Université des Sciences et Technologies de Lille
U.F.R. de Mathématiques Pures et Appliquées

IS Math 314 **Année 2009–2010**

Fiche nº 3

Ex 1. Soient $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires indépendantes et indentiquement distribuées de loi uniforme sur $[\theta, \theta+1]$. On définit la variable aléatoire Z_n par $Z_n = \min(X_1, \dots, X_n)$.

1) Montrer que Z_n converge en probabilité vers θ.

2) Montrer que $n\,(Z_n - \theta)$ convergent en loi vers une variable aléatoire Z dont on déterminera la loi.

Ex 2. Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires telles que X_n ait pour loi une loi uniforme sur l'ensemble fini

$$A_n = \left\{0, \frac{1}{n}, \frac{2}{n}, \dots, \frac{n-1}{n}\right\}.$$

Montrer que $(X_n)_{n\in\mathbb{N}^*}$ converge en loi vers une variable aléatoire X dont on déterminera la loi.

Ex 3. Échauffement du dé et des candidats (Juin 2005)
On lance 3 600 fois un dé équilibré et on s'intéresse au nombre X de « six » obtenus. Donnez en justifiant votre réponse, une valeur approchée de $P(578 \leq X \leq 622)$.

Ex 4. Pétition « Pour une vraie formation des enseignants »
On rappelle que la loi exponentielle de paramètre $a > 0$ admet pour densité

$$f : t \longmapsto a\mathrm{e}^{-at}\mathbf{1}_{[0,+\infty[}(t).$$

On rappelle aussi que la fonction Γ est définie sur $]0, +\infty[$ par

$$\Gamma(r) := \int_0^{+\infty} t^{r-1}\mathrm{e}^{-t}\,\mathrm{d}t$$

et que pour tout $n \in \mathbb{N}^*$, $\Gamma(n) = (n-1)!$.

1) Soit X une variable aléatoire de loi exponentielle de paramètre a. Montrer que $\mathbf{E}(X^n)$ est fini pour tout $n \in \mathbb{N}^*$ et calculer ce moment en fonction de a et de n. En déduire que

$$\mathbf{E}X = \frac{1}{a}, \qquad \operatorname{Var} X = \frac{1}{a^2}.$$

Une nouvelle pétition [1] contre le projet de réforme de la formation des enseignants est lancée sur le Web le 15 novembre 2008 à la suite de l'appel du 8 novembre 2008. Pour

1. http ://appeldu8novembre.fr/phpPetitions/ ?petition=2

que cette pétition ait un impact, les initiateurs espèrent recueillir les 10 000 premières signatures en moins d'une semaine, *i.e.* en moins de 7 journées de 24 heures [2].

L'expérience d'une pétition précédente permet aux initiateurs de considérer que les temps d'attente (en minutes) entre les enregistrements de deux signatures consécutives sont des variables aléatoires X_i indépendantes et de même loi exponentielle de paramètre $a = 1$ mn^{-1}, donc $\mathbf{E}X_i$ vaut une minute. Avec ces notations, l'instant d'enregistrement de la n^{e} signature est

$$T_n := \sum_{i=1}^{n} X_i.$$

2) Évaluer la probabilité que les initiateurs de la pétition gagnent leur pari.

Ex 5. Loi de Poisson de grand paramètre
Soit Y_n une variable aléatoire suivant la loi de Poisson de paramètre $n\alpha$, $\alpha > 0$.

1) Montrer que

$$Y_n^* := \frac{Y_n - n\alpha}{\sqrt{n\alpha}} \xrightarrow[n\to+\infty]{\text{loi}} Z,$$

où Z est n'importe quelle v.a. de loi $\mathfrak{N}(0,1)$.

2) Exprimer $P(Y_n \leq n)$ pour $\alpha = 1$. Puis montrer que

$$\lim_{n\to\infty} \mathrm{e}^{-n} \sum_{k=0}^{n} \frac{n^k}{k!} = \frac{1}{2}.$$

Indications : on rappelle que si X et Y sont deux v.a. indépendantes de loi de Poisson de paramètres respectifs α et β, $X+Y$ suit encore une loi de Poisson, de paramètre $\alpha+\beta$. D'autre part, l'espérance et la variance d'une loi de Poisson ont une expression simple en fonction du paramètre…

Ex 6. Cumul d'erreurs
Un programme de calcul utilise J chiffres significatifs après la virgule et arrondit tous les résultats d'opérations à ce format (donc à $\frac{1}{2} \times 10^{-J}$ près). On suppose qu'il effectue 10^6 opérations élémentaires successives, que les erreurs commises pour chacune sont indépendantes, de loi uniforme sur l'intervalle $[-\frac{1}{2} \times 10^{-J}; \frac{1}{2} \times 10^{-J}]$ et que l'erreur sur le résultat final est la somme des erreurs commises sur chaque opération.

1) Évaluer la probabilité pour que l'erreur finale soit inférieure ou égale (en valeur absolue) à $\frac{1}{2} \times 10^{-J+3}$.

2) Puisque l'on parle ici d'arrondi et de chiffres significatifs, utiliser le théorème de Berry-Esseén pour majorer l'erreur commise dans l'approximation faite ci-dessus.

Ex 7. Élection présidentielle
Entre les deux tours d'une élection présidentielle, les instituts Ipsas et Safres procèdent chacun à un sondage portant sur N =1 000 personnes pour savoir qui, de A. ou de B., va gagner. On considère qu'il n'y a pas d'abstention.
Le sondage Ipsas a pour résultat 507 intentions de vote pour A., 493 pour B., alors que le sondage Safres a pour résultat 494 intentions de vote pour A., 506 pour B.

2. Pour simplifier le modèle, nous supposons qu'il n'y a pas de ralentissement du flux des signatures la nuit.

1) En tant que statisticien, qu'en pensez-vous ?
2) En tant que citoyen, qu'en pensez-vous ?

Ex 8. Tension de rupture (sept. 2006)
Le tableau 1 page 3 donne les tensions de rupture en kilogramme-force mesurées sur cent segments de fil de nylon.

TAB. 1 – Un 100-échantillon de tensions de rupture (en kgf) pour un fil de nylon

29,446	31,717	29,818	30,373	31,178	31,146	30,505	28,746	30,037	29,993
28,330	30,734	29,131	29,492	29,554	30,211	29,849	29,147	28,474	31,252
29,003	30,102	28,598	28,552	31,802	30,788	30,060	31,180	30,175	30,556
29,405	31,623	30,786	29,135	29,954	30,782	28,228	30,015	29,856	29,814
30,200	30,688	29,790	30,639	28,846	29,853	29,834	29,140	29,654	29,300
31,195	29,784	31,686	31,001	29,572	29,275	30,160	28,139	29,739	30,869
29,443	27,991	30,020	30,784	28,378	28,554	30,897	28,299	30,723	28,422
30,091	31,060	28,507	31,758	32,052	29,924	30,780	29,086	30,664	30,329
30,922	30,666	29,872	29,035	29,195	31,790	29,412	30,374	28,115	31,782
29,437	31,409	28,381	28,275	29,242	31,132	31,060	29,099	30,127	30,564

Cette tension de rupture est une variable aléatoire Y de loi inconnue et on peut interpréter ce tableau comme les observations $Y_1(\omega), \ldots, Y_{100}(\omega)$ d'un 100-échantillon de la loi de Y. On s'intéresse à la quantité $\theta := P(Y > 31)$.

1) Proposez une valeur numérique pour estimer θ en justifiant votre choix.

2) Proposez un intervalle de confiance au niveau 95% pour θ en utilisant la méthode avec majoration de la variance inconnue d'une certaine loi de Bernoulli qui apparaît naturellement dans ce problème.

Ex 9. Taux de cholestérol
La mesure du taux de cholestérol chez 350 personnes choisies au hasard dans une population a permis de recueillir des données $x_1, \ldots, x_{350}$ dont un article médical présente une synthèse où il est dit que

$$\overline{x} = \frac{x_1 + x_2 + \cdots + x_{350}}{350} = 1,55g\,{}^0\!/_{00} \quad \text{et} \quad \sqrt{\frac{(x_1 - \overline{x})^2 + \cdots + (x_{350} - \overline{x})^2}{349}} = 0,5$$

Construire un intervalle de confiance à 90% pour le taux moyen de cholesterol dans cette population, en précisant bien quelles hypothèses sont faites.

Entraînement supplémentaire facultatif :

Ex 10. Central téléphonique
Un central téléphonique dessert 5 000 abonnés. À un instant donné, chaque abonné a une probabilité égale à 2 % d'utiliser son téléphone et les appels des abonnés sont supposés indépendants. Quel nombre minimal d'appels doit pouvoir traiter simultanément le central pour que sa probabilité d'être saturé à un instant donné soit inférieure à 2,5 % ?

Ex 11. Pixels défectueux

1) On note a un réel strictement positif donné. Montrer que la fonction f définie sur $\mathbb{R}$ par :

$$f(t) = \begin{cases} 0 & \text{si} \quad t < 0, \\ a^2 t e^{-at} & \text{si} \quad t \geq 0 \end{cases}$$

est une densité de probabilité.

2) La durée de vie T en heures d'un certain type de transistor est une variable aléatoire de densité f avec $a = 5,5 \cdot 10^{-5}$. Calculer la probabilité p que la durée de vie du transistor soit inférieure à 100 heures. Calculer la probabilité p' qu'elle soit supérieure à 10 000 heures.

3) Pour simplifier, on pose désormais $p = 1,5 \cdot 10^{-5}$ et $p' = 0{,}90$. Un écran d'ordinateur portable dit à matrice active est formé de 1 296 000 pixels (résolution 1440×900). Un pixel est le plus petit point lumineux que peut afficher l'écran. Chaque pixel est commandé par un transistor du type ci-dessus. On considère que les durées de vie de chacun de ces transistors sont des variables aléatoires indépendantes T_i de même loi à densité f. On note X la variable aléatoire égale au nombre de pixels défectueux au bout de 100 heures de fonctionnement de l'écran. Quelle est la loi exacte de X ? Par quelle loi discrète peut-on l'approximer ? En utilisant cette approximation, évaluer la probabilité qu'au bout de 100 heures de fonctionnement, l'écran ait au moins 4 pixels défectueux.

4) On s'intéresse à la proportion aléatoire U de pixels valides au bout de 10 000 heures de fonctionnement. On note Y le nombre de pixels valides : $U = Y/1\,296\,000$. Donner la loi exacte de Y, son espérance m, sa variance et son écart type σ.

5) D'après la loi des grands nombres et la valeur de p', la proportion U doit être voisine de 90%. Le théorème limite central justifie ici l'approximation de la loi de Y par une loi gaussienne ayant même espérance m et même écart type σ. En déduire les probabilités suivantes :

$$P(U \leq 0{,}90), \quad P(U \leq 0,8995), \quad P(U > 0,9005).$$

Ex 12. Surcharge au décollage

Un avion a une charge maximum au décollage (hors kérosène) de 25 tonnes. Il embarque 318 personnes et leurs bagages (équipage compris). La masse X_i du i^e individu embarqué est une v.a. d'espérance 65 kg et d'écart-type 10 kg. La masse de ses bagages est une v.a. Y_i d'espérance 12 kg et d'écart-type 2 kg.

1) Évaluez la probabilité d'une surcharge au décollage en précisant les hypothèses d'indépendance que vous serez amenés à utiliser.

2) On rappelle que la somme de deux v.a. indépendantes Z_1 et Z_2 de lois gaussiennes respectives $\mathfrak{N}(m_1, \sigma_1)$ et $\mathfrak{N}(m_2, \sigma_2)$ est encore une v.a. gaussienne. Déterminez ses paramètres.

3) La charge maximale de 25 tonnes a été déterminée en tenant compte d'une masse volumique moyenne du kérosène de 0,8 kg/l. En réalité la masse volumique du kérosène peut varier entre 0,755 kg/l et 0,845 kg/l. Du fait des avitaillements successifs d'un aéroport à l'autre, il est impossible de prévoir assez à l'avance la masse volumique du kérosène embarqué pour un vol donné. On est donc amené à modéliser cette masse

volumique par une variable aléatoire gaussienne d'espérance 0,8 kg/l et d'écart-type 0,011 kg/l. Commentez ce choix.

4) Les données techniques concernant l'avion sont les suivantes. Masse maximale au décollage 269 tonnes, masse à vide 120 tonnes, volume des réservoirs de kérosène 152 500 litres. En supposant que la densité du kérosène et la masse des personnes embarquées et de leurs bagages sont indépendantes, recalculez la probabilité de surcharge au décollage.

Ex 13. Un modèle simple pour le mélange de deux gaz

Un récipient hermétique de 2 litres est séparé en deux parties symétriques « gauche » et « droite » par une cloison hermétique munie d'une vanne à large ouverture. La vanne étant fermée, la partie gauche du récipient contient au départ 1 litre d'oxygène et sa partie droite 1 litre d'azote le tout à la pression atmosphérique. On ouvre la vanne de la cloison intermédiaire et on laisse s'effectuer le mélange, puis au bout d'un temps suffisant on ferme la vanne. On mesure alors la proportion d'azote et d'oxygène dans la partie gauche et on constate *expérimentalement* qu'elles sont égales. Le but de cet exercice est l'étude d'un modèle simple permettant de prévoir ce phénomène.

Les molécules étant agitées d'un mouvement incessant, on suppose qu'après fermeture de la vanne, chaque molécule a une probabilité 1/2 de se trouver dans la partie gauche du récipient. On dispose de n molécules d'oxygène et n d'azote. On indexe les molécules d'oxygène de 1 à n et celles d'azote de $n+1$ à $2n$. On note X_i la variable de Bernoulli indicatrice de l'événement *la i-ème molécule se trouve dans la partie gauche après fermeture de la vanne*. On suppose les X_i *indépendantes*. On pose :

$$S_n = \sum_{i=1}^{n} X_i, \qquad T_n = \sum_{i=n+1}^{2n} X_i.$$

La variable S_n est donc le nombre aléatoire de molécules d'oxygène et T_n celui des molécules d'azotes dans la partie gauche après fermeture.

1) Quelle est la loi exacte de S_n, son espérance et sa variance (en fonction de n) ? On a clairement les mêmes résultats pour T_n.

2) Soit $x > 0$ un nombre fixé. On considère l'événement

$$A = \left\{ \frac{n}{2} - \frac{x}{2}\sqrt{n} \leq S_n \leq \frac{n}{2} + \frac{x}{2}\sqrt{n} \right\}.$$

En utilisant le théorème de De Moivre-Laplace, montrer que l'on peut approximer $P(A)$ par $2\Phi(x) - 1$ où Φ est la fonction de répartition de la loi normale $\mathcal{N}(0,1)$. On a bien sûr le même résultat pour $P(B)$, avec

$$B = \left\{ \frac{n}{2} - \frac{x}{2}\sqrt{n} \leq T_n \leq \frac{n}{2} + \frac{x}{2}\sqrt{n} \right\}.$$

3) On suppose désormais que $n - x\sqrt{n} > 0$. On s'intéresse à l'événement

$$C = \left\{ \frac{n - x\sqrt{n}}{n + x\sqrt{n}} \leq \frac{T_n}{S_n} \leq \frac{n + x\sqrt{n}}{n - x\sqrt{n}} \right\}.$$

Montrer que $A \cap B \subset C$.

4) En négligeant l'erreur due à l'approximation gaussienne, proposer à l'aide de $\Phi(x)$ une majoration de $P(A^c \cup B^c)$. En déduire une *minoration* de $P(C)$. On exprimera simplement le résultat final en fonction de la quantité $R(x) = (1 - \Phi(x))$.

5) Application numérique : $n = 10^{22}$, $x = 10$. On utilisera la formule d'encadrement de $1 - \Phi(x)$ pour les « très grandes valeurs de x » fournie à la fin des tables. Commentez le résultat obtenu.