**Centre de Ressources
Comptabilité Finance**

Lycée MARIE CURIE

Avenue du 8 mai 1945 - BP 348

38435 ECHIROLLES cedex

[**http://crcf.ac-grenoble.fr/**](http://crcf.ac-grenoble.fr/)

**2.3 La prise en compte de données aléatoires**

**Variables aléatoires et lois de probabilité**

***Ce que dit le projet de programme de l’UE 8 :***

**2.3 La prise en compte de données aléatoires**

**Sens et portée de l’étude**

Il s’agit d’introduire l’aléa dans les modèles de contrôle de gestion en présentant les outils qui permettent de répondre à des problèmes de gestion en avenir aléatoire : risque d’exploitation, calcul du chiffre d’affaires, d’une marge et d’un résultat.

|  |  |
| --- | --- |
| **Compétences attendues** | **Savoirs associés** |
| - Calculer et interpréter une espérance et un écart-type sur ventes, coûts, marge et résultat, pour un ou plusieurs produits.- Identifier la loi de probabilité adaptée à une situation de gestion donnée puis calculer et interpréter les probabilités. - Déterminer et interpréter le seuil de rentabilité en avenir aléatoire. | - Variables aléatoires discrètes et continues : fonctions de distribution et de répartition, espérance mathématique, variance et écart type.- Propriétés de l’espérance et de la variance pour le seul cas de variables aléatoires indépendantes. - Caractéristiques et modalités d'application des lois suivantes : binomiale, de Poisson, normale.- Approximation des lois. |

1. Découverte de la notion de variable aléatoire

**1er exemple :**

Reprenons l’exemple de la pêche aux canards.

Le magasin Au bon marché désire réaliser une animation commerciale à l’occasion du printemps.

Le directeur propose de faire une pêche aux canards à destination des enfants de moins de 8 ans. Le magasin achète le matériel nécessaire dont 10 canards numérotés de 1 à 10 sous ceux-ci (lors de la pêche, le numéro n’est donc pas visible). Chaque canard a la même chance d’être tiré.

À chaque numéro sera attribué un bon d’achat dont la valeur est égale à cinq fois le numéro pêché.

Les bons d’achat à gagner ont donc une valeur de 5, 10, 15, etc., jusqu’à 50 euros.

***La valeur du bon d’achat à gagner constitue une variable aléatoire***. À chaque valeur de bon d’achat est attachée une probabilité.

Nous pouvons dresser le tableau suivant dans lequel xi désigne la valeur du bon d’achat et Pi la probabilité de gagner cette valeur :



Ce tableau constitue la loi de probabilité de la variable aléatoire « valeur du bon d’achat ». Les valeurs prises par la variable sont « isolées », dénombrables, on dit alors que c’est une **variable aléatoire discrète**.

**Deuxième exemple :**

Un magasin de fruits et légumes vend en juillet et août des abricots par caisse de 5 kg. Le nombre de caisses vendues varie de 0 à 30 caisses.

Chaque jour, on relève le nombre de caisses vendues. Ce nombre dépend de la demande et est donc soumis au hasard.

La variable « nombre de caisses vendues au cours d’une journée » constitue une **variable aléatoire discrète**. À chaque valeur « nombre de caisses vendues au cours d’une journée » est attachée une probabilité.

**Troisième exemple :**

La SVR, Société Vendéenne de Roulements, fabrique différents types de roulement à bille pour l’industrie.

Vous êtes chargé(e) de contrôler le diamètre des roulements dont la référence est R005. Ces roulements doivent avoir un diamètre de 5 centimètres, mais une tolérance est admise. Un roulement à bille de référence R005, est commercialisable quand son diamètre est compris entre 4,95 et 5,05 centimètres. Quand la chaîne de fabrication qui le produit est correctement réglée, le diamètre se trouve dans cet intervalle.

Vous contrôlez toutes les heures le diamètre (contrôle qualité). En une heure, il est produit environ 5 000 roulements.

Vous choisissez au hasard un roulement R005 de la production ; la valeur du diamètre de ce roulement prélevé est donc soumise au hasard.

On appelle X, la variable qui à tout roulement prélevé associe son diamètre. X est appelée variable aléatoire. **X peut prendre toutes les valeurs possibles dans un intervalle**. ***X est appelée variable aléatoire continue***. Le nombre de valeurs que peut prendre X n’est pas dénombrable (il y a une infinité de valeurs dans un intervalle).

**Précision** : Dans le cas d’une variable continue le calcul d’une probabilité pour une valeur donnée n’a pas de sens. Par exemple la probabilité que le diamètre soit exactement égal à 4,995 cm est nulle (ou quasi), le diamètre pouvant prendre une infinité de valeur dans un intervalle. En revanche calculer la probabilité que le diamètre soit inférieur à une certaine valeur a du sens.

**Formalisons un peu et généralisons**

Une variable aléatoire est une variable qui peut prendre un certain nombre de valeurs, dénombrables ou non. À chaque valeur d’une variable aléatoire discrète, est attachée une probabilité (on utilise le terme de loi de probabilité). Pour ce qui est des variables continues, nous aurons l’occasion de les définir un peu plus loin.

***Une variable aléatoire est dite discrète quand elle prend des valeurs réelles isolées (elles sont dénombrables).***

***Une variable aléatoire est dite continue quand elle peut prendre n’importe quelle valeur sur un intervalle (elles ne sont pas dénombrables).***

Pour faciliter l’écriture, nous utiliserons par la suite l’acronyme VA pour désigner le terme variable aléatoire.

1. Loi de probabilité (ou fonction de distribution) et fonction de répartition d’une VA discrète

Nous partirons d’un exemple pour définir les notions de loi de probabilité (ou fonction de distribution) et de fonction de répartition.

Exemple – (reprise des données de l’application 1 du document « des statistiques aux probabilités).

Rappel du contexte : Parmi ses activités, l’EARL (Exploitation Agricole à Responsabilité Limitée) des Terres Noires élève en plein air des volailles BIO qu’elle vend sur place ou sur les marchés. Chaque début de semaine sont abattues les volailles qui seront mises à la vente dans la semaine. Toute volaille non vendue dans la semaine est perdue et envoyée à l’équarrissage (destruction de la volaille).

Les ventes possibles au cours d’une semaine (univers des possibles ou ensemble des résultats élémentaires possibles) avec leur probabilité attachée sont résumées dans le tableau suivant :



Appelons X le nombre de volailles vendues au cours d’une semaine. **X est une variable aléatoire discrète**.

***Les probabilités pi = P(X = xi) associées aux valeurs xi de la VA X constituent la loi de probabilité de X, encore appelée fonction de distribution de X.***

La fonction de distribution peut être représentée par le diagramme suivant (diagramme en bâtons) :

**Formalisons et généralisons**

La loi de probabilité ou fonction de distribution d’une variable aléatoire discrète X est donnée par :

* L’ensemble des valeurs x1, x2, x3, etc. prises par la variable aléatoire ;
* Et les probabilités pi= P(X = xi) pour toutes les valeurs xi prises par X.

Et la somme des probabilités pi est égale à 1.

Pour bâtir la fonction de répartition à partir de l’exemple sur les ventes de volailles, nous avons besoin de calculer le cumulé croissant des probabilités des ventes de volailles. Nous obtenons le tableau suivant :



Interprétation du cumulé croissant : le cumulé croissant des probabilités donne la probabilité que les ventes de volailles soient inférieures à une certaine valeur. Par exemple la probabilité que les ventes de la semaine soient inférieures ou égal à 22 volailles est de 0,44, et l’on note cette probabilité P(X <= 22).

***Les probabilités pi = P(X <= xi) associées aux valeurs xi de la VA X constituent la fonction de répartition de la VA X.***

La fonction de répartition peut être représentée par le diagramme suivant (histogramme en cumulé) :

**Formalisons et généralisons**

La fonction de répartition d’une variable aléatoire discrète X est donnée par :

* Les probabilités pi= P(X <= xi) pour toutes les valeurs xi prises par X.

P(X <= xi) représente la probabilité que la variable aléatoire X prenne une valeur inférieure ou égale à xi.

1. Espérance mathématique et variance d’une VA discrète

**L’espérance mathématique** correspond tout simplement à la **moyenne (arithmétique)** que l’on peut calculer à partir d’une série statistique (la somme des probabilités faisant 1, le dénominateur est inutile dans la formule de calcul). Le terme d’espérance est utilisé car c’est « l’espoir » d’atteindre au moins cette moyenne qui est attendu. Dans notre exemple précédent, la vente moyenne attendue de volaille ou espérance mathématique est égale à :

(15x0,02) + (16x0,02) + (17x0,04) + (18x0,04) + (19x0,04)+ (20x0,10)+ (21x0,10)+ (22x0,08)+ (23x0,12)+ (24x0,10)+ (25x0,12)+ (26x0,10)+ (27x0,04)+ (28x0,02)+ (29x0,02)+ (30x0,04) = 22,82 ***soit environ 23 volailles.***

**Interprétation et limites** : si on fait des statistiques sur le long terme, c’est-à-dire si l’on note chaque semaine les ventes pendant un très grand nombre de semaines, et que l’on fait la moyenne de ces ventes, on se rapprochera sensiblement de l’espérance mathématique. L’espérance est souvent utilisée comme outil de comparaison (comparaison des moyennes entre plusieurs stratégies par exemple). Mais à court terme, la moyenne doit être accompagnée d’un autre indicateur qui est l’écart moyen et que l’on nomme écart-type. L’écart-type étant la racine carrée de la variance : *la moyenne est en effet une valeur que l’on atteint rarement*.

**Notation : l’espérance mathématique est notée E(X) ou parfois (**X barre).

**Formalisons et généralisons**

On considère une variable aléatoire discrète notée X. L’univers.

* On note x1, x2, x3, etc. l’ensemble des valeurs prises par la variable aléatoire,
* Et pi= P(X = xi) les probabilités pour toutes les valeurs xi prises par X.

**L'espérance mathématique de la VA X est la moyenne arithmétique des valeurs de X pondérées par leur probabilité, elle se calcule donc de la façon suivante : E(X) = p1.x1 + p2.x2+ p3.x3 + et. = .**

Les valeurs de la VA s’écartent plus ou moins fortement de son espérance (la moyenne). Il est intéressant de connaitre cet écart moyen pour avoir une idée de la dispersion des valeurs autours de l’espérance (la moyenne). Cet écart moyen doit bien sûr tenir compte des probabilités de chaque valeur de la variable.

Par ailleurs, les valeurs de la VA sont soient supérieures, soit inférieures à la moyenne. Pour neutraliser le signe de l’écart, ceux-ci sont élevés au carré. Donc pour calculer la variance, il nous faut dans un premier temps calculer pour chaque valeur de la VA les écarts suivant : (xi – E(X))². Ces écarts sont ensuite multipliés par la probabilité pi attachée à chaque valeur xi.

Le calcul décrit ci-dessus correspond à la moyenne arithmétique des écarts (xi – E(X))² pondéréspar leur probabilité. Pour l’analyse, il sera bien sûr nécessaire de prendre la racine, puisque les valeurs ont été élevées dans un premier temps au carré. La racine de la variance est appelée écart-type.

**Le recours aux statistiques pour expliquer ce que représente l’écart-type** :

Imaginons une entreprise où 30 % des salariés gagnent 2 000 euros, 50 %, 3 000 euros et 20% 4 000 euros.

La moyenne est de (30% \* 2000) + (50% \* 3000) + (20%\*4000) = 2 900 euros.

Mais aucun salarié ne gagne effectivement 2 900 euros. Ce qui est en revanche intéressant c’est l’écart moyen autours de cette moyenne, calculé comme précisé précédemment :

Variance = 30% \*(2000-2900)² + 50% \* (3000-2900)² + 20% \*(4000-2900)² = 490 000

Écart-type = Racine carré (variance) = racine carré (490 000) = 700 euros.

Donc dans cette entreprise l’écart moyen des salaires est de 700 euros, ce qui permet d’apprécier l’amplitude des salaires. C’est tout l’intérêt de cet indicateur qu’est l’écart-type en matière d’analyse.

***Nous pouvons étendre aux probabilités ce qu’il ressort de cet exemple.***

**Calcul de la variance et de l’écart-type dans l’exemple des ventes de volailles :**

Les calculs peuvent être présentés dans le tableau suivant : (calculs faits avec un tableur, dont tous les chiffres après la virgule ont été conservés)



La variance est égale à 11,99.

**Remarque** : On peut également la calculer de la façon suivante : Σ pi \* xi² - E(X)² = 533 – 22,82² = 11,99. Ce calcul est plus rapide quand il est effectué « à la main ».

L’écart-type est égale à la racine carrée de 11,99 soit environ 3,46.

**Interprétation des résultats** : la moyenne des ventes de volailles au cours d’une semaine est d’environ 23, et l’écart moyen autour de cette moyenne est d’environ 3. Ce qui laisse à penser qu’en moyenne les ventes seront assez resserrées autour de la moyenne. Le risque de ventes basses est donc faible, ce qui donne une bonne visibilité pour l’entreprise dans le nombre de volailles à abattre en début de semaine.

***Les calculatrices permettent d’obtenir facilement l’espérance (la moyenne), la variance et l’écart-type.***

***Il appartient à chacun de maîtriser ces calculs sur sa calculatrice.***

**Formalisons et généralisons**

On considère une variable aléatoire discrète notée X. L’univers.

* On note x1, x2, x3, etc. l’ensemble des valeurs prises par la variable aléatoire,
* Et pi= P(X = xi) les probabilités pour toutes les valeurs xi prises par X.

**La variance de la VA X, notée V(X) se calcule comme suit :**

**- E(X)]² ou encore de manière équivalente**

**L’écart-type noté σ(X) est égal à la racine carré de la variance : σ(X) =**

**Complément : Pourquoi faut-il souvent accompagner le calcul de l’espérance mathématique du calcul de l’écart-type ?**

Vous êtes un fidèle client d’une entreprise. Pour vous remercier elle vous propose l’alternative suivante :

* Un bon d’achat de 1000 euros ou,
* Un billet de tombola qui ne comporte qu’un lot : un bon d’achat de 100 000 euros. Le nombre de billets de tombola distribué est de 100.

**Que choisissez-vous ?**

Si nous excluons dans un premier temps tout recours à une formalisation du problème et à un outil d’aide à la décision, votre décision dépendra de votre aversion au risque.

Formalisons le problème : nous avons le choix entre un gain certain de 1 000 euros et un gain incertain. Si nous utilisons l’espérance mathématique, dans le premier cas l’espérance est de 1 000 euros. Dans le second cas, nous avons 1/100 de chance de gagner 100 000 euros et 99/100 de (mal)chance de gagner 0. L’espérance est donc (1/100 \* 100 000) + (99/100 \* 0) = 1 000 euros.

**Les espérances sont identiques**, selon ce critère les choix sont donc équivalents. Notons toutefois de nouveau l’aspect théorique de la notion d’espérance, pour la tombola, c’est un montant qui ne sera jamais atteint : c’est 100 000 ou rien. L’espérance ou moyenne n’est pas toujours un critère pertinent.

Mais calculons l’écart-type. Dans le premier cas, il est bien sûr nul. Dans le second cas il est égal à :

Racine [(1/100 \* 100 000²)+(99/100 \* 0²)] – 1 000² = 9 949,87 euros.

**L’écart type de la loterie est donc très élevé**, ce qu’intuitivement vous aviez perçu dès le départ. Mais dans des cas plus complexes, le calcul de l’écart-type pour mesurer le risque est nécessaire et apporte un élément indispensable pour prendre une décision.

***La décision finale dépend du risque que vous êtes prêt à prendre. L’écart-type est ici une mesure du risque.***

1. Loi de probabilité et fonction de répartition d’une VA continue

Rappelons qu’une variable aléatoire est dite continue quand elle peut prendre n’importe quelle valeur sur un intervalle (elles ne sont pas dénombrables).

Lorsque la variable est continue, il n’est pas possible de donner les mêmes définitions que pour les VA discrètes pour lesquelles nous donnions les probabilités de chaque résultat élémentaire possible (évènement élémentaire) avec pour contrainte que la somme des probabilités faisait 1. Le nombre de résultats possibles est en effet infini, non dénombrable.

***Avec une VA continue, la probabilité d’un résultat élémentaire est paradoxalement égal à 0***, calculer une telle probabilité n’a pas de sens. Prenons l’exemple d’une entreprise dont le chiffre d’affaires prévisionnel est entre 100 000 et 200 000 euros. Se poser la question suivante : quelle est la probabilité d’atteindre un chiffre d’affaires de 127 846,52 euros n’a pas de sens, cette probabilité est nulle. En revanche se poser la question : quelle est la probabilité d’atteindre un chiffre d’affaires au moins égal à 120 000 euros a du sens.

***Avec une VA continue, c’est donc la fonction de répartition qui est essentiellement utilisée. Elle permet notamment de calculer P(X< = a) où a est un réel (une valeur), c’est-à-dire la probabilité que la VA soit inférieure à la valeur a.***

**Il n’est pas nécessaire d’entrer dans les détails mathématiques des lois de probabilité des VA continues, nous étudierons essentiellement celle que nous rencontrons le plus souvent en gestion à savoir la loi normale (voir plus loin).**

**Nous proposons toutefois dans l’encadré qui suit quelques développements.**

**Pour aller un peu plus loin**

Une VA continue peut prendre une infinité de valeurs dans un intervalle. La somme des probabilités de cet infini est finie et égale à 1. Pour illustrer ce paradoxe, prenons l’exemple d’une entreprise dont le chiffre d’affaires prévisionnel est compris entre 500 000 et 1 000 000 euros et, nous faisons l’hypothèse qu’il y a équiprobabilité de chaque valeur. Nous pouvons représenter cela par le graphique suivant :

La fonction représentée est la fonction constante f(x) = 1 / (1 000 000 – 500 000) = 0,000002 sur l’intervalle [500 000 , 1 000 000]. Si on calcule l’aire sous la droite, c’est l’aire d’un rectangle, nous obtenons : (1 000 000 – 500 000) \* 0,000002 = 1.

Nous avons représenté ce qu’on appelle la **fonction de densité** (ou **densité de la probabilité**) de la variable aléatoire chiffre d’affaires de notre entreprise. ***Cette fonction ne permet pas de donner la probabilité d’une valeur donnée***. En termes mathématiques, la fonction densité correspond à la dérivée de la fonction de répartition (notion définie dans l’étude des VA discrètes).

À noter qu’une fonction définie sur un intervalle est une **fonction de densité** dès l’instant où **l’aire comprise entre sa représentation graphique et l’axe des abscisses est égal à 1**

La VA aléatoire définie dans l’exemple suit ce qu’on appelle une loi uniforme sur un intervalle. L’aire entre la fonction et l’axes des abscisses est bien égale à 1, c’est donc bien une fonction de densité (ou densité de la probabilité).

Nous admettrons que dans notre cas la fonction de répartition est la suivante, quand x est compris dans l’intervalle [500 000 , 1 000 000] : F(x) = (x – 500 000) / (1 000 000 – 500 000) = (x – 500 000) / 500 000.

Nous pouvons par exemple calculer la probabilité que le chiffre d’affaires soit inférieur ou égal à 700 000 € :

P(X<=700 000) = F(700 000) = (700 000 – 500 000) / 500 000 = 0,40 = 40 %.

La représentation graphique de la fonction répartition se présente comme suit :

On remarque que logiquement F(1 000 000) = P(X<=1 000 000) = 1. Puisque 1 000 000 est le chiffre d’affaires maximum que l’on peut atteindre, la probabilité de réaliser un chiffre d’affaires inférieur ou égal à 1 000 000 est donc certaine soit 1 ou 100 %.

1. Espérance mathématique et variance d’une VA continue

Le calcul de l’espérance et de la variance d’une VA continue nécessite l’utilisation du calcul intégral (mesure des aires en particulier). Il ne nous appartient donc pas de développer ici ces calculs.

Nous préciserons simplement que l’espérance mathématique et la variance (ainsi que l’écart-type) sont des indicateurs similaires à ceux étudiés pour les variables discrètes. ***Ils ont la même signification***.

Par exemple si l’on considère le chiffre d’affaires prévisionnel d’une entreprise, l’espérance mathématique représente le chiffre d’affaires moyen que l’on obtiendrait avec des statistiques sur plusieurs périodes (suffisamment nombreuses) et l’écart-type, la dispersion moyenne autour de l’espérance.

**Nous développerons plus loin l’utilisation de l’espérance et de la variance (et écart-type) d’une VA continue au travers du cas de la loi normale. La loi normale étant la loi de probabilité la plus fréquemment rencontrée.**

1. Les propriétés de l’espérance et de la variance des VA

***Les propriétés développés dans ce qui suit sont vérifiées à la fois pour les VA discrètes et pour les VA continues. Elles sont à bien maîtriser.***

Ces propriétés seront très utiles pour l’étude des distributions d’échantillonnage et de l’estimation. Ces notions font partie du programme de l’UE 11.

**Propriétés de l’espérance mathématique d’une VA**

Soient X et Y deux VA de même nature (soit discrètes, soit continues). Soient a et b deux valeurs (des réels).

Il est démontré que :

**E(X+Y) = E(X) + E(Y) L’espérance d’une Somme de VA est égale à la Somme des espérances de ces VA.**

**E(aX) = a E(X) L’espérance du produit d’un réel par une VA est égale au produit du réel par l’espérance de la VA**

**E(a X + b) = a E (X) + b**

Nous pouvons généraliser ces propriétés en considérant n variables aléatoires notées X1, X2, ..., Xn. Nous avons notamment :

**E(X1 + X2 +…….+ Xn) = E(X1) + E(X2) + …..+ E(Xn)**

Complément : si, et seulement si les deux VA X et Y sont indépendantes alors nous avons également la propriété suivante : **E(X.Y) = E(X).E(Y)**

**Propriétés de la variance d’une VA**

Nous rappelons que l’écart-type est la racine carrée de la variance. Et ***pour éviter toute erreur de calcul, il est fortement conseillé de calculer dans un premier temps la variance, puis l’écart-type, quand bien même la question serait « déterminer l’écart-type »***. L’une des raisons est que la racine d’une somme n’est pas la somme des racines (par exemple racine carrée de 2+2 soit 4, n’est pas égal à racine carrée de 2 + racine carrée de 2).

Soient X et Y deux ***VA indépendantes*** de même nature (soit discrètes, soit continues). Soient a et b deux valeurs (des réels). ***À noter que les propriétés qui suivent sur la variance nécessitent l’hypothèse d’indépendance des VA***.

Il est démontré que :

**V(a X + b) = a2 V (X)**

**V (X + Y) = V (X) + V (Y)**  et donc

Nous pouvons généraliser ces propriétés en considérant ***n variables aléatoires indépendantes*** 2 à 2 notées X1, X2, ..., Xn. Nous avons notamment :

**V(X1 + X2 +…….+ Xn) = V(X1) + V(X2) + …..+ V(Xn)**

Application 1 – Espérance et Variance d’une somme de VA

Une entreprise fabrique et vend des vélos (VTT, VTC, vélos de course sur route, vélos pour enfants,…).

Vous êtes contrôleur-e- de gestion et établissez chaque année un budget des ventes. Vous travaillez actuellement sur la prévision des ventes de VTC, l’un des produits phares de l’entreprise.

A partir de statistiques de ventes, vous avez obtenu les résultats suivants :

Vous retenez l’hypothèse que les nombres annuels de VTC vendus sont des variables aléatoires **indépendantes** 2 à 2 dont les espérances mathématiques (les moyennes) et les écarts-types sont les suivantes :



On désigne par CA la variable aléatoire représentant le chiffre d’affaires annuel HT réalisé sur les ventes de VTC. Et on désigne par R la variable aléatoire représentant la marge sur coût de revient réalisée sur les ventes de VTC.

Travail à faire :

1. Ecrire la variable aléatoire CA en fonction des variables aléatoires VTC1, VTC2 et VTC3.
2. Calculer l’espérance mathématique et l’écart-type de la variable aléatoire CA.
3. Interpréter les résultats.
4. Ecrire la variable aléatoire R en fonction des variables aléatoires VTC1, VTC2 et VTC3.
5. Calculer l’espérance mathématique et l’écart-type de la variable aléatoire R.
6. Interpréter les résultats.
7. Peut-on écrire la variable R en fonction de la variable CA ?
8. Étude d’une loi de probabilité d’une VA discrète : la loi de Bernoulli

**Un peu d’histoire : qui était Bernoulli ?**

L’histoire des sciences comporte « plusieurs Bernoulli ». Celui auquel il est fait référence ici est Jacques (ou Jacob) Bernoulli. Nous citons ci-dessous un extrait de l’encyclopédie universelle éditée chez Larousse.

« Les Bernoulli – famille suisse de savants, originaire d’Anvers, réfugiée à Bâle depuis la fin du XVIe siècle. Jacques 1er (Bâle 1654 – 1705), maîtrisant parfaitement le calcul infinitésimal leibnizien, le compléta en de nombreux points, publia la première intégration d’une équation différentielle et fut à l’origine du calcul des variations. ***Son ouvrage posthume Ars conjectandi (1713) est fondamental pour la théorie des probabilités*** […].

Nous pouvons qualifier la loi de Bernoulli de loi de base car elle permet d’en définir beaucoup d’autres.

Pour la définir, nous partirons d’une l’application.

Application 2

L’EARL (Exploitation Agricole à Responsabilité Limitée) des Terres Noires élève en plein air des volailles BIO et a d’autres activités agricoles. Elle vend également des œufs produits par ses poules élevées en plein air.

Les œufs ont des diamètres variables. Quand le diamètre est inférieur ou égal à 5 cm, il n’est pas commercialisé. Il est utilisé pour la fabrication de biscuits.

Un œuf pondu sur dix dans cette exploitation a un diamètre inférieur ou égal à 5 cm.

On considère la VA notée X suivante « un œuf pondu pris au hasard est commercialisable ».

Travail à faire :

1. Quelles sont les valeurs possibles prises par la VA X ?
2. Préciser la loi de probabilité ou fonction de distribution de X (c’est-à-dire la probabilité de chaque valeur prise par la VA X).
3. Calculer l’espérance mathématique de la VA X.
4. Calculer la variance et l’écart-type de la VA X.

**Généralisons :**

*Soit une VA qui ne prend que deux valeurs possibles 1, pour un succès, ou 0 pour un échec.*

*La probabilité d’obtenir la valeur 1 (ou un succès) est notée p et la probabilité d’obtenir 0 (ou un échec) est notée q. Nous savons que la somme des probabilités est égale à 1 donc p + q = 1 et donc* ***q = 1 – p****.*

*Une telle variable aléatoire est appelée* ***variable de Bernoulli*** *et suit la loi de probabilité dite* ***loi de Bernoulli*** *suivante :*

***P(X=1) = p*** *Probabilité de succès*

***P(X=0) = q = 1 – p*** *Probabilité d’échec*

***E(X) = p*** *Espérance mathématique d’une VA de Bernoulli*

***V(X) = p . q = p . (1-p)*** *Variance**d’une VA de Bernoulli*

*Et donc* ***σ(X) = =****Écart-type**d’une VA de Bernoulli*

***Remarque : une loi de Bernoulli ne dépend donc que d’un seul paramètre, noté p, la probabilité de succès.***

Application 3

La SVR, Société Vendéenne de Roulements, fabrique différents types de roulement à bille pour l’industrie. Le roulement à bille dont la référence est R005 doit avoir un diamètre au moins égal à 5 centimètres pour être jugé conforme. Si la machine est bien réglée, 99 % des roulements sont conformes.

On considère la VA notée X suivante «un roulement à bille prélevé au hasard est conforme».

Travail à faire :

1. Préciser la loi de probabilité suivie par la VA X.
2. Calculer l’espérance mathématique, la variance et l’écart-type de la VA X.
3. De la la loi de Bernoulli à la loi binomiale (loi de probabilité d’une VA discrète)

Nous partirons d’un exemple.

Exemple : Une société imprime pour le compte d’éditeurs différents ouvrages scolaires. La production est en général réalisée par lot de 1 000. En moyenne, sur un lot de 1 000 ouvrages, il y en a 10 qui doivent être mis au rebut (mauvais encrage, pages non imprimées, etc.).

On considère un ouvrage imprimé pris au hasard et on appelle **X1 la VA suivante : « l’ouvrage imprimé prélevé au hasard est défectueux »**. X1 prend donc la valeur 1 si l’ouvrage est défectueux et 0 s’il ne l’est pas.

On répète 5 fois l’épreuve suivante : prélèvement au hasard avec remise d’un livre dans un lot de 1000. On note **X la variable aléatoire suivante : « nombre d’ouvrages défaillants obtenus parmi les cinq »**.

**X1 suit une loi de Bernoulli de paramètre p = 0,01 (1 %), en effet :**

X1 prend deux valeurs possibles, 0 ou 1.

P(X1=1) = 10 / 1000 = 0,01 = p Probabilité que le livre soit défectueux

P(X1=0) = 1 – 0,01 = 0,99 = q = (1-p) Probabilité que le livre ne soit pas défectueux.

E(X1) = 0,01 = p Espérance mathématique ou moyenne

σ(X1) = = = = 0,0995 (arrondi à 4 décimales)

**Etude de la variable aléatoire X :**

Prélever cinq fois au hasard avec remise un livre dans un lot de 1000, c’est répéter cinq fois l’épreuve élémentaire « prélever un ouvrage imprimé pris au hasard parmi 1000 ». Nous pouvons donc écrire que la variable aléatoire D est la somme de cinq variables aléatoires semblables à X1, que nous noterons X1, X2, X3, X4 et X5 : **X = X1 + X2 + X3 + X4 + X5**

Les variables X1, X2, X3, X4 et X5 sont indépendantes entre-elles (on dit en mathématique 2 à 2) car le tirage est effectué sans remise. **Chaque VA Xi suit une loi de Bernoulli de paramètre 0,01**.

Avant d’aller plus loin, apportons une précision sur les notions de tirage avec ou sans remise :

**Tirage avec ou sans remise, quelle différence ?**

Pour que les VA soient indépendantes, il faut que les tirages soient effectuées avec remise, afin que le tirage suivant ne soit pas influencé, c’est-à-dire dépendant, du premier.

Toutefois la population, ici le nombre de livres, est important, la probabilité de tirer le même livre que la première fois est faible, donc si le premier livre tiré n’est pas remis, cela ne change fondamentalement pas le tirage suivant (car nous n’aurions de toute façon probablement pas tiré le même livre).

***En conséquence quand l’univers est grand (nombre d’éléments importants), on pourra considérer qu’un tirage sans remise est équivalent à un tirage avec remise, et nous pourrons donc faire l’hypothèse d’indépendance des VA quelle que soit la nature du tirage.***

**Pour définir la loi de probabilité de X**, précisons dans un premier temps les valeurs possibles que peut prendre X. Les valeurs possibles sont 0, 1, 2, 3, 4 ou 5. En effet, au cours des cinq tirages on pourra n’avoir tiré aucun livre défectueux ou jusqu’à cinq défectueux. Pour définir la loi de probabilité de X, il nous faut trouver la probabilité de chacune des valeurs possibles.

Pour faciliter la présentation des calculs, nous noterons S (pour succès) quand le livre tiré est défectueux et E (pour échec) dans le cas contraire. Par exemple s’il y a aucun livre défectueux le tirage sera donc noté EEEEE, s’il y a un défectueux au premier tirage et 4 non défectueux ensuite, cela donne SEEEE.

Le tableau suivant dresse toutes les combinaisons possibles (nous préciserons le calcul de nombre ensuite afin de pouvoir envisager une généralisation du problème) :



Le tableau ci-dessus nous permet de déterminer la probabilité de chaque résultat élémentaire (ce qui est bien la définition d’une loi de probabilité) :

* **Probabilité que le nombre de défaillant soit égale à 0 sur les 5 tirages ?**

Une seule combinaison possible : EEEEE (voir tableau).

La probabilité d’avoir 0 défaillant à chaque tirage est 0,99 (valeur précisée plus haut).

Comme les VA sont indépendantes, les probabilités se multiplient et donc la probabilité d’obtenir aucun défaillant sur les cinq tirages est de 0,99 \* 0,99 \* 0,99 \*0,99 \* 0,99 \* 0,99 = 0,99^5 (0,99 puissance 5). Donc :

**P(X=0) = 1 \* 0,010 \* 0,995 = 0.9509900499**

* **Probabilité que le nombre de défaillant soit égale à 1 sur les 5 tirages ?**

Cinq combinaisons possibles (voir tableau).

Pour une combinaison, il y a 1 succès et 4 échec, la probabilité d’une combinaison est donc 0,01 \* 0,99^4.

Nous avons donc :

**P(X=1) =** **5 \* 0,011 \* 0,994 = 0.0096059601**

Nous raisonnons de la même façon pour les autres probabilités, nous obtenons :

**P(X=2) =** **10 \* 0,012 \* 0,993 = 0.000970299**

**P(X=3) =** **10 \* 0,013 \* 0,992 = 0.000009801**

**P(X=4) =** **5 \* 0,014 \* 0,991 = 0.0000000495**

**P(X=5) =** **1 \* 0,015 \* 0,990 = 0 (nombre très faible arrondi à 0)**

La loi de probabilité de X est désormais entièrement définie.

Les calculs précédents sont fastidieux, il serait intéressant de disposer d’une formule générale permettant de déterminer la probabilité dans tous les cas par simple application de cette formule. Des mathématiciens se sont penchés sur la question, ont généralisé le problème et ont mis au point cette formule permettant de déterminer la loi de probabilité. ***Cette loi est connue sous le nom de loi binomiale***.

**Pour expliquer cette formule, nous préciserons juste le calcul des combinaisons puisque nous avons quelques notions sur le dénombrement.**

Prenons une combinaison : SSEEE. Elle est composée de deux succès notés et 3 échecs notés. Il y a 5 ! (factorielle 5) façons d’agencer 5 lettres, mais il y a des lettres identiques donc des combinaisons identiques. Vous pouvez montrer facilement que le nombre de combinaison est en fait 5 ! / (2 ! \* 3 !) = = 10. Il ne sera donc pas étonnant de retrouver les combinaisons dans la formule.

**Formalisons et généralisons :**

Nous répétons n fois de suite une même épreuve de Bernoulli, de manière indépendante. Cette épreuve comporte deux résultats possibles, un succès avec une probabilité p ou un échec avec une probabilité q=1-p. Soit X la variable aléatoire définie par le nombre de succès obtenus à l’issue de ces n épreuves.

***La loi de probabilité de X est la loi binomiale de paramètres n et p et notée B (n,p). X prend des valeurs de 0 à n.***

***On écrira : X ~> B (n,p) (X suit une loi binomiale de paramètre n et p***

**Propriétés de la loi binomiale :**

Pour chaque valeur de k (dont la valeur représente le nombre de succès obtenus lors des n épreuves), nous avons : (P(X=k) est la probabilité d’avoir k succès parmi n épreuves)

**** Ou encore avec une autre notation des combinaisons :

****

**Remarque** : ***la seconde notation est la plus utilisée.***

**Espérance : E(X) =n p** n : nombre d’épreuves

**Variance : V(X) = n p q** p : probabilité de succès pour une épreuve

**Ecart-type : σ(X) =** q = 1 – p = probabilité d’échec pour une épreuve

**Complément : Somme de variables aléatoires binomiales indépendantes**

Si deux **variables aléatoires indépendantes** notées X1 et X2 suivent respectivement les lois binomiales suivantes **B** (n, p) et **B** (m, p) alors X1 + X2 suit la loi binomiale **B** (n+m, p).

***Attention****: pour chaque variable la probabilité de succès p doit être la même.*

***Un peu d’histoire : pourquoi le nom « binomiale » ? (pour aller plus loin)***

Tout le monde a eu l’occasion d’ânonner les identités dites remarquables comme (a+b)² = a² + 2ab + b² (en général nous prenions a, à la place de p, et b, à la place de q).

Cette identité ou propriété remarquable a été généralisée avec une puissance quelconque et est connue sous le nom de binôme de Newton :



Si l’on prend un exemple, on obtient par exemple en prenant n = 5 (on note p à la place de a et q à la place de b) : (p+q)5 = p5 + 5p4q + 10p3q2 + 10p2q3 + 5pq4 + q5

On remarque que les coefficients de chaque membre de la somme ci-dessus sont respectivement 1, 5, 10, 10,5 et 1. ***Ce sont les combinaisons que nous avons trouvé dans l’exemple développé plus haut, et que nous retrouvons dans la loi binomiale***.

Vous avez compris, de binôme à binomiale, il n’y avait qu’un pas qui a été franchi. En clair la loi binomiale doit son nom au binôme de Newton.

Un autre mathématicien, Pascal, a développé un outil connu sous le nom de triangle de Pascal et qui permet de trouver facilement toutes les valeurs du nombre de combinaisons lorsque l’on ne dispose pas d’une calculatrice ; Ce triangle se présente comme suit (extrait) :



Pour comprendre l’élaboration et le fonctionnement de ce triangle, vous trouverez facilement sur Internet des sites et des vidéos explicatifs. Vous pouvez déjà dans un premier temps essayer de décrypter.

Application 4 – Calcul de probabilités et représentation de la loi binomiale

L’entreprise ISIS fabrique un composant destiné à une entreprise aéronautique (ISIS est sous-traitant et dépend d’un donneur d’ordre). Ce composant doit être de qualité irréprochable et notamment respecter les dimensions au dixième de millimètre près. Le cahier des charges fourni par le donneur d’ordre est très précis.

L’entreprise ISIS contrôle donc l’intégralité de sa production et rejette systématiquement tout composant ne répondant pas aux attentes du client. Le **taux de rejet** est donc particulièrement élevé puisqu’il est de **10 %**. Il est produit **100 composants par jour**. (La matière des composants rejetés est réutilisée).

On prend un composant au hasard dans la production d’une journée, on appelle Xi la VA suivante : « le composant est rejeté ».

Travail à faire :

1. Préciser la loi de probabilité suivie par la VA Xi.
2. On appelle X la VA représentant le nombre de composants rejetés dans la production d’une journée. Préciser la loi de probabilité suivie par la VA X.
3. (Pour répondre à cette question, l’utilisation d’un tableur est fortement conseillée). Calculer les probabilités pour X=0 à 30.
4. Représenter graphiquement les probabilités calculées précédemment.
5. Commenter la forme de la courbe.

Application 5 – Utilisation des tables de la loi binomiale

Il est souvent plus aisé de consulter les tables de la loi binomiale que d’utiliser sa calculatrice pour calculer les probabilités. Cette application a donc pour objectif de se familiariser avec ces tables.

Travail à faire :

Soit X une variable aléatoire qui suit la loi binomiale B(10 ; 0,1) (c’est-à-dire p=0,1 et n=10).

Calculer les probabilités suivantes en utilisant notamment les tables de la loi binomiale :

1. P(X=0) ; P(X=1) ; P(X=2) ; P(X=3) ; P(X=9) ; P(X=10).
2. P(X<=3) ; P(X>3) ; P(X<9).
3. De la loi binomiale à la loi de Poisson (loi de probabilité d’une VA discrète)

Les calculs de probabilité avec la loi binomiale peuvent vite devenir fastidieux, aussi il peut être intéressant d’utiliser une approximation par d'autres lois de probabilité, et notamment la loi de Poisson.

La loi de Poisson peut en effet être vue également comme l’approximation d’une loi binomiale.

**Remarque** : La loi de Poisson a été définie dans un premier temps indépendamment de la loi binomiale. Il s’est avéré par la suite que la loi binomiale convergeait vers la loi de Poisson. Et donc lorsque certaines conditions sont réunies (elles sont précisées plus loin), la loi binomiale peut être approximée à une loi de Poisson.

***Un peu d’histoire : qui était Poisson ? (source : Encyclopédie Larousse)***

« Siméon Denis Poisson est un mathématicien français (1781 – 1840). Il est un des créateurs de la physique mathématique. **Il a appliqué l’analyse mathématique** à la mécanique céleste et à la théorie de l’attraction, à celle de la chaleur, à l’électricité, à l’élasticité, à la lumière, au magnétisme et **au calcul des probabilités** ».

En probabilité, on lui doit la loi du même nom.

**La loi de poisson – Loi de probabilité d’une VA discrète**

La loi de Poisson s'applique à des variables aléatoires X prenant des valeurs discrètes k (et non continues ; *Rappel* : k ne peut pas prendre toutes les valeurs dans un intervalle).

**Soit X une VA, on dit qu’elle suit une loi de Poisson de paramètre λ** (**notée : X ~> P (λ)**), **si et seulement si pour toute valeur entière de k (autrement dit k appartient à l’ensemble des entiers naturels), nous avons :**



P(X=k) désigne la probabilité que la variable X prenne la valeur k. Il existe des tables permettant de donner les probabilités pour différentes valeurs de k. (voir applications).

*Le paramètre λ est toujours strictement positif*.

**Espérance : E(X) = λ**

**Variance : V(X) = λ**

**Écart-type : σ(X) =**

**Complément : Somme de variables aléatoires indépendantes suivant des lois de Poisson**

Si deux **variables aléatoires indépendantes** notées X1 et X2 suivent respectivement les lois de Poisson suivantes **P** (**λ1**) et **P** (**λ2**) alors X1 + X2 suit la loi de Poisson **P** (**λ1 + λ2**).

**Conditions d’approximation d’une loi binomiale par une loi de Poisson**

***Si X suit une loi binomiale B (n,p), X peut être approchée par une loi de Poisson avec = np :***

***Si n>= 30 et p < 0.1 et np<15***

*Remarque : ces conditions d’approximations ne sont pas universelles et peuvent varier en fonction des auteurs. Elles sont toutefois en général assez proches.*

Application 6 – Calcul de probabilités à partir d’une loi binomiale, approximation à la loi de Poisson et comparaison

La SVR, Société Vendéenne de Roulements, fabrique différents types de roulement à bille pour l’industrie. En moyenne 5% des roulements sont défectueux. La production quotidienne est d’environ 50 000 roulements. Chaque jour, on prélève un roulement au hasard parmi cette production et on répète 100 fois ce prélèvement (en clair vous constituez chaque jour un échantillon de 100 roulements). Comme la production est très grande, **l’hypothèse d’indépendance des prélèvements peut être faite**.

On appelle X la variable aléatoire suivante : « nombre de roulements défectueux sur un prélèvement de 100 roulements ».

Travail à faire :

1. Préciser la loi de probabilité suivie par la VA X.
2. Calculer les probabilités suivantes :
	1. P(X=4)
	2. P(X=6)
	3. P(X=10)
	4. P(X<=5)

**Remarque** : les calculs pourront être menés à partir d’une calculatrice, d’une table ou d’un tableur.

1. Préciser pourquoi la loi de X peut être approximée par une loi de Poisson. Préciser le paramètre de cette loi.
2. Calculer de nouveau les probabilités précédentes et dresser un tableau de comparaisons et commenter.

La loi de Poisson utilisée indépendamment de la loi binomiale

La loi de Poisson est souvent utilisée comme approximation de la loi binomiale. Toutefois son utilisation se justifie dans de nombreux problèmes de hasard indépendamment de la loi binomiale.

La distribution (loi) de Poisson est une distribution utilisée lorsqu’il s’agit de déterminer des probabilités de succès relatives à des évènements que l’on considère comme répartis dans le temps ou dans l’espace.

« On utilise l’expression processus de Poisson pour désigner un processus qui, en tous points, est analogue à celui de Bernoulli, sauf qu’ici les évènements sont étudiés dans leur apparition sur un intervalle continu au lieu d’être relié à des épreuves discontinues. »

**Exemples :**

* L’arrivée de camions à un quai pour chargement ;
* L’arrivée de clients à une caisse ;
* Le nombre d’erreurs sur une page d’un document ;
* Les ventes d’un produit donné dans un magasin ;
* Les pannes de machine.

Dans les exemples ci-dessus, on s'intéresse à un grand nombre d'événements indépendants et on observe qu’ils se produisent en moyenne un certain nombre de fois (le λ de la loi de Poisson) durant un intervalle de temps donné. La loi de Poisson indique la probabilité que l'événement se produise seulement k fois exactement durant cette période.

**Pour donner du sens** : Cette loi peut par exemple être utilisée pour gérer les caisses d’un hypermarché. Si l’on connait les fréquentations moyennes horaires du magasin, grâce à des statistiques (étude des fréquentations passées), on peut en déduire à partir de la loi de Poisson le nombre de caisses à ouvrir pour limiter le temps d’attente du client. En fonction des horaires de la journée, le nombre de caisses ouvertes sera modulé. Quelle que soit l’heure de la journée, le temps d’attente sera similaire pour la satisfaction du client. Nous aurions également pu prendre un cas d’attente à un guichet de l’administration et nombre de guichets ouverts.

Application 7 – Les pannes de machine

La SVR, Société Vendéenne de Roulements, qui fabrique différents types de roulement à bille pour l’industrie s’interroge sur l’organisation de son service maintenance. Cette entreprise fonctionne 16 heures par jour par équipes de 2 fois 8 heures. Elle travaille 7 jours sur 7. Les équipes qui acceptent de travailler les samedis et dimanches bénéficient d’un aménagement avantageux du temps de travail.

Pour limiter les arrêts de production, la société dispose d’un service de maintenance capable d’intervenir sur les machines à tout moment. La société envisage de recruter pour accélérer la réparation des machines et réduire fortement les temps d’arrêt des machines en maintenance.

Par le passé, il a été constaté qu’au cours d’une période de 8 heures de travail, le nombre moyen de pannes nécessitant l’intervention du service de maintenance était de 4. Ces pannes surviennent totalement au hasard et sont indépendantes entre-elles.

On appelle X la VA « nombre de pannes sur une période de 8 heures ». Le responsable de la maintenance estime que X suit une loi de Poisson

Travail à faire :

1. Calculer le paramètre de la loi de Poisson suivie par X.
2. Calculer l’espérance mathématique, la variance et l’écart-type.
3. Calculer la probabilité que le nombre de pannes sur une période de 8 heures soit :
	1. Égal à 2 ;
	2. Égal à 4 ;
	3. Inférieur ou égale à 4 ;
	4. Strictement supérieur à 4.

**Remarque** : pour le calcul des probabilités, vous pouvez utiliser votre calculatrice ou les tables de la loi de Poisson.

1. Des lois binomiale et de Poisson à la loi normale (loi de probabilité d’une VA continue)

La loi normale encore appelée Loi de Laplace-Gauss (du nom de ces deux mathématiciens) ou parfois simplement loi de Gauss est une loi de probabilité très utilisée dans de nombreux domaines et particulièrement par les chercheurs (en chimie, en physique, en médecine, en économie, en gestion, etc.).

***Vous trouverez plus loin un peu d’histoire sur cette loi et l’origine de son nom.***

***La loi normale est d’abord une loi de probabilité d’une VA continue, c’est-à-dire que les valeurs prises par la variable sont comprises dans un intervalle et ne sont pas dénombrables. Pourquoi alors indiquer dans le titre, « de la loi binomiale et de poisson à la loi normale » dans la mesure où les deux premières sont des lois de probabilités discrètes ? Nous verrons plus loin que sous certaines conditions, il est possible d’approximer les lois binomiales et de Poisson par la loi normale.***

**Nous approcherons dans un premier temps la loi normale à partir de la représentation graphique de statistiques dans des exemples.**

Application 8 – Représentation graphique des notes de contrôle de gestion obtenues par des étudiants à un examen

Les notes obtenues à l’épreuve de contrôle de gestion au cours d’un examen national s’échelonnent de 4 à 17. Il y a eu 1000 copies corrigées. Il a été dressé le tableau statistique suivant :



Lecture du tableau : par exemple 2 copies ont obtenu la note de 4, 86 copies la note de 11, 4 copies la note de 17, etc.

Travail à faire :

1. Calculer les fréquences. Puis calculer la moyenne et l’écart-type des notes obtenues à l’épreuve de contrôle de gestion. (Les calculs pourront être menés à partir d’un tableur).
2. Faire une représentation graphique permettant de visualiser le profil des notes obtenues. Mettre en abscisse la note obtenue et en ordonnées la fréquence. (Il est conseillé d’utiliser un tableur pour réaliser le graphique).
3. Commenter la forme de la courbe.

Application 9 – Représentation graphique des ventes d’un produit au cours d’une année

Une entreprise a relevé pendant 300 jours ouvrables consécutifs, le chiffre d’affaires réalisé par l’un de ses points de vente. Les données sont présentées dans le tableau ci-dessous.



Nous remarquons que les chiffres d’affaires réalisés vont de 145 K€ à 210 K€. La colonne nombre de jours indique le nombre de fois où le chiffre d’affaires a été observé dans l’intervalle de la 3ème colonne.

Comme il s’agit de l’examen de statistiques, la précision par intervalle est suffisante pour l’analyse. Pour le calcul de la moyenne et de l’écart-type, nous pouvons par approximation nous baser sur le centre de l’intervalle (4ème colonne).

Travail à faire :

1. Calculer les fréquences (arrondir à 3 chiffres après la virgule). Puis calculer la moyenne et l’écart-type des chiffres d’affaires observés. (Les calculs pourront être menés à partir d’un tableur).
2. Faire une représentation graphique permettant de visualiser le profil des chiffres d’affaires. (Il est conseillé d’utiliser un tableur pour réaliser le graphique).
3. Commenter la forme de la courbe.

Quelles observations et commentaires pouvons-nous retirer des deux applications précédentes ?

Dans le premier cas, nous avons travaillé sur une variable discrète, les notes peuvent aller de 0 à 20 mais sont arrondies au demi-point. Donc l’univers des possibles (nombre de notes possibles) est dénombrable.

Dans le second cas, nous avons travaillé sur une variable continue, le chiffre d’affaires pouvant prendre n’importe quelle valeur dans un intervalle. Le nombre de chiffre d’affaires possibles n’est pas dénombrable.

Mais ce qu’il y a de commun entre les deux, c’est la forme de la représentation graphique, elles sont similaires quant à leur forme.

Nous pourrions trouver multiples exemples dans la réalité et dans divers domaines (gestion, physique, chimie, etc.) ou la représentation des fréquences des variables conduit à la même forme de courbe.

Il n’est donc pas surprenant que les mathématiciens se soient penchés sur l’élaboration d’une fonction de densité (on sait que des statistiques aux probabilités, il n’y a qu’un pas) qui permet d’obtenir une telle représentation graphique. Nous donnons un peu plus loin cette fonction de densité.

Il n’est pas non plus surprenant de pouvoir approximer des lois discrètes telles que Binomiale ou Poisson à une loi continue, la loi normale.

***La forme de la courbe est souvent appelée courbe en cloche ou courbe en chapeau de gendarme.***

**Un peu d’histoire – Pourquoi « loi normale » ?**

**Source :** <https://fr.wikipedia.org/wiki/Histoire_de_la_loi_normale>

« La première attribution du terme « normal » pour la loi est attribuée à [Henri Poincaré](https://fr.wikipedia.org/wiki/Henri_Poincar%C3%A9) qui énonça pendant un de ses cours en 1893 : « Je dirai, pour abréger, que la loi de probabilité est normale, lorsque la valeur de la probabilité est représentée par cette intégrale. ». En 1894, [Irving Fisher](https://fr.wikipedia.org/wiki/Irving_Fisher) écrivit une phrase sensiblement différente : « Lorsqu'une loi d'écarts est exprimée par cette intégrale nous disons que la probabilité est normale ». ***Le terme « normal » vient du fait que cette loi apparaît souvent dans la nature et que de toutes les lois qui apparaissent naturellement, la loi normale est la plus courante et la plus adaptée aux observations***. [Karl Pearson](https://fr.wikipedia.org/wiki/Karl_Pearson) explique en 1893 le choix du terme « normal » pour la loi et la courbe par la facilité de ne pas fixer de paternité.

Puisque la première apparition de la loi normale s'est faite par l'observation de la courbe de sa densité de probabilité, le nom de la courbe sert parfois à définir la loi. Pour mieux donner une image de sa forme, cette courbe est parfois imagée en « chapeau de gendarme vu de face », « cloche plate » ou encore « boa qui a avalé un dromadaire ».

Autour de 1950, la commission de terminologie statistique de l'[Afnor](https://fr.wikipedia.org/wiki/Afnor), probablement emmené par [Fréchet](https://fr.wikipedia.org/wiki/Maurice_Ren%C3%A9_Fr%C3%A9chet), décida de normaliser le terme « loi de Laplace ». Le polytechnicien [André Pallez](https://fr.wikipedia.org/w/index.php?title=Andr%C3%A9_Pallez&action=edit&redlink=1) ajoute :

« La Commission, considérant que Laplace a découvert la loi qui devrait porter son nom et qui porte celui de Gauss, à une époque où Gauss était encore un jeune enfant, a rétabli la vérité en rendant à Laplace l’hommage qui lui était dû. ».

Cependant, aujourd'hui au XXIe siècle, les deux noms les plus utilisés sont « loi de Gauss » et « loi normale ». Le nom de Gauss est resté plus que les autres, grâce, entre autre, à l'influence de l'ouvrage pédagogique de [Joseph Bertrand](https://fr.wikipedia.org/wiki/Joseph_Bertrand) publié à la fin du XIXe siècle. ».

Il nous faut désormais définir cette loi continue appelée loi normale

« Dès qu'un phénomène est la superposition d'un grand nombre de causes aléatoires indépendantes, une cloche se présente. »

Soit X une variable aléatoire continue. X prend donc n’importe quelle valeur dans un intervalle.

On dit que X suis une loi normale (ou loi de Gauss ou loi de Laplace-Gauss ou loi Gaussienne) de paramètre **m (la moyenne ou espérance mathématique de X)** et **σ (l’écart-type de X)** quand elle a pour densité de probabilité la fonction suivante :

**Notation** : X → N(m ; σ)

**Représentation graphique générale** :



**Rappel** : la fonction densité ne permet pas de calculer la probabilité pour une valeur donnée de la variable X, cette probabilité est nulle car les valeurs prises par X ne sont pas dénombrables. Nous ne calculerons que des probabilités par exemple du type P(X<=a), a étant un nombre réel quelconque. Nous utiliserons donc la fonction de répartition de X. ***Nous donnons ci-dessous la définition de la fonction de répartition sachant que nous pourrons nous en passer dans les calculs (voir plus loin)*** :

Précisions importantes

***Pour rassurer nos lecteurs, nous tenons à préciser que pour faire les calculs de probabilité, nous n’aurons pas besoin d’utiliser les définitions des fonctions de densité ou de répartition. Des tables de loi normale ou la calculatrice viendront à notre secours.***

Une loi normale particulière : la loi normale centrée réduite N(0;1)

**Il est d’usage de noter T la variable aléatoire notée N (0,1) qui suit la loi normale centrée réduite. Elle a pour moyenne 0 et pour écart-type 1. E (T) = 0 σ (T) =1**

Si X → N(m ; σ) alors nous avons T = (X-m) / σ

Quel est l’intérêt de la loi normale centrée réduite ?

* Toute loi normale quelconque peut être « transformée » en la loi normale centrée réduite grâce à la relation T = (X-m) / σ. Or il existe des tables de la loi normale centrée réduite.
* Conséquence : ***pour calculer une probabilité avec une loi normale quelconque, il suffit d’opérer une transformation pour se ramener à la loi normale centrée réduite afin de pouvoir utiliser ces tables***.

**Propriétés fondamentales de la loi normale centrée réduite : (voir représentations graphiques en dessous de cet encadré)** (t est une valeur quelconque) (et voir remarque à la fin de cet encadré)

* **P(T<=0) = P(T>0) = 0,5** Autre notation : **Π(0) = 0,5 = 1 - Π(0)**
* **P(T>t) = 1 – P(T<=t)** Autre notation : **1 – P(T<=t) = 1 - Π(t)**
* **P(T<-t) = 1 – P(T<=t)** Autre notation : **Π(-t) = 1 - Π(t)**
* **P(T>-t) = P(T<t) = Π(t)** (cette propriété résulte logiquement des deux précédentes)
* On considère l’intervalle [t1 ; t2], t1 et t2 sont deux réels (t2 est bien supérieur à t1), nous pouvons écrire que : **P(t1<T<t2) = P(T<t2) – P(T<t1)** ou encore = **Π(t2) – Π(t1)**
* Conséquence : **P(-t<T<=t) = P(T<=t) – P(T<-t) = P(T<=t) – (1 – P(T<=t)) = 2 P(T<=t) – 1 = 2Π(t) – 1**
* **P (-t < T < t) = P (T < t) – P (T < -t) = P (T < t) – (1 – P (T < t) = 2 P (T < t) – 1 = 2Π( t ) - 1**

***Fondamental : apprenez à utiliser la courbe, vous retrouverez facilement toutes ces propriétés grâce à la symétrie. Il vous suffit de tracer schématiquement la courbe et de retenir que sous la courbe, la probabilité totale est de 1 et que de part et d’autre de l’axe des ordonnées, le cumul des probabilités est de 0,5 (voir schéma page suivante).***

***Remarque : comme la probabilité d’une valeur particulière est nulle, on pourra écrire que P(X<t) = P(X<=t) ou encore que P(X>t) = P(X>=t). Cela facilitera l’application des propriétés et les calculs dans les applications.***

***Voici une représentation graphique de la loi normale centrée réduite. La courbe est symétrique, comme l’aire sous la courbe est égale à 1 (la somme des probabilité de l’ensemble des valeurs fait toujours 1), l’axe des ordonnées partage la probabilité en deux partie égales, donc la partie hachurée qui représente P(T<=0) est égale à 0,5 (T étant la VA qui suit N(0 ;1)).***



Ci-dessous la représentation graphique qui permet de visualiser des propriété de la loi normale centrée réduite (ces propriétés résultent directement de la symétrie de la courbe) :



**Remarque : nous donnons ci-dessous les fonctions à titre indicatif, nous ne les utiliserons pas dans la résolution des problèmes)**

****

La fonction densité de la loi normale centrée réduite :



La fonction de répartition de la loi normale centrée réduite :

Il est d’usage de noter la fonction de répartition de la loi normale centrée réduite **Π(t) = P(T<=t) t étant une valeur réelle.**

Application 10 – Avec corrigé

Suite une VA X qui suit une loi normale de moyenne 500 et d’écart-type 50 : **X → N(500 ; 50)**.

Travail à faire :

1. Calculer les probabilités suivantes : (les calculs pourront être faits dans un premier temps à la calculatrice puis en utilisant la loi normale centrée réduite)
	1. P(X<=550)
	2. P(X<=600)
	3. P(X<=650)
	4. P(X<=700)
	5. P(X>550)
	6. P(X<=450)
	7. P(450<X<=500)

Proposition de correction de l’application 10 (utilisation des tables de la loi normale centrée réduite)

On désigne par T la loi normale centrée réduite.

1. P(X<=550) = P((X-500)/50 <= (550-500)/50) = P(T <= 1) = Π(1)

Dans la table nous lisons : P(T <= 1) = Π(1) = **0,8413 (soit 84,13 %)**

1. P(X<=600) = P((X-500)/50 <= (600-500)/50) = P(T <= 2) = Π(2)

Dans la table nous lisons : P(T <= 2) = Π(2) = **0,9772 (soit 97,72 %)**

1. P(X<=650) = P((X-500)/50 <= (650-500)/50) = P(T <= 3) = Π(3)

Dans la table nous lisons : P(T <= 3) = Π(3) = **0,99865 (soit 99,865 %)**

**Remarque** : la probabilité qu’une VA normale s’écarte de plus ou moins trois écart-type peut être considérée comme quasi nulle.

1. P(X<=700) = P((X-500)/50 <= (700-500)/50) = P(T <= 4) = Π(4)

Dans la table nous lisons : P(T <= 4) = Π(4) = **0,999968 (soit 99,9968 %) Quasi 1 (100 %)**

1. P(X>550) = P((X-500)/50 > (550-500)/50) = P(T > 1) = 1 - P(T <= 1) = 1 - Π(1)

Soit 1 - 0,8413 = **0.1587 (15,87 %)**

***Les tables ne donnent que P(X<=t), il faut donc utiliser les propriétés.***

1. P(X<=450) = P((X-500)/50 > (450-500)/50) = P(T <= -1) = 1 - P(T <= 1) = 1 - Π(1)

Soit 1 - 0,8413 = **0.1587 (15,87 %)**

Remarque : P(T>-t) = P(T<t) = Π(t)

1. P(450<X<=500) = P(450-500)/50<(X-500)/50<550-500)/50) = P(-1<T<1) = 2P(T<1) – 1 = 2\*0,8413 – 1 = **0,6826 (68,26 %)**

*Ou encore en utilisant les résultats précédents :*

P(450<X<=500) = 1 - P(X<=450) - P(X>550) = 1 – 0,1587 – 0,1587 = 0,6826

Application 11 – Calcul de probabilités – VA qui suit une loi normale – Utilisation des tables de la loi normale centrée réduite

Après étude statistique, il apparait que la quantité mensuelle de vélos électriques pour dames, vendus par un fabriquant et notée V, est une variable aléatoire qui suit une loi normale de moyenne 2 000 et d’écart-type 200 : **V → N(2 000 ; 200)**.

Travail à faire :

1. Calculer les probabilités suivantes : (les calculs pourront être faits dans un premier temps en utilisant la loi normale centrée réduite puis à la calculatrice)
	1. P(X<=2000)
	2. P(X<=2600)
	3. P(X<=2800)
	4. P(X>2000)
	5. P(X>2600)
	6. P(1600<X<=2400)
	7. P(1800<X<=2400)

Somme de lois normales indépendantes – Opérations sur les variables aléatoires normales – Propriétés fondamentales à retenir

Soient X1 et X2 deux lois normales **indépendantes** :

* X1 → N(m1 ; σ1)
* X2 → N(m2 ; σ2)

Soient a et b deux réels

**Produit d’une loi normale par un réel**

**aX+b → N(am+b ; ) = N(am+b ; |a|σ)**

|a| désigne la valeur absolue de a.

La variable aléatoire a X + b suit la loi normale d'espérance am+b et d'écart type |a|σ (la variance est a2 σ2).

Ajouter le réel b ne modifie pas la valeur de la variance et de l’écart-type : aX → N(am ; ).

**Somme de lois normales indépendantes**

**X1 + X2 → N(m1+m2 ; )**

L’indépendance est nécessaire pour affirmer que X1 + X2 suit une loi normale de moyenne m1+m2 et d’écart-type.

**Attention** : l’écart-type n’est pas la somme des écarts-types. Pour les calculs, il faut toujours passer par la variance. La variance est σ12 + σ22, l’écart-type est obtenu en prenant la racine carrée.

**Différence de lois normales indépendantes**

**X1 - X2 → N(m1-m2 ; )**

Remarque : l’écart-type est le même que dans le cas de la somme. Seule la moyenne est différente (m1-m2).

Application 12 – Utilisation de la loi normale pour calculer la probabilité d’atteindre un seuil de rentabilité

Un ancien coureur cycliste s’est reconverti dans la fabrication artisanale de vélos de course sur route. Sa clientèle est composée d’amateur-e-s qui pratiquent la compétition.

Il propose deux types de vélo :

* un vélo de course classique noté V1 et vendu 4 000 euros HT ;
* et un vélo plutôt destiné aux triathlètes noté V2 et vendu 4 500 euros HT.

Le taux de marge sur coût variable sur chaque vélo est en moyenne de 30 % (les coûts variables sont sensiblement les mêmes pour chaque type de vélo).

Les charges de structure (fixes) annuelles s’élèvent à environ 150 000 euros HT.

Des études statistiques ont montré qu’il était possible de considérer que les ventes annuelles Q1 et Q2 de cette entreprise suivent les lois normales suivantes : (le premier paramètre étant la moyenne et le second l’écart-type).

* Q1 → N(200 ; 40) Q1 est la VA représentant le nombre de vélos V1 vendus ;
* Q2 → N(140 ; 50) Q2 est la VA représentant le nombre de vélos V2 vendus.

Les variables aléatoires Q1 et Q2 sont indépendantes.

On désigne par R la variable aléatoire représentant le résultat réalisé sur les ventes de vélos.

Travail à faire :

1. Écrire la variable aléatoire R en fonction des variables aléatoires Q1 et Q2.
2. Préciser la loi de probabilité suivie par la variable aléatoire R et calculer ses paramètres.
3. Calculer la probabilité d’atteindre le seuil de rentabilité.
4. Trouver un intervalle centré autour de la moyenne dans lequel nous trouverons dans 95 % des cas le résultat de l’entreprise.

Approximation des lois binomiale et de Poisson par la loi normale (à connaître)

Nous avons pu vérifier sur des exemples que :

* la représentation graphique d’une loi binomiale (la fonction de distribution d’une VA qui suit une loi binomiale) à l’allure de la courbe d’une loi normale (fonction densité de la loi normale) ;
* la loi binomiale peut être approximée par une loi de Poisson.

Il n’est donc pas surprenant que des mathématiciens aient montré que sous certaines conditions, nous pouvions approximer une loi binomiale et une loi de Poisson par une loi normale.

Remarque :

* ***Il s’agit bien ici d’approximer des lois discrètes (binomiale et Poisson) par une loi continue (normale). Toutefois la correction de continuité n’est pas au programme.***
* ***Nous vous conseillons de visionner la vidéo que vous trouverez sur le net à partir du lien suivant :***

[***https://www.youtube.com/watch?v=jCcHWzI0uqA***](https://www.youtube.com/watch?v=jCcHWzI0uqA)

***Cette vidéo explique également l’utilisation des calculatrices de marque Texas et Casio pour le calcul des probabilités avec la loi normale.***

**Nous pouvons résumer ces approximations par le schéma suivant :**

*Remarque : ces conditions d’approximations ne sont pas universelles et peuvent varier en fonction des auteurs. Elles sont toutefois en général assez proches.*



**Ces approximations peuvent également être résumées sous la forme d’un tableau :**

|  |
| --- |
| **Conditions d’approximation d’une loi binomiale** |
| **Si X suit une loi binomiale B(n,p)** | **X peut être approchée par une loi** |
| Si n>= 30 et p < 0.1 et np<15 | de Poisson P() avec = np |
| n>= 30 et p **et** q non voisins de 0**ou**np> et nq>15 | Normale N(np,) |
| **Condition d’approximation d’une loi de Poisson** |
| **Si X suit une loi de Poisson P()**  | **X peut être approchée par une loi** |
| Si >15 | Normale N(,) |

Application 13 –Approximation d'une loi binomiale par une loi normale

Une entreprise sous-traitante fabrique des composants de haute précision pour l’industrie aéronautique. Chaque produit est contrôlé à la sortie de l’atelier de fabrication. La probabilité qu’un produit soit défectueux est de 5 % (0,05). La production s’élève à 500 composants par jour.

On appelle X la variable aléatoire qui représente le nombre de produits défectueux par jour de production.

Travail à faire :

1. Déterminer la loi de probabilité suivie par X.
2. Calculer l’espérance mathématique E(X), la variance V(X) et l’écart-type σ(X) arrondi à 2 décimales.
3. Vérifier que la loi de X peut être raisonnablement approximée par une loi normale.
4. Calculer la probabilité que le nombre de produits défectueux par jour soit inférieur ou égal à 35.

Application 14 – Approximation d'une loi de Poisson par une loi normale

Par le passé, il a été constaté qu’au cours d’une période de 90 jours de travail, le nombre moyen de pannes de machines nécessitant l’intervention du service de maintenance d’une entreprise était de 20. Ces pannes surviennent totalement au hasard et sont indépendantes entre-elles.

On appelle X la VA « nombre de pannes sur une période de 90 jours ». Le responsable de la maintenance estime que X suit une loi de Poisson.

Travail à faire :

1. Calculer l’espérance mathématique E(X), la variance V(X) et l’écart-type σ(X) arrondi à 2 décimales.
2. Vérifier que la loi de X peut être raisonnablement approximée par une loi normale.
3. Calculer la probabilité que le nombre de pannes sur une période de 90 jours soit :
	1. supérieur ou égal à 25 ;
	2. inférieur à 15.

**Complément :**

Un théorème fort utile : le théorème central limite (complément)

Nous compléterons cette partie par un théorème qui nous sera fort utile quand nous aborderons les notions de distribution d’échantillonnage (également au programme du DCG dans le cadre du chapitre sur la gestion de la qualité). Nous n’aurons pas à utiliser le théorème proprement dit, mais les conséquences de celui-ci. Il est donc intéressant de l’évoquer.

Nous pouvons résumer ce théorème comme suit (sans l’exprimer dans toute sa rigueur, ce qui ne nous serait pas utile) :

Soit *X*1, *X*2, …, Xn, n variables aléatoires indépendantes qui suivent la même loi de probabilité notée L(m, σ), m étant la moyenne (espérance mathématique) et σ l’écart-type.

Soit la variable aléatoire **Sn = X1 + X2 + … + Xn**

**Lorsque n est suffisamment grand (en pratique n >= 30) alors Sn suit une loi normale de moyenne nm et d’écart-type σ (racine de n multiplié par σ).**

Autrement dit la connaissance de la loi suivie par chaque variable Xi n’est pas nécessaire pour dire que la somme des n variables aléatoires suit une loi normale quand le nombre de variables est important.