Table des matières

1 TD1 : Rappels (Stats. descrt. / Theo. probabilités)

1.1 Exercice : Stats. Descriptives / indicateurs, visualisation

On observe 100 fois le nombre d'arrivées (variable X) de clients à un bureau de poste pendant un intervalle de temps (10 minutes) et on obtient les valeurs suivantes :

1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2
2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
3	3	3	3	3	3	4	4	4	4	4	4	4	4	4	4	4	4	4	4
4	4	4	4	4	4	5	5	5	5	5	5	5	6	6	6	6	6	6	6

1. Dresser le tableau statistique de la distribution de la variable X (effectifs cumulés, …).
2. Calculer les valeurs de tendance centrale de la distribution : la moyenne, le mode et les trois quartiles Q1, Q2 et Q3.
3. Calculer les valeurs de la dispersion de la distribution : variance, l'écart type et l'intervalle interquartile.
4. Tracer le diagramme en bâtons et la boite à moustaches de cette distribution.

1.2 Exercice : Notion de probabilité

Trois joueurs α, β et γ jettent une pièce à tour de rôle. Le premier qui obtient pile a gagné. Nous admettrons que α joue d'abord, puis β et enfin γ. L'univers Ω associé à cette expérience aléatoire peut être décrit comme suit :

$$\Omega = \{1,01,001,0001,\ldots\} \cup \{0000\ldots\}.$$

1. Donnez une interprétation des points de Ω.

2. Décrivez les événements suivants en terme de ces points :
a) premier événement : A = « α gagne » ;
b) deuxième événement : B = « β gagne » ;
c) troisième événement : $(A \cup B)^c$.

1.3 Exercice : Variables aléatoires discrètes

Soit X une v.a.d. telle que $\mathbb{P}([X = -1]) = \mathbb{P}([X = 0]) = \mathbb{P}([X = 1]) = 1/3$ et $\mathbb{P}([X = x]) = 0$ pour tout $x \in \mathbb{R} \setminus \{-1,0,1\}$.
Soit $Y = X^2$.
Montrez que les variables aléatoires X et Y ne sont pas corrélées. Sont-elles indépendantes ?

1.4 Exercice : Variables aléatoires continues

Soit f la fonction définie par

$$f(t) = \begin{cases} 0 & \text{si } t \notin [0,1] \\ 2t & \text{si } t \in [0,1] \end{cases}$$

1. a) Démontrez que la fonction f est une densité de probabilité.
b) Démontrez que la loi de probabilité définie par f admet une espérance mathématique et une variance que vous préciserez.
c) Déterminez la fonction de répartition associée à f.

2. Soit X une variable aléatoire de densité de probabilité f définie ci-dessus. Nous définissons $Y = 1 + X^2$. Déterminez la fonction de répartition de Y.

1.5 Exercice : Loi normale

Une variable aléatoire X suit la loi normale $\mathcal{N}(5{,}2; 0{,}8)$.

1. Calculez $\mathbb{P}(X \geqslant 4 | X \leqslant 5{,}2)$.

2. Une variable aléatoire Y indépendante de la variable X suit la loi normale $\mathcal{N}(\mu; 0{,}8)$.

a. Quelle est la loi de la variable aléatoire $Y - 2X$?

b. Déterminez μ sachant que $\mathbb{P}(Y \geqslant 2X) = 0{,}516$.

2 TD2 : lois et limites

2.1 Exercice (mise en pratique des lois)

Un questionnaire comporte dix questions. Pour chacune, il y a quatre réponses possibles. Un individu répond au hasard à toutes ces questions.

a) Quelle est la probabilité qu'il réponde juste à toutes les questions ?

b) Quelle est la probabilité qu'il réponde juste à 5 questions ?

c) Quelle est la probabilité qu'il réponde juste à au moins 8 questions ?

d) En moyenne, à combien de questions va-t-il répondre juste ?

2.2 Exercice (mise en pratique des lois)

Une entreprise dispose d'une machine dont la probabilité de tomber en panne en une journée est de 1/500.

a) Quelle est la probabilité que cette machine tombe en panne au moins une fois en 20 jours ?

b) Quelle est la probabilité que cette machine tombe strictement plus d'une fois en panne sur cette période ?

2.3 Exercice (mise en pratique des lois)

Une entreprise dispose d'une machine dont la probabilité de tomber en panne en une journée est de 1/500.

a) Quelle est la probabilité que cette machine tombe en panne au moins une fois en 20 jours ?

b) Quelle est la probabilité que cette machine tombe strictement plus d'une fois en panne sur cette période ?

2.4 Exercice (mise en pratique des lois)

Un magasin possède 4 caisses. Le nombre de clients qui attendent à chacune de ces caisses suit une loi de Poisson de paramètre 3,2. Un nombre de clients supérieur ou égal à 4 par caisse risque de ne pas plaire à la clientèle. Il faudrait alors ouvrir une nouvelle caisse.

Quelle est la probabilité d'ouvrir une nouvelle caisse ?

2.5 Exercice (mise en pratique des lois)

La moyenne des notes à un examen suit une loi normale de moyenne 9,5 et d'écart-type 4,5. Les étudiants ayant une moyenne inférieure à 7 ne sont pas admissibles. Les étudiants ayant une moyenne entre 7 et 9 doivent repasser une des épreuves pour être admissible. Les étudiants ayant entre 9 et 13 sont admissibles et doivent passer un oral. Les étudiants ayant plus de 13 sont directement admis.

Sur un ensemble de 400 000 étudiants qui se présentent à cet examen, déterminer une estimation du nombre d'étudiants de chaque catégorie.

2.6 Exercice (mise en pratique des lois)

Un avion décolle à 5h25 du matin. On note que 10 % des passagers arrivent avant l'heure d'enregistrement des bagages et 5 % d'entre eux arrivent après la fermeture de l'enregistrement et ne peuvent pas embarquer. L'heure d'ouverture de l'enregistrement est 3h50 et l'heure de fermeture de l'enregistrement est 5h.

a) Déterminer les paramètres de la loi normale qui représente l'heure d'arrivée d'un passager.

b) Combien la compagnie doit-elle enregistrer de réservations pour s'assurer qu'un avion de 200 places soit complet ?

c) Que se passe-t-il si l'avion a 10 minutes de retard et que la fermeture de l'enregistrement prend aussi 10 minutes de retard ?

2.7 Exercice

Le nombre de personnes à une caisse suit une loi de Poisson de moyenne 12.

Quelle est la probabilité de trouver plus de 20 personnes à cette caisse ?

2.8 Exercice

Une usine a produit des pièces dont 2 % ont un défaut. On teste un ensemble de 10 000 pièces.

a) Que dire de l'utilisation de la loi Binomiale pour cette étude ?

b) À l'aide d'une approximation de loi, déterminer la probabilité de trouver plus de 2,2 % de pièces défectueuses dans l'ensemble des 10 000 pièces.

2.9 Exercice

Soit $X_1, \ldots, X_{200}$ des variables aléatoires indépendantes et identiquement distribuées telles que $X_i \sim$ Poisson(0,5). Considérons la variable aléatoire Y, qui représente la somme de tous les X_i, c'est-à-dire $Y = X_1 + \cdots + X_{200}$.

a) Quelle est la loi exacte de la variable Y ?

b) En utilisant la loi exacte de Y, on peut calculer les probabilités suivantes : $P(Y = 100) = 0{,}0399$ et $P(89 < Y \leq 120) = 0{,}831$. Calculez une approximation de ces deux probabilités en utilisant le théorème central limite.

c) Quelle est la probabilité (approximative, sans correction de continuité) que la moyenne de ces 200 valeurs soit inférieure à 0,42 ?

3 TD3 : Estimation

3.1 Exercice

Un vendeur d'appareils électroménagers assure un remboursement de l'appareil s'il tombe en panne avant 5 ans. Sur un échantillon de 120 appareils, on compte une durée de vie moyenne de 5 ans et demi avec un écart-type de 2 ans. Quel est le risque pour le vendeur de voir ses appareils tomber en panne en moyenne avant 5 ans ?

3.2 Exercice

On compare les baisses du cours des actions réalisées au moment de la crise grecque, le vendredi 14 avril 2010, par des entreprises américaines du Dow Jones 30 et des entreprises françaises du CAC40.

Dow Jones 30	Variations	CAC40	Variations
Caterpillar	– 3,03 %	Alcatel Lucent	– 5,55 %
Coca-Cola	– 0,28 %	Danone	– 3,15 %
Mc Donald's	– 1,28 %	Bouygues	– 4,04 %
Bank of America	– 3,14 %	EDF	– 4,91 %
Hewlett Packard	– 2,65 %	Michelin	– 3,24 %
Microsoft	– 1,06 %	Sanofi-Aventis	– 2,44 %

a) Donner l'intervalle de confiance au risque 5 % pour la différence des variations relatives moyennes des deux populations.

b) Dans quelle mesure peut-on affirmer que le marché français souffre davantage de la crise grecque que les États-Unis ?

3.3 Exercice

Une entreprise effectue des tests de qualité sur les productions de deux de ses usines ; sur un échantillon de 400 produits de l'usine A, 2 % sont défectueux et sur un échantillon de 350 produits de l'usine B 1,5 % sont défectueux.

a) Quel est l'intervalle de confiance au risque 5 % pour la différence de proportions de pièces défectueuses entre les deux usines ?

b) Que peut-t-on conclure ?

4 TD4 : Test d'hypothèses

4.1 Exercice

La résistance moyenne à la rupture d'une fibre utilisée pour fabriquer du tissu doit être d'au moins 160 lb/po2. Selon l'expérience acquise, l'écart-type de la résistance à la rupture est de 3 lb/po2, et la loi normale est un bon modèle pour cette variable. On soupçonne que la résistance moyenne n'est pas suffisamment élevée, mais il faudrait le démontrer avec une conclusion forte pour apporter des changements à la production. On veut effectuer un test d'hypothèse à partir d'un échantillon aléatoire de quatre morceaux de tissu en utilisant un seuil de = 0,05. L'hypothèse nulle peut donc s'écrire H0 : $\mu = 160$

a) Quelle serait la bonne formulation pour l'hypothèse alternative ?
b) Déterminez la loi de la moyenne échantillonnai sous H0 : $\mu = 160$. Tracez grossièrement la courbe qui la représente et identifiez la zone de rejet de H0.
c) L'échantillon donne une résistance moyenne à la rupture de 158 lb/po2. Quelle est votre conclusion ?
d) Sur le graphique tracé en b), ajoutez la courbe représentant la loi de la moyenne échantillonnai lorsque la vraie valeur de μ est de 156 lb/po2. Hachurez la zone correspondant à la probabilité d'erreur de deuxième espèce.
e) À partir de la figure tracée en d), estimez la probabilité de ne pas rejeter H0 (à tort) si la fibre a une résistance réelle à la rupture de 156 lb/po2. Vaut-elle environ 1 %, 15 %, 50 %, 85 % ou 99 % ?

4.2 Exercice

On étudie le rendement (en pourcentage) d'un traitement chimique. Selon une expérience antérieure liée à ce procédé, la variance du rendement est de 5 %2, et la loi normale modélise bien la distribution. Les cinq derniers essais de ce traitement ont donné les rendements suivants (en pourcentage) : 91,60 ; 88,75 ; 90,80 ; 89,85 ; 91,30.

a) Y a-t-il une raison de croire que le rendement moyen soit inférieur à 90 %, au seuil de 5 % ?
b) Quelle est la puissance du test si le rendement réel est de 87 % ?
c) Calculez (à partir de la définition de β) la taille de l'échantillon nécessaire pour détecter un rendement moyen réel de 87 % avec une probabilité de 0,95.

4.3 Exercice

Un ingénieur veut tester l'hypothèse que le point de fusion moyen d'un alliage est de 1000 °C. Son équipe et lui conviennent de rejeter l'hypothèse nulle si le point de fusion moyen obtenu avec 5 mesures échantillonnais est supérieur à 1010 ou inférieur à 990 °C. Dans chacune des situations ci-après, déterminez si une erreur est commise (et de quelle espèce) ou non.

a) Le vrai point de fusion moyen de l'alliage est de 1012 degrés, et il obtient une moyenne échantillonnale de 1005 degrés.
b) Le vrai point de fusion moyen de l'alliage est de 1000 degrés, et il obtient une moyenne échantillonnale de 1005 degrés.
c) Le vrai point de fusion moyen de l'alliage est de 992 degrés, et il obtient une moyenne échantillonnale de 988 degrés.
d) Le vrai point de fusion moyen de l'alliage est de 1000 degrés, et il obtient une moyenne échantillonnale de 1015 degrés.

4.4 Exercice

La durée de conservation d'un film photographique intéresse son fabricant. Il observe les durées de conservation pour huit unités choisies au hasard dans la production courante. Il obtient une moyenne échantillonnale de 134,34 jours et une variance échantillonnale de 371,07 jours2. Supposez que la durée de conservation est distribuée normalement.

a) Y a-t-il une preuve que la durée de vie moyenne est supérieure à 125 jours au seuil de 5 % ?
b) Quelle est la valeur P du test ?

c) Si on recommençait l'expérience avec un échantillon de taille = 25 au lieu de n = 8, peut-on savoir d'avance comment les probabilités d'erreur de première et de deuxième espèces seraient affectées (s'il y a lieu)?

d) Si on recommençait l'expérience avec un échantillon de taille = 25 au lieu de n = 8, et qu'on obtenait une moyenne et une variance échantillonnais similaires à celles de l'expérience originale, la valeur P serait-elle supérieure, inférieure ou égale à la valeur P obtenue initialement?