

Probabilité et statistique au collège : le côté des élèves, celui des professeurs

Lycée Teilhard de Chardin Saint Maur
Octobre 2008

Claudine Schwartz

Professeur des universités.

Projet statistix : www.statistix.fr

Statistique et probabilité

- 1- Du côté des élèves : statistique et probabilité de l'école au lycée - objectifs.
- 2- Du côté des professeurs : Savoirs (*schémas de pensée*) nécessaires à un professeur pour enseigner au collège en 2008-2009.
- 3- Les élèves et les professeurs : Retour sur l'enseignement en troisième

La plupart des exemples sont sur présentés sur
www.statistix.fr

1- Du côté des élèves : statistique et probabilité de l'école au lycée - objectifs

En primaire

Dés : objets à observer, outils de simulation

Dé cubique : Lien entre numéro des faces opposées.

Carré, cube, faces, arêtes.

Egalité de longueurs et de surface.

Triangles, carrés, pentagones,
hexagone, polygones, polygones
réguliers.

Solides de Platon (noms ?).

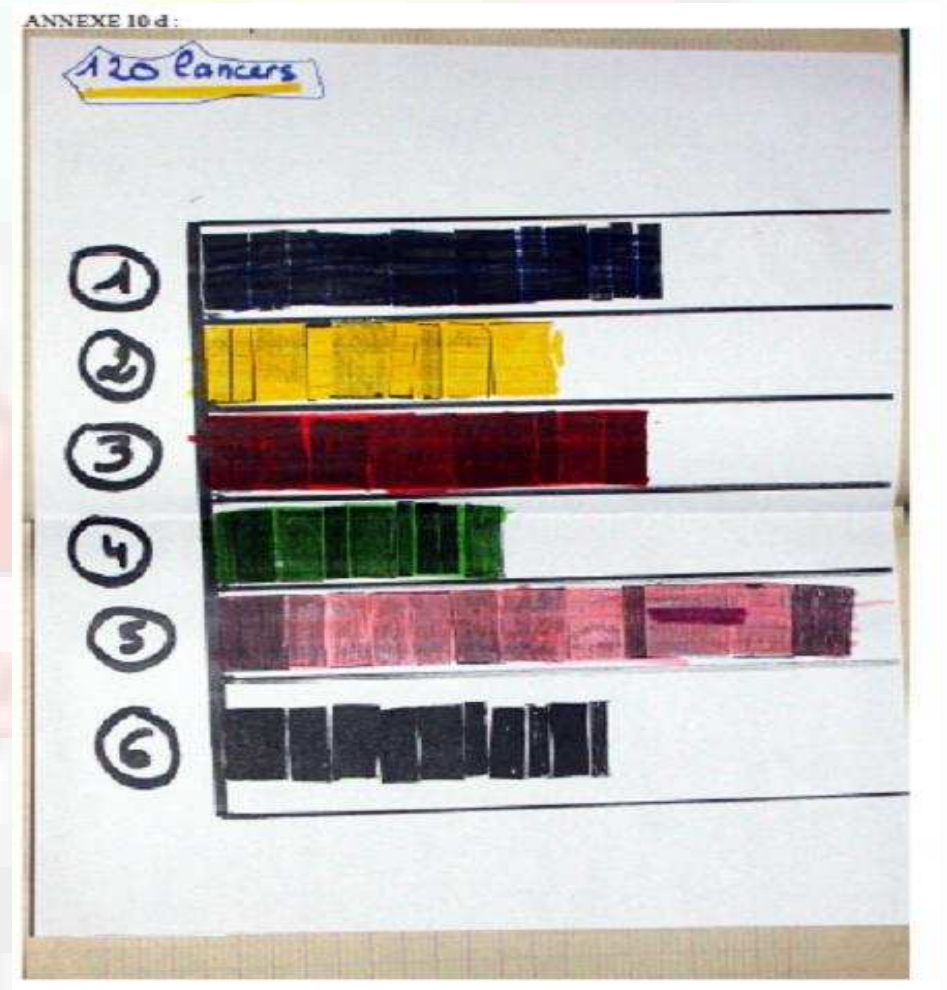
Formule d'Euler :

$$s + f = a + 2$$



Objectifs pour l'école

- Observer, expérimenter, parler écrire, mutualiser.
- Chances égales/hasard équitable.
- Le dé n'a pas de mémoire (indépendance des expériences).
- La chance aux dés n'existe que pendant un temps court.
- Histogrammes (voir illustration).



Objectifs pour le collège

(à répartir sur les disciplines)

- Calculer et comparer des pourcentages.
- Représentations graphiques.
- Moyenne, médiane, déciles, max, min, dispersion.
- Simuler.
- Langage des probabilités.

- Grand, petit, très variable, peu variable : relatif.
- Observation de la stabilisation des fréquences.
- Protocole expérimental.

Objectifs pour le lycée

* Probabilité : Lois de proba., variables aléatoires, espérance, dépendance, conditionnement, Loi des grands nombres (vulgarisation)

Processus (chaînes de Markov, processus de Poisson, mouvement brownien) - dans un avenir prochain ?

* Statistique : - modélisation ;
- prévision (ex : naissances dans une population) ;
- estimation (ex: sondages aléatoires) ;

2- Du côté des professeurs :
Savoirs nécessaires à un
professeur pour enseigner au
collège en 2008-2009

Du côté du professeur

- **Probabilité**

- *Loi des grands nombres ;*
- *Théorème central limite ;*
- *Mémoire et indépendance ;*
- *Hasard et information.*

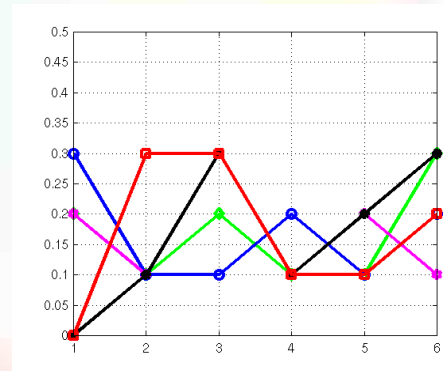
- **En statistique**

- *Les chiffres parlent-ils d'eux-mêmes ?*
- *La statistique en tant que construction sociale ;*
- *Preuve statistique.*

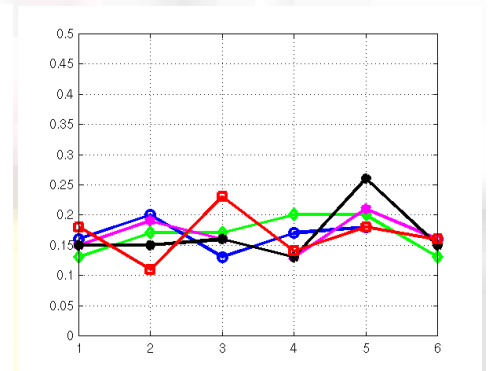
- **Histoire en probabilité et statistique**

Loi des grands nombres

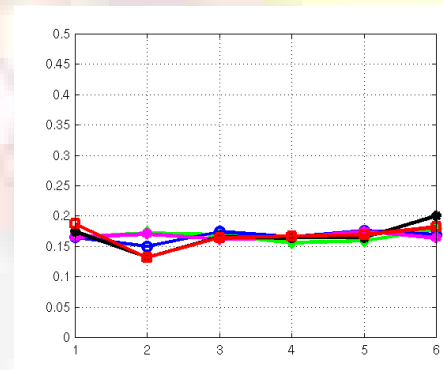
- *Les fréquences convergent vers les probabilités correspondantes.*
- Loi empirique ou théorème ?
- A propos, un dé équilibré une pièce équilibrée, c'est quoi ?



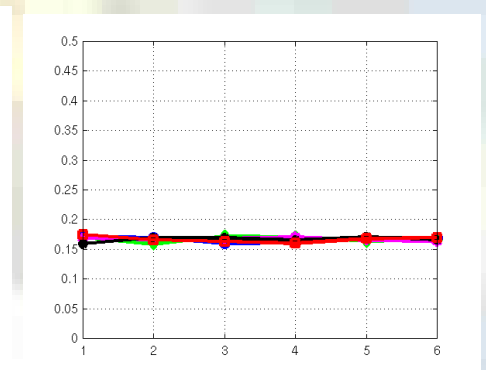
$n = 10$



$n = 100$



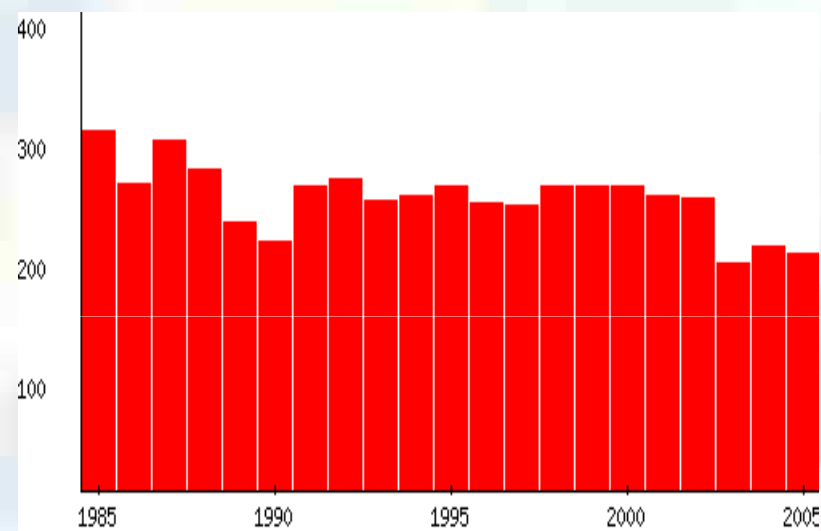
$n = 1000$



$n = 10000$

Et la liberté individuelle ?

Le prénom Henri



Année	Garçons
1995	373 409
1996	377 003
1997	373 157
1998	378 075
1999	382 132
2000	397 352
2001	394 297
2002	389 981
2003	388 780
2004	393 477
2005	396 680

Loi empirique des grands nombres et intuition

Pile -face :

on sait dire qu'il y aura environ 50% de *pile* pour N lancers, mais on ne sait pas quand.

Théorème central limite



On veut étudier les effets d'une balance pour des fruits et légumes qui arrondit automatiquement le prix de chaque vente comme indiqué dans le tableau 1.

centimes	0	1	2	3	4	5	6	7	8	9
arrondi	0	0	0	5	5	5	5	5	10	10
gain	0	-1	-2	+2	1	0	-1	-2	+2	+1

- Si prix moyen = 2 euros, le gain ou la perte sont d'au plus 1% par achat.
- Pour 20 000 pesées (un trimestre pour un commerce qui fait en moyenne 200 pesées par jour), le gain ou la perte cumulés seront sûrement dans l'intervalle $[-400, 400]$.
- Est-il rare de perdre ou gagner par exemple 100 ou 200 € ?

Une hypothèse : gains équiprobables

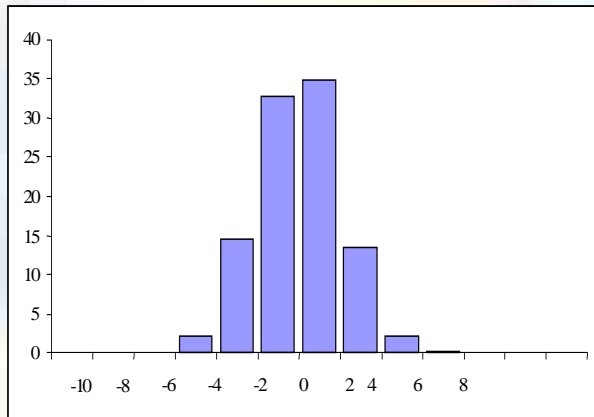


Figure 1

Gain cumulé	Nombre de trimestres
[-400 ; -8]	0
] -8 ; -6]	0
] -6 ; -4]	21
] -4 ; -2]	146
] -2 ; 0]	328
] 0 ; 2]	348
] 2 ; 4]	134
] 4 ; 6]	20
] 6 ; 8]	3
] 8 ; 400]	0

Tableau 2



$$P(-0,04\sqrt{2n} \leq S_n \leq 0,04\sqrt{2n}) \approx 0,999$$

Nombre de pesées	Intervalle de probabilité 1	Intervalle de dispersion
100	[-2,+2]	[-1 ;1]
1000	[-20 ; 20]	[-2 ; 2]
5000	[-100 ; 100]	[-4 ; 4]
10 000	[-200 ; 200]	[-6 ; 6]
20 000	[-400 ; 400]	[-8 ; 8]

Intervalles de dispersions du gain au niveau de confiance 0,999. Les calculs sont arrondis à l'entier supérieur.

Parité homme femme

L'exception française 1993 :
6% de députées

Lois :

- 2000
- 2007

N	N_F
100	43
2 500	1150 (46%)

Qu'en déduire sur le respect de la parité pour ces trois entreprises ?

Parité Homme- Femme

Parité : le sexe n'intervient pas dans le recrutement des N personnes.

Si la population de référence comporte 50% de femmes,

Alors le nombre N_F de femmes suit la loi Binomiale($N, 1/2$)

N	IO,95	$100N_F/N$
100	[40-60]	43%
2 500	[48-52]	46 %

$\text{Prob} (0,5 - 1/\sqrt{N} \leq N_f/N \leq 0,5 + 1/\sqrt{N}) \approx 0,95$

On a 95 chances sur 100 que :

- la proportion de femmes s'écarte de $1/2$ d'au plus $1/\sqrt{N}$.
- le pourcentage de femmes s'écarte de 50 d'au plus $100/\sqrt{N}$.

Théorème central limite

N grand

f_N a une proba. 0,95 d'être dans $[p - 2/\sqrt{N}, p + 2/\sqrt{N}]$

Un *grand* nombre :

pour un statisticien, 30 est souvent un grand nombre

Sondage :

- Proba. p a 95 chances sur 100 d'être dans

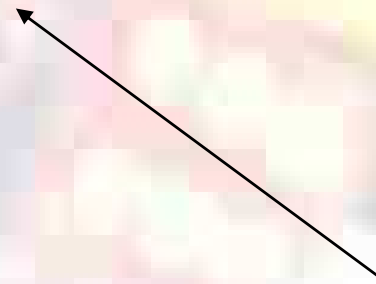
$$[f_N - 2/\sqrt{N}, f_N + 2/\sqrt{N}]$$

Calculer des pourcentages
sur des populations de taille très différente
ne permet pas de comparaison directe.
Toujours donner le nombre de données
ayant servi au calcul de pourcentages

$$\frac{\Delta N}{N}$$

ou

$$\frac{\Delta N}{\sqrt{N}}$$



Théorème central limite

Le cas de la physique

- Physique microscopique : **aléatoire**.
- Physique macroscopique : $N=6 \times 10^{23}$
déterministe.
- *exemple : proba p qu'une particule ait une propriété donnée. Indépendance des particules.
Au niveau d'une mole, variation de l'ordre de 10^{12}
inobservables
 \Rightarrow modèles déterministes.*

- *Microscopique-macroscopique, individu-population.*
- *Population : prévision.*

« Les statistiques sont vraies sont vraies quant aux populations et fausses quant à l'individu. »

Léon Schwartzberg

Qu'est-ce qu'un individu tiré au hasard ?

Statistique :
des chiffres éloquentes ?
Pays les plus pollueurs (2000)

- « *Le citoyen des Etats-Unis est le premier émetteur au Monde de gaz à effet de serre : 20 tonnes de CO2 par an, contre 10 pour un Allemand et 2,30 pour un Chinois. »*
- Les chiffres sont éloquentes! On peut ainsi ranger ces trois pays du plus pollueur au moins pollueur :
USA, Allemagne, Chine.

- « *La Chine produit 3,54 tonnes de CO2 pour la production d'un certain revenu (l'équivalent de 90\$ de PNB) tandis que l'Allemagne ne produit que 0,46 tonnes de CO2 pour la même production de revenu et les Etats-Unis 0,77 tonnes. »*
- Les chiffres sont éloquentes ! On peut ainsi ranger ces trois pays du plus pollueur au moins pollueur :
Chine, USA, Allemagne.

Les chiffres parlent d'eux-mêmes...

	1951-1955	1989-1993
Ecole polytechnique	21	7,8
ENA(Ecole Nationale de l'Administration)	18,3	6,1
ENS(Ecoles Normales Supérieures)	23,9	6,1
HEC(Hautes écoles commerciales)	38,2	11,8

	1951-1955	1989-1993
Ecole polytechnique	21	7,8
ENA(Ecole Nationale de l'Administration)	18,3	6,1
ENS(Ecoles Normales Supérieures)	23,9	6,1
HEC (Hautes écoles commerciales)	38,2	11,8
Grandes écoles	29	8,6
Dans population des 20-24ans	90,8	68,2

Avant d'aller plus loin, remarquons, à propos de l'ascenseur social, que l'entrée à une « grande école » n'est qu'une des voies de réussite possibles pour les jeunes. Cette réussite est bien identifiée pour la tranche d'âge considéré alors qu'à cet âge là, il n'y a pas encore de médecins, de chercheurs ou de chefs d'entreprises renommés, peu d'écrivains ou d'artistes connus

Un indice pour mesurer l'ascenseur social dans les grandes écoles

- Rapport des probabilités d'entrée entre un jeune d'origine sociale favorisée et un jeune d'origine sociale populaire.
- *Un jeune issu des classes favorisées a ainsi, en 1951-1955, 37 fois plus de chances de rentrer à l'école polytechnique qu'un jeune issu des classes populaires.*

	1951-1955	1989-1993
Ecole polytechnique	37	25
ENA	44	33
ENS	31,4	33
HEC	16	16
Grandes écoles	24	23

Classes sociales

- Revenons maintenant à un autre aspect des données du tableau 1 : la notion de classe populaire ou favorisée ne repose pas sur un critère intrinsèquement dichotomique (comme homme, femme)

	<i>A</i>	<i>B</i>	<i>C</i>
Nombre de reçus	8	10	50
Effectif population	100	100	100

Dans la période 1, le groupe défavorisé est composé de *A* et *B*.
Dans la période 2, le groupe *B* a changé de statut et est passé dans *F*.

Période 1 : $D=A \cup B$ et $F=C$
 $\tau = 0,5/0,09=5,5$

Période 2 : $D=A$ et $F= B \cup C$
 $\tau' = 0,3/0,08=3,75$



La statistique construite Mortalité en Inde et en France

En 2007

France : 8,6 décès pour 1000 habitants

Inde : 8,1 décès pour 1000 habitants.

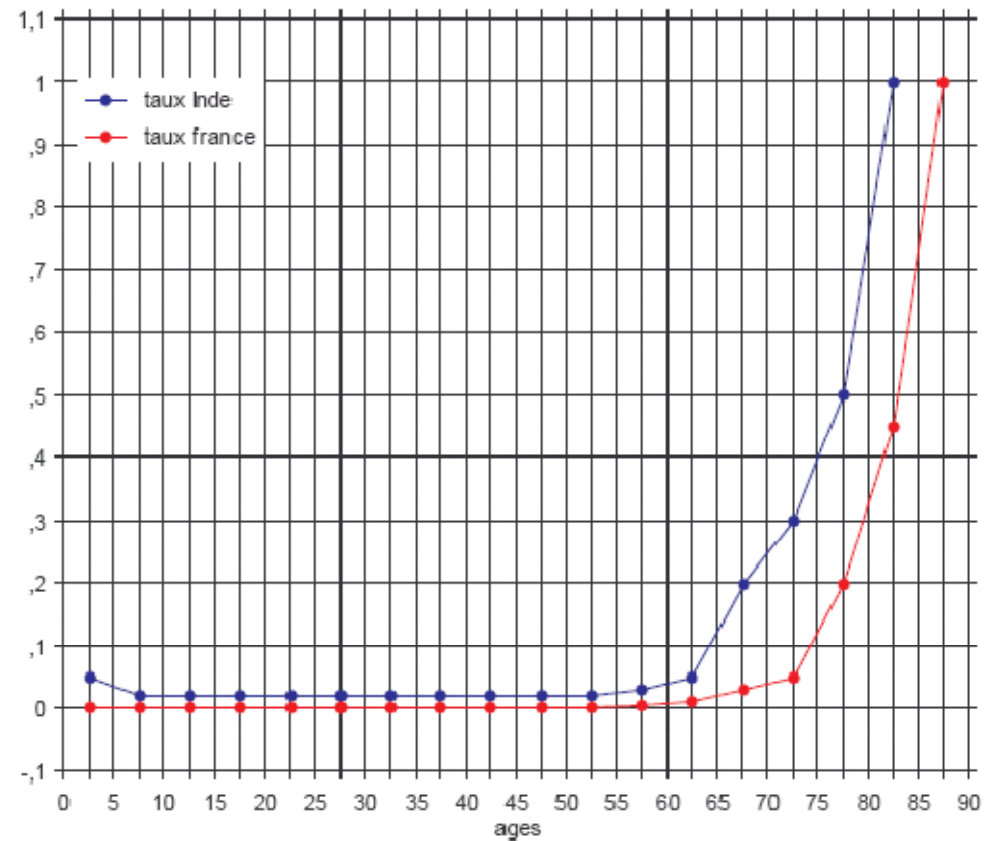
**En 2007, on meurt plus en
Inde qu'en France.**

De plus près :

Taux de mortalité
annuelle (quotient de
mortalité) $\tau_i, i=1 \dots n$

$\tau_i = \text{nbre de décès de}$
gens d'âge i en janvier/
Effectif N_i dans la
tranche d'âge i

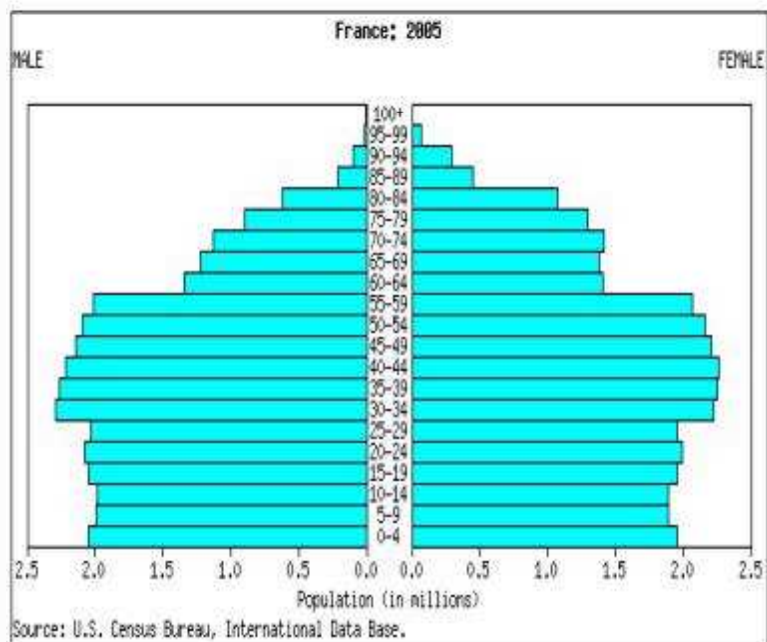
$(\tau_1 \dots \tau_{max})$: **mortalité**



Pyramides des âges

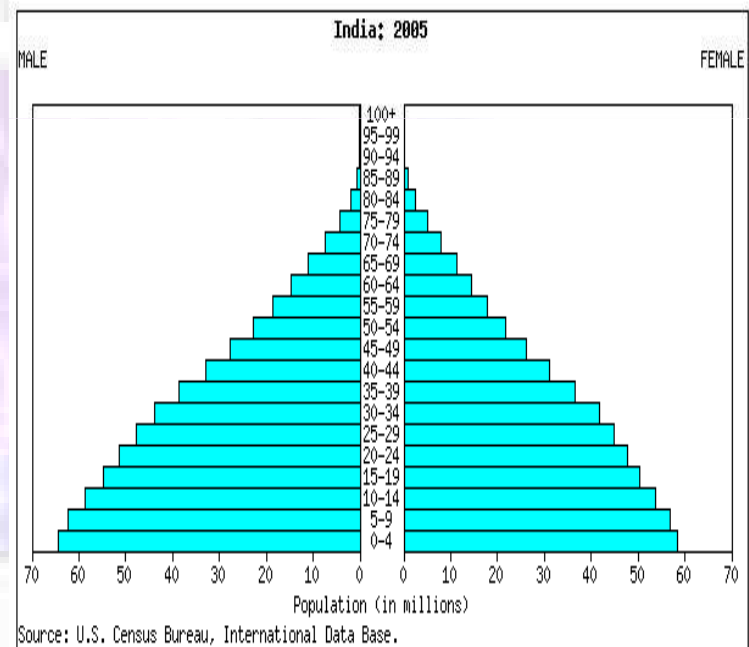
U.S. Census Bureau

Population Pyramids for France



U.S. Census Bureau

Population Pyramids for India



Nombre de décès pour 1000 habitants

$(N_1 \dots N_{max})$: pyramide des âges France

$$**$D_{France} = 1000 \sum \tau_i (N_i / N_{total})$**$$

$(N'_1 \dots N'_{max})$: pyramide des âges Inde

$$**$D_{Inde} = 1000 \sum v_i (N'_i / N'_{total})$**$$

Taux standardisés par l'âge WHO (OMS)

<http://www.who.int/whosis/indicators/2007MortAgeStandardized/en/index.html>

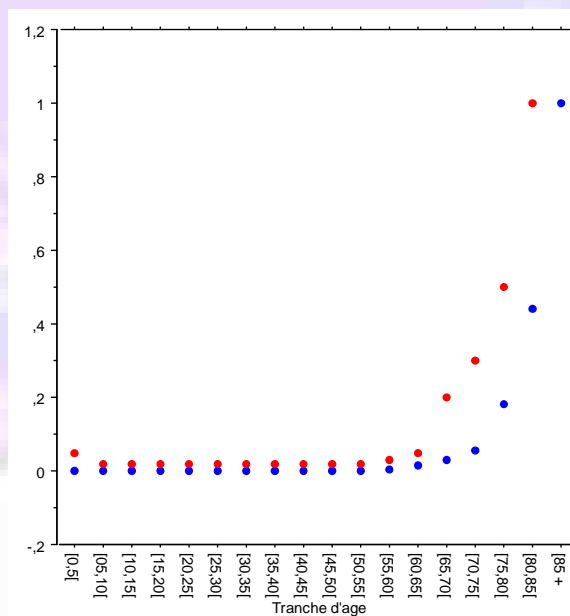
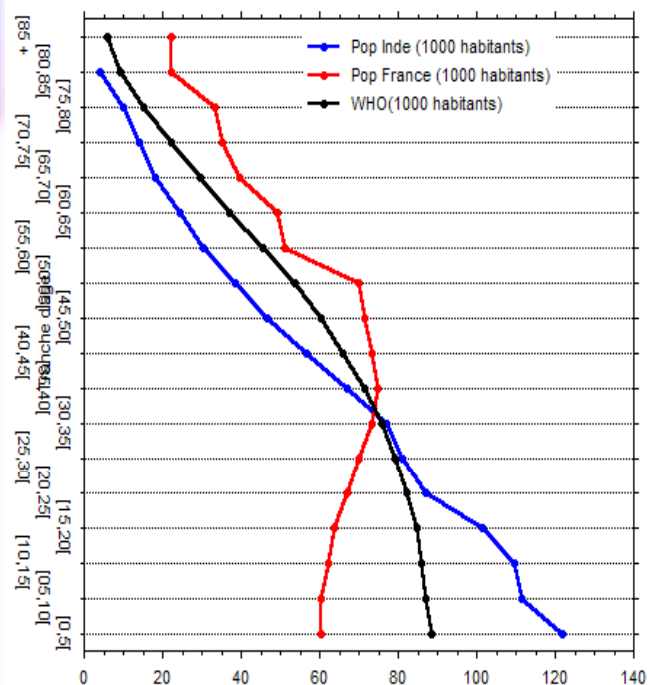
Table 4. WHO World Standard Population Distribution (%), based on world average population between 2000-2025

Age group	World Average 2000-2025
0-4	8.86
5-9	8.69
10-14	8.60
15-19	8.47
20-24	8.22
25-29	7.93
30-34	7.61
35-39	7.15
40-44	6.59
45-49	6.04
50-54	5.37
55-59	4.55
60-64	3.72
65-69	2.96
70-74	2.21
75-79	1.52
80-84	0.91
85-89	0.44
90-94	0.15
95-99	0.04
100+	0.005
Total	100

WHO Standard Population

- It is based on the average world population structure for the period 2000-2025. Estimates from the United Nations 1998 assessment
- The use of an average world population removes the effects of historical events such as wars and famine.

Retour sur l'Inde et la France



Taux de mortalité standardisés (OMS, âge), pour 1000 habitants :

Inde : 10,4

France : 3,3

Histoire de la discipline

- Quelle partie de l'histoire des idées d'une discipline éclaire ce qu'elle est aujourd'hui ?

On admet que dans ce cas, cette partie peut aider à l'enseignement de cette discipline.

- Quelle partie de l'histoire d'une discipline est un obstacle à la compréhension de ce qu'elle est aujourd'hui ?

On admet que dans ce cas, cette partie peut constituer un obstacle à l'enseignement de cette discipline.

Pour l'enseignement d'une discipline, vaut-il mieux la développer plutôt dans le temps (son histoire) ou dans l'espace (ce qu'elle couvre aujourd'hui) ?

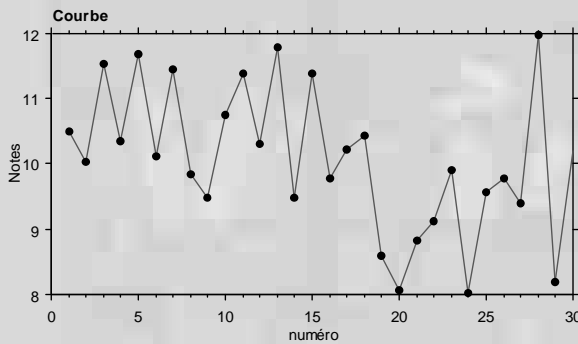
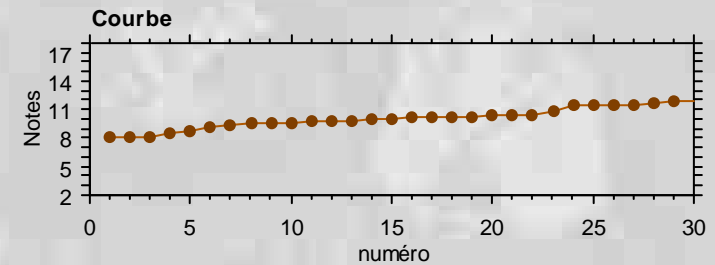
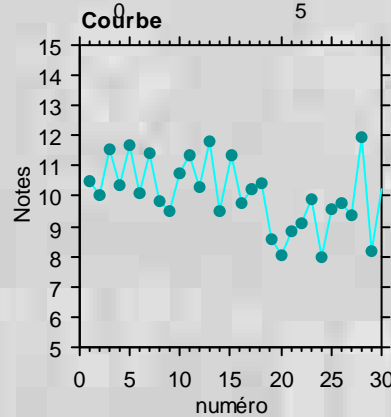
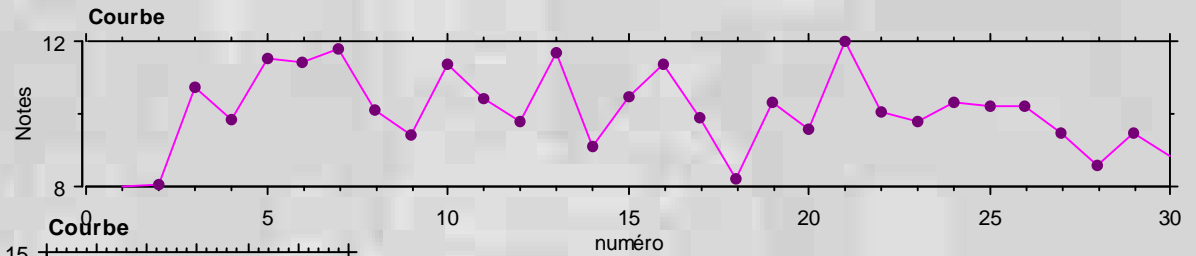
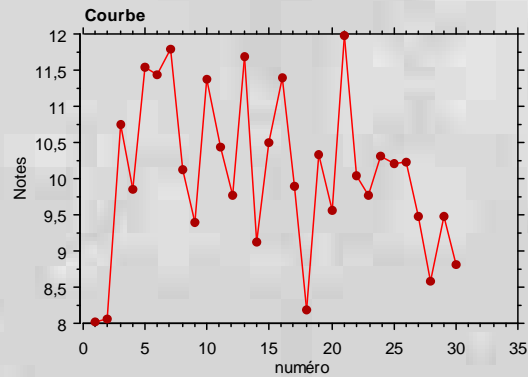
Mathématiques, physique, finance, économie

- **Second principe de la thermodynamique XIXsiècle.**
- **Mouvement brownien : 19ème, Einstein 1905.**
- **Hasard bénin, hasard sauvage.**
- **XV,XVI,et XVII siècle :émergence.**
- *Probabilis : digne d'approbation.*
- *Dualité : croyance/fréquence.*
- **XVIII-XIXsiècle : Laplace,Gauss...**
- **Kolmogorov : XXème siècle.**
- **On n'a plus de probabilité « épistémique », « personnelle », « subjective ». On a la théorie des probabilités et la statistique.**
- **Chaos déterministe .**
- ***Existence du hasard et modèle.***
- **Informatique.**

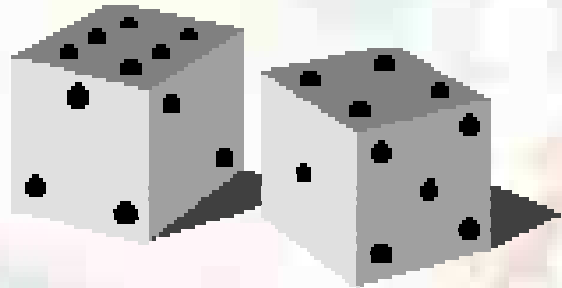
3- Les élèves et les professeurs : Retour sur l'enseignement en troisième

Collège : Représentations graphiques Grand petit, variable... (pas le même pour toute la classe)

A l'oeil nu, diriez-vous que certains graphiques donnent une impression de plus grande variabilité que d'autres ?
30 notes entre 0 et 20 données par des consommateurs à une lessive



Collège : du dé au dé électronique, un incontournable !



Lancer de deux dés

(le concevoir soi-même)

- stabilisation des fréquences
- non uniformité

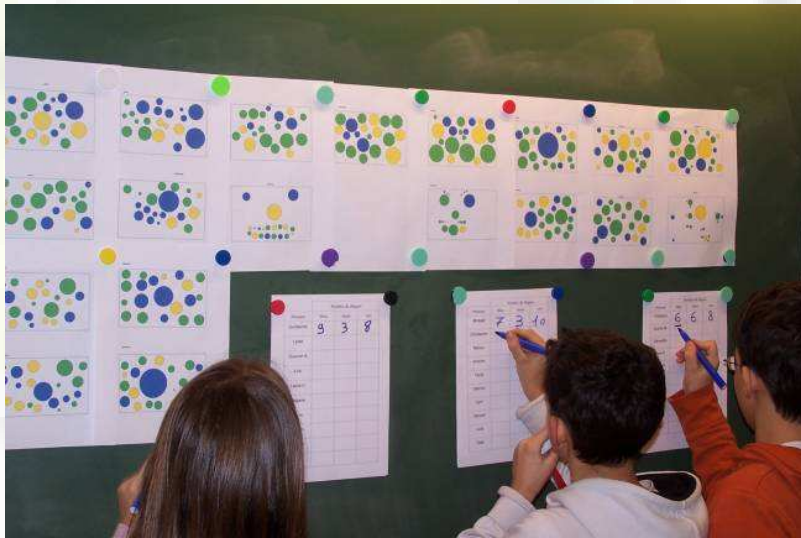
Somme de trois dés

*dés électroniques : un exemple
historique intéressant*

10 (6 possibilités)

9 (6 possibilités)

Collège :
Expérimenter et simuler
Somme de deux lois uniformes sur $\{0,1\}$.



- La forme est au choix
- Ici *Art et hasard*

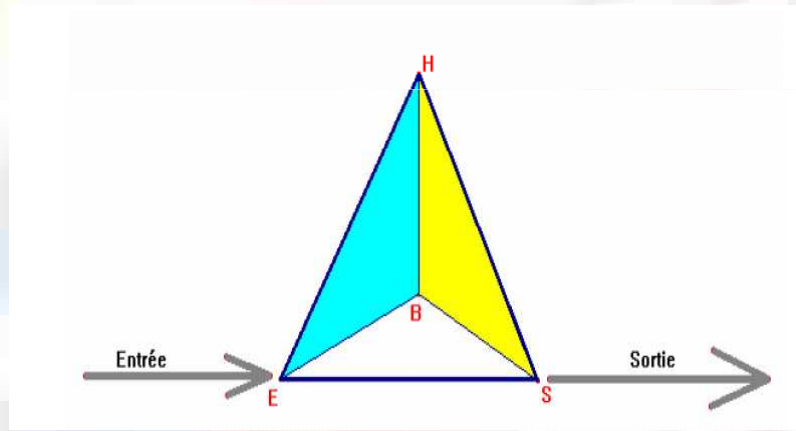
On travaille sur :

- 1 chance sur 2 et 3 chances sur 6 ;
- loi non uniforme ;
- « chance » est un mot polysémique , préférer « probabilité » ;
- stabilisation des fréquences.

Un jeu de l'oie

1 L'aire du cercle est πr^2	2 Augmenter un prix de 10%, puis diminuer de 10% ce nouveau prix le ramène au prix initial.	3 Si le carré n^2 d'un nombre n est pair, alors n est pair	4 Le litre est une unité de volume identique au dm^3
8 La somme de deux nombres consécutifs est un nombre impair	7 Si un nombre est divisible par 5 mais pas par 25, ce n'est pas un carré parfait.	6 Un décilitre est une unité de volume identique à 100cm^3	5 Si le carré n^2 d'un nombre n est divisible par deux, alors n^2 est divisible par 4
9 Deux rectangles ayant même périmètre ont même surface.	10 Le produit de deux nombres positifs n'est pas toujours plus grand que chacun d'eux	11 Les trois médiatrices d'un triangle sont concourantes	12 Euréka

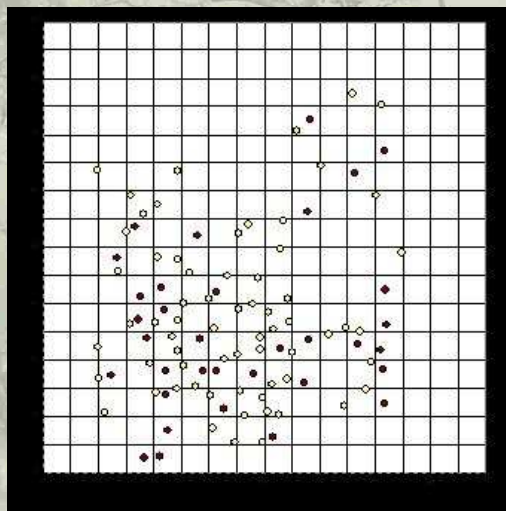
Piège à fourmis



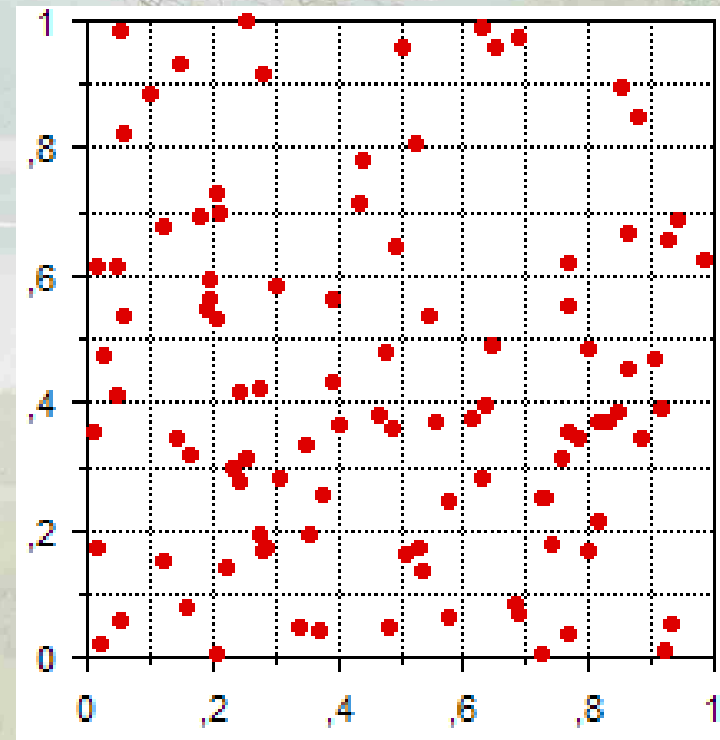
Protocole expérimental expériences

réelles

- Lancer une pièce de 5 centimes d'euros sur papier millimétré. Faire des lancers sans viser jusqu'à en avoir 100 sur la feuille.
- Question : dépend du lanceur, chances égales ??
- (ici $f=0,32$).



virtuelles

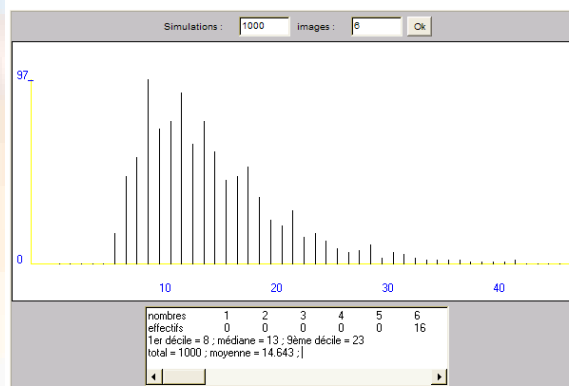


Collège

- **Elargir le champ des questions.**
- *Et encore passer des dés à la simulation sur calculatrice ou ordinateur.*
- *Cf. fiches de seconde*
- Une classe de 6 groupes. Le prof. choisit chaque jour au dé le groupe dont il relève les cahiers.
Au bout de combien de jours le prof aura-t-il vu au moins une fois chacun des cahiers ?

Nombre moyen de jours, nombre médian, histogramme

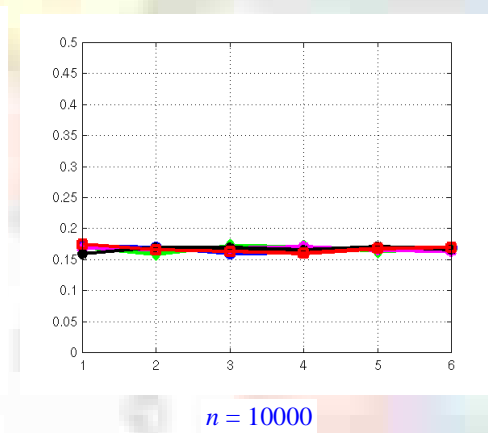
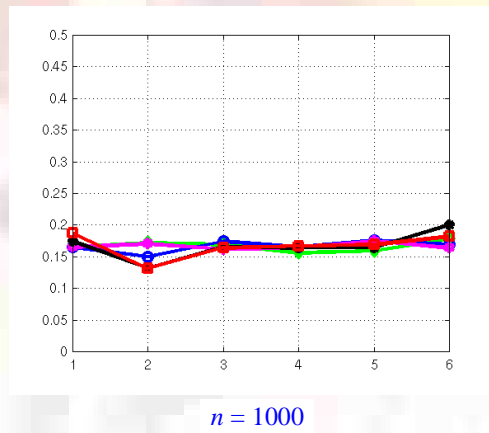
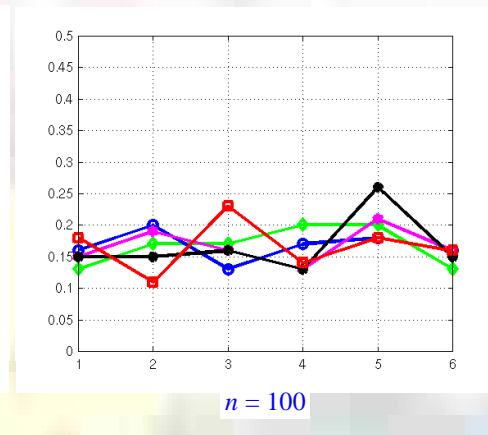
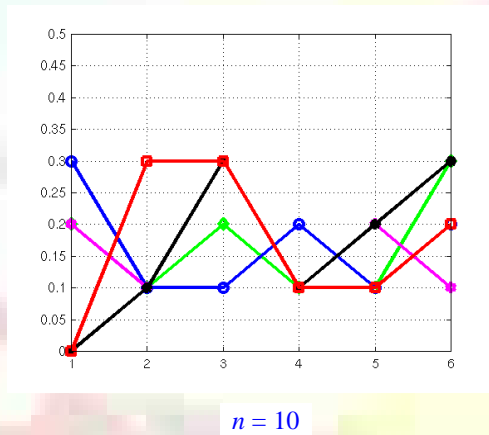
- Au bout de 10,15,20 jours quelles sont les chances d'avoir vu tous les cahiers ,
- Au bout de combien de jours a-t-on 99 chance sur 100 d'avoir vu tous les cahiers
- <http://pagesperso-orange.fr/jpq/proba/cereales/index.htm>
- *(pour le prof. savoir reconnaître la même situation)*



Collège

Langage des probabilités

- Laisser de coté le mot de chance.
- Proba = invariant (approche modèle).
- **Distribution de probabilité.**
- **Choix au hasard (random choice).**



La statistique...ça s'apprend

Diffuser des chiffres et les interpréter, c'est détenir une parcelle de pouvoir. La statistique, ayant envahi tout le champ social, appelle des citoyens mieux avertis, entraînés à exercer leur esprit critique sur des sujets variés, où la mesure quantifiée est inséparable de la pensée qualitative.

En l'absence d'une culture statistique largement partagée, le champ est en effet libre pour toute interprétation erronée, intentionnellement ou non, des données chiffrées, qui vient nourrir la méfiance entre le corps social et les décideurs de tous niveaux.



Centre de ressources, lieu de partage et de mutualisation pour l'enseignement de la statistique. Pour qui ? Les enseignants des écoles, des collèges et des lycées de toutes les disciplines.

Lire : Les 7 principes du projet Statistix

- Nous écrire
- Plan du site
- Qui sommes-nous ?
- Guide
- Faq

Rechercher avec google Ok

- **Concours 2009**
- à l'école primaire
- Divers
- Economie-Démographie
- Mathématiques
- Physique-Chimie
- Sciences de la Vie et de la Terre
- Sur le web

RSS

Parrainage :



INSTITUT DE FRANCE
Académie des sciences

dimanche 5 octobre 2008

Le hasard fait-il n'importe quoi ?

Un expérimentation en mathématiques, classe de seconde

Quelques thèmes peuvent suffire à traiter une grande partie du programme de statistique de la classe de seconde. [\[lire la suite...\]](#)

jeudi 2 octobre 2008

Série noire dans l'aviation

En août 2005, on a dénombré cinq accidents aériens graves en l'espace de 22 jours : est-ce exceptionnel ? [\[lire la suite...\]](#)

vendredi 19 septembre 2008

Concours Statistix 2009

Trois concours sont organisés par Statistix, le premier à la rentrée et Sciences et technologie au...

Concours Statistix 2009

Un concours pour les professeurs des établissements secondaires, publics ou privés sous contrats (collèges, LGT,LP) portant sur une activité en classe avec mise en oeuvre d'une démarche statistique.

Un concours pour les professeurs des écoles publiques ou privées sous contrat, portant sur le hasard.

Un concours pour les professeurs de français des établissements secondaires, publics ou privés sous contrats, et leurs élèves : production de textes sur des coïncidences.

Il y a des ordinateurs à gagner !

