



## **Lucien Le Cam (1924-2000)**

**Modernisateur de la statistique asymptotique**

*A renewer of asymptotic statistics*

**Paris, Institut Henri Poincaré**

**Amphithéâtre Hermite**

**19 novembre 2024 – *November 19th, 2024***

Colloque organisé par Laurent Mazliak (Sorbonne Université) et Catherine Vermandele (Université libre de Bruxelles) à l'initiative du groupe « Socio-histoire de la statistique et des probabilités » de la SFdS

## PROGRAMME/ PROGRAM

**9h-9h15 :** accueil par Jean-Jacques Droesbeke (président du groupe « Socio-histoire de la statistique et des probabilités » de la SFdS)

**9h15-10h :** Laurent MAZLIAK (Sorbonne Université, LPSM)

### Bref regard sur la préhistoire: l'implantation difficile de la statistique mathématique en France

Lucien Le Cam est né en 1924 et a donc commencé ses études supérieures à un moment où la statistique mathématique était tout juste stabilisée dans l'enseignement mathématique en France après des débuts compliqués. Je reviendrai sur la façon dont cette émergence a pris place au lendemain de la Première Guerre mondiale notamment sous l'action résolue d'Émile Borel, et de ses deux bras droits Maurice Fréchet et Georges Darmon, avec la création de l'ISUP en 1922 puis de l'IHP en 1928, et les nombreuses initiatives qui s'y sont développées. J'évoquerai notamment les conférences de l'IHP où furent invités des ténors mondiaux de la discipline et les doctorats de jeunes normaliens brillants consacrant l'entrée des statistiques dans le champ des mathématiques académiques.

**10h-10h45 :** Didier DACUNHA-CASTELLE (Paris-Saclay)

### Lucien Le Cam, la France de l'après-guerre et la statistique, puis ses séjours à Orsay

Lucien Le Cam a été pour moi d'abord un ami avec qui je partageais d'entrée deux choses, venir d'un milieu où l'on n'imaginait pas faire des études supérieures et d'avoir fait des maths par hasard.

J'évoquerai le cours de statistique fait par Lucien à l'IHP, ses relations avec Schwartz, sa réaction immédiate quand je lui ai demandé des conseils et de l'aide pour créer *ex nihilo* une équipe de statistique à Orsay au milieu des années 70. Enfin j'évoquerai son travail à EDF tel qu'il me l'avait raconté ainsi que les visites d'oncologues français et sa volonté de faire avancer avec sa modestie habituelle la recherche scientifique en insistant sur notre responsabilité de chercheurs. En conclusion, j'évoquerai l'évolution de la place de chercheurs très originaux comme Le Cam entre les années 50 et aujourd'hui.

**10h45-11h00 :** pause-café

**11h00- 11h45 :** Stephen M. STIGLER (University of Chicago)

### The History of Le Cam

A discussion of Le Cam's work on the history of statistics, and of my 37 years as student and correspondent.

**11h45- 12h30 :** Grace YANG (Maryland)

[Recount of my experience working with Professor Le Cam](#)

In the early 80s, I began a long-term collaboration with Dr. Charles E. Swenberg who was a biophysicist at the Armed Forces Radiology Research Institute. Two of our studies led to Le Cam's involvement. One problem was to develop dose-response curves for single yeast cells in radiation experiments. We used the Neyman-Puri model for our study. The second was a neurophysiological problem of estimating parameters with response data of cell membrane channels to a change in membrane voltage. Le Cam was my Ph.D. thesis advisor in the mid 60s. These problems restarted my work with him. I will discuss Le Cam's involvement. The first problem was in the domain of applied statistics in which Le Cam made fundamental contributions. The second problem came from our difficulties in applying Le Cam's estimation procedure in his LAN theory. We asked him what to do. With his enormous achievement and profound impact, I asked him "Shouldn't you be a member of the U.S. National Academy of Sciences(NAS) ?". He said that it is very difficult to be elected to NAS as a foreign member. After the retirement in 1991, Le Cam continued working and writing papers with unbelievable high speed. A few days before he passed away, I received his additions to the 2nd edition of our book "Asymptotics in Statistics:Some Basic Concepts".

**12h30-14h :** pause-déjeuner

**14h-14h45 :** Marc HALLIN (Université libre de Bruxelles)

[Le Cam, Hájek, and ranks](#)

Lorsque Jaroslav Hájek s'est rendu pour la première fois à Berkeley en 1961-1962, le concept de normalité asymptotique locale (LAN) n'avait pas encore atteint sa forme définitive. Bien que les premières définitions (sous la terminologie *Differential (Asymptotic) Normality*) remontent à l'article de Le Cam publié en 1960 dans *University of California Publications in Statistics*, leurs implications considérables, découlant de la faible convergence des expériences statistiques (Le Cam 1972, dans *Proceedings of the Sixth Berkeley Symposium*), n'étaient pas tout à fait claires. L'article de 1962 de Hájek dans les *Annals of Statistics* et sa monographie *Theory of Rank Tests* (publiée en 1967 mais écrite en 1963, en collaboration avec Zbynek Šidak) utilisent essentiellement la notion de contiguïté et le soi-disant troisième lemme de Le Cam, une terminologie inventée par Hájek lui-même. La lecture attentive du dernier chapitre permet toutefois de comprendre comment Hájek, en exploitant pleinement la théorie de Le Cam, aurait probablement développé sa théorie des tests de classement s'il n'était pas décédé prématurément en 1974.

**14h45-15h30 :** Dominique PICARD (Université Paris-Cité)

[Lucien Le Cam : distance entre expériences, normalité asymptotique locale](#)

C'est grâce à des mathématiciens comme Lucien Le Cam, que la statistique est devenue une branche à part entière des mathématiques. Bien que considéré comme l'un des statisticiens les plus importants du XX<sup>e</sup> siècle, Lucien Le Cam est pourtant peu connu en France, que ce soit du grand public ou même des mathématiciens. Nous aborderons dans cet exposé le travail de Lucien Le Cam et parmi ses abondants résultats, nous expliciterons certains des concepts majeurs sur lesquels il a travaillé, comme celui de « distance entre expériences statistiques » et de « normalité asymptotique locale ». On se concentrera plus particulièrement sur l'article de 1964 : « Sufficiency and approximate sufficiency » (Exhaustivité et exhaustivité approchée) paru dans les *Annals of Mathematical Statistics*. En effet, ce travail, profond et difficile mathématiquement, met en évidence une notion de proximité entre deux expériences statistiques en fonction de l'information qu'elles contiennent. Cette notion a eu des répercussions importantes et encore très actuelles si l'on pense à la notion de « privacy » (confidentialité), très étudiée ces dernières années, ou encore celle de « jumeau numérique ».

**15h30-16h00 :** pause-café

**16h00-16h45 :** Ester MARIUCCI (Université de Versailles Saint-Quentin)

**Comparaison des expériences statistiques et applications en statistique des processus**

Le mathématicien français Lucien Le Cam a imaginé une pseudo-distance, la distance  $\Delta$ , permettant de comparer deux modèles statistiques partageant le même espace de paramètres. L'idée derrière la définition d'une telle pseudo-distance est très naturelle. On dira que les modèles statistiques  $(P_{1;\theta})_{\theta \in \Theta}$  et  $(P_{2;\theta})_{\theta \in \Theta}$  sont équivalents au sens de Le Cam s'il existe un mécanisme qui ne dépend pas de  $\theta \in \Theta$ , permettant de reconstruire  $P_{1;\theta}$  à partir de  $P_{2;\theta}$  et vice-versa. Reformulé en termes de théorie de l'information, si la distance de Le Cam entre deux expériences est nulle, ces modèles contiennent la même information sur le paramètre inconnu. En plus d'avoir une interprétation claire en termes de théorie de l'information, la distance de Le Cam peut également être interprétée en termes de théorie de la décision, avec un lien clair avec la théorie minimax. Par exemple, une équivalence (asymptotique) est un outil puissant pour dériver des bornes inférieures (asymptotiques). Dans cet exposé, j'explorerai les implications que la théorie de Le Cam sur la comparaison d'expériences a eues en statistique non paramétrique et comment elle a été utilisée en statistique des processus stochastiques.