**« à propos d’une expérience de G. DERAMECOURT sur des paris d’élèves de CE1 » (1972) Analyse de la variance**

**Guy Brousseau (2015)**

Evocation d’une Note de Travail simple et incomplète : Pourquoi ?

Il est difficile de comprendre aujourd’hui comment fonctionnait le COREM dans les années 70. Les tâches d’enseignement, de recherche expérimentale et théorique sur l’enseignement et de la formation des chercheurs devaient s’imbriquer et se compléter pour déterminer à la fois ce qui pouvait et devait s’accomplir à un instant précis, et ce qui devait rester indécise dans l’attente de réalisations à venir : une sorte de méthode PERT dynamique, où non seulement les travaux mais aussi leur conception dépendait d’évènements à venir. Comment de petits groupes se chargent d’une étape précise dans un vaste projet visant plusieurs objectifs interdépendants et complémentaires. Chaque document était nécessaire, immédiatement, mais il ne pouvait pas souvent être achevé et publié rapidement car il devait attendre que d’autres travaux viennent le préciser, le compléter ou l’infirmer. Les gestions des diverses tâches du COREM, présentaient de ce point de vue des similitudes étonnantes.

La note de travail présentée ci-après ne présente qu’un fragment sans grand intérêt d’une recherche inaboutie, transformé en exercice de statistique… mais de ce fait même il illustre bien ce fonctionnement :

Dans le cadre du plan de recherches sur « l’enseignement des probabilités » (1971-72) évoqué dans plusieurs dossiers sur ce site, il était intéressant de savoir si - et si oui, à partir de quel âge - les élèves pouvaient manifester spontanément (au moins implicitement) quelque sensibilité aux tendances centrales et à la dispersion d’une statistique. Ainsi dans l’expérience des machines éducables, les élèves manifestent l’apparition d’une connaissance implicite des probabilités en modifiant le contenu d’une urne pour favoriser la fréquence d’un évènement.

Dans la même perspective, Gérard Deramecourt a conçu l’expérience, beaucoup plus facile et puissante, décrite ci-après. L’agencement subtil d’un jeu de hasard original pourrait provoquer et/ou révéler l’existence et le développement de connaissances implicites à ce propos. Il réalisa l’expérience décrite plus précisément ci-après dans une classe de CE1 à Périgueux. Nous avons hélas perdu de vue le rapport qu’il rédigea à cette occasion à l’intention de l’IPN.

A cette époque, avant que les grandes analyses factorielles n’arrivent à notre portée, nous programmions, implantions et réimplantions toutes sortes de programmes de calculs statistiques au fur et à mesure de l’acquisition de machines nouvelles : analyse de la variance, Statistiques non paramétriques, régression polynomiale… dans le but de pouvoir les appliquer à toutes les données que nous recueillions. Nous[[1]](#footnote-2) avons donc réalisé une analyse de la variance des données de Périgueux et rédigé une note de travail pour de futures utilisations (1973). Cette fiche a été publiée dans un cahier de l’IREM (n°13) sans doute pour mettre la méthode à la disposition d’autres utilisateurs. Elle a été intégrée plus tard dans le « cours de Statistique du DEA » (1977). Hélas les données que nous avons utilisées ont aujourd’hui disparu, perdues de vue au cours d’un des déménagements successifs de nos bureaux. Elles n’avaient pas été tapées, ni imprimées dans le cahier, ni enregistrées sur un autre support…

Nous publions aujourd’hui ce fragment archéologique incomplet seulement pour témoigner de l’organisation et de la vraie nature de nos activités, de nos difficultés et de l’ampleur de nos ambitions véritables.

Note de Travail (1972-73, complétée en 2015)

**1. - L’EXPERIENCE, LES DONNEES ET LES FACTEURS ETUDIES**

* 1. – CIRCONSTANCES DE L'EXPERIENCE

L’expérience conçue par Deramecourt porte sur la perception des probabilités. Elle a été réalisée dans un CE1, (séance du 4 juin 1973) à Périgueux. Elle avait pour but de déterminer si l’observation et la prévision des résultats d’une suite de tirages (avec remise) fait évoluer spontanément les prévisions des enfants. Les distributions des paris et leur évolution pourraient manifester par leurs variations la perception et la sensibilité progressive des élèves

- d’une part à une tendance centrale (fréquence ou probabilité)

- et d’autre part à une dispersion des résultats.

1.2 – 1ere EXPERIENCE : Présentation

L’expérience est composée de 12 parties. Chaque partie consiste à prévoir le résultat d’une expérience de 50 tirages avec remise dans une urne de composition inconnue. Les élèves de la classe sont répartis dans 5 équipes concurrentes ( I, II, ...V) .

Avant chaque partie, chaque équipe fait 10 paris sur le nombre (à venir) d’apparitions d’une boule noire au cours des 50 tirages : ils s’accordent ainsi pour écrire 10 nombres entre 0 et 50 : un pari dans chacune des 10 cases d’une bande de papier qui constitue ainsi le pari de l’équipe. Après les 50 tirages d’une même partie, chaque équipe augmente son score du nombre de paris gagnants figurant sur sa bande. Ils préparent dix nouveaux paris pour la partie suivante. Le relevé d’observation des échanges verbaux des élèves ont été sans doute rapportés dans un autre cadre (Institut pédagogique national) et n’est pas parvenu ou n’a pas été conservé à l’IREM de Bordeaux.

1.3 - LES OBSERVATIONS

Les paris de chacune des 5 équipes forment une matrice de 10 colonnes et de 12 lignes (600 résultats). Il s’agit de savoir si cette matrice manifeste une évolution des paris et laquelle.

Il est tout à fait regrettable les 12 distributions des tirages effectifs réalisés au cours de l’expérience n’ait pas été confrontés aux prévisions des élèves.

La comparaison des distributions de paris à ceux des tirages devait figurer en annexe du rapport final de Deramecourt. Il aurait du servir après1975 pour l’obtention d’un DEA que Deramecourt ne brigua jamais.

1.4 LES MODELISATIONS

La *stratégie optimale* *objective*, pour ce jeu, consiste évidemment à parier dans chacune des cases sur une seule et même valeur : celle que l’on croit « la plus probable ». (25 dans le cas de l’équiprobabilité entre deux couleurs seulement). Mais si cette stratégie « jouer dans chaque case la valeur que l’on croit la plus probable » peut être conçue et justifiée par un calcul simple, elle n’a qu’une probabilité très faible d’apparaître dans les conditions de l’expérience à des élèves de 8/9 ans.

En effet, chaque essai dans cette stratégie ne rapporte aucune information sur les autres choix possibles : soit assez souvent zéro point et zéro information, soit rarement dix points et l’information que la valeur la plus probable pourrait être voisine de cette valeur sans plus). L’information retirée à chaque épreuve est nulle ou faible.

Les joueurs ne la connaissant pas, et ne pouvant pas la soupçonner a priori, ne peuvent tirer aucune information de leurs insuccès. Elle n’offre aucune place à une « démarche » mathématique « expérimentale ».

Au contraire en centrant correctement leurs paris autour de la valeur soupçonnée (la valeur centrale 25, ils obtiennent quelques points à chaque partie et peuvent resserrer leur « filet » autour de la moyenne. Cette deuxième stratégie est-elle bien celle qu’ils utilisent ?

Le calcul mathématique des propriétés des deux modèles est possible et c’est l’un des agréments de la recherche en Didactique des Mathématiques. Les modèles de situations y ont des propriétés significatives calculables.

1.5 L’OBJET DE L’ANALYSE

L’objet de la présente analyse est d’établir

* Si les caractéristiques des distributions de paris (moyenne, étendue,…) varient *significativement*
  + en fonction du groupe d’élèves ?
  + en fonction du rang de la partie dans la suite de 12
  + en fonction du groupe et du rang

D’autres questions comme

* la suite de ces distributions dénote-t-elle une évolution
* si ces évolutions se stabilisent ou « aboutissent »
* les choix correspondent-ils à un modèle « correct » ou « erroné »

L’observation directe des résultats a montré qu’aucun groupe n’est parvenu à la stratégie optimale qui consisterait après avoir perçu une valeur de probabilité maximale à miser cette valeur dans toutes les cases.

NB La confrontation de ces résultats avec les échanges verbaux des élèves dans leurs groupes (relevés par ailleurs) devait faire l’objet d’une autre étude.

**2- ETUDE DES FACTEURS "RANG" ET "GROUPE"**

* 1. - Les questions :

1. Est-ce que les paris faits par les enfants (la kième estimation d’un pari) évoluent significativement sous l'effet des expériences successives (i : rang de l'expérience)
2. Est-ce que les paris et leur évolution sont significativement différents suivant les groupes auxquels appartiennent les enfants ? (j : facteur groupe).

c) Est-ce que leur évolution est différente suivant les groupes ? (interaction rang - groupe).

**2.2 – TEST**

*2.2.1 –* **Principe de l’analyse de variance,**

Le chapitre 6 de l’ouvrage de Viallet cité en référence contient un exposé détaillé et très élémentaire des calculs de la variance en une, deux, ou trois dimensions.

Dans notre exemple, la partition de X en 10 colonnes, selon le rang de la case où est inscrite une prévision[[2]](#footnote-3), crée 10 classes équipotentes : case1, de la case 2,…case 10 (des 12 jeux et des 5 équipes confondus) : X1 sont les premières prévisions, X2 les secondes, … X10 les dernières (de chaque prévision). (NB : les effectifs de chaque classe sont les mêmes : nombre total d’élèves multiplié par le nombre de parties).

Notations : Chaque groupe « j » d’élèves, effectue 10 prévisions par partie « i ».

Chaque « pari »xijk du groupe i est identifiée par k, son rang dans sa ligne j. 1 k  10

Le facteur i « rang de l’estimation » présente 12 modalités 1 i  12

Le facteur j « groupe d’élèves» présente 5 modalités 1  j  5

*2.2.2* - **Modèle**

xijk = m + ai + bj + cij + ij k

où  ai = 0 bj = 0 c ij = 0 \_ c ij = 0 et

où les ij k sont des variables aléatoires indépendantes, obéissant à une loi de Laplace Gauss

Rappel : La variation totale est la somme des variations selon ces trois facteurs

***2.2.3* – Hypothèses nulles**

Aucun des deux facteurs (facteur rang, facteur groupe) qui nous intéressent n'a une contribution moyenne à la variance totale, « significativement » plus forte que la contribution moyenne sur l’ensemble des paris.

***2.2.3* - Tableau des résultats**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Sources de variation | Somme des  carrés | Degrés des  Liberté | Carrés moyens | F | S |
| Rang de la partie  Groupe d'élèves  Interaction  Résidu  Total | 1 828  2 622  3 799  22 992  31 241 | 11  4  44  540  599 | 166,18  655,50  86,34  42,57  52,15 | Fa = 3,90  F.01 = 2,74  (8 ; 80)  Fb = 15,39  F.01 = 3,56  (4 ; 80)  Fab = 2,02  F01 = 2,03  (24 ; 80) | S.01  TS.01  S.02 |

Les trois valeurs de F sont trop élevées : extérieures à l'intervalle de 0,01. Nous devons rejeter les trois hypothèses nulles.

**2.3 - CONCLUSIONS**

**1 -** Les paris faits par les enfants varient sensiblement au cours des expériences

**2 -** Les paris diffèrent d'un groupe à l'autre (on ne peut pas les regrouper)

**3 -** Les évolutions ne sont pas les mêmes suivants les groupes (il y a une interaction des deux facteurs).

**3 - LES ENFANTS ARRIVENT ILS A PARIER CORRECTEMENT ?**

L'analyse doit donc être faite par équipe et par rang

**3.1 - PRINCIPE**

**3.1.1 -** Nous allons tester l'hypothèse nulle que les paris des élèves, d'un groupe donné à un rang donné, peuvent être considérés comme un échantillon pris au hasard dans un ensemble parent qui suit la loi de l'expérience qu'ils regardent.

**3.1.2 -** Loi suivie par les tirages de boules dans l’urne : chaque expérience est composée de la répétition de 50 épreuves de Bernoulli, indépendantes, dont les deux évènements sont équiprobables, de probabilité p = q = ½.

La variable aléatoire x = nombre de boules noires sortant à chaque expérience suit une loi binominale B (½, 50) :

La probabilité de {tirer p boules d’une couleur en 50 épreuves successives} est :

proba {X = p} = (

E (X) = 25 o2 = mq (1 - q) = 12,5

X

Cette loi constitue le modèle que se donne l'observateur pour rendre compte de l'expérience.

**3.1.3 -** Les paris sont la manifestation du modèle d'un élève hypothétique (qui ferait autant de paris que le groupe).

Pour tester l'écart entre ces 2 modèles, (modèle de l'observateur et modèle de l'élève) nous calculerons pour chaque case du tableau de données

t = | m - m | où m = 25 o = \_12,5 = 3,53

o /  m est la moyenne de l'échantillon et N sa taille.

| t | suit la loi de Student à y = N - 1 degré de liberté.

**3.2 - RESULTATS**

**3.2.1.** Le tableau (I) donne pour chaque case, c'est-à-dire pour chaque équipe et à chaque rang de l'expérience, la moyenne et l'erreur type des paris :

m + 

N

De plus, nous indiquons si l'écart est significatif :

à 0,01 à l'aide d'un rond noire : "O"

à 0,05 à l'aide d'un rond blanc : "O"

**m + S**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| N° des paris | Equipe 1 | Equipe 2 | Equipe 3 | Equipe 4 | Equipe 5 |
| 1 | 34,1 + 1,42  • | 26,5 + 3,69 | 26,7 + 4,12 | 23 + 2,43 | 31,6 + 5,85  • |
| 2 | 29,6 + 0,80  • | 30,2 + 2,96  • | 31,9 + 4,28  • | 23,6 + 1,12 | 34,2 + 4,10  • |
| 3 | 30,7+1,58  • | 31,7 + 3,10  • | 33,3 + 3,33  • | 26,3 + 0,94 | 35,2 + 3,13  • |
| 4 | 30,3 + 1,55  • | 30,3 + 2,93  • | 29,7 + 3,32  • | 25,9 + 0,72 | 30,7 + 2,38  • |
| 5 | 32 + 1,57  • | 27,3 + 2,51 | 22,7 + 1,15 | 26,4 + 0,89 | 33,9 + 2,96  • |
| 6 | 28,6 + 0,65  o | 27,9 + 1,79  o | 23,7 + 2,34 | 25,8 + 0,85 | 40,5 + 4,54  • |
| 7 | 29 + 0,59  • | 25,7 + 1,66 | 24,4 + 0,99 | 24,1 + 0,76 | 33,4 + 2,94  • |
| 8 | 29,3 + 1,11  • | 25,8 + 0,71 | 24,9 + 0,92 | 25,7 +0,74 | 28,2 + 1,50  o |
| 9 | 27,2 + 0,53 | 24,7 + 1,05 | 24,5 + 0,95 | 27,7 + 0,42 | 30,3 + 1,55  • |
| 10 | 27,7 + 0,53  o | 25,2 + 0,84 | 25,1 + 0,82 | 27 + 0,25 | 28,4 + 1,23  o |
| 11 | 26,9 + 0,80 | 26,3 + 0,65 | 25,3 + 0,86 | 25,7 + 0,25 | 25,2 + 0,81 |
| 12 | 27,3 + 0,61 | 25,8 + 0,61 | 25,3 + 0,70 | 24 + 0,25 | 23,2 + 0,57 |

 . + n = 27,50   =  = 3,53

   =  = 3,16 / N = 1,11

**Tableau I**

**3.2.2 - Interprétation des résultats du tableau I (ci-dessus)**

**-** L'équipe IV a, dès le début, parié correctement pour la moyenne, et resséré constamment la dispersion de ses paris.

- Les équipes II et III, après 2 paris corrects au premier coup, ont suivi la tendance générale jusqu'au cinquième ou sixième coup, à partir duquel ils ont parié correctement pour la moyenne et resserré eux aussi leur dispersion.

- Les équipes I et V ont corrigé progressivement leur moyenne et réduit leur dispersion. Les paris sont corrects à partir de la 11ème partie.

**3.3 - CONCLUSIONS**

La position des valeurs significatives (ronds noirs et blancs sur le tableau) montre que dans les 4 équipes où se produit une évolution des paris, cette évolution marque un progrès vers des paris corrects. Les 5 groupes ont, pour les parties 11 et 12, un modèle de pari qui leur donne la même moyenne que celle du modèle de l'observateur.

**4 - CONCLUSIONS PROVISOIRES**

*Cette étude n'est pas complète ; les résultats obtenus par les enfants au cours de l'expérience réalisée en CE par G. DERAMECOURT devaient continuer à être dépouillés et analysés. L’ont-ils étés ?*

**4.1 – REPRÉSENTATION DES TIRAGES**

Les tirages effectués constituent un échantillon valable de la loi théorique (différence non significative au seuil de 0.1). Donc, au début de l'expérience, on ne peut pas dire que les enfants avaient un modèle implicite qui était celui de l'observateur.

**4.2 - ETUDE DE LA DISPERSION :**

Il apparaît, sur le tableau I, que les enfants dispersent leurs paris au début : (la valeur théorique de  est 1,11) et les resserrent sans cesse après, sans toute fois les refermer sur un seul nombre. Si ces écarts types s'éloignaient trop des valeurs théoriques, cela pourrait infirmer l'hypothèse d'indépendance des paris que nous avons dû accepter en faisant l'analyse de la variance. Toutefois, si l'effet de resserrement des paris est évident et général, la moyenne des dispersions paraît être voisine de 1,11.

**4.4 - NOUS NOUS PROPOSONS AUSSI D'ETUDIER** la manière dont les paris évoluent, dont ils sont corrigés de façon à en déduire des lois sur l'apprentissage, et, si possible, à un modèle d'apprentissage.

**Exemple** : - Les enfants encadrent-ils seulement le dernier coup sorti ?

- Ou y ajustent-ils leur distribution à la distribution des coups déjà tirés ?

N. ET G. BROUSSEAU (Août 1973)

Référence : François Viallet, Statistique et recherche appliquée, Chotard et associés Editeurs, 1970.

1. signé “Nadine et Guy Brousseau [↑](#footnote-ref-2)
2. Il n’y a pas de pertes ni de gains : On ne parle pas de paris mais de prévisions [↑](#footnote-ref-3)