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

IS-Math314 **Année 2005-2006**

Quelques corrigés et indications pour la fiche n° 3

Exercice 2. *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 % ?

Correction. On fixe un instant donné et on note X le nombre d'appels qui auront lieu à cet instant. Si on appelle m le nombre minimal d'appels que le central peut traiter simultanément, le central est saturé dès que $X > m$, et le but de l'exercice est de déterminer m de telle sorte qu'on ait $P(X > m) < 0,025$.

Il y a $n = 5000$ abonnés indépendants et chacun a une probabilité $p = 0,02$ d'utiliser son téléphone à l'instant fixé, donc le nombre X d'appels a cet instant suit la loi binomiale $\mathrm{Bin}(n, p)$.

On remarque que le nombre moyen d'appels est $np = 100$, ce qui est trop grand pour approcher correctement cette loi binomiale par une loi de Poisson. Par contre, $n = 5000$ est assez grand pour qu'on obtienne une bonne approximation de la loi de X en utilisant le théorème de De Moivre-Laplace (ou le théorème limite central appliqué à des v.a. de Bernoulli, ce qui revient au même).

D'après le théorème de de Moivre-Laplace, si X_n a pour loi $\mathrm{Bin}(n, p)$, alors $X_n^* = \frac{X_n - np}{\sqrt{np(1-p)}}$ converge en loi quand n tend vers l'infini vers Z de loi $\mathfrak{N}(0, 1)$. En tenant compte du fait que $np = 100$ et $np(1 - p) = 98$ ici, on a :

$$P(X \le m) = P(\frac{X - 100}{\sqrt{98}} \le \frac{m - 100}{\sqrt{98}}) \simeq P(Z \le \frac{m - 100}{\sqrt{98}})$$

Avoir $P(X > m) < 0,025$ équivaut à avoir $P(X \le m) < 0,975$, et d'aprés la table de valeurs numériques, c'est pour $t \simeq 1,96$ qu'on a $P(Z \le t) \simeq 0,975$. Par conséquent il faut choisir m tel que

$$\frac{m - 100}{\sqrt{98}} < 1,96$$

c'est-à-dire $m > 100 + 1,96\sqrt{98} \simeq 119,4$. Il faut prévoir un standard capable de traiter simultanément au moins 120 appels.

Exercice 4. *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. Evaluer la probabilité pour que l'erreur finale soit inférieure ou égale (en valeur absolue) à $\frac{1}{2} \times 10^{-J+3}$.
Indication : La probabilité demandée est $2\Phi(\sqrt{3}) - 1 \simeq 0,9164$ (où Φ est la fonction de répartition de la loi normale centrée réduite).

Exercice 5. *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.
Indication :

1) Par l'approximation normale on évalue la probabilité de surcharge au décollage à $0,0024$.

2) Question de cours.

3) Evaluer par exemple la probabilité qu'une variable aléatoire gaussienne d'espérance 0,8 kg/l et d'écart-type 0,011 kg/l appartienne à l'intervalle $[0,755; 0,845]$.

4) On obtient $6,8\%$.

Exercice 6. On effectue un sondage auprés d'un échantillon de 930 personnes représentatif de la population française agée de 18 ans et plus. Il indique que 484 d'entre elles vont voter pour A (au deuxième tour d'une élection présidentielle). Donner un intervalle de confiance de niveau au moins égal à 95% pour la proportion d'électeurs votant pour A, en utilisant,

1) l'inégalité de Tchebytchev,

2) une amélioration de l'inégalité de Tchebytchev (obtenue à l'aide de l'inégalité de Chernoff) : si S_n est une variable aléatoire de loi binomiale $\mathrm{Bin}(n,p)$ alors

$$P\left(\left|\frac{S_n}{n}-p\right|>\epsilon\right)\leq 2\exp(-2n\epsilon^2).$$

3) le théorème de de Moivre- Laplace.

Correction. Définissons la variable aléatoire S_n égale au nombre de votant pour A parmi les n personnes interrogées avec $n=930$. On peut considérer que la loi de S_n est une loi binomiale $\mathrm{Bin}(n,p)$ où p est la proportion inconnue d'électeurs votant pour A[1]. On a observé une réalisation de la variable S_n : $S_n(\omega)=484$ et donc la proportion d'électeurs votant pour A observée sur 930 personnes est $484/930=52\%$. On peut estimer la valeur inconnue p par 52% mais on commet une erreur. On va construire un intervalle $I_n(\omega)$ centré autour de $\frac{S_n}{n}(\omega)=52\%$ de telle sorte que $P(\{\omega\in\Omega\,:\,p\in I_n(\omega)\})$ soit controllée, par exemple ici on demande à ce qu'elle soit d'au moins 95%. Donc, en posant $I_n=[\frac{S_n}{n}-\epsilon,\frac{S_n}{n}+\epsilon]$

$$P(\{\omega\in\Omega\,:\,p\in I_n(\omega)\})=1-P\left(\left|\frac{S_n}{n}-p\right|>\epsilon\right).$$

Nous allons maintenant majorer $P\left(\left|\frac{S_n}{n}-p\right|>\epsilon\right)$ de trois manière différentes ce qui va nous permettre de trouver une valeur de ϵ telle que $P\left(\left|\frac{S_n}{n}-p\right|>\epsilon\right)\leq 0,05$.

1) Par l'inégalité de Tchebytchev,

$$P\left(\left|\frac{S_n}{n}-p\right|>\epsilon\right)\leq\frac{p(1-p)}{n\epsilon^2}\leq\frac{1}{4n\epsilon^2}.$$

On prend donc $\epsilon\geq\sqrt{\frac{1}{4n\times 0,05}}$ cad $\epsilon=0,073$ et on obtient comme intervalle de confiance $[0,446;0,593]$.

2) Par l'inégalité de Hoeffding,

$$P\left(\left|\frac{S_n}{n}-p\right|>\epsilon\right)\leq 2\exp(-2n\epsilon^2),$$

on choisit $\epsilon\geq\sqrt{\frac{-\ln(0,05/2)}{2n}}$, soit $\epsilon=0,045$ et on obtient comme intervalle de confiance $[0,475;0,565]$ (ce qui est un peu plus précis que le précédent).

3) Par l'approximation normale, on sait que la loi de $\sqrt{n}p(1-p)(\frac{S_n}{n}-p)$ est proche de la loi normale centrée réduite et comme $p(1-p)\leq\frac14$, on a

$$P\left(2\sqrt{n}\left|\frac{S_n}{n}-p\right|>t\right)\leq P\left(\sqrt{\frac{n}{p(1-p)}}\left|\frac{S_n}{n}-p\right|>t\right)\simeq 2(1-\phi(t)),$$

[1]En réalité c'est une loi hypergéométrique $\mathcal{H}(N,N_A,n)$ où N est le nombre total d'inscrits sur les listes élétorales et N_A le nombre des inscrits votant pour A. Mais, N est suffisament grand pour qu'on puisse approximer la loi hypergéométrique $\mathcal{H}(N,N_A,n)$ par une loi binomiale $\mathrm{Bin}(n,p)$ où $p=\frac{N_A}{N}$.

où ϕ est la fonction de répartition d'une loi normale centrée réduite. On choisit donc t tel que $2(1-\phi(t)) \leq 0,05$ soit $\phi(t) \geq 0,975$ ce qui donne $t = 1,96$ d'où $\epsilon = \frac{t}{2\sqrt{n}} \simeq 0,032$ et on obtient comme intervalle de confiance $[0,487; 0,552]$.

Exercice 7. Pour déterminer la concentration d'un certain produit dans une solution, on effectue des dosages à l'aide d'une technique expérimentale donnée. On admet que le résultat de chaque dosage est une variable aléatoire normale dont l'espérance m est la valeur que l'on cherche à déterminer et dont l'écart-type est égal à $0,05$ mg/l. Six dosages indépendants d'une solution de concentration inconnue ont été effectués. Les résultats suivants ont été obtenus :

$$2,97 \quad 3,01 \quad 2,98 \quad 2,94 \quad 3,03 \quad 2,95 \quad \text{mg/l.}$$

Donner un intervalle de confiance de la concentration cherchée de niveau 95%.
Indication : On obtient $I = [2,94; 3,02]$ (sans utiliser le TLC).