

Série noire ?

Groupe statistique et citoyenneté
IREM de Paris-Nord

Objectifs pour la classe de troisième

- Réfléchir sur le hasard, les coïncidences et la prétendue « loi des séries noires »
- Élaborer une simulation (le modèle est fourni).
- Estimer une probabilité.

La série noire d'août 2005 dans le transport aérien

En août 2005, on a dénombré cinq accidents aériens graves en l'espace de 22 jours : voilà de quoi frapper l'opinion ! On parle alors d'une « inquiétante série noire dans l'aviation civile » et on évoque une « augmentation du risque » et « des problèmes accrus de maintenance des avions ».

Date	Compagnie	Lieu de l'accident
2 août 2005	Air France	Toronto, Canada
6 août 2005	Tuninter	Palerme, Italie
14 août 2005	Helios Airways	Grammatikos, Grèce
16 août 2005	West-Caribbean	Machiques, Venezuela
23 août 2005	Trans Peru	Pucallapa, Pérou

(http://www.1001crash.com/index-page-crash_year-lg-2-numpage-2005.html#accidents)

Bien sûr, il est indispensable d'analyser chaque accident, de confronter éventuellement les analyses pour repérer des analogies, des causes communes. Mais il convient aussi par ailleurs d'étudier cette « série noire » : est-elle aussi extraordinaire qu'elle en a l'air ?

Nous proposons deux approches de cette situation, basée sur des visions et des hypothèses différentes. La première approche concerne la classe de troisième, la deuxième est proposée en complément pour les enseignants. On pourrait en imaginer d'autres.

Première approche

On va considérer qu'il ne peut y avoir qu'un seul accident d'avion par jour dans le monde, que les accidents sont indépendants les uns des autres et que chaque jour de l'année, la probabilité d'avoir un accident est la même.

Bien sûr, ce modèle est grossier, mais il va nous permettre de nous faire une idée de la rareté de la série observée.

Une recherche sur Internet (<http://www.1001crash.com/index-page-statistique-lg-1-numpage-2.html>) permet d'obtenir la statistique suivante : de 1995 à 2004, on dénombre 376 accidents, soit une moyenne d'environ 0,1 accident par jour. On peut prendre cette valeur comme probabilité journalière d'accident.

Pour simuler le modèle, imaginons qu'on dispose d'une roulette dont le secteur « rouge » correspond à 10 % de la surface du disque. Chacun des 365 jours d'une année, on fait tourner la roulette.

On veut estimer la probabilité p que, sur une année quelconque, la roulette est tombée au moins 5 fois sur le secteur rouge en 22 jours

La roulette parle à l'imagination, mais il est aujourd'hui techniquement plus simple de simuler le modèle avec un tableur pour estimer la probabilité de cet événement.

Programme de simulation

C25 =SOMME(B4:B25)			
	A	B	C
1	La loi des séries		
2			
3	jour		Simulation 1
4	1	0	
5	2	0	
6	3	0	
7	4	0	
8	5	1	
9	6	0	
10	7	0	
11	8	0	
12	9	0	
13	10	0	
14	11	1	
15	12	0	
16	13	0	
17	14	0	
18	15	0	
19	16	0	
20	17	0	
21	18	0	
22	19	0	
23	20	0	
24	21	0	
25	22	0	2
26	23	0	2
27	24	0	2
28	25	0	2

La roulette peut-être simulée par l'instruction =ENT(ALEA()+0,1) , ou le résultat « 1 » correspond au secteur rouge et « 0 » au reste de la roulette.

À partir du 22^e jour, on peut sommer les valeurs « 1 » entre les jours 1 et 22, par une instruction du type =SOMME(B4:B25)

si le premier jour correspond à la cellule B4.

Il suffit ensuite de recopier le tout vers le bas pour obtenir 365 jours.

On obtient la somme maximale sur une période de 22 jours consécutifs par une instruction du type =MAX(C4:C368) .

Il suffit ensuite de sélectionner, puis de recopier cette première simulation, pour obtenir, par exemple, 100 simulations d'une année.

Un graphique en nuage de points peut permettre d'observer la situation.

En faisant plusieurs fois F9, on peut estimer p .

On observera que p est de l'ordre d'au moins 90%.

Quelles « leçons » ?

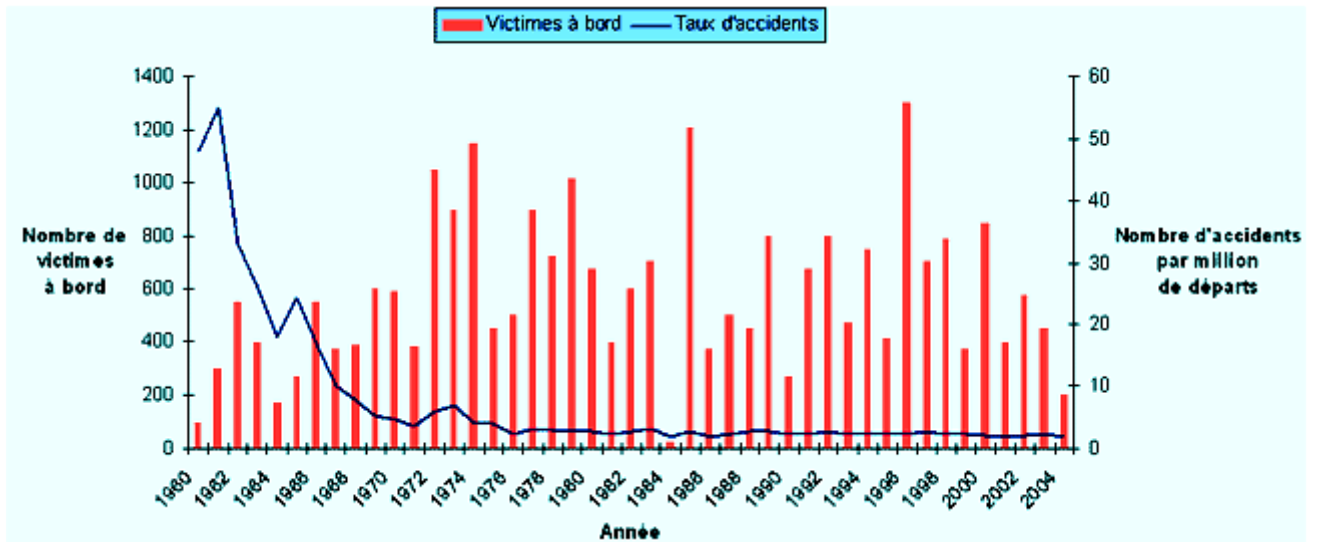
La probabilité obtenue relativise le côté « exceptionnel » de cette « série noire » : ce n'est pas « extraordinaire ». Cette probabilité est estimée dans le cadre d'un modèle où la probabilité d'avoir un accident est la même chaque jour de l'année : ce n'est sans doute pas le cas, mais donne un ordre de grandeur en l'absence de toute autre information. Dans ce modèle, les calculs montrent qu'une telle série d'accidents ne peut à elle seule justifier des affirmations rapides sur un risque accru pour les passagers.

Devoir

Rédiger un commentaire sur cette série d'accidents, comme si vous étiez un journaliste...

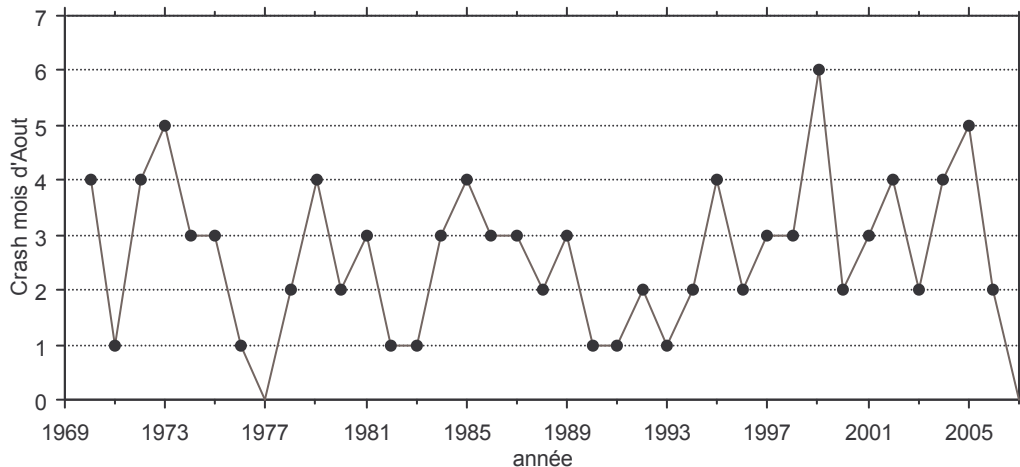
Prolongements

On peut faire réfléchir les élèves aux statistiques à faire (accidents par kilomètre, heures de vol, nombre de vols, accidents par compagnie), faire analyser l'histogramme du nombre annuels de crashes (il y en a peu au début car il y a peu de vols, etc.).



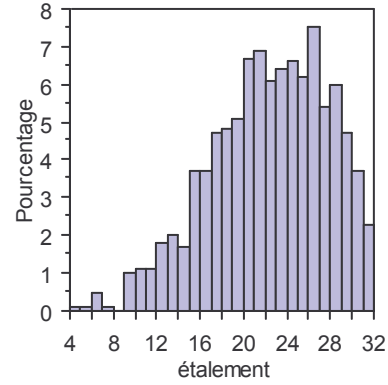
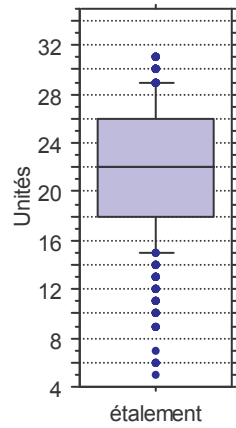
Deuxième approche

On peut trouver sur internet les statistiques du nombre d'accidents pour chaque mois d'Août, entre 1950 et 2008 (<http://www.airdisaster.com/cgi-bin/database.cgi>). Il y a eu 5 accidents au mois d'Août 2005: la courbe ci-dessous montre que c'est élevé, mais sans plus.



On peut se demander s'il est extraordinaire que ces 5 accidents se soient déroulés sur une durée de 22 jours. Nous avons fait des simulations sous l'hypothèse que chaque accident à la probabilité $1/31$ d'arriver un jour donné du mois d'Août, les accidents étant indépendants les uns des autres (sous ces hypothèses, il peut y avoir plus d'un accident un jour donné). On a simulé pour 1000 mois d'Août, la durée qui sépare le premier du dernier accident (variable nommée *étalement*). On peut voir ci-dessous l'histogramme, le diagramme en boîte et des résumés statistiques : la médiane des 1000 données est...22 jours !

Moyenne : 21,9
Médiane : 22
Minimum : 4
Maximum : 31
Mode : 26



Pour finir, signalons que les modèles sont ici des outils qui permettent la réflexion et la quantification ; ils sont basés sur des hypothèses que l'on peut faire avec l'information et les données dont on dispose. Ils n'ont pas été soumis à des procédures de validation.

Sources :

É. Janvresse et T. de la Rue – « *La loi des séries, hasard ou fatalité ?* » – Collection *Les petites pommes du savoir* – Le Pommier 2007.

<http://www.1001crash.com/index-page-statistique-lg-1-numpage-2.html>

<http://www.airdisaster.com/cgi-bin/database.cgi>