquement significatifs et dont la contribution à d^2est forte, sont ceux dont la valeur est supérieure à 2 ou inférieure à -2, qui correspondent à au moins deux écart type dans la loi normale centrée réduite.

Le tableau des résidus de Pearson s'obtient sous R avec la commande >khi$residuals où khi est la variable informatique contenant le tableau des effectifs.

type de couvert végétal / conséquence environnementale	moutarde	phacélie	raygrass
Faible	-2.0619652	-0.787336	4.335924
Moyenne	2.5106124	-1.439753	-1.472003
Forte	-0.7666865	2.301752	-2.506447

Le rejet de l'indépendance peut s'expliquer, par exemple, par une surreprésentation de la valeur observée de raygrass et faible conséquence environnementale et aussi par la sous-représentation de la valeur observée de raygrass et forte conséquence environnementale par rapport aux valeurs attendues.

<u>Remarque :</u>
Les lignes suivantes avec R permettent d'obtenir la légende du graphique avec les différentes lois du χ^2 :

```
graphics.off()
library(latex2exp)#écrire en latex dans les titres de graphiques
x=seq(0,30,0.1)
hist(d,breaks=100,prob=T,main=TeX(r'(Histogramme de la variable $d^2$)'),xlab='',ylim=c(0,0.3))
for (i in 2:6){
  lines(x,dchisq(x,i),col=i,lwd=2)
  legend(x=7,y=0.2+0.025*i,col=i,text.col=i,legend=(TeX(sprintf(r'($\chi_{%i}^{2}$)',i))),
         cex=1,bty='n',lwd=2)
}
```

- **Utilisation de notions de statistique et de probabilité en vue d'une estimation.**

Distribution d'échantillonnage.

Pour être en mesure de proposer une estimation d'un paramètre, l'apprenant doit appréhender la notion de distribution d'échantillonnage.
Après avoir manipuler la notion d'échantillon aléatoire simple, on observe la distribution d'échantillonnage d'un paramètre d'une population (moyenne, variance, proportion) à l'aide des outils numériques avant de définir les variables aléatoires $\overline{X}$, S^2et F associées.
Dans le cas de grands échantillons, les lois de $\overline{X}$ et F sont approchées par des lois normales.

Estimation ponctuelle ou par intervalle de confiance.

Dans le cadre, si possible, d'une situation professionnelle, l'apprenant doit être capable de donner une estimation ponctuelle ou par intervalle de confiance symétrique en probabilité à un niveau de confiance de 0,95 ou 0,99 d'une moyenne, d'un écart-type ou d'une proportion.
L'approche expérimentale de cette estimation peut avantageusement faire l'objet d'une simulation qui souligne l'importance de la taille de l'échantillon et de la variabilité du phénomène étudié.
Le cas de l'estimation par un intervalle de confiance d'une moyenne d'un caractère distribué selon une loi normale de variance inconnue donne l'occasion d'introduire et d'étudier la loi de Student. L'intervalle de confiance d'une variance, qui n'est pas symétrique, est construit à partir de la variable aléatoire $\frac{nS^2}{\sigma^2}$.

- **Mise en œuvre d'un test statistique en vue d'une prise de décision.**

Dans ce module C4, il s'agit uniquement de récolter les résultats d'analyse après avoir choisi celle qui convient dans le cadre de la capacité C7.3. La justification et la démarche du test statistique ont donc déjà été explicitées. Son interprétation dans le contexte de la situation professionnelle sera l'objet de la capacité C8.2.
On attend de l'apprenant, conformément à l'intitulé du bloc, qu'il accomplisse les mesures nécessaires à la réalisation du test statistique choisi en amont afin qu'il puisse prendre une décision ultérieurement une fois les mesures effectuées.

- **Mise en place d'automatismes (ou de rituels).**

En complément de ces études en contexte, la pratique d'automatismes vise à construire et entretenir des aptitudes dans le domaine mathématique. L'ensemble des automatismes doit être pratiqué quelles que soient les thématiques travaillées. On pourra rencontrer les situations suivantes, sans caractère exhaustif ni obligatoire :

-Cellule de Malassez ou cellule de comptage pour dénombrer des bactéries, des levures, des cellules, … dans une solution.
-Maîtrise des puissances de 10.
- Acceptabilité des résultats avec incertitude lors d'un dosage (l'idée est de se contenter ici du calcul de moyennes à partir par exemple d'un logigramme, d'un écart-type de répétabilité et de résultats initiaux donnés. Une réflexion sera menée sur le nombre de chiffres significatifs.).
-Calcul de vitesse moyenne en cinétique chimique et biochimique.
-Calcul de prix TTC à partir d'un prix HT et inversement, avec des taux de TVA différents, dans le cadre de l'estimation du coût des analyses.
- Calcul de concentrations et de dilutions diverses pour préparer des solutions (.exemples : concentration molaire d'une solution d'HCl à 38%, volume d'acide sulfurique concentré à prélever pour préparer 1L d'une solution, titre alcoométrique en % à partir d'une expression en g/L, taux de concentration de 1 L de solution sur une cartouche SPE et récupération avec 5 mL de solvant, masse à peser de chlorure de sodium pour l'obtention d'une solution de chlorure à une concentration massique ou molaire donnée, …)

C7.3 - Choisir les analyses et contrôles adaptés aux objectifs fixés -

On attend de l'apprenant qu'il soit capable de proposer et de justifier, pour un point de contrôle, une méthode d'analyse adaptée au contexte. Il doit être capable de faire des choix de tests statistiques pertinents adaptés au besoin de l'analyse. Pour cela, l'apprenant mettra en place la :

-Réalisation d'un plan d'échantillonnage adapté prenant en compte les risques de première (fournisseur) et deuxième (client) espèce.

Cette mise en place doit répondre aux normes en vigueur et/ou cahiers des charges et sera en lien avec le bloc C4.

-Comparaison des méthodes d'analyse.

L'outil mathématique est utilisé pour mesurer la justesse, la fidélité, la répétabilité, la reproductibilité… en tenant compte de différents critères (coûts, spécificité, limite de détection, temps de réponse, appareillage, fréquence des analyses…).

-Validation du choix final en adéquation avec la hiérarchisation des exigences du client.

Ce choix alimente les arguments apportés dans la sous-capacités C8.2

Cette sous-capacité C7.3 est le point central du cours de mathématiques.

Mobilisation de l'outil mathématique

- Mise en place d'un rituel (ou automatismes).
- Mise en œuvre d'une démarche statistique pour analyser des résultats expérimentaux.

- **Mise en œuvre d'une démarche statistique pour analyser des résultats (expérimentaux).**

La mise en place d'un test statistique devra toujours répondre à une situation contextualisée, issue le plus souvent possible d'un cas concret du domaine professionnel. La mise en place des mesures relève du bloc de compétences C4 et l'interprétation dans le contexte initial pour le commanditaire relève de la capacité C8.2. En ce qui concerne le cours de mathématiques, la compréhension de la démarche d'un test statistique classique correspond donc à l'objectif principal de la capacité C7.3. L'apprenant doit être capable de comparer des méthodes d'analyse en évaluant justesse, fidélité, répétabilité, reproductibilité … et de répondre aux questionnements qu'il pourra rencontrer le plus fréquemment dans sa vie professionnelle.

Le travail sur ce module est conduit sur un temps long, il paraît donc essentiel de développer des méthodes statistiques à partir de simulations, le théorème central limite étant le théorème sous-jacent. Il n'est pas nécessaire de l'énoncer mais par contre il est indispensable de l'illustrer pour diverses situations avec différentes lois. L'importance de la loi normale doit alors apparaître. Les outils numériques ont dans leur grande majorité les lois normales implémentées, il est donc impératif de se séparer des tables de lois normales et du recours systématique au changement de variable. Le théorème central limite amènera à s'interroger sur le passage du discret au continu et donc de développer la notion de loi continue, majoritairement inconnue des étudiants.

L'enseignement doit concourir à développer la capacité à repérer des situations de référence de mise en œuvre de tests statistiques. L'objectif est moins de faire apprendre un catalogue de tests statistiques que de faire comprendre la méthodologie des tests et la construction de règles de décision s'appuyant sur la fluctuation d'échantillonnage de certaines grandeurs obtenue en premier lieu par simulation. La connaissance de certaines lois de probabilité de grandeurs lors de la variabilité des échantillons est l'aboutissement d'un travail préparatoire effectué par des simulations. Les tests doivent être adaptés aux situations rencontrées par les élèves, l'enseignant veillera à avoir des exemples suffisamment diversifiés.

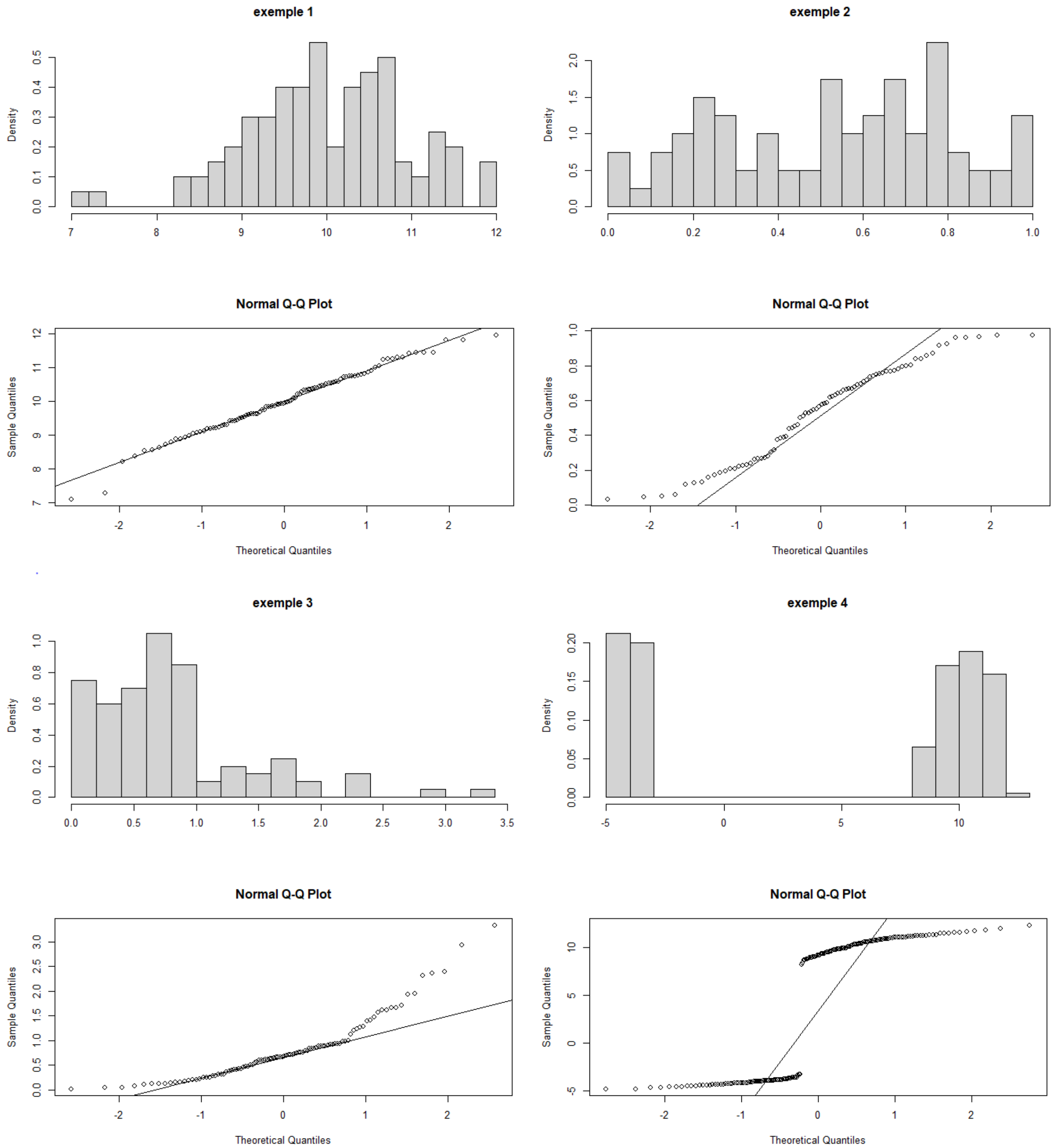
L'apprenant doit être capable de clairement identifier le cadre dans lequel se situe son travail : échantillons indépendants ou appariés, normalité des grandeurs ou non (test paramétrique ou non paramétrique, en comprenant les conditions et les limites de chaque situation), recherche d'une conformité d'une grandeur ou comparaison entre deux grandeurs ou comparaison de moyennes entre plusieurs séries. Une fois le cadre posé, le choix du test statistique s'effectuera en fonction des besoins et des habitudes professionnelles des secteurs concernés. Par conséquent, l'apprenant doit être capable de mener un test de conformité d'une proportion, d'une moyenne ou d'une variance ; un test de comparaison de deux proportions, de deux moyennes ou de deux variances ; une analyse de la variance à un facteur pour comparer la moyenne de plusieurs séries. Enfin, pour répondre aux habitudes en entreprise, l'apprenant doit être capable d'établir une carte de contrôle de la moyenne ou de l'écart type et de prendre les mesures nécessaires en cas de dysfonctionnement à l'aide de ses connaissances acquises dans les enseignements des matières professionnelles.

Les situations variées rencontrées dans le cadre de projets expérimentaux peuvent amener à mettre en place d'autres tests ou d'autres démarches (corrélation entre deux variables quantitatives, ANOVA à deux facteurs, test de Newman et Keuls, de Cochran, Analyse en Composante Principale,,…) Il n'est pas possible d'envisager tous les cas de figure dans le cadre de ce document et de l'horaire imparti. L'apprenant doit être capable de comprendre l'esprit de la mise en place d'un test statistique, puis avec l'accompagnement de l'enseignant ou du formateur et à l'aide de ses recherches personnelles, d'élargir ce qu'il est possible d'étudier lorsque cela s'avère nécessaire pour répondre à un problème initial.

Tous les calculs sont laissés à l'outil numérique. Même si son apprentissage peut être laborieux, le logiciel R fournit un grand nombre d'outils permettant de répondre à toutes les demandes et en particulier de travailler avec des données provenant de véritables expérimentations. Le travail est centré sur la reconnaissance des situations et le choix des méthodes.

Exemple 1 : Tester la normalité

Un préalable à beaucoup d'études est la normalité des grandeurs en jeu. Parfois, la situation impose de fait la normalité des grandeurs, d'autres fois il sera peut-être nécessaire de débuter par un test de normalité. Pour une première approche on pourra s'appuyer sur la forme des histogrammes des échantillons et exposer la méthode de la droite de Henry. Tous les graphiques sont obtenus à l'aide de l'outil numérique. Par exemple, la commande qqnorm() du logiciel R permet de tracer le graphique quantile-quantile qui confronte les quantiles de la loi normale en abscisse et les quantiles empiriques de l'échantillon en ordonnée. La commande qqline() construit la droite joignant le couple des quantiles 0,25 et le couple des quantiles 0,75.



Cette approche graphique pourra être complétée par le test de Shapiro-Wilk obtenu directement par la commande shapiro.test() du logiciel R. On ne rentrera pas dans les détails de ce test. Il s'agit de développer un questionnement sur l'hypothèse de normalité au regard de son importance dans les conclusions du théorème central limite et de la somme de variables aléatoires indépendantes suivant une loi normale.

On trouve pour les exemples ci-dessus :

Exemple 1	**Exemple 2**	**Exemple 3**	**Exemple 4**
W = 0,98982	W = 0,93316	W = 0,93141	W = 0,7231
p-value = 0,6503	p-value = 0.0004218	p-value = 5.993e-05	p-value < 2.2e-16

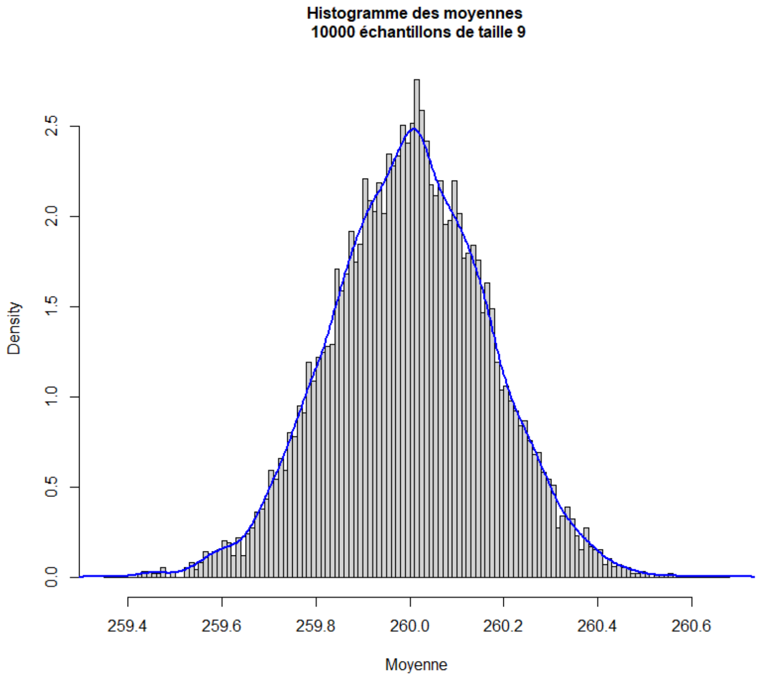
Pour le test de Kolmogorov-Smirnov, voir la commande ks.test() du logiciel R

Exemple 2 : Introduire le test de conformité d'une moyenne

Vous êtes technicien(ne) dans un laboratoire d'analyse et de conseil œnologique appartenant à un groupement de viticulteurs produisant du vin blanc doux. Pour se conformer au cahier des charges et à la réglementation (norme INAO), la teneur en sulfites maximale est fixée à 260 mg/L. On décide de mettre en place un contrôle régulier en prélevant des échantillons de taille $n = 9$. On admet que la distribution des teneurs est gaussienne et d'écart type $\sigma = 0{,}5 mg/L$.
Un échantillon est prélevé et on obtient par exemple les valeurs :

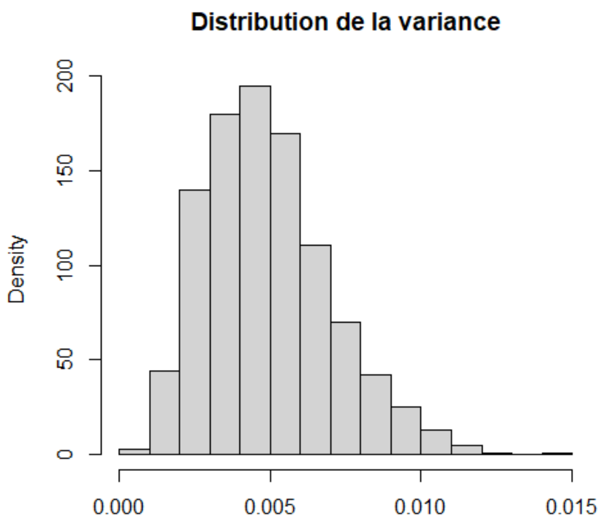
260,25 259,34 260,90 259,74 259,30 259,80 259,67 259,00 259,87

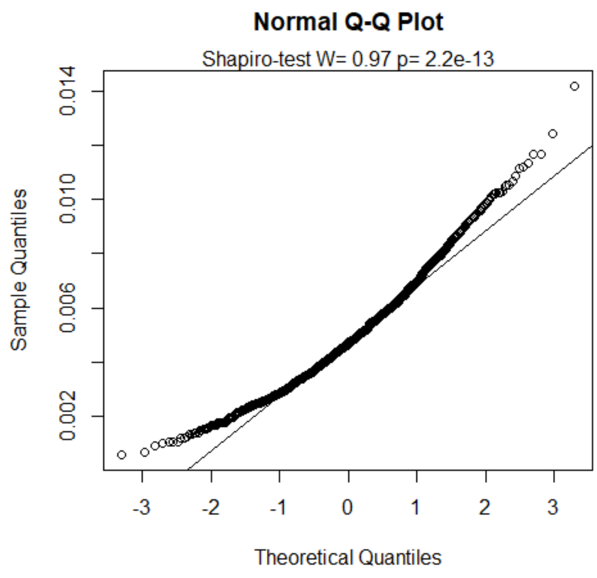
On simule par exemple 10000 échantillons de taille $n = 9$ d'une loi normale de paramètres $\mu = 260$ et $\sigma = 0{,}5$ et on étudie la distribution de la moyenne des teneurs de chaque échantillon. L'outil numérique permet d'obtenir un graphique du type ci-dessous. En s'appuyant sur ce graphique, on pourra alors s'interroger sur la distribution d'échantillonnage d'une moyenne en particulier la normalité de la distribution puis amener le questionnement de la conformité d'une moyenne. La mise en place du test de conformité d'une moyenne et l'utilisation de la loi normale sont l'aboutissement de l'étude. On discutera de l'unilatéralité ou la bilatéralité du test.

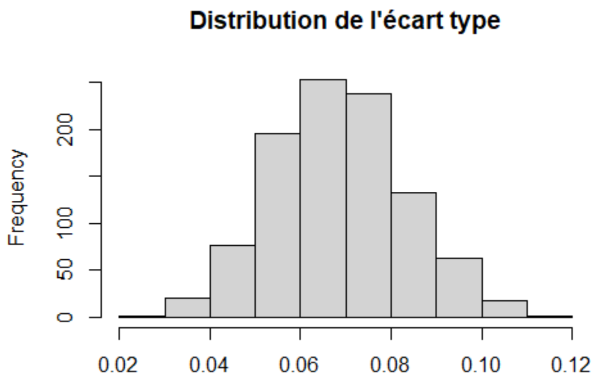


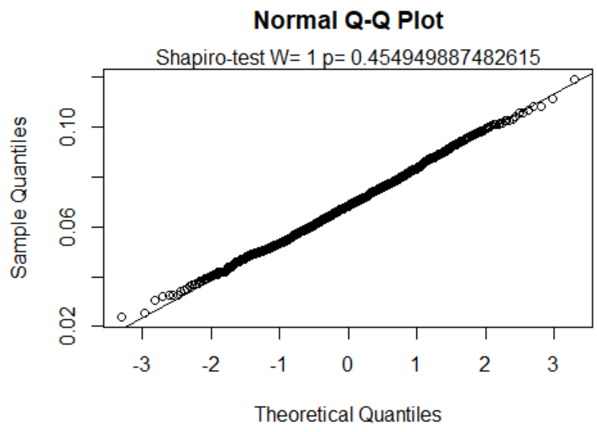
Exemple 3 : Introduire une nouvelle variable statistique

On considère que, pour un cépage donné, un vin rouge peut paraître acide si le pH est inférieur à 3,4.
On prélève un échantillon de 12 bouteilles dont les pH sont les suivants :
Echantillon 0 : 3.43 3.49 3.47 3.36 3.42 3.56 3.32 3.46 3.43 3.39 3.30 3.42
Après avoir posé la question de la normalité des valeurs de l'échantillon, l'estimation de l'écart type sera au cœur des discussions. On sera amené à étudier graphiquement les distributions des variances et des écarts types pour des échantillons gaussiens de taille 12 obtenus par simulation de la loi normale de moyenne 3,4 et d'écart type égale à l'écart type de l'échantillon 0. On pourra au passage se poser la question de la normalité de ces distributions.

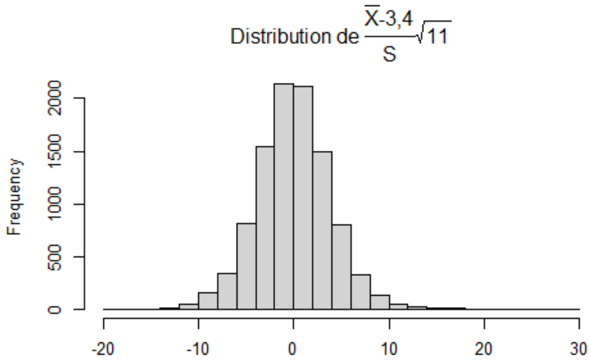




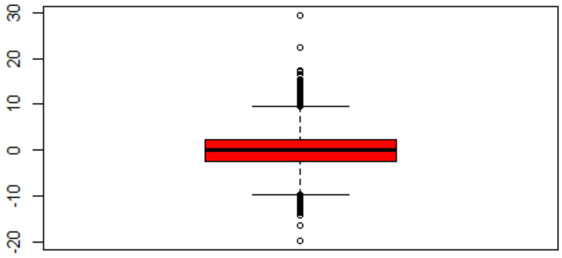




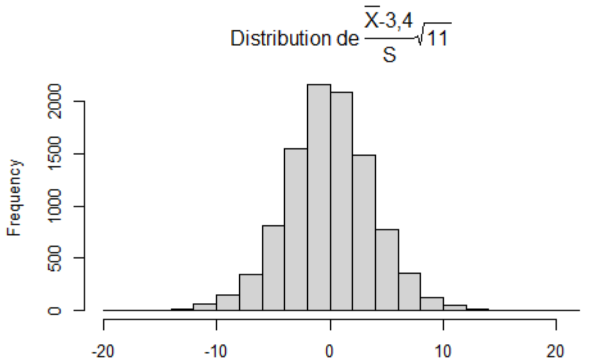
La variabilité de l’écart type des échantillons et le choix arbitraire de la valeur de l’écart type de l’échantillon 0 plutôt qu’une autre valeur amènera à considérer une autre variable statistique $\frac{\bar{X}-3,4}{S}\sqrt{11}$ où chaque échantillon simulé a la même importance dans le choix de l’écart type. On pourra constater que si l’on fait varier l’écart type autour de la valeur de l’échantillon 0, la loi de cette variable reste stable, ce qui peut se constater en observant les histogrammes et les diagrammes en boîte ci-dessous.



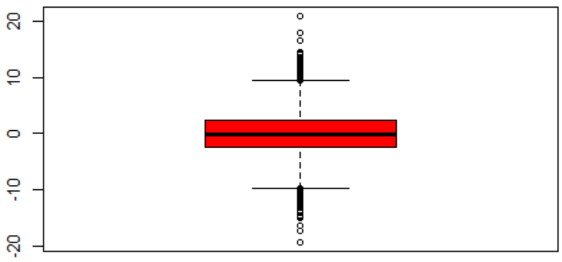
σ = 0.04



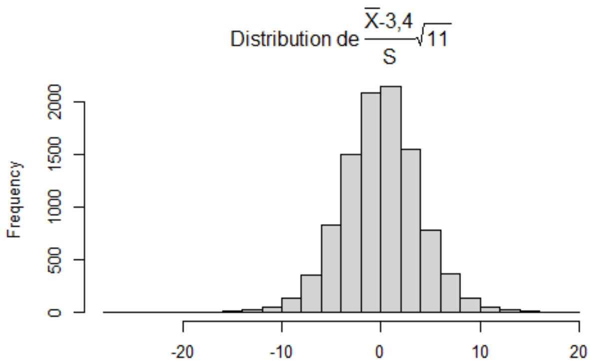
σ = 0.04



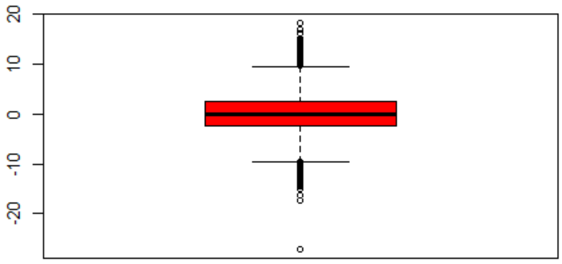
σ = 0.05



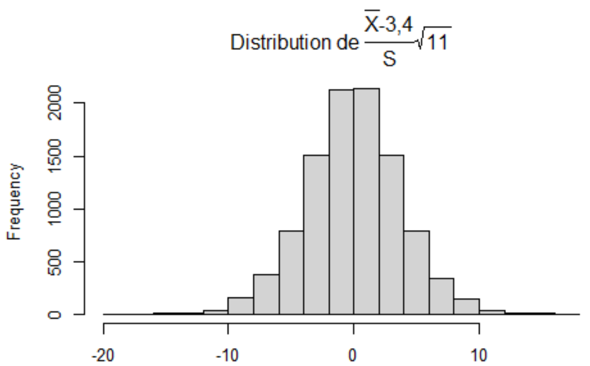
σ = 0.05



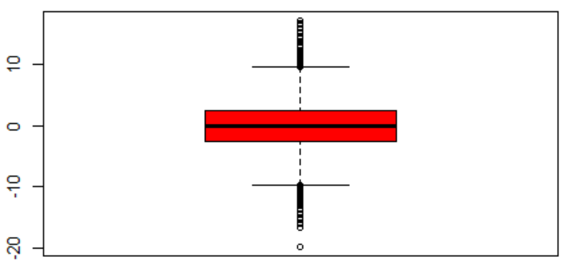
σ = 0.06



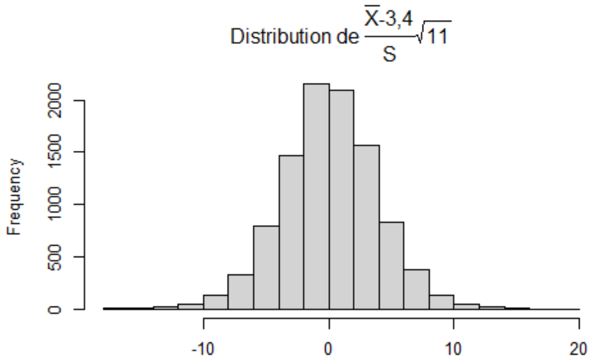
σ = 0.06



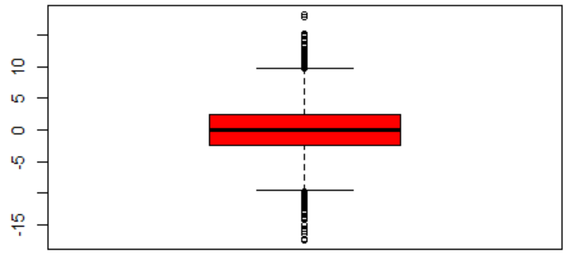
σ = 0.07



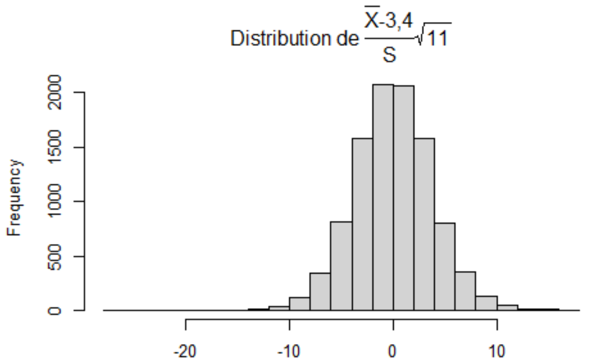
σ = 0.07



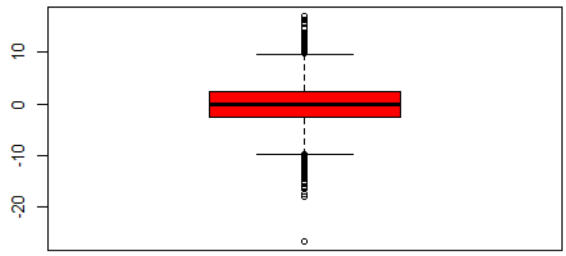
σ = 0.08



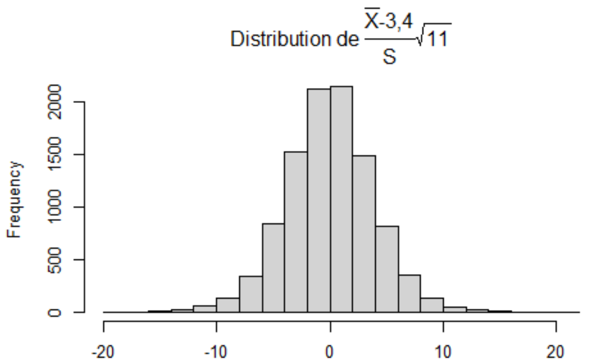
σ = 0.08



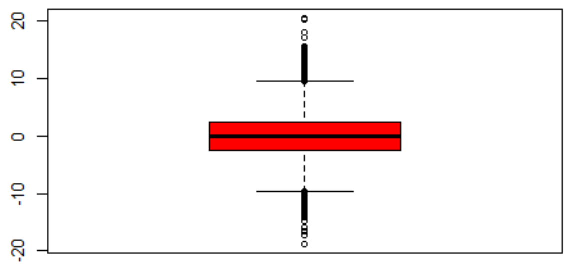
σ = 0.09



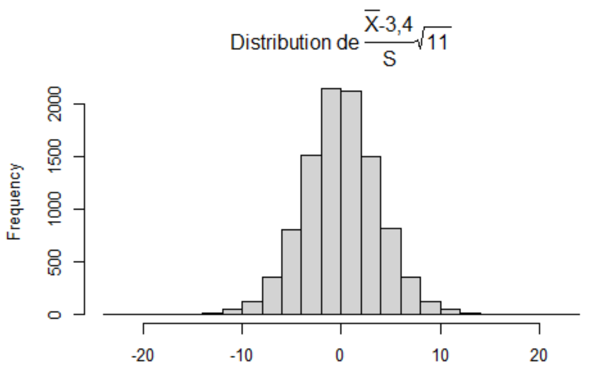
σ = 0.09



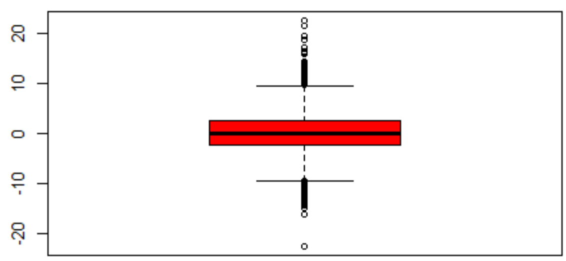
σ = 0.10



σ = 0.10

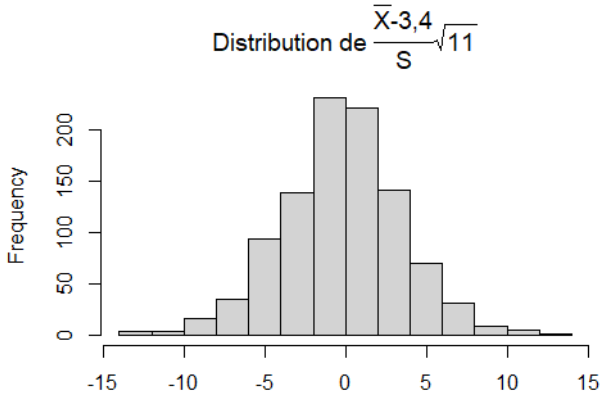


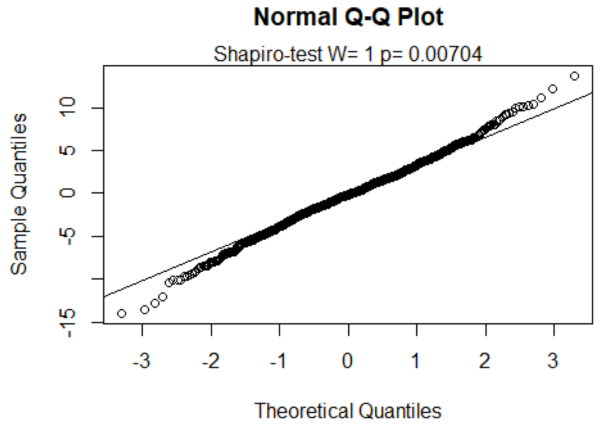
σ = 0.11



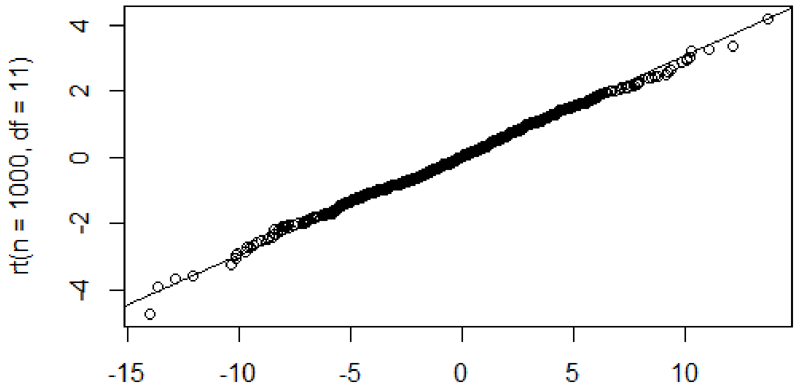
σ = 0.11

Le rejet de la loi normale pour cette nouvelle variable statistique induira l'introduction de la loi de Student.





On pourra pour confirmer construire le graphique quantile-quantile entre $\frac{\bar{X}-3,4}{S}\sqrt{11}$ et la loi de Student T_{11}.



La mise en place du test avec la loi de Student conclura l'exemple. On pourra discuter de l'unilatéralité ou de la bilatéralité du test.

- **Mise en place d'automatismes (ou de rituels).**

En complément de ces études en contexte, la pratique d'automatismes vise à construire et entretenir des aptitudes dans le domaine mathématique. L'ensemble des automatismes doit être pratiqué quelles que soient les thématiques travaillées. On pourra rencontrer les situations suivantes, sans caractère exhaustif ni obligatoire :

-Résolution d'une équation à une inconnue pour déterminer la concentration en réactif dans une réaction chimique par exemple.
-Application des formules dans le cas de mesure de cinétiques microbiennes.
-Calcul de proportionnalité en lien avec la quantité de matière (atomes, molécules, …). Calcul de pourcentages massiques.
-Maîtrise des échelles en microscopie.
-Manipulation de formules dans des cas de dissolution ou de dilution.
-Concentrations en mg/mL ou en ppm.
-Dosage par chromatographie (l'intégration de chaque pic, faite informatiquement, permet de déterminer sa surface, qui est proportionnelle à sa concentration.)
-Vitesse instantanée de réaction cinétique (tracé de tangente, détermination de son équation et de son coefficient directeur).

-Détermination de l'équation d'une droite dans le cadre d'un étalonnage ou dans le cadre d'un dosage colorimétrique. Interprétation de la pente et de la déviation à l'origine.
-Loi de désintégration radioactive.

C8.2 Valider des résultats -

On attend de l'apprenant qu'il soit capable d'interpréter des résultats obtenus à l'aide d'outils statistiques et de proposer une analyse critique des résultats compréhensible par un professionnel ou le grand public en fonction de l'interlocuteur destinataire.

Les outils mathématiques et statistiques utilisés dans les autres blocs et les outil numériques ont été mobilisés pour:

La présentation et le traitement des résultats d'activités

L'apprenant, qui a déjà enregistrer les données, utiliser une méthode adaptées au contexte et utiliser des outils numérique, a désormais pour but de faire comprendre au commanditaire la signification de ces résultats. Il est donc important d'avoir une réflexion sur :

L'interprétation et l'analyse critique de résultats après avoir traité les résultats d'activités

en mettant en évidence les limites éventuelles de la méthode utilisée, en élaborant une conclusion compréhensible par le commanditaire si besoin à l'aide de graphiques et en émettant des propositions d'amélioration.

Mobilisation de l'outil mathématique

- Mise en place d'un rituel (ou automatismes).
- Interprétation et analyse critique de résultats après avoir traité les données récoltées.

- **Interprétation et analyse critique de résultats après avoir traité les données récoltées.**

Une démarche mathématique ayant été mise en place au service d'une analyse, d'un essai ou d'un contrôle, elle sera traitée dans le cadre du module M7 (capacité C7.3) et du module M4. L'apprenant doit, pour cette capacité C8.2, pouvoir rendre compréhensibles les résultats obtenus, que ce soit à destination d'un professionnel ou du grand public, et adapter son message en fonction des interlocuteurs pour rendre l'interprétation accessible. Les limites et les précautions à prendre quant à l'interprétation des résultats sont expliquées. L'analyse critique qu'il fait de ces résultats doit ainsi être comprise et plus facilement acceptée par le commanditaire qui reste décisionnaire. On veille donc à ce que les conclusions formulées n'emploient pas de termes scientifiques qui ne soient pas aisément intelligibles et que la prise de décision future soit éclairée par les explications fournies par l'apprenant. Ce partage d'informations doit pouvoir éventuellement générer des propositions d'amélioration de la part de l'apprenant en fonction des situations rencontrées.

- Mise en place d'automatismes (ou de rituels).

En complément de ces études en contexte, la pratique d'automatismes vise à construire et entretenir des aptitudes dans le domaine mathématique. L'ensemble des automatismes doit être pratiqué quelles que soient les thématiques travaillées. On pourra rencontrer les situations suivantes, sans caractère exhaustif ni obligatoire :

-Détermination des concentrations d'ions par photométrie de flamme et des rapports de concentration. La lumière émise est proportionnelle à la concentration de l'échantillon.
-Application des formules de dénombrement microbien.
-Conversion cm3 en L, ha en km²,…)
-Détermination de l'ordre d'une réaction enzymatique par la méthode d'intégration (dans un cas où la loi de la vitesse est simple à intégrer : ordre 0, 1 ou 2).
-Calcul de rendement instantané, de productivités...
-Détermination de vitesses spécifiques de croissance, de temps de génération à partir de données expérimentales et de représentation graphique.
-Détermination graphique du coefficient de transfert volumétrique du dioxygène (KLa) par la méthode de désoxygénation statique en recherchant le coefficient directeur d'une droite.
-Conversion d'un logarithme de base a, en particulier le logarithme décimal.

-Utilisation de la loi de Beer-Lambert pour des lectures graphiques de l'absorbance qui est proportionnelle à la concentration de l'entité chimique ou à la longueur du trajet parcouru par la lumière dans une solution.
-Calcul du pH en fonction de la concentration en ions H+, lien avec le pKa (équations du 1er et 2nd degré), systèmes tampons (constantes de la droite tampon).
-Temps de génération nécessaire pour qu'une population microbienne double sa taille pendant la phase exponentielle.

Annexe 1

Exemple: Introduire le test d'indépendance du Khi2
Introduire la distribution du χ^2par simulation de l'adéquation à une loi. Par exemple, appuyons-nous sur l'expérience de Mendel sur la transmission de caractères sur les pois.

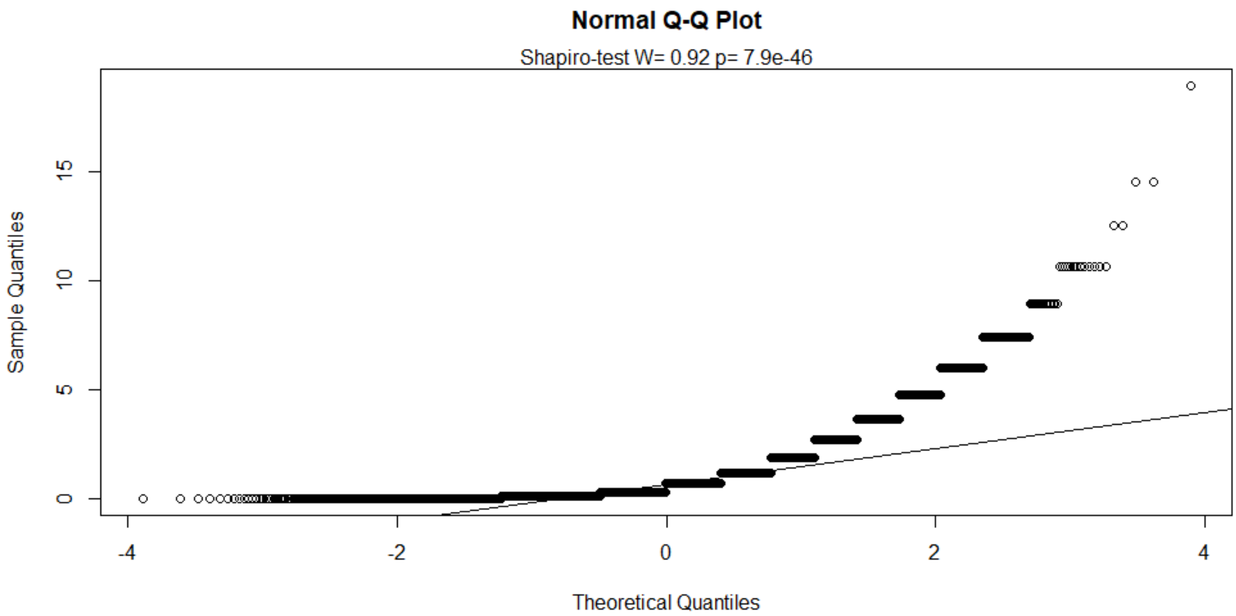
Pour comprendre la transmission d'un caractère d'une génération à l'autre, Mendel féconde artificiellement deux variétés de pois de lignée pure. L'un avec le caractère « graines lisses », l'autre avec le caractère « graines ridées ». La descendance obtenue (F1) ne possède que des graines lisses. Il poursuit l'expérience en réalisant l'autofécondation de la génération (F1). Il obtient la répartition suivante pour la génération (F2).

Caractère	**Graines ridées**	**Graines lisses**	**Total**
Effectifs	21	51	72

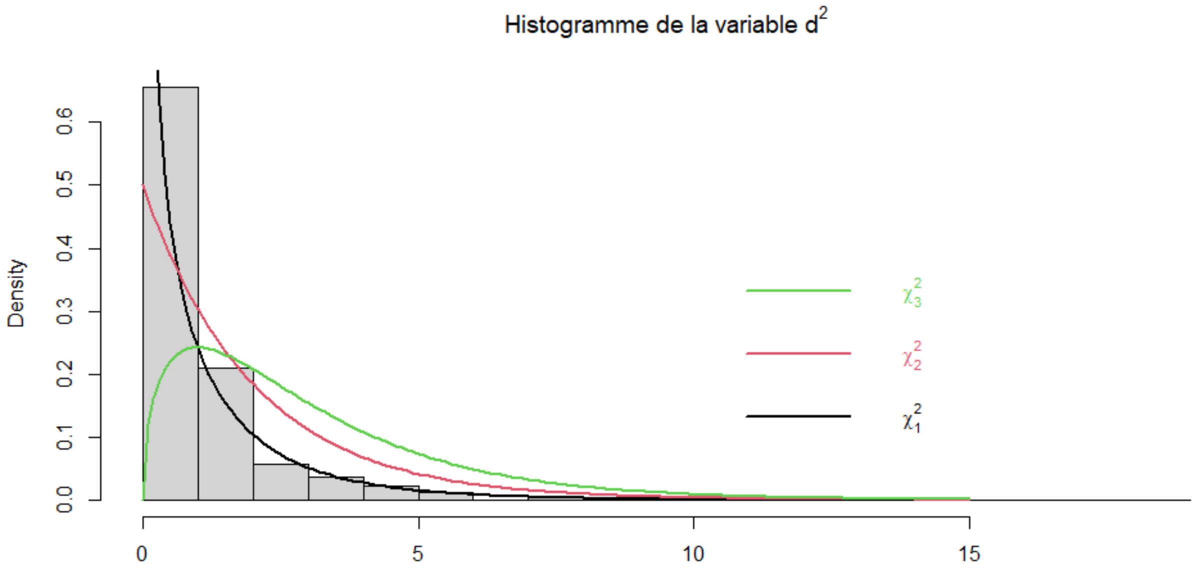
Ces résultats expérimentaux confirment-t-ils l'hypothèse de Mendel qui prévoit une répartition de 25% et 75% ?

On simule la variable statistique d^2 et on étudie sa répartition. En notant $O_{i,j}$ les effectifs observées et $E_{i,j}$ les effectifs attendus calculés à partir du modèle de Mendel.

$$d^2 = \sum_{i,j} \frac{\left(O_{i,j} - E_{i,j}\right)^2}{E_{i,j}}$$



La forme de l'histogramme indique que la distribution de d^2 ne s'apparente pas à une loi normale, hypothèse qui pourra être confirmée avec un Q-Q Plot et un test de Shapiro-Wilk. On est donc amené à chercher une autre loi. Ce qui permettra d'introduire les lois du χ^2.

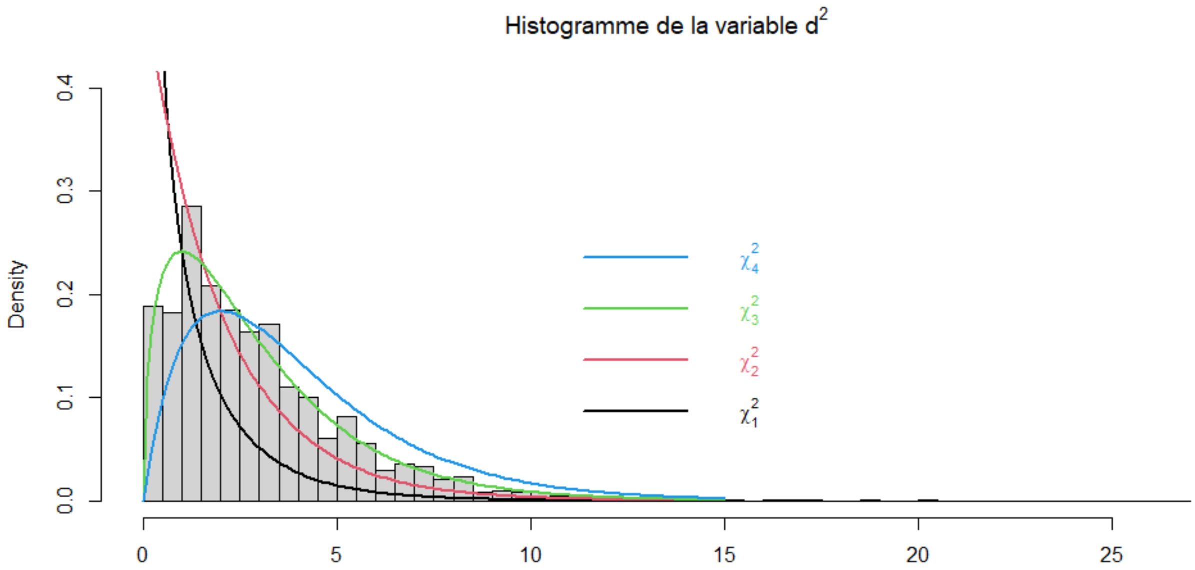


On peut alors s'intéresser à l'expérience de Mendel avec deux caractères exprimés par des gènes comportant deux allèles (l'un dominant A, B et l'autre récessif a, b) sur des chromosomes différents.
On obtient pour la génération (F2) le tableau suivant :

Caractères	Ab	aB	Ab	AB	Total
Effectifs	3	15	13	33	64

Ces résultats expérimentaux confirment-t-ils l'hypothèse de Mendel qui prévoit la distribution $\left(\frac{1}{16}, \frac{3}{16}, \frac{3}{16}, \frac{9}{16}\right)$?

De la même manière, on obtient :



Bibliographie I :

-Les normes ISO (https://www.iso.org/fr/standards.html) ou AFNOR (https://www.afnor.org/) servent de référence.

-« Mesure et incertitudes au lycée », mai 2021, Eduscol.

https://eduscol.education.fr/document/7067/download

Des Mooc (Massive Open Online Course ou FLOT) peuvent permettre d'approfondir certaines notions. Par exemple, Statistiques pour l'ingénieur de l'Institut Mines-Télécom ou Analyse de données multidimensionnelles d'Agrocampus Ouest sont régulièrement ouverts sur la plateforme Fun.