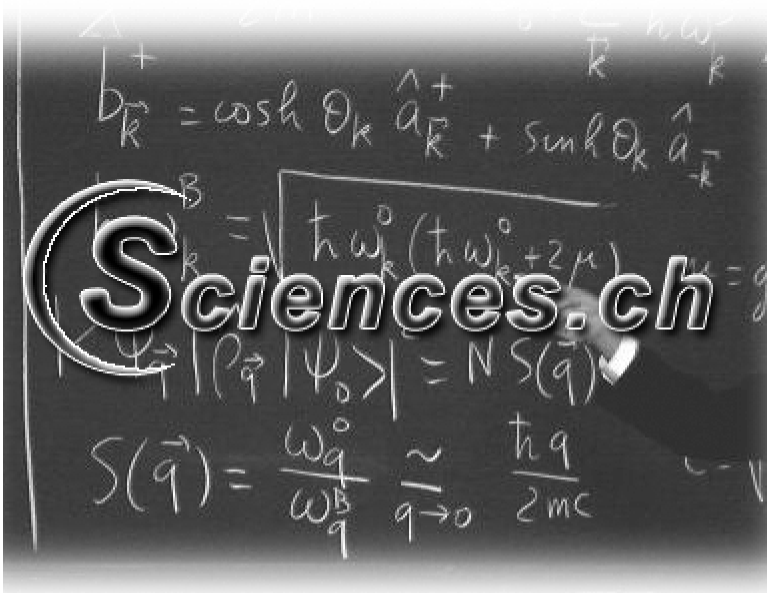


EXERCICES DE

PROBABILITÉS ET STATISTIQUES

EXERCICE 1.

Niveau : Université
Auteur : Dhyne Miguël (01.09.04, miguel.dhyne@win.be)
Mots-clés : concept de probabilité

Enoncé :

On tire au hasard une carte dans un jeu de 52 cartes. Avec quelle probabilité cette carte est-elle soit une dame soit un cœur ?

Solution :

Un jeu de 52 cartes présente :

- 13 cartes de pique
- 13 cartes de carreau
- 13 cartes de trèfle
- 13 cartes de cœur

Et dans chaque catégorie, nous retrouvons les cartes de 1 à 10 ainsi que le valet, la dame et le roi.

Nous avons alors 13 "chances" de tirer une carte de cœur. Et nous avons 4 "chances" de tirer une dame. Mais il ne faut pas oublier que dans les cartes de cœur, il y a une dame, donc, il nous reste, après avoir considérer les cartes de cœur que 3 "chances" d'avoir une dame.

Soit A, le fait que nous avons tiré une dame ou une carte de cœur.

Sa probabilité est donc égale à :

$$P(A)=\frac{\textit{nombre de chances}}{\textit{nombre de cartes total}}=\frac{13+3}{52}=\frac{16}{52}=\frac{4}{13}$$

EXERCICE 2.

Niveau : Université
Auteur : Dhyne Miguël (01.09.04, miguel.dhyne@win.be)
Mots-clés : concept de probabilité

Enoncé :

Un barman propose au client le jeu suivant : on lance trois dés ; si le total des points obtenus est inférieur à 10 il donne au client une somme $S = 4$ €, s'il est supérieur à 10 le client lui donne une somme égale à la différence entre 10 et le résultat obtenu, s'il est égal à 10 aucune somme n'est échangée.

Le jeu est-il équitable ? Sinon quel devrait être le montant de la somme S pour qu'il soit équitable ?

Solution :

Classons dans un tableau les différents résultats que nous pouvons obtenir en lançant trois dés, nous allons également poser que ces événements sont équiprobables :

Total des points des trois dés lancés	**Somme reçue (+) ou donnée (-) par le client**	*Probabilité*
3	$+4$	$1/216$
4	$+4$	$3/216$
5	$+4$	$6/216$
6	$+4$	$10/216$
7	$+4$	$15/216$
8	$+4$	$21/216$
9	$+4$	$25/216$
10	0	$27/216$
11	-1	$27/216$
12	-2	$25/216$
13	-3	$21/216$
14	-4	$15/216$
15	-5	$10/216$
16	-6	$6/216$

17	-7	$3/216$
18	-8	$1/16$

Le jeu est équitable si et seulement si la moyenne est nulle.

Calculons cette moyenne :

$$Moyenne = \left[4\cdot\left(1+3+6+10+15+21+25\right)-27-2\cdot 25-3\cdot 21-4\cdot 15-5\cdot 10-6\cdot 6-7\cdot 3-8\cdot 1\right]\frac{1}{216}$$

$$Moyenne = -\frac{9}{216}$$

Le joueur a donc tendance à perdre.

Pour que ce jeu soit équitable, calculons la valeur que doit prendre S. Pour cela remplaçons par x la somme donnée par le barman.

$$Moy = ((1+3+6+10+15+21+25)x-1\cdot 27-2\cdot 25-3\cdot 21-4\cdot 15-5\cdot 10-6\cdot 6-7\cdot 3-8\cdot 1)\frac{1}{216}=0$$

$$\Leftrightarrow$$

$$x = 3.90\,€$$

Pour que le jeu soit équitable, le barman doit donner 3.90 € au client si le résultat est inférieur à 10.

EXERCICE 3.

Niveau : Université
Auteur : Dhyne Miguël (02.09.04, miguel.dhyne@win.be)
Mots-clés : concept de probabilité

Enoncé :

Quelle est la probabilité que dans une famille de trois enfants les deux plus jeunes soient des garçons ?

Solution :

Soit G, un garçon et F, une fille. Regardons les combinaisons possibles :

$$GGG \quad GGF \quad GFF \quad GFG \quad FFF \quad FGG \quad FFG \quad FGF$$

Nous avons donc huit configurations possibles.

(nous aurions pu savoir cela à l'aide de l'analyse combinatoire, on choisit trois éléments (les trois enfants de la famille) par mi un ensemble de deux (garçon et fille), il s'agit donc d'un arrangement multiple : $2^3 = 8$)

Sur ces huit configurations, nous en avons 2 qui présentent les deux plus jeunes enfants comme garçon.

La probabilité de cet événement est donc de :

$$P(2 \; jeunes \; garçons) = \frac{2}{8} = \frac{1}{4}$$

EXERCICE 4.

Niveau : Université
Auteur : Dhyne Miguël (02.09.04, miguel.dhyne@win.be)
Mots-clés : concept de probabilité

Enoncé :

Calculez les probabilités de gagner aux différents rangs en jouant au LOTTO belge. (le LOTTO belge consiste à cocher 6 numéros dans une grilles de 42 nombres).

Solution :

Rang 1 → 6 chiffres corrects

Rang 2 → 5 chiffres corrects + nombre complémentaire

Rang 3 → 5 chiffres corrects

Rang 4 → 4 chiffres corrects

Rang 5 → 3 chiffres corrects

Pour le Rang 1, nous avons une chance de trouver la bonne combinaison, c'est-à-dire 6 chiffres parmi 42 :

$$P(Rang1)=\frac{1}{C_{42}^{6}}=\frac{1}{42!/[(42-6)!\cdot 6!]}=\frac{1}{5245786}\cong 0.0000002$$

(le nombre qui se trouve sous la barre de fraction est donc le nombre de grilles totales possibles.)

Pour le Rang 2, nous avons du cocher 5 chiffres sur les 6 »principaux » tirés, et un autre qui doit correspondre au numéro complémentaire tiré.

$$P(Rang2)=\frac{C_{6}^{5}\cdot 1}{C_{42}^{6}}=\frac{6}{5245786}\cong 0.000001$$

Pour le Rang 3, nous avons donc cochés 5 chiffres parmi les 6 tirés, et nous avons également coché un chiffre qui ne correspond pas à ceux tirés (n'oublions pas que 7 numéros sont tirés, car nous avons la combinaison des 6 et le numéro complémentaire), ce chiffre peut être un des 35 restant :

$$P(Rang3)=\frac{35\cdot C_{6}^{5}}{C_{42}^{6}}=\frac{35\cdot 6}{5245786}\cong 0.00004$$

Pour le Rang 4, nous avons cochés 4 numéros sur les 6 tirés, et nous avons cochés deux autres chiffres ne figurant pas dans la combinaison gagnante du jour, les possibilités respectives sont de l'ordre de 35 et 34 :

$$P(rang 4) = \frac{\cdot C_{35}^{2} \cdot C_{6}^{4}}{C_{42}^{6}} = \frac{595 \cdot 15}{5245786} \cong 0.0017$$

Pour le Rang 5 , nous avons donc cochés 3 nombres gagnants et trois autres étrangers au tirage du jour :

$$P(Rang 5) = \frac{C_{35}^{3} \cdot C_{6}^{3}}{C_{42}^{6}} = \frac{6545 \cdot 20}{5245786} \cong 0.025$$

EXERCICE 5.

Niveau : Université
Auteur : Dhyne Miguël (01.09.04, miguel.dhyne@win.be)
Mots-clés : concept de probabilité

Enoncé :

Le sexe d'un enfant est imprévisible. Dans le cas de la naissance d'un enfant, certaines personnes pensent qu'il y a autant de chance d'avoir une fille que d'avoir un garçon. Voici un tableau reprenant de 1982 à 1992 le nombre de naissance de chaque sexe :

Année	Nombre de naissance total	Nombre de naissance « garçons »	Nombre de naissance « filles »
1982	120241	61831	58410
1983	117145	60209	56936
1984	115651	59331	56320
1985	114092	58572	55520
1986	117114	60493	56621
1987	117354	60390	56964
1988	119779	61520	58259
1989	120904	61948	58656
1990	123776	63421	60335
1991	125924	64586	61338
1992	124774	63883	60891

Ces données confirment-elles cette opinion ? Estimez les probabilités de « naître filles » et de « naître garçon » pour chaque année, donnez-en un petit commentaire.

Solution :

Construisons ce même type de tableau, avec les probabilités correspondantes.

Notons que la probabilité sera calculée de cette manière :

$$P(garçon\ ou\ fille) = \frac{nombre\ naissance\ garçon\ ou\ fille}{nombre\ total\ naissance}$$

Economisons de l'énergie, ne calculons que la probabilité pour les garçons (par exemple), ensuite soustrayons ce résultat de 1 pour connaître la probabilité de « naître fille ».

Année	$P(garçon)$	$P(filles)$
1982	0.51	0.49
1983	0.51	0.49
1984	0.51	0.49
1985	0.51	0.49
1986	0.52	0.48

1987	0.51	0.49
1988	0.51	0.49
1989	0.51	0.49
1990	0.51	0.49
1991	0.51	0.49
1992	0.51	0.49

Nous remarquons tout de suite que l'opinion de certaines personnes n'est pas tout à fait correcte. En effet, nous avons 51% de naissance de garçons contre 49% de naissance de filles. L'opinion du $50\% - 50\%$ n'est, cependant, pas très loin de la « réalité ».

Nous pouvons dégager une année « spéciale », celle de 1986 où le taux de naissance de garçon est supérieur aux autres années. Mais ce n'est rien d'exceptionnel…

EXERCICE 6.

Niveau : Université
Auteur : Dhyne Miguël (01.09.04, miguel.dhyne@win.be)
Mots-clés : probabilité conditionnelle

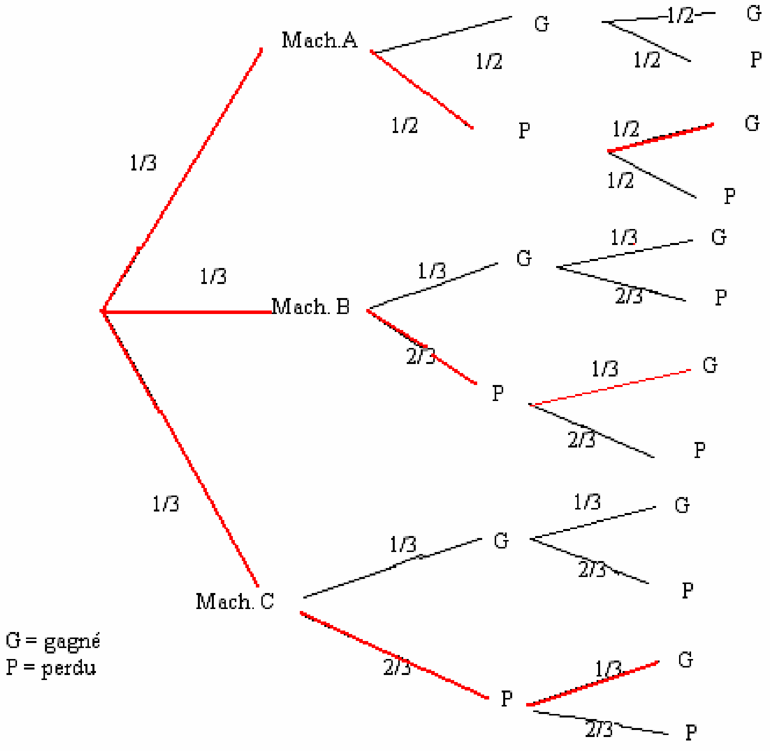
Enoncé :

Un joueur apprend que sur trois machines à sous, l'une permet de gagner avec une probabilité de $\frac{1}{2}$ et chacun des deux autres avec une probabilité de $\frac{1}{3}$. Si le joueur choisit une machine au hasard et joue deux fois, quelle est la probabilité qu'il perde la première fois et gagne la seconde fois ?

Solution :

Soit A, l'événement suivant : le joueur perd la première fois et gagne la seconde.

Considérons que les probabilités de choix d'une ou l'autre cabine soient égales. Etablissons un arbre résumant la situation et mettons en évidence les « chemins » qui nous intéressent :



$$P(A)=\frac{1}{3}\cdot\left(\frac{1}{2}\cdot\frac{1}{2}\right)+\frac{1}{3}\cdot\left(\frac{2}{3}\cdot\frac{1}{3}\right)+\frac{1}{3}\cdot\left(\frac{2}{3}\cdot\frac{1}{3}\right)=0.23$$

EXERCICE 7.

Niveau : Université
Auteur : Dhyne Miguël (01.09.04, miguel.dhyne@win.be)
Mots-clés : probabilité conditionnelle

*** Exercice tiré de « Que sais-je ? Les probabilités », Albert Jacquard, N°1571, p.53-54 ***

Enoncé :

Devant un malade souffrant d'une angine, le médecin A pense que la cause peut en être un streptocoque. Compte tenu de l'ensemble des symptômes, il pense qu'il s'agit plutôt de streptocoques, mais il n'en conclut par d'autres causes, par exemple un virus. Toute réflexion faite, il admet que la probabilité du streptocoque est quatre fois plus élevée que celle des autres causes ; ce qui peut s'écrire :

$$_{A}P(S) = 0.8 \qquad _{A}P(\overline{S}) = 0.2$$

Mais, selon le diagnostique adopté, le traitement est tout différent, une erreur peut avoir de graves conséquences pour le malade. Il s'adresse alors à un confrère B qui confirme les deux causes possibles, mais qui, toute réflexion faite, pense que dans ce cas, le virus est quatre fois plus probable que le streptocoque, ce qui peut s'écrire :

$$_{B}P(S) = 0.2 \qquad _{B}P(\overline{S}) = 0.8$$

Ils décident alors tous deux de faire une analyse de laboratoire pour déceler la présence éventuelle de streptocoques. Mais les techniques de laboratoire comportent des risques d'erreur ; nous admettrons que, dans le laboratoire disponible, le streptocoque, lorsqu'il est présent chez le malade a environ trois chances sur dix de n'être pas décelé et, lorsqu'il n'est pas présent chez le malade, une chance sur dix d'être décelé dans les préparations, par suite d'erreur ou de contamination.

En représentant par « oui » la réponse positive (il y a des streptocoques) et par « non » la réponse négative (il n'y en a pas), les informations peuvent être exprimées par les probabilités :

$$P(oui/S) = 0.7 \qquad P(oui/\overline{S}) = 0.1$$
$$P(non/S) = 0.3 \qquad P(non/\overline{S}) = 0.9$$

Les médecins font cinq prélèvements ; la réponse du laboratoire est :

PRELEVEMENT N°	1	2	3	4	5
Réponse R	oui	non	oui	non	oui

Quel est le diagnostic de chaque médecin compte tenu de cette observation ?

Solution :

Dans le cas où le malade souffre effectivement d'une infection à streptocoque, la probabilité de la réponse du laboratoire est :

$$P(R/S)=P(oui/S)\cdot P(non/S)\cdot P(oui/S)\cdot P(non/S)\cdot P(oui/S)$$

Nous avons obtenu cette expression en appliquant le principe des probabilités composées au cas d'évènements indépendants, d'où :

$$P(R/S)=0.7^3\cdot 0.3^2$$

Dans le cas où le malade n'a au contraire pas d'infection à streptocoque, la réponse du laboratoire est, en appliquant toujours le même principe :

$$P(R/\overline{S})=0.1^3\cdot 0.9^2$$

L'opinion du médecin A, une fois connue cette observation, peut s'exprimer par la probabilité qu'il accorde maintenant à la présence d'un streptocoque ${}_A P(S/R)$ qui s'écrit, d'après le théorème de Bayes :

$${}_A P(S/R)=\frac{{}_A P(S)\cdot P(R/S)}{{}_A P(S)\cdot P(R/S)+{}_A P(\overline{S})\cdot P(R/\overline{S})}=99.3\%$$

L'opinion du médecin B devient de la même façon :

$${}_B P(S/R)=\frac{{}_B P(S)\cdot P(R/S)}{{}_B P(S)\cdot P(R/S)+{}_B P(\overline{S})\cdot P(R/\overline{S})}=90.5\%$$

Malgré l'opposition de leurs opinions initiales et en dépit de l'imprécision des résultats du laboratoire, les deux médecins sont maintenant d'accord : il faut soigner le malade en admettant que son mal est dû aux streptocoques. Ils ne divergent plus que leur évaluation du risque d'erreur de diagnostic : 0.7% selon A (au lieu de 20% avant l'examen de laboratoire), 9.5% selon B (au lieu de 80% s'il laissait la décision à A, 20% si A s'en remettait à lui).

EXERCICE 8.

Niveau : Université
Auteur : Dhyne Miguël (06.09.04, miguel.dhyne@win.be)
Mots-clés : probabilité conditionnelle

Enoncé :

Deux joueurs A et B jouent à une série de jeux pour lesquels A a une probabilité de gagner. Chaque jeu a un gagnant unique. L'un des joueurs gagne l'ensemble soit s'il gagne deux jeux consécutifs, soit s'il gagne 3 jeux au total.

1. Dressez un arbre traduisant l'ensemble du problème.

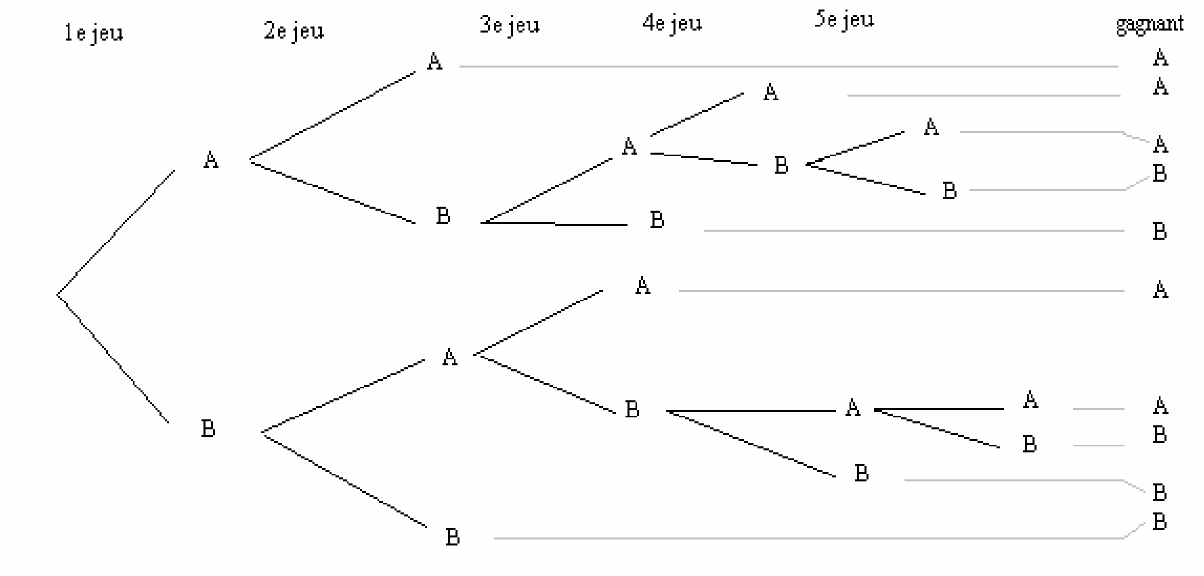
2. Quelle est la probabilité que A gagne ?

3. Quelle est la probabilité que les joueurs doivent jouer plus de 3 jeux avant de connaître le gagnant ?

4. Quelle est la probabilité que ce soit B qui gagne l'ensemble si les joueurs ont dû participer à plus de trois jeux ?

Solution :

1.



2. $P(A\ gagne) = \left(\dfrac{1}{3}\right)^2 + \left(\dfrac{1}{3}\right)^3 \cdot \dfrac{2}{3} + \left(\dfrac{1}{3}\right)^2 + \left(\dfrac{2}{3}\right)^2 + \left(\dfrac{1}{3}\right)^2 \cdot \dfrac{2}{3} + \left(\dfrac{1}{3}\right)^3 \cdot \dfrac{2}{3} = 0.24$

3. $P(+de\ 3\,jeux) = 1 - P(3\,jeux\,\max imum)$

$$P(+de\ 3\,jeux)=1-\left[\left(\frac{1}{3}\right)^2+\left(\frac{1}{3}\right)\cdot\left(\frac{2}{3}\right)^2+\left(\frac{2}{3}\right)\cdot\left(\frac{1}{3}\right)^2+\left(\frac{2}{3}\right)^2\right]$$

$$P(+de\ 3\,jeux)=0.30$$

4. $P(B\ gagne si+de\ 3\,jeux)=P(B/T)=\dfrac{P(B\cap T)}{P(T)}=0.55$

EXERCICE 9.

Niveau : Université
Auteur : Isoz Vincent (20.09.07)
Mots-clés : Questionnaires, Théorème centrale limite

Enoncé :

Un examen se présente sous forme de questionnaire à choix multiples. Notons *N* le nombre de questions et *n* le nombre de choix pour chacune des questions. On suppose qu'il n'y a qu'une réponse correcte pour chacun des *n* choix. On fixe *t* vérifiant $0 < t < 1$ et représentant la proportion de questions correctes à fournir pour passer l'examen (par exemple $t = 0.8$). De plus on se donne α vérifiant $0 < \alpha < 1$ (par exemple $\alpha = 0.05$).

Le problème consiste à déterminer les couples (N, n) pour lesquels la probabilité de réussir l'examen sachant qu'on a répondu au hasard est inférieure à α .

[Dit autrement on désire déterminer (N, n) de sorte que parmi les personnes ayant répondu au hasard à toutes les questions de l'examen la proportion de ceux qui le réussissent soit inférieure à α].

Solution :

Notons *X* le nombre de réponses correctes fournies par une personne répondant au hasard à toutes les questions de l'examen. *X* est une variable aléatoire suivant la loi binomiale de paramètres *N* et $p = \dfrac{1}{n}$. Où *p* est la probabilité de répondre correctement à une question.

Le problème se traduit ainsi : on cherche les couples (N, n) satisfaisant l'inégalité

$$P(X \geq Nt) \leq \alpha .$$

Où bien de façon équivalente :

$$P(X < Nt) \geq 1 - \alpha .$$

Si on suppose que le nombre Nt est entier:

$$P(X < Nt) = \sum_{k=0}^{N \cdot t - 1} \binom{N}{k} p^k (1-p)^{N \cdot t - 1 - k} .$$

Il faudrait donc déterminer les couples (N, n) qui vérifient:

$$\sum_{k=0}^{N \cdot t - 1} \binom{N}{k} p^k (1-p)^{N \cdot t - 1 - k} \geq 1 - \alpha$$

où $p=\dfrac{1}{n}$.

Il est très difficile d'aboutir à une relation "simple" reliant N à n à partir de cette inégalité.

Nous allons par conséquent faire une approximation en utilisant le Théorème central limite de Moivre-Laplace qui pour rappel nous dit:

Soit X_n une suite de variables aléatoires indépendantes de même loi de Bernoulli de paramètre p, $0<p<1$. Notons :

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

Remarque : S_n suit la loi binomiale de paramètres *n*, *p*.

Alors:

$$F_n(x)=P\left(\frac{S_n-np}{\sqrt{np(1-p)}}\leq x\right)$$

tend uniformément vers $\Phi(x)$ sur $\mathbb{R}$ lorsque $n\rightarrow\infty$.

Rappel : Φ est la fonction de répartition de la loi normale centrée réduite:

$$\Phi(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-\frac{t^2}{2}}dt\ .$$

Revenons-en à notre inégalité :

$$P(X<Nt)\geq 1-\alpha\Leftrightarrow P\left(\frac{X-Np}{\sqrt{Np(1-p)}}<\frac{N(t-p)}{\sqrt{Np(1-p)}}\right)\geq 1-\alpha\ .$$

Par le théorème précédent,

$$P\left(\frac{X-Np}{\sqrt{Np(1-p)}}<\frac{N(t-p)}{\sqrt{Np(1-p)}}\right)\approx\Phi\left(\frac{N(t-p)}{\sqrt{Np(1-p)}}\right).$$

L'erreur que l'on commet avec cette approximation tend vers zéro lorsque $N\rightarrow+\infty$ car dans le théorème la convergence est **uniforme**.

Remarque : Le fait qu'ici l'inégalité est stricte ne change rien.

Nous allons par conséquent résoudre l'inégalité

$$\Phi\left(\frac{N(t-p)}{\sqrt{Np(1-p)}}\right) \geq 1-\alpha .$$

On a :

$$\frac{N(t-p)}{\sqrt{Np(1-p)}} \geq \Phi^{-1}(1-\alpha) \Leftrightarrow \sqrt{N} \geq \frac{\Phi^{-1}(1-\alpha)\sqrt{p(1-p)}}{(t-p)}$$

et donc:

$$N \geq \frac{\left(\Phi^{-1}(1-\alpha)\right)^2 p(1-p)}{(t-p)^2} .$$

Exemple numérique:

On considère que l'examen est réussi si au moins le 80% des réponses sont correctes. On choisit par conséquent : $t = 0.8$.

On désire de plus que parmi les personnes ayant répondu au hasard à toutes les questions de l'examen, la proportion de celles qui le réussissent soit inférieure à 1%. On choisit donc $\alpha = 0.01$.

Dans les tables on lit $\Phi^{-1}(0.99) \approx 2.326$.

Et :

$$N \geq \frac{2.326^2 p(1-p)}{(0.8-p)^2} .$$

Si on décide que pour chaque question il y aura trois choix possibles alors $p = 1/3$ et on obtient $N \geq 6$.

EXERCICE 10.

Niveau : Université
Auteur : Anonyme (04.11.07)
Mots-clés : loi binomiale

Enoncé :

Charles-Basile passe un test de statistique et probabilités. Le test est un questionnaire à choix multiple comportant 6 questions. Chaque question a trois réponses possibles, dont une est juste. Charles-Basile réussi le test s'il répond correctement à au moins 4 questions.

1. Si Charles-Basile répond au hasard, quelle est l'espérance du nombre de réponses correctes ? Quelle est sa variance ? Quelle est la probabilité que Charles-Basile réussisse le test ?

2. Etant un peu mieux préparé, Charles-Basile est capable d'éliminer une réponse incorrecte. Il choisit sa réponse au hasard parmi les deux possibilités restantes. Trouver l'espérance et la variance de nombre de réponses correctes dans ce cas. Quelle est la probabilité que Charles-Basile réussisse l'examen ?

Solution :

1. Notons X la variable aléatoire égale au nombre de réponses correctes. La probabilité de répondre correctement à une question donnée est de $\dfrac{1}{3}$ et par suite celle de répondre faux de $\dfrac{2}{3}$. La probabilité que Charles-Basile réponde correctement à n questions $(0 \leq n \leq 6)$ est:

$$P(X=n)=\left(\frac{1}{3}\right)^n \cdot \left(\frac{2}{3}\right)^{6-n} \cdot C_n^6$$

On reconnaît ici la loi binomiale $Binomiale(6,1/3)$. Donc:

$$E[X]=6\cdot\frac{1}{3}=2 \text{ et } V[X]=6\cdot\frac{1}{3}\cdot\frac{2}{3}=\frac{4}{3}$$

La probabilité que Charles-Basile réussisse le test est:

$$P(X\geq 4)=P(X=4)+P(X=5)+P(X=6)=\left(\frac{1}{3}\right)^4\cdot\left(\frac{2}{3}\right)^2\cdot C_4^6+\left(\frac{1}{3}\right)^5\cdot\left(\frac{2}{3}\right)\cdot C_5^6+\left(\frac{1}{3}\right)^6\cdot C_6^6$$

$$=\frac{4}{3^6}\cdot 15+\frac{2}{3^6}\cdot 6+\frac{1}{3^6}=\frac{73}{3^6}\approx 0.1$$

2. A présent la probabilité de répondre correctement à une question donnée est de 1/2 et:

$$P(X=n)=\left(\frac{1}{2}\right)^{n}\cdot\left(\frac{1}{2}\right)^{6-n}\cdot C_n^6=\frac{C_n^6}{2^6}$$

X suit la loi $Binomiale(6,1/2)$. L'espérance et la variance sont:

$$E[X]=6\cdot\frac{1}{2}=3 \text{ et } V[X]=6\cdot\frac{1}{2}\cdot\frac{1}{2}=\frac{3}{2}.$$

La probabilité que Charles-Basile réussisse le test est:

$$P(X\geq 4)=P(X=4)+P(X=5)+P(X=6)=\frac{C_4^6}{2^6}+\frac{C_5^6}{2^6}+\frac{C_6^6}{2^6}=\frac{22}{2^6}\approx 0.34.$$

EXERCICE 11.

Niveau : Université
Auteur : Anonyme (04.11.07)
Mots-clés : loi de Poisson, formule de Bayes

Enoncé :

Le nombre de rhumes attrapés par un individu en l'espace d'un an est une variable aléatoire de Poisson de paramètre $\lambda = 5$. Un remède-miracle (basé sur l'effet de la vitamine C à haute dose) a été lancé sur le marché. Il abaisse le paramètre λ à 3 pour 75% de la population. Pour les 25 derniers pourcent de la population le remède n'a pas d'effet appréciable.

1. Si 200 individus essaient le remède, quel est le nombre moyen de rhumes qu'ils attrapent durant une année ?

2. Sachant qu'un individu essaie ce médicament pendant un an et attrape deux rhumes, quelle est la probabilité que le remède ait eu un effet sur lui ?

Solution :

1. $0.75 \cdot 200 \cdot 3 + 0.25 \cdot 200 \cdot 5 = 700$.

2. Notons $A =$ « le médicament a eu un effet sur l'individu » et soit X la variable aléatoire égale au nombres de rhumes attrapés en un an par un individu ayant pris le médicament. Il s'agit de calculer $P(A \mid X = 2)$. Par la formule de Bayes on a :

$$P(A \mid X = 2) = \frac{P(X = 2 \mid A) \cdot P(A)}{P(X = 2 \mid A) \cdot P(A) + P(X = 2 \mid \overline{A}) \cdot P(\overline{A})}.$$

Par l'énoncé on sait que:

$$P(A) = 0.75,\ P(X = 2 \mid A) = e^{-3}\frac{3^2}{2!},\ P(X = 2 \mid \overline{A}) = e^{-5}\frac{5^2}{2!}.$$

Ainsi:

$$P(A \mid X = 2) = \frac{e^{-3}\frac{3^2}{2!} \cdot 0.75}{e^{-3}\frac{3^2}{2!} \cdot 0.75 + e^{-5}\frac{5^2}{2!} \cdot 0.25} \approx 0.89.$$

EXERCICE 12.

Niveau : Université
Auteur : Anonyme (04.11.07)
Mots-clés : loi Binomiale, loi de Poisson, formule de Bayes

Enoncé :

Lors du tour de France (TDF) 1999 qui comporte 25 étapes, on sait que 7 fois sur 10, l'étape (n'importe laquelle) est remportée par un cycliste dopé.

1. Calculer la probabilité que sur les quatre premières étapes du TDF, au moins 3 soient remportées par des cyclistes non dopés.

2. Sur l'ensemble du TDF, combien d'étapes peuvent espérer gagner les cyclistes non dopés ?

3. A la fin de chaque étape, le vainqueur est soumis à un contrôle anti-dopage. Si le contrôle révèle que le cycliste est dopé, ce dernier est disqualifié. Malheureusement, ce contrôle n'est pas fiable à 100%. En effet, la probabilité qu'on disqualifie un cycliste réellement dopé est de 0.6. En revanche la probabilité de disqualifier un cycliste non dopé est de 0.05.

- **(a)** Calculer la probabilité qu'un cycliste dopé gagne une des étapes du TDF et qu'il ne soit pas disqualifié.
- **(b)** Lors de la 4$^{\text{ème}}$ étape, le vainqueur a été disqualifié, calculer la probabilité qu'il soit réellement dopé.

4. En moyenne 2 coureurs chutent chaque jour. Quelle est la probabilité que 5 cyclistes chutent en trois jours.

Solution :

1. Le nombre d'étapes remportées par des cyclistes non dopés lors des quatre premières étapes est une v.a X suivant la loi $Binomiale(4,3/10)$. On a donc:

$$P(X \geq 3) = P(X = 3) + P(X = 4) = C_3^4 \cdot 0.3^3 \cdot 0.7 + C_4^4 \cdot 0.3^4 \approx 0.084 .$$

2. Si X est la v.a représentant le nombre d'étapes gagnées par des cyclistes non dopés, nous savons que X suit la loi $Binomiale(25,3/10)$. Les cyclistes non dopés peuvent espérer gagner:

$$E[X] = 25 \cdot \frac{3}{10} = 7.5 \text{ étapes}$$

3(a). La probabilité qu'une étape donnée soit gagnée par un cycliste dopé et que ce cycliste ne soit pas disqualifié est $0.7 \cdot (1-0.6) = 0.28$.

3(b). Notons $A =$ « le vainqueur de la 4$^{\text{ème}}$ étape a été disqualifié » et $B =$ « à la 4$^{\text{ème}}$ étape le vainqueur était dopé ». On veut calculer $P(B|A)$.

On a:

$$P(A \mid B) = 0.6\,,\; P\left(A \mid \overline{B}\right) = 0.05$$

Nous pouvons utiliser la formule de Bayes :

$$P(B \mid A) = \frac{P(A \mid B) \cdot P(B)}{P(A \mid B) \cdot P(B) + P\left(A \mid \overline{B}\right) \cdot P\left(\overline{B}\right)} = \frac{0.6 \cdot 0.7}{0.6 \cdot 0.7 + 0.05 \cdot 0.3} \approx 0.97$$

4. En trois jours, la moyenne de coureurs qui chutent est de 6. Le nombre de cyclistes qui chutent en trois jours est une v.a. X suivant une loi de Poisson de paramètre $\lambda = 6$. Par suite:

$$P(X = 5) = e^{-6} \cdot \frac{6^5}{5!} \approx 0.16\,.$$

EXERCICE 13.

Niveau : Université
Auteur : Anonyme (04.11.07)
Mots-clés : loi Binomiale, loi de Poisson

Enoncé :

Remarque : les parties A et B sont indépendantes

A. Marsida part à la cueillette des champignons. Elle ne sait pas faire la différence entre un champignon comestible et un champignon toxique. On estime que la proportion de champignons toxiques se trouvant dans les bois s'élève à 0.7.

1. Marsida ramasse six champignons au hasard. Calculer la probabilité qu'elle ramasse exactement quatre champignons toxiques.

2. Marsida invite Richard à une cueillette. Richard connaît bien les champignons ; sur dix champignons qu'il ramasse, neuf sont comestibles. Ce jour-là, il ramasse quatre champignons et Marsida en ramasse trois. Calculer la probabilité que tous les champignons soient comestibles.

B. Richard cueille en moyenne 12 champignons par heure.

1. Calculer la probabilité qu'il ramasse exactement 8 champignons en une heure.

2. Calculer la probabilité qu'il ramasse au moins un champignon en 20 minutes.

Solution :

A.1 : Le nombre de champignons toxiques ramassés par Marsida suit une loi binomiales de paramètres $n=6$ et $p=0.7$. Par conséquent la probabilité cherchée est $C_4^6 \cdot 0.7^4 \cdot 0.3^2 \approx 0.32$.

A2: Les événements $A=$ « les 3 champignons de Marsida sont comestibles » et

$B=$ « les 4 champignons de Richard sont comestibles » sont indépendants. Par conséquent:

$$P(A \cap B) = P(A) \cdot P(B) = 0.3^3 \cdot 0.9^4 \approx 0.018$$

B.1 : Notons *X* la v.a. égale au nombre de champignons ramassés par Richard en une heure. *X* suit une loi de Poisson de paramètre $\lambda = 12$. Donc:

$$P(X=8) = e^{-12}\frac{12^8}{8!} \approx 0.066$$

B.2 : On considère à présent une période 3 fois moins importante, c'est-à-dire 20 minutes. Le paramètre λ devient $\lambda = 12/3 = 4$ et:

$$P(X \geq 1) = 1 - P(X=0) = 1 - e^{-4} \approx 0.98.$$

EXERCICE 14.

Niveau : Université
Auteur : Anonyme (04.11.07)
Mots-clés : loi uniforme, probabilité conditionnelle

Enoncé :

Vous arrivez à un arrêt de bus sachant que le temps d'attente est une variable aléatoire distribuée uniformément entre 0 et 30 minutes.

La fonction de densité de cette variable s'écrit comme :

$$f(t)=\begin{cases} 0,\ t\leq 0 \\ 1/30,\ 0<t<30 \\ 0,\ t\geq 30 \end{cases}$$

1. Déterminer la fonction de répartition F de cette variable.

2. Quelle est la probabilité que vous deviez attendre plus de 10 minutes.

3. Si après 15 minutes d'attente le bus n'est pas encore arrivé, quelle est la probabilité que vous deviez attendre au moins 10 minutes supplémentaires ?

Solution :

1. $F(x)=\int\limits_{-\infty}^{x} f(t)dt$. $F(x)=\begin{cases} 0 \text{ si } x\leq 0 \\ \dfrac{x}{30} \text{ si } 0<x<30 \\ 1 \text{ si } x\geq 30 \end{cases}$

2. La probabilité d'attendre plus de 10 minutes est

$$\int\limits_{10}^{+\infty} f(t)dt=\int\limits_{10}^{30} f(t)dt=F(30)-F(10)=1-\frac{10}{30}=\frac{2}{3}.$$

3. Notons X la v.a égale au temps d'attente. Il faut calculer:

$$P(X\geq 25\,|\,X\geq 15)=\frac{P(X\geq 25\cap X\geq 15)}{P(X\geq 15)}=\frac{P(X\geq 25)}{P(X\geq 15)}=\frac{1-F(25)}{1-F(15)}=\frac{5/30}{15/30}=1/3.$$

EXERCICE 15.

Niveau : Université
Auteur : Anonyme (04.11.07)
Mots-clés : espérance, variance

Enoncé :

On appelle variable de Pareto de paramètres *a* et x_0 la variable aléatoire *X* dont la fonction de répartition est :

$$F(x)=\begin{cases} 0 & , \text{ si } x \leq x_0 \\ 1-\left(\dfrac{x_0}{x}\right)^a & , \text{ si } x > x_0, \text{ avec } a > 2 \text{ et } x_0 > 0 \end{cases}$$

1. Déterminer la fonction de densité *f*. Vérifier qu'il s'agit bien d'une fonction de densité.

2. Calculer $E[X]$ et $E\left[X^2\right]$.

3. Déduire du point 2. la variance de *X*.

4. Cas pratique : on suppose que *X* est une variable aléatoire représentant le revenu d'un ménage choisi au hasard et ayant pour fonction de répartition F . Déterminer en fonction de *a* et x_0, la proportion des ménages d'une population dont le revenu est :

(a) supérieur à 10'000.-

(b) compris entre 5'000 et 10'000.-

Solution :

1. On sait que *f* est la dérivée de *F* donc, $f(x)=0$ si $x < x_0$ et:

$$f(x)=-a\left(\frac{x_0}{x}\right)^{a-1}\cdot\left(-\frac{x_0}{x^2}\right)=a\cdot\frac{x_0^a}{x^{a+1}}$$

si $x > x_0$. *f* n'est pas définie en $x = x_0$ car *F* n'est pas dérivable en x_0.

Vérifions que *f* est bien une fonction de densité. Tout d'abord, $f(x) \geq 0$ pour tout $x \in \mathbb{R}$.
Puis:

$$\int_{-\infty}^{+\infty} f(x)dx=\int_{x_0}^{+\infty} a\cdot\frac{x_0^a}{x^{a+1}}dx=a\cdot x_0^a\int_{x_0}^{+\infty}\frac{1}{x^{a+1}}dx=a\cdot x_0^a\cdot\left(-\frac{1}{a}x^{-a}\right)\Bigg|_{x_0}^{+\infty}=a\cdot x_0^a\cdot\frac{1}{a}x_0^{-a}=1$$

Ceci prouve que *f* est une fonction de densité.

2.

$$E[X]=\int_{x_0}^{+\infty} a\cdot x\cdot\frac{x_0^a}{x^{a+1}}dx=a\int_{x_0}^{+\infty}\frac{x_0^a}{x^a}dx=ax_0^a\cdot\int_{x_0}^{+\infty}\frac{1}{x^a}dx=ax_0^a\left(\frac{x^{-a+1}}{-a+1}\right)\Bigg|_{x_0}^{+\infty}=ax_0^a\cdot\left(-\frac{x_0^{-a+1}}{-a+1}\right)=\frac{ax_0}{a-1}.$$

$$E\left[X^2\right]=\int_{x_0}^{+\infty} a\cdot x^2\cdot\frac{x_0^a}{x^{a+1}}dx=a\int_{x_0}^{+\infty}\frac{x_0^a}{x^{a-1}}dx=ax_0^a\cdot\int_{x_0}^{+\infty}\frac{1}{x^{a-1}}dx=ax_0^a\left(\frac{x^{-a+2}}{-a+2}\right)\Bigg|_{x_0}^{+\infty}=ax_0^a\cdot\left(-\frac{x_0^{-a+2}}{-a+2}\right)=\frac{ax_0^2}{a-2}$$

3. $V[X]=E\left[X^2\right]-\left(E[X]\right)^2=\frac{ax_0^2}{a-2}-\left(\frac{ax_0}{a-1}\right)^2=\frac{ax_0^2}{(a-2)(a-1)}$.

4.(a) Il faut calculer $P(X>10'000)$. On a:

$$P(X>10'000)=1-P(X\leq 10'000)=1-F(10'000).$$

Donc:

$$P(X>10'000)=\begin{cases}1 \text{ si } 10'000\leq x_0\\ 1-\left(\frac{x_0}{10'000}\right)^a \text{ si } 10'000>x_0\end{cases}$$

4.(b) Il faut calculer $P(5'000\leq X\leq 10'000)$. On a:

$$P(5'000\leq X\leq 10'000)=P(X\leq 10'000)-P(X\leq 5'000)=F(10'000)-F(5'000).$$

Donc:

$$P(5'000\leq X\leq 10'000)=\begin{cases}0 \text{ si } 10'000\leq x_0\\ 1-\left(\frac{x_0}{10'000}\right)^a \text{ si } 5'000<x_0<10'000\\ \left(\frac{x_0}{5'000}\right)^a-\left(\frac{x_0}{10'000}\right)^a \text{ si } x_0\leq 5'000\end{cases}$$

EXERCICE 16.

Niveau : Université
Auteur : Anonyme (04.11.07)
Mots-clés : loi normale

Remarque : On note Φ la fonction de répartition de la loi $N(0,1)$.

Enoncé :

200 skieurs professionnels ont descendu une célèbre piste de ski des Alpes suisses. Le temps X en secondes que chacun a réalisé suit une distribution normale $N(140;100)$.

1. Déterminer le nombre de skieurs ayant réalisé un temps supérieur à 150 secondes.

2. Quel temps t faut-il réaliser pour être parmi les 20 premiers ?

Solution :

1. Il faut calculer $P(X>150)$. La v.a $\dfrac{X-140}{10}$ suit la loi $N(0,1)$.

Donc :

$$P(X>150)=1-P(X\leq 150)=1-P\left(\frac{X-140}{10}\leq\frac{150-140}{10}\right)=1-\underbrace{P\left(\frac{X-140}{10}\leq 1\right)}_{=\Phi(1)}\approx 1-0.8413=0.1587$$

Ceci signifie qu'il y a 15.87% de skieurs ayant réalisé un temps supérieur à 150 secondes. Sur 200 skieurs il y a donc $0.1587\cdot 200\approx 32$ skieurs ayant réalisé un temps supérieur à 150 secondes.

2. 20 skieurs sur 200 représentent une proportion de $20/200=1/10$. Le temps t vérifie donc l'égalité:

$$P(X\leq t)=1/10\text{. On a } P\left(\frac{X-140}{10}\leq\frac{t-140}{10}\right)=\Phi\left(\frac{t-140}{10}\right)=1/10$$

Donc:

$$\Phi\left(-\frac{t-140}{10}\right)=1-\Phi\left(\frac{t-140}{10}\right)=1-1/10=0.9$$

(cette ligne est nécessaire si vous ne disposez pas d'une table contenant les petites valeurs de Φ^{-1}). Dans la table on lit $-\dfrac{t-140}{10}=1.2816$. Donc:

$$t=127.184.$$

EXERCICE 17.

Niveau : Université
Auteur : Anonyme (04.11.07)
Mots-clés : loi normale, espérance, variance

Remarque : On note Φ la fonction de répartition de la loi $N(0,1)$.

Enoncé :

Soit X_i le prix en francs d'une action *i* dans un an à partir d'aujourd'hui.

Supposons que X_A, le prix de l'action A, suive une loi normale d'espérance mathématique 15 et de variance 100 et que X_B, le prix de l'action B, suive une loi normale d'espérance mathématique 20 et de variance 2025. On achète aujourd'hui un portefeuille composé de 3 actions A à 12 francs chacune et de 2 actions B à 17 francs chacune. On suppose également que X_A et X_B sont des variables aléatoires indépendantes.

1. Notons $Y = 3X_A + 2X_B$ la valeur du portefeuille. Trouver l'espérance et la variance de la valeur du portefeuille dans un an.

2. Sachant que Y suit une distribution normale, quelle est la probabilité que dans un an ce portefeuille ait rapporté plus de 30% ?

Solution :

1. Nous avons donc:

$$E[Y] = E[3X_A + 2X_B] = 3E[X_A] + 2E[X_B] = 3\cdot 15 + 2\cdot 20 = 85\,.$$

$$V[Y] = V[3X_A + 2X_B] = 9V[X_A] + 4V[X_B] = 9\cdot 100 + 4\cdot 2025 = 9000\,.$$

2. La valeur du portefeuille au moment de l'achat est de $3\cdot 12 + 2\cdot 17 = 70$ francs. On veut calculer $P(Y > 70 + 0.3\cdot 70) = P(Y > 91)$. On a:

$$P(Y > 91) = 1 - P(Y \leq 91)$$

Y suit la loi $N(85, 9000)$. Donc:

$$1 - P(Y \leq 91) = 1 - P\left(\frac{Y-85}{\sqrt{9000}} \leq \frac{91-85}{\sqrt{9000}}\right) = 1 - P\left(\frac{Y-85}{\sqrt{9000}} \leq 0.0632\right) = 1 - \Phi(0.0632) \approx 1 - 0.5239 = 0.4761$$

EXERCICE 18.

Niveau : Université
Auteur : (04.11.07)
Mots-clés : loi normale, probabilité conditionnelle

Remarque : On note Φ la fonction de répartition de la loi $N(0,1)$.

Enoncé :

Le nombre de cas de malaria chez les touristes non-vaccinés par année au Laos suit une loi normale d'espérance 1340 et d'écart-type de 40.

1. Quelle est la probabilité pour une année que le nombre de cas de malaria pour les touristes non-vaccinés soit inférieur à 1460.

2. Quel est le nombre maximum de cas de Malaria admissible, si l'on veut que la proportion de cas ne dépasse pas 25% ?

3. Sachant que cette année-là, le nombre de cas a été supérieur à 1300, quelle est la probabilité que le nombre de cas soit compris entre 1260 et 1460.

Solution :

1. Notons X la v.a égale au nombre de cas de malaria chez les touristes non-vaccinés par année. Il faut calculer $P(X \leq 1460)$:

$$P(X \leq 1460) = P\left(\frac{X-1340}{40} \leq \frac{1460-1340}{40}\right) = \Phi\left(\frac{1460-1340}{40}\right) = \Phi(3) \approx 0.99865 .$$

2. Notons Q_1 le nombre maximum de cas de Malaria admissible pour que la proportion de cas ne dépasse pas 25%. Q_1 vérifie $P(X \leq Q_1) = 0.25$. En fait Q_1 est le quartile inférieur de la distribution de X. On a donc:

$$P\left(\frac{X-1340}{40} \leq \frac{Q_1-1340}{40}\right) = \Phi\left(\frac{Q_1-1340}{40}\right) = 0.25$$

C'est-à-dire:

$$\Phi\left(-\frac{Q_1-1340}{40}\right) = 1-0.25 = 0.75$$

Dans les tables on trouve :

$$-\frac{Q_1-1340}{40} = 0.6745$$

et donc:

$$Q_1 = 1340 - 40 \cdot 0.6745 \approx 1313.$$

3.Il faut déterminer la probabilité conditionnelle $P(1260 \leq X \leq 1460 \mid X \geq 1300)$.

$$P(1260 \leq X \leq 1460 \mid X \geq 1300) = \frac{P((1260 \leq X \leq 1460) \cap (X \geq 1300))}{P(X \geq 1300)} = \frac{P(1300 \leq X \leq 1460)}{P(X \geq 1300)}$$

Commençons par calculer $P(1300 \leq X \leq 1460)$:

$$P(1300 \leq X \leq 1460) = P\left(\frac{1300-1340}{40} \leq \frac{X-1340}{40} \leq \frac{1460-1340}{40}\right) = P\left(-1 \leq \frac{X-1340}{40} \leq 3\right)$$
$$= \Phi(3) - \Phi(-1) = \Phi(3) - 1 + \Phi(1) \approx 0.83995$$

A présent, calculons $P(X \geq 1300)$:

$$P(X \geq 1300) = 1 - P(X \leq 1300) = 1 - P\left(\frac{X-1340}{40} \leq \frac{1300-1340}{40}\right)$$
$$= 1 - P\left(\frac{X-1340}{40} \leq -1\right) = 1 - \Phi(-1) = \Phi(1) \approx 0.8413$$

Pour finir:

$$P(1260 \leq X \leq 1460 \mid X \geq 1300) \approx \frac{0.83995}{0.8413} \approx 0.9984.$$

EXERCICE 19.

Niveau : Université
Auteur : (04.11.07)
Mots-clés : fonction de densité, espérance, probabilité conditionnelle, loi binomiale

---

Enoncé :

La durée de vie (en heures) d'un téléviseur est une variable aléatoire continue dont la fonction de densité est donnée par :

$$f(x)=\begin{cases}\dfrac{\lambda}{x^2} & x>150\\ 0 & \text{sinon}\end{cases}$$

1. Trouver la valeur du paramètre λ.

2. Quelle est la durée de vie moyenne d'un téléviseur ?

3. Quelle est la probabilité que la durée de vie du téléviseur soit supérieure à 300 heures ?

4. Un téléviseur fonctionne depuis 200 heures. Quelle est la probabilité qu'il fonctionne 200 heures supplémentaires ?

5. Un magasin vend 20 téléviseurs du même type. Quelle est la probabilité qu'après 200 heures il faille en remplacer au moins 2 ?

Solution :

1. *f* est une fonction de densité. Par suite elle vérifie:

$$\int_{-\infty}^{+\infty} f(x)\,dx=1$$

On a:

$$\int_{-\infty}^{+\infty} f(x)\,dx=\int_{150}^{+\infty}\frac{\lambda}{x^2}\,dx=-\frac{\lambda}{x}\bigg|_{150}^{+\infty}=\frac{\lambda}{150}$$

Donc:

$$\frac{\lambda}{150}=1 \text{ et } \lambda=150\,.$$

2. Notons X la v.a égale à la durée de vie d'un téléviseur. On a:

$$E[X]=\int_{-\infty}^{+\infty} x\cdot f(x)dx=150\int_{150}^{+\infty}\frac{1}{x}dx=150\ln(x)\Big|_{150}^{+\infty}=+\infty$$

Dans ce cas on voit que l'espérance (la moyenne) n'existe pas.

3. Il faut calculer $P(X\geq 300)$:

$$P(X\geq 300)=150\int_{300}^{+\infty}\frac{1}{x^2}dx=-\frac{150}{x}\bigg|_{300}^{+\infty}=\frac{150}{300}=0.5$$

4. Il faut calculer la probabilité conditionnelle $P(X\geq 400\,|\,X\geq 200)$. On a:

$$P(X\geq 400\,|\,X\geq 200)=\frac{P((X\geq 400)\cap(X\geq 200))}{P(X\geq 200)}=\frac{P(X\geq 400)}{P(X\geq 200)}=\frac{150\int_{400}^{+\infty}\frac{1}{x^2}dx}{150\int_{200}^{+\infty}\frac{1}{x^2}dx}=\frac{\int_{400}^{+\infty}\frac{1}{x^2}dx}{\int_{200}^{+\infty}\frac{1}{x^2}dx}$$

$$=\frac{-\frac{1}{x}\Big|_{400}^{+\infty}}{-\frac{1}{x}\Big|_{200}^{+\infty}}=\frac{1/400}{1/200}=0.5$$

5. La probabilité qu'un téléviseur arrête de fonctionner durant les 200 heures est:

$$p=P(X\leq 200)=150\int_{150}^{200}\frac{1}{x^2}dx=-\frac{150}{x}\bigg|_{150}^{200}=-\frac{150}{200}+1=1/4\,.$$

Notons Y la v.a égale au nombre de téléviseurs à remplacer sur les 20 vendus après 200 heures. Y suit la loi binomiale $Binomiale(20,1/4)$. Nous voulons calculer $P(Y\geq 2)$.

On a:

$$P(Y\geq 2)=1-P(Y\leq 1)=1-P(Y=0)-P(Y=1)=1-(1-p)^{20}-20p(1-p)^{19}$$
$$=1-0.75^{20}-20\cdot 0.25\cdot 0.75^{19}\approx 0.976$$

EXERCICE 20.

Niveau : Université
Auteur : (04.11.07)
Mots-clés : loi normale

Remarque : On note Φ la fonction de répartition de la loi $N(0,1)$.

Enoncé :

MailCopy vient d'acheter deux photocopieuses de deux marques différentes : Minulta et Canyon. Les deux machines sont dotées d'un système qui les fait s'arrêter lorsqu'elles ont besoin d'un entretien. La durée de fonctionnement en jours avant qu'une machine ne s'arrête suit une loi normale. L'espérance de la Minulta est de 200 et sa variance est de 81. Pour la Canyon, l'espérance est égale à 210 et la variance est égale à 144.

D'autre part, on suppose que les durées de fonctionnement des photocopieuses sont indépendantes l'une de l'autre.

a) Quelle est la probabilité que la machine Canyon ait besoin d'un entretien dans les 7 mois qui suivent (on suppose que 1 mois contient 30 jours) ? Refaire le calcul pour la Minulta.

b) Quelle est la probabilité que la machine Canyon s'arrête pour la première fois avant la Minulta ?

c) Quelle est la probabilité que la machine Minulta s'arrête pour la première fois avant la Canyon ?

d) Quelle est la probabilité que les machines s'arrêtent pour la première fois en même temps ?

Solution:

a) Notons X la v.a égale à la durée de fonctionnement avant un entretien pour la Canyon et Y la v.a correspondante pour la Minulta.

Il faut calculer $P(X \leq 210)$ et $P(Y \leq 210)$:

$$P(X \leq 210) = P\left(\frac{X-210}{12} \leq \frac{210-210}{12}\right) = P\left(\frac{X-210}{12} \leq 0\right) = \Phi(0) = 0.5$$

$$P(Y \leq 210) = P\left(\frac{Y-200}{9} \leq \frac{210-200}{9}\right) = P\left(\frac{Y-200}{9} \leq \frac{10}{9}\right) = \Phi\left(\frac{10}{9}\right) \approx 0.8665$$

b) La probabilité cherchée est $P(X < Y)$. On a $P(X < Y) = P(X - Y < 0)$. Or nous savons que les v.a X et Y sont indépendantes. Par suite la v.a $Z = X - Y$ suit une loi normale d'espérance $210 - 200 = 10$ et de variance $81 + 144 = 225$. Donc:

$$P(Z<0)=P\left(\frac{Z-10}{15}<\frac{-10}{15}\right)=\Phi\left(-\frac{2}{3}\right)=1-\Phi\left(\frac{2}{3}\right)\approx 1-0.7454=0.2546\,.$$

c) La probabilité cherchée est $P(Y<X)$. On a $P(Y<X)=1-P(X\leq Y)$ et $P(X\leq Y)=P(X<Y)=0.2546$ (par la question précédente).

Donc:

$$P(Y<X)=1-0.2546=0.7454\,.$$

d) Cette probabilité est $P(X=Y)$. Mais $P(X=Y)=P(X-Y=0)$. Or la v.a $Z=X-Y$ suit la loi normale qui est absolument continue. Par suite, $P(Z=0)=0$.

EXERCICE 21.

Niveau : Université
Auteur : (04.11.07)
Mots-clés : loi normale

Remarque : On note Φ la fonction de répartition de la loi $N(0,1)$.

Enoncé :

On a observé dans une population que 35% des personnes entrent dans le monde professionnel avant 20 ans, et que 55% d'entre elles y parviennent avant leurs 24 ans. En supposant que l'âge d'entrée dans le monde professionnel suit une loi normale, quelle est la probabilité, pour une personne, le jour de ses 25 ans, de n'être toujours pas entrée dans le monde professionnel ?

Indication : Déterminer d'abord les paramètres de la loi normale.

Solution :

Notons X la v.a égale à l'âge d'entrée dans le monde professionnel d'une personne. X suit la loi $N(\mu,\sigma^2)$. Il faut calculer $P(X > 25)$. Pour cela nous devons d'abord déterminer la valeur des paramètres μ et σ.

Par l'énoncé, nous savons que $P(X \leq 20) = 0.35$ et que $P(X \leq 24) = 0.55$.

Or:

$$P(X \leq 20) = P\left(\frac{X-\mu}{\sigma} \leq \frac{20-\mu}{\sigma}\right) = \Phi\left(\frac{20-\mu}{\sigma}\right) = 0.35$$

Donc:

$$\Phi\left(-\frac{20-\mu}{\sigma}\right) = 1 - 0.35 = 0.65$$

La table nous donne:

$$-\frac{20-\mu}{\sigma} = 0.3853$$

C'est-à-dire:

$$-0.3853\sigma + \mu = 20 \text{ (*).}$$

Nous avons obtenu une équation à deux inconnues.

A présent déterminons la deuxième équation à partir de l'égalité $P(X \leq 24) = 0.55$.

$$P(X \leq 24) = P\left(\frac{X-\mu}{\sigma} \leq \frac{24-\mu}{\sigma}\right) = \Phi\left(\frac{24-\mu}{\sigma}\right) = 0.55$$

La table nous donne $\dfrac{24-\mu}{\sigma} = 0.1257$. C'est-à-dire

$$0.1257\sigma + \mu = 24 \text{ (**)}.$$

Nous avons obtenu un système de deux équations à deux inconnues (*) et (**) :

$$\begin{cases} -0.3853\sigma + \mu = 20 \\ 0.1257\sigma + \mu = 24 \end{cases}$$

En résolvant ce système nous allons déterminer la valeur des paramètres μ et σ.

En soustrayant la première équation de la deuxième nous obtenons :

$$0.511\sigma = 4$$

c'est-à-dire :

$$\sigma \approx 7.828$$

En remplaçant la valeur de σ dans la première équation nous obtenons :

$$\mu \approx 23.016.$$

Nous pouvons à présent déterminer $P(X > 25)$. On a:

$$P(X > 25) = 1 - P(X \leq 25) = 1 - P\left(\frac{X-\mu}{\sigma} \leq \frac{25-\mu}{\sigma}\right) \approx 1 - P\left(\frac{X-\mu}{\sigma} \leq 0.25\right) = 1 - \Phi(0.25)$$
$$= 1 - 0.5987 = 0.4013$$

EXERCICE 22.

Niveau : Université
Auteur : (04.11.07)
Mots-clés : loi normale

Remarque : On note Φ la fonction de répartition de la loi $N(0,1)$.

Enoncé :

Dans une population donnée, un taux sanguin T suit une loi normale. 4% des

sujets de cette population possèdent un taux supérieur à 2.7 et 79% possèdent un taux supérieur à 1.68.

Calculer le taux moyen et l'écart-type de T.

Solution :

T suit la loi $N(\mu,\sigma^2)$, il s'agit de déterminer μ et σ.

Le fait que 4% des sujets possèdent un taux supérieur à 2.7 se traduit par $P(T > 2.7) = 0.04$.

Donc:

$$P\left(\frac{T-\mu}{\sigma} > \frac{2.7-\mu}{\sigma}\right) = P\left(\frac{T-\mu}{\sigma} < -\frac{2.7-\mu}{\sigma}\right) = \Phi\left(-\frac{2.7-\mu}{\sigma}\right) = 0.04$$

Car:

$$\frac{T-\mu}{\sigma} \sim N(0,1)$$

Donc:

$$\Phi\left(\frac{2.7-\mu}{\sigma}\right) = 1-\Phi\left(-\frac{2.7-\mu}{\sigma}\right) = 1-0.04 = 0.96$$

Ccette ligne est nécessaire si vous ne disposez pas d'une table contenant les petites valeurs de Φ^{-1}.

Dans la table on lit $\dfrac{2.7-\mu}{\sigma} = 1.7507$ c'est-à-dire,

$$\mu + 1.7507\sigma = 2.7 \quad (*)$$

Nous obtenons une deuxième équation à partir de l'égalité $P(T > 1.68) = 0.79$ de l'énoncé.

On a:

$$P\left(\frac{T-\mu}{\sigma} > \frac{1.68-\mu}{\sigma}\right) = P\left(\frac{T-\mu}{\sigma} < -\frac{1.68-\mu}{\sigma}\right) = \Phi\left(-\frac{1.68-\mu}{\sigma}\right) = 0.79$$

Dans la table on lit $-\dfrac{1.68-\mu}{\sigma} = 0.8064$.

Donc,

$$\mu - 0.8064\sigma = 1.68 \quad (**)$$

Nous avons le système suivant :

$$\begin{cases} \mu + 1.7507\sigma = 2.7 \\ \mu - 0.8064\sigma = 1.68 \end{cases}$$

La résolution de ce système nous donne $\mu \approx 2.00$ et $\sigma \approx 0.40$.

EXERCICE 23.

Niveau : Université
Auteur : (04.11.07)
Mots-clés : loi normale, loi d'une somme

Remarque : On note Φ la fonction de répartition de la loi $N(0,1)$.

Enoncé :

Un ascenseur peut porter une charge de 500 kg. On admet que le poids d'un individu suit une loi normale de moyenne 75 et d'écart-type 4. Quel est le nombre maximum de personnes que l'on peut autoriser à monter ensemble dans l'ascenseur si on veut que le risque de surcharge ne dépasse pas 0.001 ?

Solution :

Notons *n* ce nombre maximum de personnes. Notons $X_1,...,X_n$ les variables correspondant aux poids de *n* personnes montant dans l'ascenseur. On sait que pour $i=1..n$, $X_i \sim N(75,16)$ et que les X_i sont indépendantes.

La v.a $Y=\sum_{i=1}^{n} X_i$ représente le poids total de ces *n* personnes. On a $Y \sim N(75n,16n)$ et on veut $P(Y>500)=0.001$.

Donc:

$$P\left(\frac{Y-75n}{4\sqrt{n}}>\frac{500-75n}{4\sqrt{n}}\right)=1-P\left(\frac{Y-75n}{4\sqrt{n}}<\frac{500-75n}{4\sqrt{n}}\right)=1-\Phi\left(\frac{500-75n}{4\sqrt{n}}\right)=0.001.$$

C'est-à-dire:

$$\Phi\left(\frac{500-75n}{4\sqrt{n}}\right)=0.999.$$

Dans les tables on lit:

$$\frac{500-75n}{4\sqrt{n}}=3.0902.$$

On aboutit à l'équation:

$$75n+12.3608\sqrt{n}-500=0.$$

Cette équation est du deuxième degré en $\sqrt{n}$. Sa résolution nous donne,

$$\sqrt{n} = 2.501 \text{ (la solution positive est l'unique acceptable)}$$

Pour finir $n = 6$ personnes.

EXERCICE 24.

Niveau : Université
Auteur : (04.11.07)
Mots-clés : probabilité conditionnelle

Enoncé :

La société ASR, un grand producteur d'ordinateurs, sait que 3 machines sur mille sortant de son usine sont défectueuses. Afin de détecter ces ordinateurs défectueux, les ingénieurs de ASR ont mis au point un test de fiabilité qu'ils ont éprouvé sur un échantillon test connu d'avance.

Parmi les ordinateurs en parfait état, 94% réussissent le test (donc 6% de ceux qui sont en bon état ne le réussissent pas) et parmi ceux défectueux, seulement 2% réussissent le test (donc il y a 98% d'ordinateurs qui ne passent pas le test sachant qu'ils sont défectueux).

Il a été décidé que les ordinateurs n'ayant pas passé le test avec succès seront détruits. Le directeur général de ASR, qui vient de vous engager comme proche collaborateur, vous demande votre avis sur cette stratégie de sélection.

Répondez aux questions suivantes :

1. Quelle est la probabilité qu'un ordinateur soit défectueux et qu'il ne réussisse pas le test ?

2. Quelle est la probabilité qu'un ordinateur ne réussisse pas le test ?

3. Quelle est la proportion d'ordinateurs vraiment défectueux qui sont détruits ?

Solution :

1. Notons $A=$ « l'ordinateur est défectueux » et $B=$ «l'ordinateur ne réussi pas le test ». Il faut calculer $P(A \cap B)$. On sait par l'énoncé que la probabilité que l'ordinateur ne réussisse pas le test sachant qu'il est défectueux est de:

$$P(B \mid A)=1-0.02=0.98$$

et que $P(A)=0.003$. Donc:

$$P(A \cap B)=P(B \mid A) \cdot P(A)=0.98 \cdot 0.003=0.00294=0.294\%$$

Ce résultat peut paraître étonnant mais il montre que dans l'énoncé, l'information que nous avons sur l'état des ordinateurs avant de faire sur le test à une grande influence. Ceci dit, nous remarquons déjà que le test n'est pas très bon.

2. Il faut calculer $P(B)$. Or:

$$B = (B \cap A) \cup (B \cap \overline{A})$$

où l'union est disjointe. Donc:

$$P(B) = P(B \cap A) + P(B \cap \overline{A}) = 0.00294 + P(B \mid \overline{A}) P(\overline{A}) = 0.00294 + 0.06 \cdot 0.997 = 0.06276 .$$

3. Il faut calculer la proportion d'ordinateurs défectueux parmi ceux qui ne réussissent pas le test. Il s'agit de la probabilité conditionnelle $P(A \mid B)$:

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{0.00294}{0.06276} \approx 0.0468$$

Parmi les ordinateurs qui sont détruits seulement le 4.68% sont vraiment défectueux. Nous pouvons conclure que la stratégie de sélection est très mauvaise.

EXERCICE 25.

Niveau : Lycée
Auteur : Isoz Vincent (12.11.07)
Mots-clés : Probabilités

Enoncé :

Un parc informatique doit comporter *N* machines (machine irréparable). Nous avons les informations suivantes:

- Une machine neuve a une probabilité de 1/6 de tomber en panne durant l'année courante

- Une machine d'une année de vie à une probabilité de 13/20 de tomber en panne

- Une machine de deux années de vie à une probabilité 6/7 de tomber en panne

- Une machine de trois ans à 100% trois ans de vie (i.e tombe certainement en panne durant l'année courante)

Q1.Calculer l'espérance de vie d'une machine

Q2.Donner relation de récurrence donnant le nombre de machines à acheter au bout de l'année courante en fonction du nombre des machines neuves, d'un an d'âge, de deux ans d'âge et de trois ans d'âge.

Q3. Déduire l'expression de $N(k)$ le nombre de machines neuves à acheter à l'année *k*.

Solutions:

S1. L'espérance de vie d'une machine:

$$P_1 = \frac{1}{6}, P_2 = \frac{13}{20}, P_3 = \frac{6}{7}, P_4 = 1$$

Soit $P(X)$ la probabilité de tomber en panne lors de l'année *X* (sachant que la machine est neuve).

$$\begin{aligned}
&P(1) = P_1 = \frac{1}{6}\\
&P(2) = (1-P_1)P_2 = \frac{5}{6}\frac{13}{20} = \frac{13}{24}\\
&P(3) = (1-P_1)(1-P_2)P_3 = \frac{1}{4} =\\
&P(4) = (1-P_1)(1-P_2)(1-P_3)P_4 = \frac{1}{24}
\end{aligned}$$

L'espérance mathématique est alors:

$$E(X)=\sum_{i=1}^{n} X_i P(X_i)=2.16$$

L'espérance de vie d'une machine est donc de 2.16 ans, soit 2 ans et 2 mois.

S2. Soit *a* le nombre de machines neuves, *b* celle d'un an d'âge, *c* celle de deux ans d'âge, et *d* celle de trois ans d'âge.

Nous remplaçons $P_1 \cdot a$ machines neuves, $P_2 \cdot b$ machines de deux ans d'âge, $P_3 \cdot c$ machines de trois ans d'âge et $P_4 \cdot d$ pour les plus anciennes.

Donc le nombre de machine à acheter au bout de l'année courante est:

$$N=\frac{1}{6}a+\frac{13}{24}b+\frac{1}{4}c+\frac{1}{24}d$$

S3. Si nous désignons par $N(k)$ le nombre de machines neuves à l'année *k* :

$$N(k)=\frac{1}{6}N(k-1)+\frac{13}{24}N(k-2)+\frac{1}{4}N(k-3)+\frac{1}{24}N(k-4)$$

avec $k>4$ et:

$$\begin{aligned}
&N(0)=N\\
&N(1)=\frac{1}{6}N(0)\\
&N(2)=\frac{1}{6}N(1)+\frac{13}{24}N(0)\\
&N(3)=\frac{1}{6}N(2)+\frac{13}{24}N(1)+\frac{1}{4}N(0)\\
&N(4)=\frac{1}{6}N(3)+\frac{13}{24}N(2)+\frac{1}{4}N(1)+\frac{1}{24}N(0)
\end{aligned}$$

EXERCICE 26.

Niveau : Université
Auteur : (04.11.07)
Mots-clés : probabilités conditionnelles

Enoncé :

Dans une population, 7% des individus sont contaminés par un virus. Nous disposons d'un test de dépistage qui présente les propriétés suivantes:

- Parmi les individus contaminés, le test est positif à 99%

- Par les individus non contaminés le test est tout de même positif à 3% (il y a donc des risques de mauvais diagnostic)

Les questions sont alors les suivantes:

1. Quelle est la probabilité, que le test appliqué à un individu pris au hasard soit positif?

2. Sachant, pour un individu donné, le test est positif, quelle est la probabilité que cet individu soit contaminé?

3. Calculez la probabilité qu'une personne soit réellement non contaminée sachant que son test est négatif.

4. Calculer la probabilité qu'une personne contaminée ne soit pas dépistée par le test.

Pour répondre aux questions, notez T l'événement "le test est positif" et C "l'individu est contaminé".

Les données de l'énoncé s'écrivent alors:

$$P(C) = 7\% \Rightarrow P(C^c) = P(\overline{C}) = 93\%$$
$$P(T \mid C) = 99\% \Rightarrow P(T^c \mid C^c) = P(\overline{T} \mid \overline{C}) = 1\%$$
$$P(T \mid C^c) = P(T \mid \overline{C}) = 3\% \Rightarrow P(T^c \mid C) = P(\overline{T} \mid C) = 97\%$$

Solutions:

S1. Nous devons donc déterminer quelle est la probabilité, que le test appliqué à un individu pris au hasard soit positif.

Or:

$$T = (T \cap C) \cup (T \cap \overline{C})$$

où l'union est disjointe. Donc:

$$P(T) = P(T \cap C) + P(T \cap \overline{C})$$

soit en utilisant la relation des probabilités composées:

$$P(T) = P(T \mid C) \cdot P(C) + P(T \mid \overline{C}) \cdot P(\overline{C}) = 9.7\%$$

D'où:

$$P(\overline{T}) = 90.3\%$$

S2. Pour un individu donné, si le test est positif, calculons quelle est la probabilité que cet individu soit réellement contaminé:

$$P(C \mid T) = \frac{P(C \cap T)}{P(T)} = \frac{P(T \mid C)P(C)}{P(T)} = 72\%$$

D'où:

$$P(\overline{C} \mid \overline{T}) = 28\%$$

S3. Calculons la probabilité qu'une personne soit réellement non contaminée sachant que son test est négatif:

$$P(\overline{C} / \overline{T}) = \frac{P(\overline{C} \cap \overline{T})}{P(\overline{T})} = \frac{P(\overline{T} \mid \overline{C})P(\overline{C})}{P(\overline{T})} = 99.92\%$$

S4. Calculons la probabilité qu'une personne contaminée ne soit pas dépistée par le test:

$$P(C \mid \overline{T}) = \frac{P(C \cap \overline{T})}{P(\overline{T})} = \frac{P(\overline{T} \mid C)}{P(\overline{T})} P(C) = 0.07\%$$

EXERCICE 27.

Niveau : Lycée
Auteur : Isoz Vincent (21.01.08)
Mots-clés : Intervalles

Enoncé :

Une société de service vend des prestations dont l'analyse montre qu'elle suite une loi normale de moyenne 550.- et d'écart-type de 27.-.

Quel est l'intervalle de prix représentant 80% des ventes et combien de sigma cela représente-t-il?

Solution:

Dans MS Excel il suffit de saisir:

=NORMINV(10%;550;27)

pour avoir la borne inférieure qui est de 515.39.

De même pour la borne supérieure:

=NORMINV(90%;550;27)

dont le résultat est de 584.60.

Nous avons alors en Sigma:

$$\frac{584.6-515.39}{27}=2.56\sigma$$

EXERCICE 28.

Niveau : Lycée
Auteur : Isoz Vincent (12.11.07)
Mots-clés : Droite de Henry, Test d'indépendance du khi-deux (Pearson)

Enoncé :

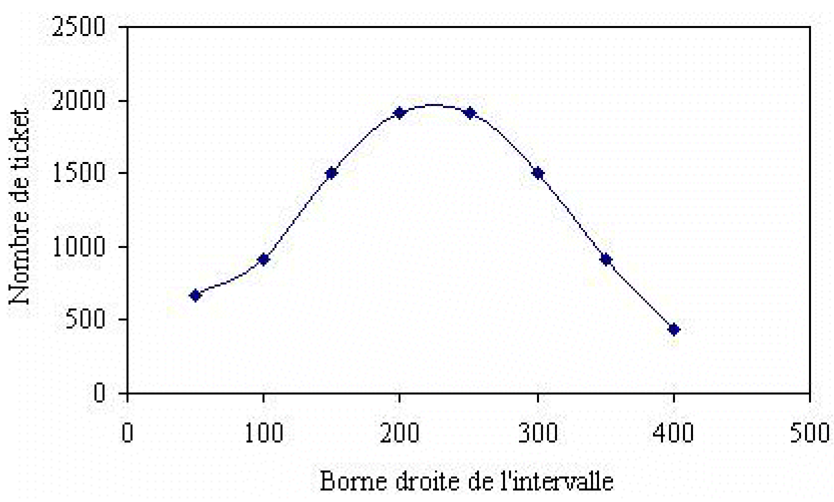
Supposons que nous ayons l'analyse fréquentielle suivante de 10'000 tickets de caisse dans un supermarché :

Montant des tickets	Nombre de tickets	Nombre cumulés de tickets	Fréquences relatives cumulées
[0;50[	668	668	0.068
[50,100[	919	1'587	0.1587
[100,150[	1'498	3'085	0.3085
[150,200[	1'915	5000	0.5000
[200,250[	1'915	6'915	0.6915
[250,300[	1'498	8'413	0.8413
[300,350[	919	9'332	0.9332
[350,400[	440	9'772	0.9772
[400 et +	228	10'000	1

Déterminez avec MS Excel l'écart-type et l'espérance (moyenne) de cette distribution selon l'hypothèse que les données suivent une loi Normale.

Solution:

S1. Si nous traçons maintenant cela sous MS Excel nous obtenons:



Ce qui ressemble terriblement à une loi Normale d'où l'autorisation, sans trop de risques, d'utiliser dans cet exemple la technique de la droite d'Henry.

Mais que faire maintenant? Eh bien connaissant les fréquences cumulées, il ne nous reste plus qu'à calculer pour chacune d'entre elles k^* (la variable centrée réduite) avec la fonction NORMSINV( ) de MS Excel.

Ceci nous donnera les valeurs de la loi Normale centrée réduite $N(0,1)$ de ces mêmes fréquences respectives cumulées (fonction de répartition). Ainsi nous obtenons:

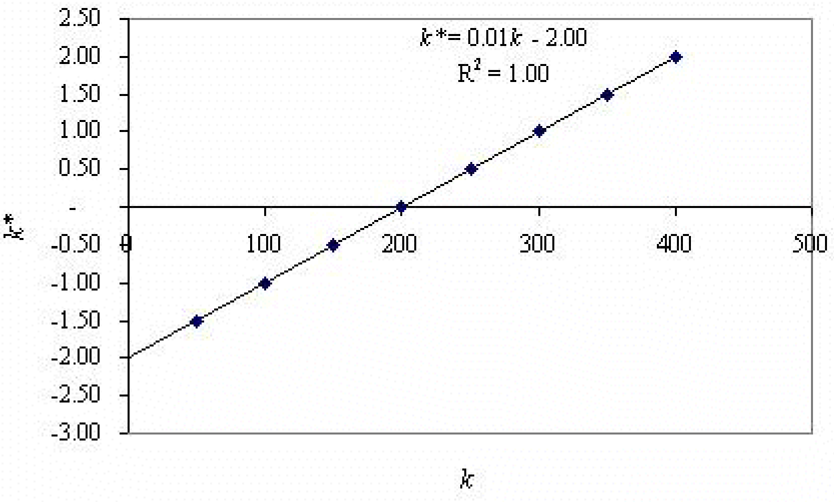
Borne de l'intervalle	Fréquences relatives cumulées	Correspondance pour k^* de $N(0,1)$
50	0.068	-1.5
100	0.1587	-1
150	0.3085	-0.5
200	0.5000	0
250	0.6915	0.5
300	0.8413	1
350	0.9332	1.5
400	0.9772	2
-	1	-

Signalons que dans le tableau ci-dessus, dans MS Excel, les valeurs de fréquences cumulées nulles et unitaires (extrêmes) posent problèmes. Il faut alors jouer un petit peu...

Comme nous l'avons spécifié plus haut, nous avons:

$$k_i^* = f(k_i) = \frac{k_i}{\sigma} - \frac{\mu}{\sigma}$$

Donc graphiquement sous MS Excel nous obtenons grâce à notre tableau le graphique suivant:



Donc à l'aide de la régression donnée par MS Excel soit avec l'équation de la droite qui est sur le graphique soit avec les fonctions ORDONNEE.ORIGINE( ), PENTE( ) et l'ensemble avec la fonction DROITEREG( ) il vient alors sous forme continue:

$$k^* = \frac{k}{\sigma} - \frac{\mu}{\sigma} = 0.01k - 2$$

Donc nous avons immédiatement :

$$\sigma = 100 \quad \mu = 200$$