Stage « Probabilités au lycée », recueil d’exercices

# Statistiques descriptives (Seconde, Première)

## Exercice 1 : fréquences cumulées, diagramme en bâtons, utilisation du tableur*Source : Manuel de Seconde*

Les biologistes s’intéressent à certaines caractéristiques des perce-oreilles et en particulier à la longueur des pinces. Pour faire cette étude, ils mesurent la longueur (en mm) des pinces de 585 perce-oreilles adultes. Les résultats bruts sont donnés dans le tableau joint.

1. Que peut-on dire de ces données ?
2. Compléter le tableau des effectifs cumulés croissants.
3. Ranger ces données en 8 classes de 1mm de largeur (compléter les effectifs).
4. Calculer les fréquences correspondantes.
5. Construire le diagramme en bâtons des effectifs rangés en classe.
6. Faire un autre rangement, en classes de 0,5 mm de largeur, et comparer avec le précédent. Que peut-on penser finalement ?

## Exercice 2 : médianes et quartiles, diagramme en boîte, utilisation du tableur*Source : Document ressource de la classe de Seconde,*

On donne dans la feuille de calcul jointe le nombre d’habitants de chaque commune des départements de Charente-Maritime (17) et des Yvelines (78).

Donner des éléments de comparaison entre les communes de ces départements.

## Exercice 3 : notion de variable, résumé d’une série de données, utilisation du tableur ou GeoGebra*Source : Atelier Probabilités Statistiques (2011) présenté aux JDI (Académie de Grenoble)*

Question : le tabagisme de la mère a-t-il une influence sur le poids du bébé à la naissance ?

L’étude sur une population de 1226 bébés, les données sont présentées dans la feuille de calcul jointe.

1. Identifier les individus statistiques et les variables.
2. Distinguer les variables qualitatives des variables quantitatives.
3. On distingue la sous-population des mères non-fumeuses de celle des mères fumeuses. Résumer les deux séries de données « poids des bébés » à l’aide des indicateurs du cours.
4. Comparer les deux boîtes à moustaches.
5. Calculer les quartiles et ainsi que la médiane du poids des enfants dans la population générale.
6. On dit que le poids P du bébé est :
* faible lorsque
* plutôt faible lorsque
* plutôt élevé lorsque
* élevé lorsque
1. Compléter le tableau à double-entrée avec les effectifs :

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | faible  | plutôt faible  | plutôt élevé  | élevé  | Total  |
| Non fumeuse |  |  |  |  |  |
| Fumeuse |  |  |  |  |  |
| Total  |  |  |  |  |  |

1. Pourquoi ne retrouve-t-on pas des effectifs égaux dans les classes de poids ?
2. Calculer la fréquence des bébés ayant un poids faible dans chacune des populations. Commenter.

(Suite de l’exercice dans la partie Echantillonnage de Première)

# Probabilités (Seconde)

## Exercice 4 : simulation sur tableur ou calculatrice*Source : Document ressource Bac Professionnel*

Deux points A et B sont pris « au hasard » sur un segment de longueur 1.

Quelle est la probabilité de l’événement : «  la longueur AB est supérieure à 0,5 » ?

## Exercice 5 : Modélisation à l’aide d’un arbre, simulation sur tableur*Source : Document ressource Bac Professionnel*

Lors de certains contrôles de qualité en cours de fabrication dans l’industrie (diamètre d’une pièce par exemple), des cartes de contrôle reposent sur la procédure suivante : la moyenne de la cote surveillée (le diamètre de la pièce par exemple) est calculée sur des échantillons aléatoires prélevés régulièrement en fin de fabrication. Ces moyennes sont reportées sur une carte de contrôle. Si une série de sept points consécutifs se trouve du même côté de la « moyenne attendue » (la norme visée), le processus doit être surveillé pour déceler une éventuelle « dérive » dans le processus de fabrication.

L’explication du choix du nombre 7 se trouve dans la résolution du problème de probabilités suivant : une pièce de monnaie équilibrée est lancée 7 fois, quelle est la probabilité de l’événement A : « la pièce est tombée 7 fois sur pile » ?

1. Estimer la valeur de cette probabilité à l’aide de simulations sur un tableur.
2. On considère l’expérience qui consiste à lancer 3 fois de suite une pièce de monnaie équilibrée.
a) Dessiner un arbre figurant tous les résultats possibles de l’expérience.
b) À l’aide de l’arbre précédent, calculer la probabilité de l’événement « la pièce est tombée trois fois sur pile ».
3. Par analogie avec le cas de 3 lancers, donner la probabilité de l’événement A et comparer avec l’estimation de la question 1.

# Loi géométrique tronquée (Première)

## Exercice 6 : Loi d’une variable aléatoire, algorithmique, approfondissement*Source : Stage « Nouveaux programmes de Première » (Académie de Poitiers)*

Une société d’installation de panneaux photovoltaïques effectue une campagne de publicité téléphonique. Elle appelle 30 jours de suite M. Eole à la même heure jusqu’à ce qu’il réponde au téléphone. Chaque jour, la probabilité que M. Eole réponde est 0,3.

**Première partie**

1. A l’aide d’arbres pondérés, déterminer la probabilité que M. Eole réponde au téléphone au bout de 5 jours ? de 10 jours ?
2. On note la variable aléatoire égale au numéro du jour où M. Eole répond ; on conviendra que prend la valeur 0, si après les 30 jours M. Eole n’a toujours pas répondu.
	1. Quel est l’ensemble des valeurs que peut prendre  ?
	2. Déterminer .
	3. Donner , puis avec .
	4. Peut-on retrouver la valeur de ?

**Deuxième partie**

Il s’agit ici en activité de construire un algorithme modélisant cette expérience qui renvoie le numéro du jour où M. Eole répond au téléphone et qui renvoie 0 si au bout des 30 jours M. Eole n’a pas répondu.

* Chaque élève peut tester 10 fois l’algorithme et effectuer des constats sur les valeurs qu’il retourne.
* Puis on mutualise les résultats de toute la classe et on effectue la moyenne des temps d’attente.
* On répète ainsi l’expérience afin de constater que la moyenne des temps d’attente est relativement stable.

**Troisième partie**

On note et .

1. Montrer que .

On pose alors et on peut écrire plus simplement :

1. Soit la fonction définie sur par .
	1. Pour tout réel , écrire sous la forme d’un quotient.
	2. Vérifier que est dérivable sur et calculer deux expressions différentes de pour tout réel .
	3. En déduire le calcul de la somme :
	4. Calculer et vérifier que .
	5. Utiliser une calculatrice ou un tableur pour émettre une conjecture sur la limite de lorsque tend vers .

## Exercice 7 : Simulation sur tableur, espérance d’une variable aléatoire*Source : Document ressource Première*

Pour limiter le nombre de filles dans un pays (imaginaire ?), on décide que :

* chaque famille aura au maximum 4 enfants ;
* chaque famille arrêtera de procréer après la naissance d’un garçon.

On considère que chaque enfant a une chance sur deux d’être un garçon ou une fille et que, pour chaque couple de parents, le sexe d’un enfant est indépendant du sexe des précédents.

Ce choix a-t-il la conséquence attendue, à savoir de diminuer le nombre de filles dans la population ?

# Loi binomiale (Première)

## Exercice 8 : Loi d’une variable aléatoire, découverte des coefficients binomiaux*Source : Stage « Nouveaux programmes de Première » (Académie de Poitiers)*

**Partie A**

M. Fingourmet dîne dans le même restaurant 3 soirs de suite. La carte de ce restaurant propose uniquement deux menus au choix : le menu Découverte et le menu Tradition.

Chaque soir, le choix de M. Fingourmet s’effectue indépendamment des soirs précédents et chaque soir, M. Fingourmet choisit le menu Découverte avec la probabilité , .

1. Représenter cette situation à l’aide d’un arbre pondéré.
2. Quelle est la probabilité que M. Fingourmet ne choisisse jamais le menu Découverte ?
3. Quelle est la probabilité qu’il choisisse le menu Découverte 1 soir seulement ?
4. On note la variable aléatoire égale au nombre de soirs que M. Fingourmet choisit le menu Découverte.
	1. Quelles valeurs peut prendre ?
	2. Compléter le tableau suivant :

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | 0 | 1 | 2 | 3 |
|  |  |  |  |  |

* 1. Vérifier que
	2. Calculer l’espérance de .
	3. Combien y a-t-il de possibilités que M. Fingourmet choisisse exactement deux fois le menu Découverte au cours des 3 soirées ? On note ce nombre.
	4. En raisonnant de manière analogue, donner les valeurs de , et .
1. M. Fingourmet décide de retourner deux soirs supplémentaires dans ce restaurant. En utilisant les notations précédentes, quel est le nombre de possibilités que Mr Fingourmet choisisse le menu Découverte exactement 2 fois ? 3 fois ?

**Partie B**

M. Fingourmet ayant particulièrement apprécié ce restaurant, il y retourne soirs de suite. On s’intéresse une nouvelle fois au nombre de soirs où Mr Fingourmet a choisi le menu Découverte.

1. On note avec le nombre de chemins de l’arbre pondéré modélisant cette situation qui mènent exactement à succès (c’est-à-dire lorsque Mr Fingourmet choisit exactement fois le menu Découverte).
	1. Que valent et ?
	2. Comment peut-on exprimer autrement : « M. Fingourmet choisit exactement fois le menu Découverte » ? On a alors
	3. On suppose que M. Fingourmet a choisi au cours des dîners exactement fois le menu Découverte . Combien de soirs avait-il choisit le menu Découverte l’avant-dernier soir ? Quelle égalité peut-on en déduire ?
	4. Soit . Calculer à l’aide des propriétés précédentes, les coefficients , en complétant le tableau :



1. On s’intéresse de nouveau à la variable aléatoire égale au nombre de soirs où M. Fingourmet choisit le menu Découverte.
	1. Quelle sont les valeurs possibles de ?
	2. En raisonnant sur un arbre pondéré, quelle est la probabilité d’un chemin comportant exactement 2 succès ? succès ?
	3. Pour , donner alors .

## Exercice 9 : Calcul de probabilités avec la loi binomiale*Source : Document ressource Première*

1. Paul affirme : « Avec un dé régulier, on a autant de chance d’obtenir au moins un six en 4 lancers que d’obtenir au moins deux six avec 8 lancers ».
Sara objecte : « Pas du tout. Dans le premier cas, la probabilité est supérieure à 0,5, dans le deuxième cas, elle est inférieure à 0,5. ».
Qui a raison ?
2. À chaque tir, un archer atteint sa cible avec une probabilité égale à 0,7.
Combien de tirs doit-il effectuer pour que, avec une probabilité supérieure ou égale 0,99, il atteigne la cible au moins deux fois ? Au moins trois fois ?
3. Un QCM comporte 20 questions. Pour chaque question, quatre réponses sont proposées dont une seule est juste. Chaque réponse juste rapporte un point et il n’y a pas de pénalité pour une réponse fausse.
Un candidat répond au hasard à chaque question. Quel nombre total de points peut-il espérer ?
Quelle pénalité doit-on attribuer à une réponse fausse pour que le total espéré, en répondant entièrement au hasard, soit égal à 2 sur 20 ?

# Echantillonnage et prise de décision (Première)

## Exercice 10 : Intervalle de fluctuation*Source : Document ressourcePremière*

En Novembre 1976 dans un comté du sud du Texas, Rodrigo Partida est condamné à huit ans de prison. Il attaque ce jugement au motif que la désignation des jurés de ce comté est, selon lui, discriminante à l’égard des Américains d’origine mexicaine. Alors que 80 % de la population du comté est d’origine mexicaine, sur les 870 personnes convoquées pour être jurés lors des années précédentes, il n’y a eu que 339 personnes d’origine mexicaine.

Devant la Cour Suprême, un expert statisticien produit des arguments pour convaincre du bien fondé de la requête de l’accusé. En vous situant dans le rôle de cet expert, pouvez-vous décider si les Américains d’origine mexicaine sont sous-représentés dans les jurys de ce comté ?

## Exercice 11 : Intervalle de fluctuation*Source : Atelier Probabilités Statistiques (2011) présenté aux JDI (Académie de Grenoble)*

On reprend le fichier « poids des bébés à la naissance » de l’exercice 3.

Les courbes des carnets de naissance nous indiquent que 3% de bébés naissent avec un poids inférieur à 2,4kg.

Les données observées dans le fichier sont-elles conformes avec cette valeur :

* pour l’ensemble de la population étudiée ?
* pour la population des mères non-fumeuses ?
* pour la population des mères fumeuses ?

Indication : On pourra considérer les intervalles de fluctuation issus des lois B(1226 ;0,03), B(742 ;0,03) et B(484 ;0,03).

# Loi normale (Terminale) : *d’après document ressource*

## Exercice 12 : Calcul de probabilités à l’aide de la loi normale

La production laitière annuelle en litres des vaches laitières de la race FFPN peut être modélisée par une variable aléatoire à densité X, de loi normale de moyenne µ = 6000 et d’écart-type σ = 400. La fonction g désigne la fonction de densité de cette loi normale.

1. Afin de gérer au plus près son quota laitier (production maximale autorisée), en déterminant la taille optimale de son troupeau, un éleveur faisant naître des vaches de cette race souhaite disposer de certaines probabilités.
2. Calculer la probabilité qu'une vache quelconque de cette race produise entre 5900 et 6100 litres de lait par an.
3. Calculer la probabilité qu'une vache quelconque de cette race produise moins de 5800 litres par an.
4. Calculer la probabilité qu'une vache quelconque de cette race produise plus de 6250 litres par an.
5. Dans son futur troupeau, l’éleveur souhaite connaître :

a) la production maximale prévisible des 30% de vaches les moins productives du troupeau.
Il s’agit de déterminer la valeur x de X telle que P(X < x) = 0,30.

b) la production minimale prévisible des 20% des vaches les plus productives.
Il s’agit de déterminer la valeur x de X telle que P(X > x) = 0,20.

## Exercice 13 : Centrer et réduire une variable aléatoire, utilisation de

Une coopérative produit du beurre en microplaquettes de 12,5g pour des collectivités et des chaînes hôtelières. Les microplaquettes sont conditionnées dans des boîtes de 40. La masse des microplaquettes peut être modélisée par une variable aléatoire suivant une loi normale d’espérance *μ* = 12,5 et de variance *σ²* = 0,2² et on admet que la variable aléatoire *X* égale à la masse d’une boîte de 40 microplaquettes suit alors une loi normale d’espérance μ = 500 et de variance *σ*² = 1,6.

La boîte est jugée conforme si sa masse est comprise entre 496,2 g et 503,8 g.

1. Calculer la probabilité qu’une boîte prélevée aléatoirement en fin de chaîne de conditionnement soit non conforme.
2. Pour contrôler le réglage de la machine, on détermine des poids d'alerte μ − h et μ + h tels que
P(μ − h < X < μ + h) = 0,99. Ces poids d’alerte sont inscrits sur une carte de contrôle et correspondent à une marge de sécurité en lien avec des normes de conformité. Calculer les poids d'alerte.

## Exercice 14 : Centrer et réduire une variable aléatoire

Sur une chaîne d'embouteillage dans une brasserie, la quantité *X* (en cl) de liquide fournie par la machine pour remplir chaque bouteille de contenance 110 cl peut être modélisée par une variable aléatoire de loi normale de moyenne *μ* et d’écart-type *σ* = 2.

La législation impose qu'il y ait moins de 0,1% de bouteilles contenant moins d'un litre.

1. À quelle valeur de la moyenne *μ* doit-on régler la machine pour respecter cette législation?
2. La contenance des bouteilles étant de 110cl, quelle est alors la probabilité qu’une bouteille déborde lors du remplissage ?
3. Le directeur de la coopérative veut qu’il y ait moins de 1% de bouteilles qui débordent au risque de ne plus suivre la législation.
a) Quelle est alors la valeur de *μ*?
b) Quelle est dans les conditions de la question a) la probabilité que la bouteille contienne moins d’un litre ?
c) Déterminer *μ* et *σ* afin qu’il y ait moins de 0,1% de bouteilles de moins d’un litre ET moins de 1% de bouteilles qui débordent.

## Exercice 15 : Centrer et réduire une variable aléatoire, utilisation de

La durée de vie d'un certain type d’appareil est modélisée par une variable aléatoire suivant une loi normale de moyenne et d’écart-type inconnus. Les spécifications impliquent que 80 % de la production des appareils ait une durée de vie entre 120 et 200 jours et que 5% de la production ait une durée de vie inférieure à 120 jours.

1. Quelles sont les valeurs de *μ* et *σ*² ?
2. Quelle est la probabilité d’avoir un appareil dont la durée de vie est comprise entre 200 jours et 230 jours ?

# Intervalle de fluctuation (Tale) : *d’après document ressource*

## Exercice 16 :

On admet que dans la population d’enfants de 11 à 14 ans d’un département français le pourcentage d’enfants ayant déjà eu une crise d’asthme dans leur vie est de 13%.

Un médecin d’une ville de ce département est surpris du nombre important d’enfants le consultant ayant des crises d’asthme et en informe les services sanitaires. Ceux–ci décident d’entreprendre une étude et d’évaluer la proportion d’enfants de 11 à 14 ans ayant déjà eu des crises d’asthme. Ils sélectionnent de manière aléatoire 100 jeunes de 11 à 14 ans de la ville.

La règle de décision prise est la suivante : si la proportion observée est supérieure à la borne supérieure de l’intervalle de fluctuation asymptotique au seuil de 95% alors une investigation plus complète sera mise en place afin de rechercher les facteurs de risque pouvant expliquer cette proportion élevée.

1. Déterminer l’intervalle de fluctuation asymptotique au seuil de 95% de la proportion de jeunes de 11 à 14 ans ayant eu une crise d’asthme dans un échantillon de taille 100.
2. L’étude réalisée auprès des 100 personnes a dénombré 19 jeunes ayant déjà eu des crises d’asthme.
Que pouvez-vous conclure ?
3. Le médecin n’est pas convaincu par cette conclusion et déclare que le nombre de personnes interrogées était insuffisant pour mettre en évidence qu’il y avait plus de jeunes ayant eu des crises d’asthme que dans le reste du département. Combien faudrait-il prendre de sujets pour qu’une proportion observée de 19% soit en dehors de l’intervalle de fluctuation asymptotique ?
4. Représenter graphiquement la taille de l’échantillon nécessaire en fonction de la valeur *psup* de la borne supérieure de l’intervalle de fluctuation au seuil de 95%.

## Exercice 17 :

Une compagnie aérienne possède des A340 (longs courriers) d’une capacité de 300 places.
Cette compagnie a vendu *n* billets pour le vol 2012.
La probabilité pour qu’un acheteur se présente à l’embarquement est *p* et les comportements des acheteurs sont indépendants les uns des autres.
On note *Xn* la variable aléatoire désignant le nombre d’acheteurs d’un billet se présentant à l’embarquement.
La compagnie cherche à optimiser le remplissage de l’avion en vendant éventuellement plus de places que la capacité totale de l’avion (surréservation ou surbooking) soit ici n > 300.
Comme il y a évidemment un risque que le nombre de passagers munis d’un billet se présentant à l’embarquement excède 300, la compagnie veut maîtriser ce risque.

1. Déterminer la loi de *Xn*.
2. On suppose que 0,5 ≤ *p ≤* 0,95 .
Écrire l’intervalle de fluctuation asymptotique *In* de  au seuil de 0,95.
3. Montrer que si  alors la probabilité que le nombre de passagers se présentant à l’embarquement excède 300 est inférieure à 0,05.
4. On cherche à déterminer la valeur de *n* maximale permettant de satisfaire la condition de l’inclusion 

**a.** Montrer que  ⇒ .

**b.** On pose *f* (*x*) = . Montrer qu’il existe un entier *n*0 unique tel que si *n* ≤ *n*0 alors *f* (*n*) ≤ 0 et si *n* > *n*0 alors *f* (*n*) > 0 .

**c.** Tracer la courbe représentative de *f* pour les valeurs *p* = 0,85 ; *p* = 0,9 ; *p* = 0,95.

**d.** Déterminer à la calculatrice les valeurs de *n*0 pour *p* = 0,85 ; *p* = 0,9 ; *p* = 0,95.

# Intervalle de confiance (Tale) : *d’après document ressource*

## Exercice 18 :

Le 18 avril 2002, l’institut IPSOS effectue un sondage dans la population en âge de voter.
On constitue un échantillon de 1000 personnes (inscrites sur les listes électorales) que l’on suppose choisies ici de manière aléatoire. Les résultats partiels en sont les suivants : sur les 1000 personnes :

* 135 ont déclaré vouloir voter pour Jean-Marie Le Pen
* 195 ont déclaré vouloir voter pour Jacques Chirac
* 170 ont déclaré vouloir voter pour Lionel Jospin.
1. Déterminer l’intervalle de confiance au niveau de confiance de 95% pour chacun des 3 candidats.
2. Que peut-on à priori imaginer sur cette élection ?
3. Les résultats le jour de l’élection étaient 16,9% pour M Le Pen, 19,9% pour M Chirac et16,2%. pour M Jospin. Comparer ces résultats avec les intervalles de confiance.

## Exercice 19 :

Un test de diagnostic rapide effectué sur des sujets ictériques (coloration jaune de la peau, des muqueuses -couche de cellules de protection recouvrant les organes creux en contact avec l’extérieur et du blanc de l’œil –sclérotique-) doit permettre d’estimer si l’ictère est d’origine virale ou non, sans avoir besoin de faire des analyses longues et compliquées. Cependant il est important de pouvoir s’assurer que ce test est de bonne qualité c'est-à-dire qu’il doit pouvoir indiquer correctement si l’ictère est viral ou non. Il doit être capable d’identifier correctement le type d’ictère : il est positif chez les sujets dont l’ictère est viral et négatif sinon.

Une étude est effectuée sur 100 personnes ayant un ictère viral et 100 personnes ayant un ictère d’origine non virale. Les résultats obtenus sont présentés dans le tableau ci-dessous :

|  |  |  |
| --- | --- | --- |
|  | Hépatite virale | Ictère d’origine non virale |
| Test positif | 85 | 20 |
| Test négatif | 15 | 80 |

1. Déterminer la proportion de sujets ayant un test positif parmi ceux ayant un ictère viral.
2. Déterminer un intervalle de confiance à 95% de la proportion de tests positifs lorsque l’ictère est viral. Cette proportion est appelée sensibilité du test diagnostic, c'est-à-dire la probabilité qu’une personne ayant un ictère viral réagisse au test. Un test diagnostic sera d’autant meilleur que la sensibilité est importante.
3. Déterminer la proportion de sujets ayant un test négatif parmi celles ayant un ictère non viral.
4. Déterminer un intervalle de confiance à 95% de la proportion de tests négatifs lorsque l’ictère est non viral. Cette proportion est appelée spécificité du test diagnostic, c'est-à-dire la probabilité qu’une personne ayant un ictère non viral ne réagisse pas au test. Un test diagnostic sera d’autant meilleur que la spécificité est importante.

# Pour aller plus loin

## Exercice 20 :

L’objectif de cet exercice est de démontrer de deux manières différentes que 

1ere méthode : par changement de variable.

*Pré-requis*

On admet le théorème suivant :

Soit f une fonction de deux variables, continue sur [a,b]×[c,d], à valeurs dans **C**. Alors



1. Démontrer que cest bien intégrable sur R
2. Soit a un réel positif. On pose I(a) = . Démontrer que I(a)² = pour tout réel a positif.
3. Soit Da le domaine défini par Da = {(rcosθ ; rsinθ) avec 0 ≤ r ≤ a et 0≤ θ ≤ π/2 }.
4. On pose J(a) =. Démontrer que J(a) =  pour tout réel a positif.
5. En déduire la limite de J(a) quand a tend vers +∞.
6. Expliquer pourquoi J(a) ≤ I(a)² ≤ J(a) pour tout réel a positif.
7. Déduire des questions précédentes la limite de I(a)² quand a tend vers +∞.
8. Déterminer alors 

2ème méthode : avec les intégrales dépendant d’un paramètre.

On admet le théorème suivant :

Soit I un intervalle de R. Si F est une fonction de deux variables, continue sur I×[a ; b] et si existe et est continue sur I×[a ; b] alors l’application f définie sur I par f(x) =  est dérivable sur I et f’(x) = 

Soit f la fonction définie sur R par f(x) = 

1. Démontrer que f est dérivable sur R
2. Démontrer que f’(x) =  pour tout réel x (on pourra utiliser un changement de variable de la forme u = tx).
3. En déduire que pour tout réel x, f’(x) = -2g’(x) g(x) où g est définie sur R par g(x) = 
4. Calculer f(0) et en déduire que pour tout réel x, f(x) = - (g(x))².
5. Démontrer que la limite de f à l’infini est égale à 0 (on pourra utiliser l’encadrement valable pour tout réel x, 0 ≤ f(x) ≤ )
6. Déterminer alors .

## Exercice 21 :

Soit la fonction définie sur par : .

L’objet de l’exercice 1 était de démontrer que : .

 étant continue, positive sur et telle que , est une densité de probabilité : c’est la densité de la loi normale centrée réduite notée **N(0,1).**

1. Etudier la fonction sur et donner l’allure de sa courbe représentative.
2. On dit qu’une variable aléatoire suit la loi normale centrée réduite lorsque sa fonction de répartition est donnée par :
3. Etablir que réalise une bijection de sur ]0 ;1[.
4. Montrer que pour tout
5. Soit une variable aléatoire qui suit la loi normale centrée réduite,

Démontrer que : , il existe un unique réel positif tel que

*(on pourra remarquer que pour tout réel u > 0, P(-u ≤ X ≤ u) = 2 P(0 ≤ X ≤ u) et appliquer un corollaire du théorème des valeurs intermédiaires)*

1. Soit une variable aléatoire qui suit la loi normale centrée réduite, démontrer que :
	1. admet une espérance et que
	2. admet une variance et que

*(on remarquera que V(X) = E((X-E(X))²) puis utiliser une intégration par parties)*

## Exercice 22 :

*On rappelle les définitions suivantes :*

**Définition 1 :** Soit *X* une variable aléatoire correspondant à la durée de vie d'un individu ou d'un objet.

Soient t et h deux réels positifs.

On dit que *X* suit la propriété de durée de vie sans vieillissement lorsque la probabilité que l'individu (ou l'objet) soit vivant (ou fonctionne) à l'instant *t* + *h* sachant qu'il est vivant (ou qu'il fonctionne) à l'instant *t* ne dépend pas de son âge *t.* Autrement dit, lorsque *P* (*X ≥ t*) (*X ≥ t* + *h*) = *P*(*X ≥ h*).

**Définition 2 :** Une variable aléatoire à densité X suit la loi exponentielle de paramètre λ > 0 lorsque sa densité est la fonction f définie sur [0 ; +∞[ par f(x) = λ e-λx .

1. Démontrer que si X est une variable aléatoire qui suit une loi exponentielle, alors elle suit la propriété de durée de vie sans vieillissement.
2. L’objectif de cette question est de démontrer la propriété réciproque du 1).

Soit X une variable aléatoire qui suit une propriété de durée de vie sans vieillissement.

Soit F la fonction de répartition de X et ϕ la fonction définie sur [0 : +∞[ par ϕ(t) = 1 – F(t).

**a)** Démontrer que ϕ(0)=1 et que pour tous réels h et t positifs, ϕ(t+h) = ϕ(t) ϕ(h).

**b)** En déduire une expression de ϕ, puis de F.

**c)** En déduire que X suit bien la loi exponentielle.

Annexe : Description d’une série statistique à une variable quantitative

Source : *Statistiques au lycée, volume 1* (brochure APMEP n°156)

La description d’une série statistique à une variable quantitative est faite pour synthétiser et visualiser l’information contenue dans les données. L’objectif est de comparer des distributions observées, soit entre elles, soit à une distribution de référence (modèle théorique, gaussien par exemple). On évitera donc de faire calculer aux élèves une moyenne et un écart-type sans ce souci de comparaison.

# Les 4 étapes de la description d’une série :

**Etape 1 : Validation et listage**

* Validation : pour repérer les valeurs suspectes (éventuelles erreurs de saisie, valeurs s’éloignant beaucoup du groupe)
* Listage des données : construction d’un tableau récapitulatif, tri des données.

Dans nos exercices en classe, cette étape n’est pas réalisée par les élèves, sauf lorsqu’ils construisent leur propre enquête.

**Etape 2 : Regroupement des données, graphiques des distributions, analyse**

* Regroupement pour réaliser un découpage en classes des valeurs de la série. Il existe des formules pour fixer a priori un nombre de classes, par exemple : où est l’effectif de la série mais il n’y a pas de recette miracle.
* Construction d’un graphique : diagramme en bâtons ou *tige et feuilles.*
* Analyse des distributions observées :
* Recherche de structures : est-on bien en présence d’une seule population ? Un tel mélange se manifeste en général par des distributions multimodales (plusieurs *pics*). Le calcul de paramètres tels la moyenne perd alors tout son sens. Voir l’exercice « Perce oreilles ».
* Analyse de la symétrie : on l’évalue à l’œil. Une forte dissymétrie fait que l’on ne pourra pas comparer notre série au modèle gaussien par exemple.
* Comparaison des dispersions des séries : contrairement aux deux critères précédents, l’évaluation de la dispersion ne se fait pas dans l’absolu. On compare deux séries entre elles ou une série à une distribution de référence. On l’analyse d’abord à l’œil puis, éventuellement, en calculant des paramètres (écart-type, …)
* Comparaison des positions des séries par rapport à des références ou entre elles : d’abord à l’œil puis, pour affiner, par le calcul de paramètres.

**Etape 3 : Calcul des valeurs des paramètres pour résumer en chiffres**

Au lycée, on dispose de trois couples : (mode ; étendue), (médiane ; écart-interquartile) et (moyenne ; écart-type).

**Etape 4 : Construction des boîtes à moustaches**

On présentera ces boîtes verticalement, cela facilite les comparaisons. Attention, l’information sur la structure de la dispersion n’est plus visible. Il ne faut donc jamais omettre le *tige et feuilles* ou le diagramme des effectifs.