Table des matières

1 TD0 : 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 : 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 $\mathbb{P}(Y \geqslant 2X) = 0{,}516$.

1.5 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 ?

1.6 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 ?

1.7 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 ?

1.8 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 ?

1.9 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.

1.10 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 ?

1.11 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 ?

1.12 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.

1.13 Exercice

Soit $X_1, \ldots, X_{200}$ des variables aléatoires indépendantes et identiquement distribuées telles que $X_i \sim \text{Poisson}(0{,}5)$. 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 ?

2 TD1 : Théorèmes limites

2.1 Exercice

Le nombre d'inscriptions à un cours d'économie est une variable aléatoire de Poisson de paramètre 100. Le professeur donnant ce cours a décidé que si le nombre d'inscriptions est au-delàs de 120, il créera 2 sections et donnera donc 2 cours, tandis qu'en deçà une seule classe sera formée. Quelle est la probabilité que ce professeur ait à donner 2 fois ce cours ?

2.2 Exercice

Soit X_1, ..., X_{200} des variables aléatoires indépendantes et identiquement distribuées telles que X_i~Poisson(0.5). Considérons la variable aléatoire Y, qui représente la somme de tous les Xi, c'est-à-dire $Y = X_1+X_2+\ldots+X_{200}$.

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< 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 ?

2.3 Exercice

Une compagnie aérienne doit faire embarquer 225 passagers dans un avion. La charge maximale autorisée – passagers et bagages – ne doit pas dépasser 20 tonnes. Le poids d'un passager avec ses bagages est une variable aléatoire X d'espérance mathématique 80 kilos et d'écart type 9 kilos.

1. Indiquer par quelle loi et dans quelles conditions la variable aléatoire C égale au poids des 225 passagers avec leurs bagages peut être approximée.
2. Déterminer l'espérance mathématique et l'écart type de C en kilos puis en tonnes. Déterminer la probabilité que le poids total des passagers avec leurs bagages dépasse 18 tonnes.
3. Déterminer la probabilité que le poids total des passagers avec leurs bagages soit compris entre 17,8 et 18,2 tonnes.
4. Déterminer la probabilité que le poids total des passagers avec leurs bagages dépasse la charge maximale.

3 TD2 : Estimation

3.1 Exercice : intervalle de confiance

Un fabricant de voitures fait un test d'efficacité d'un certain modèle. On mesure les litres de consommation d'essence pour 100 kilomètres :

14,60	11,21	15,56	11,37	13,68	11,06	26,58
13,37	15,98	12,07	13,22	12,01	15,07	

On connaît la variance de la population : elle vaut $\sigma^2 = 16$.

1. Calculer un intervalle de confiance à 95 % pour la moyenne.

2. Quel niveau correspond à un intervalle de longueur 3 litres/100 km ?

3. Combien d'observations additionnelles sont nécessaires pour avoir un intervalle à 99 % de longueur 2 litres/100 km ?

3.2 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 ?

3.3 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 ?

3.4 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 ?

3.5 Exercice : Estimation

Un tueur est recherché par la police, il a été signalé 23 fois dans un village à 6km du lieu du crime, 47 fois dans un supermarché à 25 km du lieu du crime et 3 fois dans une autre région à 150 km du lieu du crime.

À quelle distance du lieu du crime la police doit-elle effectuer ses recherches pour avoir 95 % de chances de trouver le tueur ?

4 TD3 : 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 $\alpha = 0.05$. L'hypothèse nulle peut donc s'écrire $H0 : \mu = 160$

a) Quelle serait la bonne formulation pour l'hypothèse alternative ?
b) Déterminez les régions de rejet et d'acceptation en utilisant le seuil de $\alpha = 0.05$
   sous $H0 : \mu = 160$.
   Tracez grossièrement la courbe qui la représente et identifiez la zone de rejet de H0.
c) Supposons que 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

Lors des dernières élections, le parti *A* a réalisé 52 % des suffrages. On réalise un sondage sur 1000 personnes et 49 % de ces personnes affirment vouloir voter pour ce parti. On s'intéresse à cette baisse de pourcentage.

1) Poser les hypothèses d'un test unilatéral.
2) Déterminer la probabilité critique liée à l'observation.
3) Déterminer la valeur critique de proportion au risque 5%.
4) Conclure.

4.3 Exercice

Le chiffre d'affaires journalier d'un magasin suit une loi normale de moyenne 1 000 DA et d'écart-type 500 DA. Le magasin change la présentation de sa vitrine et s'interroge sur l'effet de ce changement sur le chiffre d'affaires des 20 prochains jours.

a) Formulez un test d'hypothèse bilatéral sur la moyenne du chiffre d'affaires journalier pour déterminer un intervalle critique en dehors duquel il est possible d'affirmer au risque 1% que le changement de présentation a induit une modification du chiffre d'affaires.
b) On constate un chiffre d'affaires moyen de 1200 DA sur ces 20 jours. Que conclure ?

4.4 Exercice (χ2)

Dans une entreprise, on observe la répartition des salaires suivante :

Salaires	[800 ; 1 000[	[1 000 ; 1 200[	[1 200 ; 1 500[	[1 500 ; 1 800[	[1 800 ; 2 500[	[2 500 ; 3 000[	[3 000 ; 4 000[
Effectifs	10	40	100	110	30	8	2

a) Si on pouvait modéliser la répartition des salaires par une loi normale quels en serait les paramètres ?

b) À l'aide d'un test du χ2 déterminer au risque 1% de rejeter à tort le modèle, si le modèle d'une loi normale est compatible avec l'observation.

4.5 Exercice (χ2)

On a interrogé des habitants de trois villes A, B, et C sur l'appréciation de 4 stations de radio. Le croisement de ces deux variables donne le tableau de contingence suivant :

	Ville A	Ville B	Ville C	Total
Radio 1	20	17	10	47
Radio 2	17	16	12	45
Radio 3	39	24	12	75
Radio 4	12	8	13	33
Total	88	65	47	200

Déterminer si l'appréciation d'une radio dépend ou non de la localisation géographique.

4.6 Exercice

Pour déterminer si le contenu effectif d'une substance chimique d'un produit alimentaire est plus élevé que ce qui est annoncé sur le paquet (1.4 mg), on procède à un test sur un échantillon de taille n = 25 produits. Pour cet échantillon, on obtient $\bar{x} = 1.79$. On suppose que la variance d'une mesure vaut $\sigma = 1$.

1) Formuler l'hypothèse nulle et l'hypothèse alternative dans le cas de cette étude.
2) Définir le seuil de rejet ou la probabilité de rejet sous l'hypothèse nulle (dite p-valeur).
3) Est-ce que le résultat est significatif au niveau de 5 % ?
4) Est-ce que le résultat est significatif au niveau de 1 % ?

5 TD4 : Chaines de Markov

5.1 Exercice

Dans une ville, 90 % des jours ensoleillés sont suivis de jours ensoleillés, et 80 % des jours nuageux sont suivis de jours nuageux. Utilisez ces informations pour modéliser la météo de cette ville sous forme de chaîne de Markov.

Supposons que la météo de demain dépende des deux derniers jours, comme suit :

(1) Si les deux derniers jours ont été ensoleillés, il y a 95 % de chances qu'il fasse beau demain.
(2) Si hier il faisait nuageux et qu'aujourd'hui il fait soleil, il y a 70 % de chances qu'il fasse beau demain.
(3) Si hier il faisait soleil et qu'aujourd'hui il fait nuageux, il y a 60 % de chances qu'il fasse nuageux demain. (4) Si les deux derniers jours ont été nuageux, il y a 80 % de chances qu'il fasse nuageux demain.

À partir de ces informations, modélisez la météo comme une chaîne de Markov.

Si la météo de demain dépendait des trois derniers jours, combien d'états seraient nécessaires pour modéliser la météo comme une chaîne de Markov ? (Remarque : L'approche utilisée dans ce problème peut être utilisée pour modéliser un processus stochastique à temps discret comme une chaîne de Markov, même si X_{t+1} dépend d'états antérieurs à X_t, comme X_{t-1} dans l'exemple actuel.)

5.2 Exercice

Considérons la matrice de transition suivante :

$$P = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ \frac{1}{4} & \frac{1}{4} & 0 & \frac{1}{2} & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{3} & 0 & 0 & 0 & \frac{2}{3} \end{bmatrix}$$

(1) Quels états sont transitoires ?
(2) Quels états sont récurrents ?
(3) Identifier tous les ensembles fermés d'états.
(4) Cette chaîne est-elle ergodique ?

5.3 Exercice

Pour chacune des chaînes suivantes, déterminez si la chaîne de Markov est ergodique. Déterminez également pour chaque chaîne les états récurrents, transitoires et absorbants.

$$P_1 = \begin{bmatrix} 0 & .8 & .2 \\ .3 & .7 & 0 \\ .4 & .5 & .1 \end{bmatrix} \qquad P_2 = \begin{bmatrix} .2 & .8 & 0 & 0 \\ 0 & 0 & .9 & .1 \\ .4 & .5 & .1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$