Exercices de Probabilités Statistique

**Exercice 1 :**

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.

*Grâce à des échantillons prélevés en sortie de chaine ces masses d’alerte permettent de déceler des anomalies en temps réel.*

**Exercice 2 :**

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.

À quelle valeur de la moyenne *μ* doit-on régler la machine pour respecter cette législation?

**Exercice 3 :**

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 ?

**Exercice 4 :**

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 ?

*On en conclut que la règle de décision choisie ne prévoit pas de réaliser une enquête supplémentaire.*

**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 5 :**

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 proche de 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.

**Exercice 6 :**

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 7 :**

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 |

**a)** Déterminer la proportion de sujets ayant un test positif parmi ceux ayant un ictère viral.

**b)** 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.

**c)** Déterminer la proportion de sujets ayant un test négatif parmi celles ayant un ictère non viral.

**d)** 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.

 **Exercice 8 :**

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.

*Pré-requis*

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 9 :** Quelques propriétés de la loi normale centrée réduite…

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 10 :**

*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 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.

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.